CESifo DICE REPORT

Journal for Institutional Comparisons

VOLUME 3, NO. 3 AUTUMN 2005

The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute for Economic Research, Munich

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
Annual subscription rate: €50.00
Editors: Wolfgang Ochel (ochel@ifo.de), Rigmar Osterkamp (osterkamp@ifo.de)
Editor of this issue: Rigmar Osterkamp
Copy editing: Anne Heritage, Paul Kremmel

DICE Database: www.cesifo.de/DICE

CESifo DICE Report

Volume 3, Number 3 AUTUMN 2005

Forum

CITY TOLL

Research Reports

Reform Models

Database

News

City Toll

Urban Congestion Pricing – Economists and Practitioners Join Hands

Tilmann Rave*

This issue of *DICE Report* was inspired by the renewed interest in congestion pricing, especially (but not exclusively) in Europe. Several scholars and practitioners were invited to take stock after several decades of research and advocacy for, as well as resistance against, congestion pricing, and to report from recent real-life experiences. While economists' pleas for congestion pricing have fallen on deaf ears for a long time (at least since the early 1960s), it is interesting to note that there now seems to be a shift in perspective with a merging of theoretical and practical reasoning. This change in perspective and some of its nuances can be witnessed throughout the articles assembled here.

Richard Arnott from Boston College, one of the world's leading urban and transport economists, opens the scene with his overview article "City Tolls – One Element of an Effective Policy Cocktail". He not only recalls the economic rationale for marginal cost-based pricing, but distinguishes several phases in the discussion which have slowly transformed the narrow views of both ardent supporters and opponents of road pricing. In contrast to some of his economist colleagues he gives a cautionary note saying that welfare gains put forward in the textbook version of auto congestion pricing may not be as large after all, given some real-life complexities (like other distortions in the economy) and implementation problems. He also encourages other economists to look beyond auto congestion pricing in the narrow sense and analyse some outside-the-box policies to alleviate urban traffic congestion (like parking policy or bicycling).

* Tilmann Rave is a researcher at the Ifo Institute for Economic Research, Munich. rave@ifo.de.

While Arnott calls for careful reasoning, the three following articles demonstrate that congestion pricing schemes can indeed be implemented and have a substantial impact on the transport system. Kian-*Keong Chin*, chief engineer in the Singapore Land Transport Authority, reports from one of the oldest and often cited experience with road pricing worldwide. Not only does he provide a detailed picture on how the Singapore scheme works on the ground, he also illustrates, covering a period of 30 years, how the scheme evolved over time giving rise to strategic and technological refinement and learning.

This insight seems to be interesting given the more recent experience with congestion pricing in the city of London, which is dealt with in the article from *Todd Litman*, a policy-oriented transport economist from the Canadian Victoria Transport Policy Institute. Despite some drawbacks in the way the scheme was designed (e.g. no time-variability) it has turned out to be more successful than originally predicted. Litman also highlights that London has sent out a political signal showing that congestion pricing is feasible and effective, and that it is possible to overcome the political and institutional resistance to such pricing.

Catharina Sikow-Magny and *Marcel Rommerts*, working for the European Commission, put this signal into a wider European perspective. They not only remind us that the European Commission has advocated the reform of transport pricing for over a decade, they also emphasize the importance of informal and formal networking to make these reform efforts successful. Informal networks, created in successive research projects, seem to have created rather stable policy communities which share similar values and interact continuously and creatively. Formal network activities under the EU Framework Programmes have established a kind of two-way process, with research results being fed into the policy implementation process and relevant policy questions being picked up by researchers.

Given this rather enlightening experience some readers may now feel more positive about the prospects for real-life congestion pricing. Others, howev-

er, may object that congestion pricing policy and technology are not standardised across cities, which prevents low-cost and easy-to-implement solutions. Thus, different policies and technologies may be appropriate for cities of different size and with different budgetary resources and planning capacities. *Ronnie Schöb*, Professor for Public Finance at the University of Magdeburg, draws on Arnott's plea for complementary policies beyond "pure" congestion pricing and develops a simple policy that might be particularly appropriate for medium-sized cities. The policy entails charging cars that enter the city-centre a toll equal to the public transit fare. Schöb's "multi-mode ticket" reduces the opportunity costs of travelling downtown by public transport and improves the revenue situation of public transport authorities.

Overall, this issue of DICE Report demonstrates that road pricing is high on the agenda of policy makers and practitioners, given some of the unresolved transport problems in densely populated urban areas. Yet, other than some economists implicitly assume there is no cure-all to urban transportation problems in general and traffic congestion in particular.

City Tolls – One Element of an Effective Policy Cocktail*

Richard Arnott**

For almost fifty years, following William Vickrey's lead (1959, 1963), urban transport economists have been advocating congestion pricing as *the* way of dealing with urban auto congestion.[1] Their advocacy was based on the standard analysis of the Pigouvian internalization of an externality. Drivers impose an external congestion cost on other drivers by slowing them down. If this congestion cost is internalized by means of a Pigouvian tax/congestion toll, then traffic congestion will be at the right level. Vickrey's original (1959) proposal was to equip each car with a transponder that would send a signal to roadside receptors, which would be relayed to a central computer. The computer would record the driver's movements, calculate the congestion externality cost imposed by him in his travels and bill him for this amount.

This congestion pricing scheme was criticized on many grounds: it would compromise privacy; by making people pay with money rather than with time, it would hurt the poor; it would be impractical and excessively costly to administer; and it would be political poison since it would make people pay for a commodity – public road space – that had previously been provided free of charge, and would be regarded as a tax grab. For many years, congestion pricing was simply dismissed by policy makers. That was phase I.

In phase II, a number of changes occurred. Urban transport economists, acknowledging the validity of many of the criticisms of Vickrey-type schemes, modified their congestion pricing proposals. Privacy issues would be dealt with through the use of prepaid smart cards; concerns about equity and a tax grab would be addressed by transparently redistributing toll revenues in a way that would benefit all major stakeholders (Small (1992, 1993)); and congestion pricing would be applied more coarsely, through cordon pricing or through charging for travel on only urban freeways and highway, for example.

Policy makers, meanwhile, were getting increasingly frustrated by the failure of standard policies – building more roads and later encouraging mass transit – to stop the seemingly inexorable worsening of traffic congestion.[2] The results of Hong Kong's congestion-pricing experiment and the increasingly refined congestion-pricing schemes employed in Singapore were widely discussed; they demonstrated that drivers do respond predictably to congestion prices and that congestion pricing is technologically feasible and can be implemented at reasonable cost. There was also political momentum to privatize transport systems, which for highways requires road pricing. By the end of phase II, many freeways around the world had more or less sophisticated tolling structures. Singapore had a generally successful cordon pricing scheme in place, several Norwegian cities had implemented cordon tolls, and a number of jurisdictions (Hong Kong, Cambridge, Berkeley, the Ranstaad area of the Netherlands, and Stockholm) had given serious consideration to the implementation of congestion pricing but had backed down in the face of political opposition.

* Many of the themes in this article are found in Chapter 1 of *Alleviating Urban Traffic Congestion* (2005). This notwithstanding, my co-authors on that book, Tilmann Rave and Ronnie Schöb, may not share the opinions I express in this article. I would like to thank Robin Lindsey and Ronnie Schöb for providing comments on an earlier draft, and Robin Lindsey for updating me on the literature.

** Richard Arnott is Professor of Economics, Boston College. richard.arnott@bc.edu

[1] This is of course an exaggeration. Urban transport economists instead say, "Get the prices right through congestion pricing, and other elements of urban transport policy will be easy to optimize." If all prices are right, then urban land use will be efficient, and infrastructure will be chosen efficiently if the intuitive, first-best investment rules are followed. As a bonus, if as is argued in footnote 4 urban auto transportation is characterized by increasing long-run average cost, then the revenue raised from the optimal tolls will be more than enough to finance optimal capacity (Mohring and Harwitz (1962), Strotz (1965) and Arnott and Kraus (1998)).

[2] For US metropolitan areas, the steady rise in congestion levels has been documented by the Texas Transportation Institute (The "2005 Urban Mobility Report" is available at http://mobility.tamu.edu/ums/report/). The Institute measures the degree of urban auto congestion by the number of hours of vehicle delay in rush-hour traffic per year experienced by the average driver. For other countries (except perhaps the United Kingdom) the steady rise in congestion levels is not well documented but is asserted by almost all experts.

The highly visible London area pricing experiment has been more successful than almost all the experts had predicted. It has succeeded in substantially reducing the level of auto congestion in central London, without a marked increase in congestion on the Inner Ring Road just outside the congestion charging zone. And unexpectedly, by substantially reducing travel time variability on city streets, the modal shift has been from car to bus. The success of the London experiment has prompted many cities around the world to initiate a congestion pricing planning process. We are now entering phase III, in which urban auto congestion tolling is being widely discussed in policy circles and may come to be implemented on a widespread basis. The topicality of urban auto congestion pricing is the primary reason for the focus of this issue of the journal on "City Tolls".

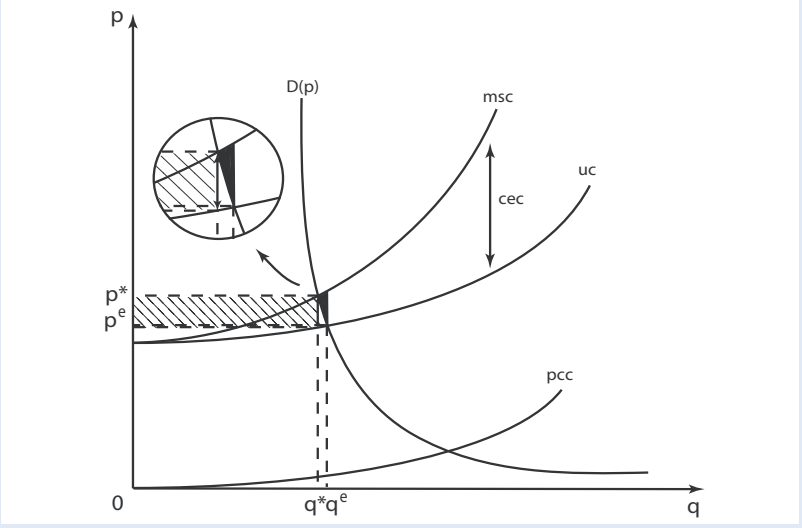
This article has three aims. The first is to summarily review the economic theory of congestion tolling as background for this and other articles in the issue. The second is to sound a cautionary and somewhat heretical note: Look before you leap; the net benefits from urban auto congestion tolling may not be as considerable as its ardent supporters claim. The third is to urge urban transport economists to move beyond their preoccupation with urban auto congestion pricing and to become more active in bringing economics to bear on the wide range of complementary policies that could be effective in alleviating urban traffic congestion.

The economic theory of congestion pricing

The Figure presents the standard diagram for the economic analysis of traffic congestion. The analysis is done in price-quantity space, where p is the full price of a trip, including money costs, time costs and where applicable the toll, and q is the number or flow rate of trips. Capacity is taken to be fixed; the analysis is therefore short run. The demand function plots the quantity of trips demanded as a function of price, D(p). There are two relevant cost functions. The first, sometimes termed the user cost function and sometimes the marginal private cost function, relates the private cost of a trip to the flow on the road; congestion increases as flow increases, increasing trip time and user cost. User cost (uc) in excess of free-flow user cost is termed private congestion cost (pcc). The second is the marginal social cost (msc), which gives the social cost of an extra trip as a function of flow. This equals the user cost plus the congestion externality cost (cec) – the cost a driver imposes on others by slowing them down.

In the absence of congestion pricing, the equilibrium occurs at that flow level, q^e, for which marginal private cost equals trip price – where the marginal private cost curve intersects the demand curve. The optimum however occurs at that level of flow, q*, for which marginal social cost equals marginal social benefit. On the assumption that marginal social benefit equals private willingness to pay, the optimum is given by the point of intersection of the demand curve and the marginal social cost curve. Imposing a congestion toll equal to the congestion externality cost, evaluated at the optimum, τ^*, causes the trip cost function to shift up such that it intersects the demand curve at the optimal flow level. The revenue from the optimal toll is shown as the cross-hatched area. At a particular flow level, marginal deadweight loss, the increase in deadweight loss from an additional driver, equals marginal social cost minus marginal social benefit at that level of flow. The total deadweight loss from imposing no congestion toll, shown as the solidly shaded area in the Figure, equals the integral of marginal deadweight

URBAN TRAFFIC CONGESTION:
THE DEMAND FOR AND COSTS OF CITY AUTO TRIPS



loss from the optimal to the equilibrium level of flow. Capacity expansion results in a downward shift of both the user cost and marginal social cost curves.

Some cautionary notes

In this section, I explain why I am more restrained in my advocacy of urban congestion tolling than most other urban transport economists.

The efficiency gains from even first-best congestion tolling may not be large

When we think about traffic congestion, our instinctive reaction is to think about all the time miserably wasted in traffic, and how great the benefits would be if this time could be substantially reduced. But this is not thinking like a well-trained economist. Traffic congestion is so high because of the spatial concentration of economic activity in cities. Everyone benefits from this spatial concentration through new and more varied products, lower prices for many consumer goods and higher economic growth, and city residents additionally through higher wages, ready access to experts, urban amenities and a richer set of social contacts. Traffic congestion is simply one of the costs we pay to enjoy these benefits. It is excessive congestion that should be our principal concern. Due to underpriced auto travel, traffic congestion is indeed excessive. Efficient congestion pricing would therefore reduce the amount of time wasted in traffic, but its implementation might not reduce congestion by much or result in sizeable efficiency gains or slow down the growth in congestion.[3]

As the diagram is drawn, the efficiency gains from congestion tolling are only a modest fraction of the toll revenues collected. By redrawing the Figure, it can be seen that the gains are increasing in the demand elasticity for rush-hour auto trips and the elasticity of private congestion cost with respect to traffic density. Empirical evidence suggests that the former elasticity is low,[4] while the magnitude of the latter is subject to considerable disagreement (Arnott, Rave and Schöb 2005, Ch. 5). Because the basic model ignores many margins of travel choice (of which the timing of trips is probably the most important (Arnott, de Palma and Lindsey (1990 or 1993)), as well as the endogeneity of urban spatial structure, its application may result in considerable underestimation of the efficiency gains from first-best congestion tolling, but we simply do not know.

The costs of implementation may be substantial

In the design of any congestion pricing scheme, there is a tradeoff between the fineness of the scheme and implementation costs. At one extreme are schemes similar to the one originally envisioned by Vickrey that attempt to charge the driver his marginal congestion externality cost at each point in time, taking into account his time and route of travel, and even actual traffic conditions. Such schemes are typically technology intensive, with high fixed costs, and can be intrusive and expensive to administer, though Singapore's electronic road pricing scheme and the responsive pricing on Interstate 15 in California have been cost effective. At close to the other extreme is a time-invariant cordon toll. Such a tolling scheme would be effective in reducing the amount of traffic crossing the cordon but would do nothing to encourage efficient driving behavior inside or outside the cordon, or efficient trip timing, and would cause traffic to divert around the cordon, perhaps to such an extent that overall congestion would increase.[5]

What proportion of the potential efficiency gains (those that would be achieved with perfect Pigouvian tolling and no administrative costs) would be achieved under various practical tolling schemes? Even when the potential efficiency gains are sizeable, the actual gains that can be achieved might be small or even negative.[6] The theoretical models

[3] The introduction of congestion tolling will result in a one-time reduction in congestion, after which congestion will likely continue to grow. Since centrally located urban land, the most important input into urban road construction, is in fixed supply, urban auto travel is likely characterized by increasing long-run average cost (Small 1999). The continuing trend towards increased urbanization and larger cities caused by technological changes in both production and consumption has given rise to a steady increase in the derived demand for urban auto trips and typically, too, in average trip cost and congestion. (Gordon, Kumar and Richardson 1989, 1990 point to a possible counterexample – that despite rapid growth, average commuting times in Los Angeles did not increase over a twenty-year period because of changes in urban spatial structure deriving from job growth in exurban areas.)

[4] The elasticity of rush-hour auto travel has probably been rising. First, an increasing proportion of the workforce has flexible work hours, allowing travel away from the peak. Second, an increasing proportion of rush-hour trips are non-commuting trips, for which the demand elasticity is higher.

[5] Apparently this phenomenon occurred when Singapore first installed cordon pricing around the central city (Chin 2002; Santos, Li and Koh 2004).

[6] There is quite a large academic literature that calculates, for both simple theoretical models and more realistic computer simulation models, the proportion of the potential efficiency gains that would be achieved under various coarse tolling schemes. This literature has considered tolling only a subset of roads – urban freeways and highways but not city streets (Verhoef, Nijkamp and Rietveld 1996; Mohring 1999), constraining the variation of tolls over time and space (Chu 1999; Mohring 1999), varying cordon placement (May, Liu, Shepherd and Sumalee 2000), and pricing on the basis of expected travel time, actual travel time and distance traveled (May and Milne 2004)). Taken as a whole, the literature has generated two valuable insights. First, ignoring implementation costs, the benefits of a congestion pricing program are sensitive to its details. Second, in evaluating benefits, it is important to treat driver heterogeneity (Small and Yan 2001; Verhoef and Small 2004).
There has also been extensive discussion and analysis of the efficacy of the gas tax in alleviating urban traffic congestion.

ignore the costs of developing and applying the technology to implement congestion pricing, and then of administering and enforcing the program. These costs can easily exceed the efficiency benefits achieved with practical congestion pricing schemes.

One issue that deserves special mention is the choice of technology. High technology systems, which tend to be advocated by engineers, are expensive to implement and rapidly become outdated; low technology systems are more flexible and easier to implement, but enforcement and administrative costs tend to be higher once the system is in place.[7] The London experiment provides an interesting case study. On one hand, the experiment has been more successful than most of the experts expected (Shafer and Santos 2004). On the other, careful ex post cost-benefit analysis (Prud'homme and Bocarejo 2005), employing standard methodology, suggests that the program's net benefits lie in the range of small to negative, principally due to unexpectedly high administrative and enforcement costs.[8] Some proportion of these costs should be amortized since they have the character of fixed costs, while another proportion is offset by the social benefit from learning from London's mistakes. Even taking these factors into account, and despite the high overall quality of the experiment's design, the net benefits are not large, however.

The political barriers to implementation are high

Congestion pricing programs were for many years regarded as political poison. Modern programs are designed to deal with many of the objections raised against the earlier programs. Nevertheless, the political attractiveness of congestion pricing remains open to question. The politics of the London scheme were, if not unique, at least particular to very large cities. Crudely put, the scheme has suburbanite car drivers pay for improvements to central city residents' mass transit. Since it is the central city residents who vote, it is not surprising that the experiment has been politically successful. In many other cities, the political calculus is not as favorable – as evidenced by the recent rejection by referendum of a double cordon scheme in Edinburgh.

Congestion pricing may exacerbate other distortions

The theory of the second best tells us that it is often desirable to distort a price in market A to offset a distorted price in market B. One of the earliest applications of this principle was in the context of urban transportation (Lévy-Lambert 1968). Setting the mass transit fare below the congestion externality cost imposed by a passenger offsets the distortion associated with underpriced auto congestion. Commuters divert from car to mass transit, with the reduction in the deadweight loss associated with underpriced auto congestion more than offsetting the deadweight loss generated by underpricing mass transit. In the context of urban auto travel, two distortions other than underpriced auto congestion are particularly important: income taxation and agglomeration externalities. The substitution effect of the income tax distorts the labor-leisure choice decision, encouraging leisure. Since labor is a complement to travel on the journey to work, raising the price of rush-hour auto travel through congestion tolling exacerbates the labor-leisure distortion. The current empirical wisdom is that this effect is quantitatively important (Parry and Bento 2002).

The nature of agglomeration externalities and how they relate to traffic congestion requires more explanation. It is widely believed that many of the productivity benefits from the spatial agglomeration of economic activity derive from non-market interaction and are external to the individual firm (Fujita and Thisse 2002). Each firm has a horizontal marginal cost curve, and the curve falls the greater the amount of non-market interaction. Because the amount of non-market interaction increases with city population size, the marginal product of labor exceeds the average product. And because the agglomeration benefits are external to the individual firm, workers are paid the average rather than the marginal product. Thus, the market generates insufficient interaction. Since people need to travel to interact, congestion pricing would exacerbate this distortion unless toll revenues were spent in ways that neutralize this effect. No one has attempted to quantify the importance of this effect,[9] but it could

[7] A related issue is technological standardization. On one hand, the best technology and the best congestion pricing system to employ depend on city size and local conditions; for example, Schöb (this issue) argues that, because of its simplicity, a multi-mode ticket, that charges drivers entering the city center a fee equal to the bus fare, is a promising policy for small cities. On the other, a proliferation of congestion pricing systems will lead to technological incompatibilities that will make the treatment of out-of-towners more difficult.

[8] London decided to employ a reliable, low technology system. The license plates of all cars driving within the charging zone are photographed and compared to the license plates of cars for which the fee has been paid. Violators are then sent warnings or are fined.

[9] Because non-market interaction does not, by its nature, leave a paper trail, its importance can be inferred only indirectly. The degree of returns to scale to population can be inferred from cross-city differences in wages. But the effect here is more subtle, entailing holding population constant and measuring the elasticity of interaction with respect to the cost of transportation.

be substantial. Taking both income taxation and agglomeration economies into account could result in second-best congestion tolls being so low that the benefits from even the best congestion-tolling scheme would be more than offset by implementation costs.

Alternatives to congestion pricing

Urban transport policy making used to be dominated by planners and engineers. Economists stood on the sidelines waving the banner of "congestion pricing". But times have changed. Planners and engineers are better trained in economics than they were a generation ago and, through an accumulation of evidence and frustration at the inability of traditional policies to stem congestion, are coming to recognize that pricing has an important role to play in the efficient management of urban traffic. They are therefore far more receptive to economists' ideas than they used to be, and economists are playing an increasingly important role in urban transport policy debates. Now that urban transport economists are getting more engaged in policy decisions, they need to develop a more balanced and sophisticated view of urban traffic management. Congestion pricing by itself is not enough. What is needed is a balanced portfolio of policies, which includes both congestion pricing and the policies traditionally favored by engineers and planners, as well as novel policies that have been unjustly neglected because they do not fit neatly into the disciplinary conceptualizations of the urban transportation problem. Urban transport economists should also recognize that they can contribute to improving road traffic management policy in many ways other than advocating congestion pricing.

Building our way out of the problem

The approach of traffic engineers in the 1960s was to build our way out of the problem, by constructing more and better-engineered freeways, highways and roads. The standard cost-benefit rule employed was to expand capacity when the discounted value of travel time savings from doing so exceeds construction costs. Application of this rule results in efficient choices when prices are right.[10] But urban auto travel was substantially underpriced.[11] In this situation, application of the standard rule results in too much capacity. To see this, consider the effects of an incremental road expansion. If traffic flow is fixed at the pre-expansion level, the expansion reduces congestion and trip price, inducing additional drivers to use the road. The increase in flow that accompanies an increase in capacity is known as latent demand. With underpriced congestion, trip price is below marginal social cost. The marginal social benefit from these additional drivers using the road is simply the trip price. Thus, the additional drivers who use the road add to the deadweight loss due to underpriced congestion, which dissipates the benefit from the road expansion. Latent demand therefore weakens the effectiveness of road building in reducing traffic congestion when congestion is underpriced[12] but not when it is properly priced. Traffic engineers should therefore support congestion pricing since it renders road construction more effective. They should also welcome economists' expertise in cost-benefit analysis and in the analysis of irreversible investment decisions under uncertainty (Dixit and Pindyck 1994).

Upgrading existing infrastructure

Confronted by strong opposition from environmental and neighborhood groups and faced with sky-rocketing land costs, traffic engineers have been pushing road building in already built-up areas less strongly and are now emphasizing expanding capacity through upgrading existing infrastructure. Here too economists can contribute valuable advice.[13] Traffic engineers' current fascination with vehicle information systems seems to me somewhat misguided, but again economists have a role to play in emphasizing that more information is better when the prices are right but not generally otherwise – congestion pricing and vehicle information systems are complements too.

[10] This can be intuited by imagining the adjustment to an incremental capacity expansion as occurring in two steps. In the first, flow is held fixed; in the second, flow adjusts to its new equilibrium level. The welfare gain from the first step is simply the value of travel time savings to existing drivers. The welfare gain from the second step is the net social benefit from the increase in traffic flow, or alternatively the net social benefit associated with those drivers who choose to travel after the capacity expansion but not before. The social benefit associated with each of these drivers is the private benefit or trip price; the social cost is simply marginal social cost. Since these are equal under Pigouvian congestion pricing, the additional welfare gain in the second step is zero.

[11] Not only was urban auto congestion directly unpriced but it was also indirectly subsidized. The excessive suburbanization resulting from underpricing urban auto travel was further encouraged by paying for suburban infrastructure out of general revenue. Most parking was provided free as well, and in the United States the price of gasoline was set below the world price.

[12] In the basic model, the proportion of the benefits dissipated through latent demand depends on the same two elasticities that determine the magnitude of the deadweight loss due to underpriced congestion, the demand elasticity and the elasticity of private congestion costs with respect to traffic density. On a network, latent demand can be so serious that the addition of a link can have negative gross benefits (the Braess Paradox; Braess 1968).

[13] In designing roads, traffic engineers draw heavily on design standards. These design standards are often drawn up on the basis of inferior statistical analysis, are rarely informed by economics and tend to be applied without consideration to traffic conditions. Economists have an important role to play in rationalizing the choice and application of design standards.

Encouraging alternative modes of travel

Environmentalists and planners have their own favored remedies. One is to encourage alternative, "less wasteful" modes of transportation, another to alter land use patterns in a way that reduces the amount of travel. Economists often find themselves at odds with planners; economists favor prices, planners regulations; economists favor consumer sovereignty, while many planners seem eager to impose their vision of the good life on the citizenry. Nevertheless, many of the policies that planners advocate merit inclusion in the policy cocktail, and economists can offer useful advice on how these policies can be more effectively designed. If pricing were efficient, governments making capacity decisions based on first-best rules, together with individuals making decisions based on full prices, would result in full efficiency. But since pricing is not efficient, in particular since car travel is underpriced, sound economic arguments can be made for encouraging travel by modes that generate less distortion. In most European cities, schemes to promote bicycling (Arnott, Rave and Schöb 2005, Ch. 3) and the high levels of transit subsidies are probably justified. In the United States, the pricing sins of the fathers have been visited upon the sons. The heavy subsidization of auto travel in the decades after World War II resulted in suburbanization at densities too low for mass transit to be viable. Because of their preoccupation with congestion pricing and auto travel, urban transport economists have rather neglected mass transit; it is time that imbalance was redressed.

Encouraging higher density and more mixed land use

Planners, especially the new urbanists, are advocating policies to increase densities and to mix land uses. Land use regulation to promote higher densities reduces the amount of travel but that travel is slower and more congested. Land use regulation to mix uses has the beneficial effect of getting people to walk rather than drive on short trips but is unlikely to have a major impact on traffic congestion. Because their training is so suitable, urban economists can contribute far more than they have done to policy issues at the intersection of transportation and land use.

Parking policy

There are many promising congestion alleviation policies that have been neglected because they do not fall neatly into the domain of economics, engineering or planning. One is parking policy.[14] Here there are vast, untapped efficiency gains. While the right level of second-best congestion pricing is much contended, it can confidently be asserted that providing free or nearly-free on-street parking in heavily-trafficked areas is grossly inefficient. The full price of on-street parking has two components, the parking fee and the cost of time spent cruising for parking. Consider the situation with no off-street parking.[15] The full price of on-street parking is determined by the intersection of the parking demand curve and the parking availability constraint. Raising the parking fee has no effect on the full price but simply replaces time wasted cruising for parking with parking fee revenue. There is the additional benefit that reducing cruising for parking reduces traffic congestion; the revenue is therefore raised with negative burden!

Freight delivery management

Another area in which potentially huge efficiency gains can be achieved is urban freight delivery management. Oversized trucks block traffic when backing into loading docks and get stuck negotiating turns on narrow streets; vans block traffic when double parking to deliver a small package. Freight delivery would contribute far less to congestion if it were required that all deliveries be made by vans in the morning hours, and if during these hours some curbside were restricted to van deliveries.

Outside-the-box policies

Other, outside-the-box policies should be explored. Since it is the value of time lost due to congestion that matters, congestion costs can be reduced by making driving more pleasant;[16] traffic noise, a curse of life in many urban areas, can be reduced by mandating quieter cars; urban road construction, a major contributor to non-recurrent congestion, can be done at night under sound-absorbent tents;[17] the time loss due to road accidents, another major con-

[14] Donald Shoup, a planner trained as an economist, is an exception. He has done excellent work documenting the high social costs of subsidizing parking (Shoup 2005).

[15] With off-street parking as well, the same argument goes through but with the modification that the full price of on-street parking is determined by the off-street parking rate rather than by the intersection of the parking demand curve and the parking availability constraint.

[16] Improved car audio systems and the expanded range of talking books have already contributed to this.
While reducing congestion costs associated with a given volume of traffic and level of capacity, making driving more pleasant will also stimulate demand and increase the deadweight loss associated with underpriced congestion.

[17] This is already done in London.

tributor to non-recurrent congestion, can be reduced by introducing no-fault insurance and immediately towing the cars involved in an accident; and aggressive driving, that not only greatly impedes traffic flow but also makes driving considerably more dangerous and less pleasant, can be discouraged by equipping all new cars with electronic devices that monitor horn-honking, frequent lane changes, tailgating, etc.

Conclusion

Urban transport economists are entering a brave new world in which – finally, after fifty years of crying in the wilderness – urban congestion pricing schemes may become commonplace. On one hand, I look forward to seeing what schemes are put in place and how well they do, and sincerely hope that urban congestion pricing proves to be worth the wait. On the other, I have my doubts that urban congestion pricing will be as effective as most other urban transport economists believe. Whether or not my doubts prove well founded, city tolls are only one element of an effective policy cocktail for dealing with urban traffic congestion. Urban transport economists should broaden their horizons beyond congestion pricing to give due attention to the myriad other congestion-relief policies whose effectiveness can only be improved by the application of sound economics.

References

Arnott, R., A. de Palma and R. Lindsey (1990), "Economics of a Bottleneck", *Journal of Urban Economics* 27, 111–30.

Arnott, R., A. de Palma and R. Lindsey (1993), "A Structural Model of Peak-period Congestion: A Traffic Bottleneck with Elastic Demand", *American Economic Review* 83(1), 161–179.

Arnott, R. and M. Kraus (1998), "Self-financing of Congestible Facilities in a Growing Economy", in D. Pines, E. Sadka and I. Zilcha, eds., *Topics in Public Economics*, Cambridge University Press, Cambridge, 161–85.

Arnott, R., T. Rave and R. Schöb (2005), *Alleviating Urban Traffic Congestion*, M.I.T. Press, Cambridge, MA.

Braess, D. (1968), "Über ein Paradoxen der Verkehrsplanung", *Unternehmensforschung* 12, 258–68.

Chin, K. K. (2002), "Road Pricing: Singapore's Experience", essay prepared for the third seminar of the IMPRINT-EUROPE: Thematic Network: Implementing Reform on Transport Pricing: Constraints and solutions: Learning from best practice. Brussels, October 23–24. http://www.imprint-eu.org/public/Papers/ IMPRINT3_chin.pdf.

Chu, X. (1999), "Alternative Congestion Pricing Schedules", *Regional Science and Urban Economics* 29, 697–722.

Dixit, A. K. and R. S. Pindyck (1994), *Investment under Uncertainty,* Princeton University Press, Princeton, NJ.

Fujita, M. and J.-F. Thisse (2002), *Economics of Agglomeration: Cities, Industrial Location and Regional Growth*, Cambridge University Press, Cambridge

Gordon, P., A. Kumar and H.W. Richardson (1989), "The Influence of Metropolitan Spatial Structure on Commuting Time", *Journal of Urban Economics* 26(2), 138–51.

Gordon, P., A. Kumar and H.W. Richardson (1990), "Peak-spreading: How Much?" *Transportation Research* A 24A(3), 165–75.

Lévy-Lambert, H. (1968), "Tarification des services à qualité variable: Application aux péages de circulation", *Econometrica* 36, 564–74.

May, A. and D. Milne (2004), "The Impact on Network Performance of Drivers' Response to Alternative Road Pricing Schemes", in G. Santos, ed., *Road Pricing: Theory and Evidence, Research in Transportation Economics* 9, Elsevier Science, 61–85.

May, A. D., R. Liu, S. P. Shepherd and A. Sumalee (2002), "The Impact of Cordon Design on the Performance of Road Pricing Schemes", *Transport Policy* 9, 209–20.

Mohring, H. (1999), "Congestion", in J.A. Gomez-Ibanez, W.B. Tye and C. Winston (eds.), *Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer*, Brookings Institution Press, Washington DC, 181–221.

Mohring, H. and M. Harwitz (1962), *Highway Benefits: An Analytical Framework*, Northwestern University Press, Evanston, IL.

Parry, I. W. H. and A. Bento (2002), "Estimating the Welfare Effect of Congestion Taxes: The Critical Importance of Other Distortions within the Transport System", *Journal of Urban Economics* 51, 339–65.

Prud'homme, R. and J. Bocarejo (2005), "The London Congestion Charge: A Tentative Economic Appraisal", *Transport Policy* 12, 279–87.

Santos, G., W. W. Li and W. T . H. Koh (2004), "Transport Policies in Singapore", in G. Santos (ed.), *Road Pricing: Theory and Evidence, Research in Transportation Economics* 9, Elsevier Science, 107–31.

Shaffer, B. and G. Santos (2004), "Preliminary Results of the London Congestion Charging Scheme", 83rd Annual Meeting of the Transportation Research Board, conference CD.

Shoup, D.C. (2005), *The High Cost of Free Parking*, APA Planners Press, Chicago.

Small, K. A. (1992), "Using the Revenues from Congestion Pricing", *Transportation* 19, 359–81.

Small, K.A. (1993), "Urban Traffic Congestion: A New Approach to the Gordian Knot", *The Brookings Review* 11(2), 6–11.

Small, K.A. (1999), "Economies of Scale and Self-financing Rules with Non-competitive Factor Markets", *Journal of Public Economics* 74, 431–50.

Small, K. A. and J. Yan (2001), "The Value of 'Value Pricing' of Roads: Second-best Pricing and Product Differentiation", *Journal of Urban Economics* 49(2), 310–36.

Strotz, R. H. (1965), "Urban Transportation Parables", in J. Margolis (eds.), *The Public Economy of Urban Communities,* Resources for the Future, Washington DC.

Texas Transportation Institute (2005), *Urban Mobility Report*, TTI Publications, College Station, TX.

Verhoef, E. T. and K. A. Small (2004), "Product Differentiation on Roads: Constrained Congestion Pricing with Heterogeneous Users", Journal of Transport Economics and Policy 38(1), 127–56.

Verhoef, E. T., P. Nijkamp and P. Rietveld (1996), "Second-best Congestion Pricing: The Case of an Untolled Alternative", *Journal of Urban Economics* 40(3), 279–302.

Vickrey, W. S. (1959), "Statement on the Pricing of Urban Car Use", Hearings: US Congress, *Joint Committee on Metropolitan Washington Problems*, November, 454–77.

Vickrey, W. S. (1963), "Pricing in Urban and Suburban Transport", *American Economic Review* 53, 452–65.

Road Pricing – Singapore's 30 Years of Experience

Kian-Keong Chin*

Singapore is an island-state with a land area of just over 690 square km, measuring 42 km across and 23 km from north to south. On this island live more than 4 million people, making it one of the most densely populated countries. Its transport needs are served with vehicles totalling 730,000, of which 420,000 are cars.

In the field of transportation, road pricing has long been associated with Singapore. Indeed, the practical application of road pricing started in Singapore in June 1975. Many changes have been made to the road pricing scheme since that time. Initially a manual scheme based on paper permits and applicable during the morning peak period only, it has evolved over the past 30 years to an electronic version that operates presently almost throughout the day.

Road pricing is an important component of Singapore's overall transportation strategy. While road capacity continues to be increased judiciously to meet rising travel demand, the strategy also calls for greater reliance on public transport usage and demand management. This is especially important in Singapore, given its limited land resource. One aspect of demand management is the restraint of vehicle ownership, either through the imposition of high upfront ownership costs or restriction on the actual growth of the car population. The former includes excise duties and vehicle registration fees, which amounted to 130 percent of the car's open market value presently, while the latter is managed through a Vehicle Quota System, where the net increase in vehicle ownership is capped at 3 percent per year. The other aspect of demand management is the restraint of vehicle usage through the levy – a charge on motorists based on the quantity, place or time of the use of their vehicles. Generally, the more one uses his car the more one has to pay. The road pricing schemes, petrol tax, diesel duty and parking charges are measured in this category.

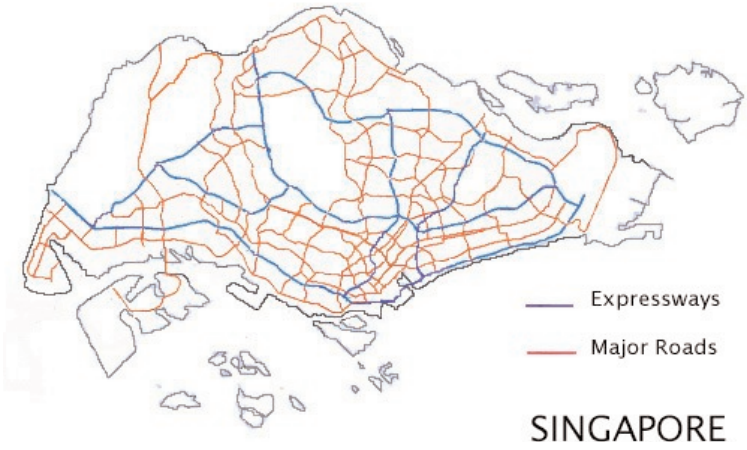
One of the goals set out in the demand management strategy of Singapore is to move away from relying predominantly on vehicle ownership costs to one of a better balance between it and usage costs. The resulting system would be a fairer and more equitable one. This paper looks at Singapore's experiences with road pricing over the past 30 years, from the manual scheme to the current electronic one.

Manual road pricing scheme

The Area Licensing Scheme (ALS) was in place for 23 years before being replaced by an electronic version called the Electronic Road Pricing System (ERP). This manual scheme was based on the need for paper licenses to be purchased prior to their passage through control points set up on the roads.

How the ALS worked

The ALS covered the more congested parts of the Central Business District (CBD), designated as the



* Dr. Kian-Keong Chin is Chief Engineer, Transportation, in the Land Transport Authority, Singapore; kian_keong_chin@lta.gov.sg. The author thanks the Chief Executive, Land Transport Authority, for permission to write this article.

Restricted Zone (RZ). When the ALS first started, the RZ had an area of only 610 hectares; but over the next 14 years, it increased to 725 hectares eventually because of the inclusion of areas that had turned commercial in nature and the availability of reclaimed land at the sea-front. Prior to its demise in 1998, the RZ was demarcated by 31 overhead gantry signs at its control points.

To gain access into the RZ during the restriction periods, non-exempt vehicles needed to purchase and display an ALS area licence. The daily area licences could be purchased at roadside sales booths located at the approach roads to the RZ, petrol stations, post offices and convenience stores.

Enforcement personnel were stationed at the control points during the restricted hours to observe whether vehicles displayed the valid licences on their windscreens, or on the handle-bars in the case of motorcycles and scooters. Violating vehicles were not stopped at the gantries, but their vehicle licence numbers were taken down and their owners would receive a summons for entering the RZ without a valid licence. The checks for licences were made only at the control points, and vehicles were free to move around or leave the RZ without having the licences.

Operating hours

The ALS started in 1975 with the restricted hours of 7.30 a.m. to 9.30 a.m. daily, except on Sundays and public holidays. Three weeks later, the restricted hours were extended to 10.15 a.m. in order to restrain the surge in vehicle entries immediately after the lifting of the ALS at 9.30 a.m. The ALS operated for two and three-quarters hours each weekday during the morning peak period until June 1989, when major fundamental changes were made. They were triggered by a rapid growth of vehicle population during that period. The restriction period was extended to cover the evening peak hours of 4.30 p.m. to 7.00 p.m. on weekdays. The evening period was later cut back by half an hour to 6.30 p.m. to accommodate requests from residents who lived inside the RZ but worked outside, although this was subsequently extended back to 7.00 p.m. because of increased traffic congestion.

In 1994, more fundamental changes were made to the ALS scheme. The restricted hours were further extended to cover the inter-peak period of 10.15 a.m. to 4.30 p.m. on weekdays and the post-peak period of 10.15 a.m. to 3.00 p.m. on Saturdays. The Saturday restriction period was subsequently cut back to 2.00 p.m. due to improved traffic conditions within the RZ.

Vehicles covered in the ALS scheme

When the ALS started in 1975, taxis, public transport buses, goods vehicles, motorcycles, and passenger cars carrying three or more passengers (excluding the driver) were exempted from the scheme. Car-pooling was allowed under the scheme to optimise vehicle usage and to counter the charge that the scheme favoured only the rich. The exemption on taxis was subsequently removed in 1975.

In the review of June 1989, more vehicles were required to purchase licences prior to their entering the RZ during the restriction period. Motorcycles and goods vehicles, which together made up about two-thirds of the traffic entering the RZ, were no longer exempted.

Furthermore, exemptions for car-pools were also abolished. This was because private cars were picking up bus commuters instead of forming genuine car-pools. This constituted a form of pinching of bus commuters by motorists and defeated the intent of the car-pool exemption.

Road pricing charges

The licence fees had been revised upwards over the years to cater for inflation and to check the gradual upward creep in the number of restricted vehicles entering the RZ. Starting at $3[1] per day for a licence for cars in 1975, this had crept up to $5 per day in 1980. However, with the major review in 1989, there was a reduction in rates – essentially because more vehicles were required to purchase licences. The daily licence fee for a car was reduced back to $3.

With the review in Jan 1994, there were two levels of fees for licences – the one that permits usage throughout the day, and the one that was applicable for use during the inter-peak period only. For cars, the daily licence fees were $3 and $2, respectively.

Traffic impact

The initial drop in traffic entering the RZ was 44 percent, but it crept up to a 31 percent drop by 1988.

[1] The fees mentioned in this paper are in Singapore dollars, and as at July 2005, the exchange rate is US$1 = S$1.65.

However, this was despite the growth by a third in employment in the city and by 77 percent in vehicle population during the same period. The drop in traffic was caused by the decanting of motorists whose destinations were not the city itself but had merely been using the city roads as a bypass, as well as by those who changed their journey start time to avoid paying the ALS fees. There was little evidence to suggest that motorists had transferred to public transport in significant quantities.

Limitations of the manual scheme

Being a manual scheme, the ALS had limitations. It was labour intensive: about 60 enforcement officers were required at all the gantry points and another 60 officers at the dedicated licence sales booths. Extending the schemes to other points would have needed even more people to run them. The enforcement job was tiring as long hours were spent under the sun and rain. The environment these people worked in was at the roadside, and this was dusty and noisy. The work also required considerable concentration because of the varied vehicle heights and categories of licences. There were 16 types of licences for the two schemes, divided into daily/monthly, peak/inter-peak and the different vehicle classes. Human enforcement by visual means was prone to error leading occasionally to wrongful summonses being issued.

Under the manual system, a licence offered a vehicle unlimited number of entries to the RZ or passage through the control points. Although it was not legally transferable between vehicles, it was tempting for such action to take place.

There was always a rush to enter the RZ just before or after the restricted hours because of the sudden change of licence fee from nothing to $3 or vice versa. This resulted in sharp and short peaks of entering traffic volume. "Shoulder-charging", or having intermediate rates, would have smoothened out the peaks, but it was difficult to implement in a manual system. Having more categories of licences would make enforcement more difficult and more prone to mistakes.

Electronic road pricing (ERP)

With the shortcomings of the manual road pricing schemes, the search for a more efficient technology began in earnest in the early 1990s. Technology for an electronic road tolling system was emerging at that time, and there were many learning visits to cities with such electronic tolling systems installed. After several years of discussions and prototype testing with potential suppliers, a contract for the installation of a Dedicated Short-Range Communication (DSRC) electronic road pricing system was awarded in 1995.

Implementing the ERP system

A comprehensive test programme was included in the contract, and this took place from December 1996 to August 1997. A section of an unopened stretch of expressway was converted into a test site, comprising 12 sets of ERP gantries. Using a fleet of 250 vehicles, each with a prototype In-vehicle Unit (IU) or transponder fitted, about 4.8 million ERP transactions were clocked before the test was considered a success. It was only after this exhaustive testing that the mass production of the IUs and the gantry equipment were allowed to commence.

The ERP system has three major groups of components. The first centred around the IU and the stored-value smart-card. While the IUs were produced specifically for the ERP system, the smart-cards were marketed by a consortium of local banks for multiple uses. Different IUs were produced for different classes of vehicles, with each type distinguished by different colours.

The second group of components comprises those installed at the on-site ERP gantries. These include the antennae, the vehicle detectors and the enforcement camera system. All these are linked to a controller box located at each of the site. Data collected is transmitted back to the Control Centre continuously through leased telecommunication lines.

The third group of components is at the Control Centre, and includes various servers, monitoring systems as well as a master-clock to ensure that the timing at all the ERP gantries are synchronised. All the financial transactions are processed here, before being sent to the banks for settlement. In addition, violation images are processed at the Control Centre, and letters and summonses are printed and sent out from here to all those that did not have valid ERP transactions.

How the ERP scheme worked

The ERP system is designed to be simple to use. Indeed, all that it required of the user is to insert the

smart-card into the IU, whereby a diagnostic check is automatically done to ensure that both the IU and the smart-card are in working condition. Should there be a problem with either the IU or the smart-card, the user will be alerted so that remedial action can be taken. The IU is also designed to have a low balance indicator, whereby the user will be alerted should the cash balance in the smart-card fall below a pre-defined amount.

When the vehicle passes through the ERP gantry, the appropriate ERP charge is deducted from the smart-card. There is a short beep, signifying a successful transaction, and the remaining cash balance in the smart-card is displayed momentarily. Should there be insufficient cash in the smart-card or should there be no smart-card in the IU, the enforcement cameras in the gantry will take a picture of the rear of the vehicle. Similarly, should the vehicle have no IU installed in it, the enforcement cameras will also take a picture of the offending vehicle. The picture is sent back to the control centre, where the vehicles' registration numbers are automatically read using OCR techniques. The owners of the vehicles are then issued with letters to ask for payments. For cases of insufficient cash balance or no smart-card in the IU, an administrative charge of $10 plus the outstanding ERP charge is imposed. To promote electronic payments, this administrative charge is $8 if payment is made on-line at a number of electronic kiosks or through the Internet. If this payment is not made within the prescribed 28-day period, a summons will follow. For cases where vehicles pass through the ERP gantry with no IU, summonses are issued immediately. Failure to pay the penalty can result in the offender being called up to appear in the courts.

IU fitting programme

There were two major programmes launched prior to the start of the ERP. The first was the installation of IUs on the then 680,000 eligible vehicles while the second was on publicity, to get motorists and motorcyclists aware and ready for the ERP system.

The IU fitting programme took 10 months, starting from September 1997. The 680,000 vehicles that were potential users of the ERP system were grouped into batches, and owners of each batch were invited to have IUs fitted at one of the 200 authorised IU fitting centres. This was spread out over the 10-month period, so that there would be no last-minute rush to fit IUs. To encourage vehicle owners to keep to the schedule, the IUs were given away at no cost to them if they had their IU fitted during the allocated time period. Otherwise, a charge of $150 was payable. The fitting of IUs onto vehicles was not compulsory – it was left to the individual to decide. Nevertheless, at the end of the IU fitting programme, more than 98 percent of the registered vehicles were fitted with IUs.

Publicity programme

Publicity was another important aspect and this started even before the start of the IU fitting programme and was in place for more than a year, all the way up to and beyond the launch date of the ERP system. All vehicle owners were sent brochures, detailing the ERP system, how it works and the differences between the ERP system and the then working ALS/RPS. Advertisements were also placed in the print media as well as on television to drum up awareness of the new road pricing system.

One of the differences between the ERP and the then ALS pointed out in the publicity programme was that, unlike the latter, the ERP system imposes a charge each time a vehicles passes through the control point. In the ALS schemes, the fee payable was a daily charge that allowed the vehicle to make unlimited number of entries into the RZ for that day.

Traffic impact

Traffic volume into the CBD had reduced by about 10 to 15 percent during the ERP operation hours, as compared to the ALS scheme. This was in spite of the road pricing charge being lower at between $0.50 and $2.50; the charge for the ALS was $3.00 for peak periods and $2.00 for the inter-peak period. The major difference is that the ERP charge is applicable for each passing, while the ALS charge allowed multiple entries for that day. Hence, the ERP had influenced particularly the behaviour of those who made multiple trips into the CBD – and this was estimated to be about 23 percent of trips that entered the CBD during the ALS days. Many of these multiple trip-makers cut down their number of trips, e.g. office workers no longer use their cars to attend mid-day meetings or lunches – more relied on the public transport system instead.

Varying the road pricing charge

The ERP system allowed more frequent changes to be made to the road pricing charges, so that it can better

optimise road usage. Speed-flow curves were established for two different types of roads – urban roads with side friction and traffic signals, and expressways. The intent of ERP is to optimise road usage, i.e. flows should be near the maximum possible. From the speed-flow curves, it was decided that speeds on selected roads should be between 20 km/h to 30 km/h while that for expressways, the speeds should be between 45 km/h to 65 km/h. When speed goes above the upper threshold, too few vehicles are deemed to be using the roads and hence, the road space available is not being optimally used. Hence, the road pricing charge can be reduced to allow more vehicles to use the roads. Conversely, if the speed falls below the lower threshold, too many vehicles are on the roads and this is a signal that the road pricing charge can be increased. Other considerations do come in when deciding on the final road pricing charge, e.g. the effect of traffic diverting to other roads.

Other issues

There was the inevitable issue of privacy. There was a lot done to allay the fears of motorists. Being an active system, there was no necessity for the central computer system to keep track of vehicle movements since all charges were deducted from the inserted smart-card at the point of use. Records of such transactions were kept in the memory chip of the smart-card that belonged to the individual. The authorities also took a further step to assure the public that all records of transactions required to secure payments from the banks were erased from the central computer system once this was done – typically within 24 hours.

Radiation from the system was also a concern. However, this again was a non-issue since the IUs are non-radiating – they reflect what it receives from the antenna, which is positioned 6 metres above the road surface. The radiated power from the antenna is more than a thousand times less than what is allowable by the International Agencies on Radiation Protection.

The ERP system allowed a finer graduation in rates and this can potentially have the ERP charges set at levels that commensurate better with the prevailing traffic conditions. This also helped to reduce, but not eliminate, the problem of motorists waiting for the road pricing to end for the day, as the savings that could be made by motorists have now became less significant.

The issue of the ERP system as a revenue tool for the Government was also raised. However, ERP has always been positioned as a traffic management tool and revenue was and is never a consideration. Indeed, with the ERP system replacing the ALS, the revenue collected today is still only about 80 percent of what used to be collected with the ALS. This is in spite of there being more ERP gantries in operation. The reduction in revenue collected is due to lower charges and operating hours that came about because the ERP system is a more flexible tool.

Extending ERP

As vehicle ownership continues to increase, more pockets of congestion on the road network becomes inevitable. Hence, more ERP gantries have been erected since its launch in 1998. Starting with 33 gantries when it was first introduced, this has increased to 45 by 2004. An additional 3 gantries are planned for 2005. One of these gantries is on the out-bound direction of a major expressway from the city and to be operational during the evening peak hours, while the other two are to create two separate pricing cordons for the CBD. Having two separate cordons allows a more targeted approach towards the setting of ERP charges as the characteristics of the zones covered by these two cordons are slightly different – one being more of a shopping zone while the other being more of an office zone.

Conclusion

Road pricing in Singapore has been effective in managing congestion on roads in the CBD since its inception in 1975, and in recent years on expressways and other major roads outside the CBD. Technology had helped to make the expansion of the original road pricing scheme possible; and the authorities are still keeping tab on new developments in technology, e.g. the GPS-based system for heavy vehicles in Germany, to further enhance and refine road pricing in Singapore.

References

Menon, A. P. G. and K.-K. Chin (1998), "The Making of Singapore's Electronic Road Pricing System", Proceedings of the International Conference on Transportation into the next Millennium, Singapore.

Lew, Y.-D. and C. L.-Y. Tee (2000), "Singapore's Experience with Road Pricing: From Manual to Electronic", Technical Report at the 5th ASEAN-Japan Workshop-cum-seminar on Urban Transportation.

Menon, A. P. G. and K.-K. Chin (2004), "ERP in Singapore – What's Been Learnt from Five Years of Operation?", Traffic Engineering & Control (tec) 45 (2).

London Congestion Pricing – Implications for Other Cities

Todd Litman*

Background

A basic economic principle is that consumers should pay directly for the costs they impose as an incentive to use resources efficiently. Urban traffic congestion is often cited as an example: if road space is unpriced traffic volumes will increase until congestion limits further growth. For decades economists have recommended road congestion pricing (special tolls for driving on congested roadways) as a way to encourage more efficient use of the transport system, and address congestion and pollution problems, providing net benefits to society.

In recent years a few cities have implemented various forms of congestion pricing, including Singapore, Orange County (California State Route 91) and the cities of Trondheim, Oslo, and Bergen in Norway, but proponents have been frustrated by the political resistance congestion pricing faced in other major cities.

Central London is a particularly suitable city for congestion pricing because of its limited road capacity (the streets network in the core area has hardly expanded since the medieval ages), and heavy travel demand result in severe congestion, plus relatively good travel alternatives, including walking, taxi, bus and subway services, which are used by most travelers. Only about 10 per cent of peak-period trips were made by private automobile. For decades transport planners have recommended congestion pricing in central London.

* Todd Litman is Executive Director of the Victoria Transport Policy Institute, Victoria, British Columbia, Canada. litman@vtpi.org.

In 2000, London's political system was restructured to create a new, elected mayor with new powers to manage the city's transport system and raise taxes to fund transport improvements. Ken Livingstone won with a platform that included congestion pricing implementation. Revenues are to be used to fund public transit improvements.

This plan was criticized by various interest groups, including politicians, motorist groups and some labor organizations. The Conservative mayor candidate promised to end the congestion pricing program if elected. There are half a dozen daily newspapers published in London, and many of them were skeptical or opposed to the program. However, Mayor Livingstone, considered a "radical" politician, proceeded with the charge.

How the program works

Since 17 February 2003 motorists driving in central London (see Figure) on weekdays between 7:00 a.m. and 6:30 p.m. are required to pay £5, increasing to £8 in July 2005. There are some exemptions, including motorcycles, licensed taxis, vehicles used by disabled people, some alternative fuel vehicles, buses and emergency vehicles. Area residents receive a 90 percent discount for their vehicles. The charging area is indicated by roadside signs and symbols painted on the roadway. The city is currently developing a plan to expand the charging zone westward in 2006 or later.

London Congestion Charging Area

Payments can be made at selected retail outlets, payment machines located in the area, by Internet and cellular telephone messaging, any time during that day. Motorists can purchase weekly, monthly and annual passes with modest (15 percent) discounts. A network of video cameras records the license plate numbers of vehicles and matches it with the paid list. The owners of vehicles that have not paid as required are sent a £80 fine. This fine is reduced to £40 if paid within two weeks, and increases to £120 if not paid after a month – the same policy for parking penalties in the inner London area.

The system is considered effective. Approximately 110,000 motorists a day pay the charge (98,000 individual drivers and 12,000 fleet vehicles), increasingly by mobile phone text message. Non-payment rates were high during the first few weeks, due to general confusion and errors (such as motorists confusing number 0 or 1 and the letter O or l), but these have declined as users and operators gain experience.

This system is not considered optimal because:

- The fee is not based on how many miles a vehicle is driven within the charging area.
- The fee is not time-variable, that is, the fee is not higher during the most congested periods and lower during less congested periods.
- The fee does not vary by location. It would be more efficient to have higher rates on more congested roads.
- The system has relatively high overhead costs.
- Transit service (particularly the underground rail system) is crowded and unreliable, although this is changing as bus service improves and pricing revenue is used to upgrade the system.

Costs and revenues

The program was predicted to have the following costs and revenues between 2001 and 2008, including three years of development and five years of operation:

Table 1 shows the charging program's originally projected costs and revenues. However, charge revenues turned out to be lower, and penalty revenues higher than anticipated – resulting in a net "gain". The 2004/05 budget year is projected to earn £190 (instead of £160) million in total revenues (£118 million in fees and £72 million in fines), with £92 million in overhead expenses, resulting in £97 million in net revenues.

Travel impacts

Transport for London, the local transport authority and part of the city of London's administration, and various academic organizations established a five-year monitoring program to evaluate the transport, economic, social and environmental impacts of congestion charging (see references for London Congestion Charging Website).

Just over a million people enter central London during a typical weekday morning peak (7–10 a.m.). Over 85 percent of these trips are by public transport. Prior to the congestion pricing program about 12 percent of peak-period trips were by private automobile. During the programs first few months automobile traffic declined about 20 percent (a reduction of about 20,000 vehicles per day), resulting in a 10 percent automobile mode share.

Table 1

Congestion charging program: Projected costs and revenues

	Total (NPV) in £ millions	Per operating year in £ millions
Start up costs	180	36
Operating costs	320	64
Total cost, 2001 – 2003	**500**	**100**
Charge revenues	690	138
Penalty revenues	110	22
Total annualized revenue, 2004 – 2008	**800**	**160**

NPV = net present value.

Source: Compilation by the author.

Most people who change their travel patterns due to the charge transfer to public transport, particularly bus. Some motorists who would otherwise drive through central London during peak periods shift their route, travel time or destination. Others shift mode to taxis, motorcycles, pedal cycles, or to walking

This has significantly increased traffic speeds within the zone. Average traffic speed during charging days (including time stopped at intersections) increased 37 percent, from 8 miles-per-hour (13 km/hr) prior to the charge up to 11 miles-per-hour (17 kms/hr) after pricing was introduced. Peak period congestion delays declined about 30 percent, and bus congestion delays declined 50 percent. Bus ridership increased 14 percent and subway ridership about 1 percent. The third-year annual report indicates that these improvements are continuing (Transport for London 2005).

Taxi travel costs declined significantly (by 20–40 percent) due to reduced delays. Vehicles can cover more miles per hour, so taxi and bus service productivity (riders per day) and efficiency (cost per passenger-mile) increased substantially. There has been some increase in motorcycle, moped and bicycle travel, and vendors have promoted these modes.

The program's net revenues will be used to improve public transit services, including more buses and major renovations to the subway ("tube") system, which is widely agreed to be in need of significant redevelopment. Bus service is being improved in many ways, including an expanded bus lane system, with enforced surveillance using a network of over 1,400 on-bus and roadside video cameras.

Public and political response

Before implementation this plan was widely criticized by various interest groups, including politicians, motorist groups and some labor organizations. The Conservative mayor candidate promised to end the program if elected. Many newspapers were skeptical or opposed to the program, and opponents produced a website titled Sod-U-Ken (see references) to promote their objections.

The congestion pricing program has since become generally accepted by the public and interest groups, including many that initially criticized it, such as automobile clubs. Within a month of its start residents of other areas in London began requesting to be included, and the Conservative candidate no longer promised to end the fee. In 2004, Mayor Livingstone was reelected, largely on the success of the road pricing program and his plans to expand the pricing zone.

London First (see references), a business group whose members account for 22 percent of the city's GDP, supports the city's congestion charge. A survey performed May 2003 found that its members consider the scheme to have overall positive impacts on business activity. The majority (69 percent) felt charging had no impact on their business, 22 percent reported positive impacts on their business, and only 9 percent reported an overall negative impact. Many industries support the charge because its direct costs are offset by savings and benefits, such as faster delivery times. Cert Logistics, a distribution company that delivers to many downtown restaurants and hotels, reports its delivery times have been cut by as much as 50 percent, and other industries find that their employees spend less time delayed in traffic, and so can attend more out-of-office meetings in a day. The £5 charge pays for itself in just 17 minutes of travel-time savings for employees earning average London wages (£34,000 per year), and many city center employees earn far more.

London Chamber of Commerce (see references) members have been more skeptical of the program. A March 2003 survey found that many city center retailers blame the charge for reduced sales (although the Iraq war and its security measures, a temporary closure of two subway lines, and a general economic downturn also impacted local business activity), and some threaten to leave the city. Opposition to the program tends to be greater among smaller retailers, and may partly reflect political ideology (the business community tends to oppose Mayor Ken Livingstone and his policies), and may represent a political strategy to gain more special treatment to benefit local retail businesses, such as lower parking fees and special discounts.

Criticism

Some issues of criticism are discussed below.

Business activity

As mentioned above, some businesses consider themselves harmed by the program, particularly bulk good retailers that rely on customers who drive private cars. However, other economic activities have benefited due to improved access by other modes, reduced delay for high-value vehicle trips, and improved environmental conditions. A shift in the location of bulk retailing is not necessarily harmful to the regional economy if the city center locations they vacate are used for equally productive activities. Economic theory suggests that congestion pricing should increase overall productivity and business activity by favoring higher-value activities over lower-value activities, and reducing congestion costs. Some criticisms were based on inaccurate assumptions (for example, theater patronage impacts are slight since the charge does not apply evenings and weekends when most performances occur), or were addressed by special treatments (such as discounts for certain types of employees). Various studies and surveys indicate minimal overall impact on economic productivity and activity, although some sectors and businesses are affected more than others, either positively or negatively (Transport for London 2005).

System accuracy

The congestion pricing system uses a network of video cameras to record license plate numbers, and optical character recognition (OCR) technology to read this information, identify "unpaid" vehicles and generate citations for violators. During the first few weeks the rate of false positives (motorists wrongly ticketed) was high, in part due to both drivers and OCR systems misinterpreting characters or using incorrect registration data. Failure rates have declined over time as motorists and operators gain experience.

Traffic spillover impacts

There was concern that congestion may increase on nearby roads due to diverted traffic. Although some diversion occurred the effect appears to be too small to measure, and may be addressed in the future by expanding the priced area and charging more variable fees (higher rates in the center and lower rates in outer zones). Although there is 10 percent more traffic on the peripheral roads, journey times on them have not increased, in part because traffic signal systems on these roads were adjusted in anticipation of these traffic shifts.

Fairness

Some critics argue that road pricing is unfair because it constitutes "double charging", since motorists already pay registration and fuel taxes and is unfair to lower-income people who must drive, such as workers who commute to central London during off-hours. Some motorists are exempt (e.g., people with disabilities) or have substantial discounts (residents within the priced area) not available to others. This criticism has raised debate concerning what pricing is equitable and how reforms can be most fair and beneficial to consumers. Some concerns have been addressed with special exemptions or policies.

Pricing efficiency

Most economists favor a variable road use fee that reflects the type of vehicle, when it is driven, where and how much it is driven within the priced area, since that most accurately reflects the costs imposed by driving and gives motorists an incentive to minimize their negative impacts, for example, by shifting from peak to off-peak periods, or by minimizing their mileage. The current system uses a flat fee applied to any vehicle that drives within the central area during weekdays, so once a motorist pays the fee they have no incentive to minimize driving. This flat-rate system was chosen because it was relatively fast and easy to implement, and simple to understand. A more sophisticated system that allows variable fees is likely to be implemented in the future.

Cost efficiency

A substantial portion of revenue is used to pay program expenses. About half of the program's revenues are spent on overhead costs (project development, equipment and operations). Critics argue that there are more cost effective ways to collect money. However, as a congestion reduction strategy and a way to improve bus service it has proven to be an effective and cost efficient investment.

Loss of privacy

There is concern that the network of video cameras and the system for tracking vehicles within London is an invasion of privacy. However, British cities already have extensive video surveillance systems, and access to vehicle location data is controlled to limit invasion of privacy. However, these factors have not eliminated this concern.

Winners and losers

Table 2 characterizes the groups which are either winners or losers of London congestion pricing.

Table 2

Congestion pricing: Winners and losers

Winners	Losers
• Downtown bus riders • All transit riders (due to increased funding for improvements) • Taxi riders and drivers • Motorists with high-value trips • Most city center businesses • Overall city productivity • Pedestrians and cyclists	• Motorists with marginal-value trips • City center businesses that depend on low-cost weekday car access • Residents and motorists in border areas who experience spillover impacts • City center parking revenue recipients

Source: Author.

Implications for other cities

London's congestion pricing project is considered an important test of the political feasibility of congestion pricing in major democratic cities. London's experience shows that congestion pricing is technically feasible and effective, and that it is possible to overcome the political and institutional resistance to such pricing. As a result, it will help put congestion pricing on the menu of transportation improvement options in other cities.

This pricing program indicates that private automobile travel is more price sensitive than most experts believed. This is good news for congestion reduction but bad news for revenue generation.

Better pricing systems are needed to optimize the incentive, with prices that vary depending on the type of vehicle, when, where and how much it is driven. These are technically feasible, and have been implemented in Hong Kong and Singapore, but they involve greater investments and potential loss of privacy.

Implementation is not easy. It requires a suitable combination of travel and political conditions, including widely dispersed benefits and the ability to overcome public skepticism. Compared with other cities London has a particularly small portion of automobile commuters, and many of them reside outside the city. As a result, a relatively large portion of voters perceive themselves as benefiting from the fee. Voters in other cities may be more skeptical that they will benefit from such a fee.

Summary

Since 17 February 2003 the city of London has charged a fee for driving private automobiles in its central area during weekdays as a way to reduce traffic congestion and raise revenues to fund transport improvements. This has significantly reduced traffic congestion, improved bus and taxi service and generates substantial revenues. Public acceptance has grown and there is now support to expand the program to other parts of London and other cities in the UK. This is the first congestion pricing program in a major European city, and its success suggests that congestion pricing may become more politically feasible elsewhere.

References

CFIT (2002), *Paying for Road Use*, Commission for Integrated Transport, www.cfit.gov.uk/reports/pfru/ index.htm).

Congestion Pricing Bibliography, www.hhh.umn. edu/centers/ slp/ conpric/bib.htm.

European Program for Mobility Management, www.epommweb.org, provides resources for transportation demand management planning and program development.

European Transport Pricing Initiatives, www.transport-pricing.ne,includes various efforts to develop more fair and efficient pricing.

Faber, O. (2000), *Fair and Efficient Pricing in Transport – The Role of Charges and Taxes*, European Commission DG TREN in association with EC DG TAXUD and EC DG ENV. Available through the European Program for Mobility Management, www.epomm web.org.

Hau, T. (1992), *Economic Fundamentals of Road Pricing*, Report Nos. TWU 1 and TWU 2, Infrastructure and Urban development, World Bank (Washington DC), available at www.worldbank.org /html/fpd/transport/publicat/pub_tran.htm and at www.econ. hku. hk/~timhau.

HHH Institute of Public Affairs (1996), *Buying Time; Guidebook for Those Considering Congestion Relief Tolls*, Value Pricing Website, www.hhh.umn.edu/centers/slp/conpric/guide1.pdf.

Litman, T. (1996), "Using Road Pricing Revenue", in *Transportation Research Record* 1558, TRB (www.trb.org), pp. 24-28, available at www.vtpi.org.

Litman, T. (2005), *Socially Optimal Transport Prices and Markets*, VTPI (www.vtpi.org).

London Chamber of Commerce, www.londonchamber.co.uk.

London Congestion Charging Website, www.cclondon.com. For information on the monitoring and evaluation program see www.tfl.gov.uk/tfl/cclondon/cc_monitoring.shtml.

London First, www.london-first.co.uk.

Shaffer, B. and G. Santos (2004), *Preliminary Results of the London Congestion Charging Scheme*, TRB Annual Meeting, (www.geog. ox.ac.uk/staff/gsantos-files/trb2004.pdf).

Sod-U-Ken, www.sod-u-ken.co.uk.

T&E (2003), *Congestion Pricing in London: A European Perspective, European Federation for Transport and Environment* (www.t-e.nu/Factsheets/2003/14-2-03-CongestionBriefing.PDF).

Transport For London, www.tfl.gov.uk/tfl, provides updates on London's pricing program.

Transport for London (2005), *Central London Congestion Charging Impacts Monitoring Third Annual Report*, Transport for London, www.tfl.gov.uk/tfl/cclondon/pdfs/ThirdAnnualReportFinal.pdf.

Vickrey, W. (1992), *Principles of Efficient Congestion Pricing*, Columbia University, available at www.vtpi.org/vickrey.htm.

VTPI (2005), "Road Pricing," *Online TDM Encyclopedia*, Victoria Transport Policy Institute, www.vtpi.org/tdm/tdm35.htm.

Policy Transfer Through Research Networks – The Example of Urban Road Pricing

Catharina Sikow-Magny and Marcel Rommerts*

The European Commission has advocated the reform of transport pricing for over a decade now. The theoretical concept of marginal cost pricing was first advocated in the green paper Towards Fair and Efficient Pricing in Transport (1996) and taken further in the white papers Fair Payment for Infrastructure Use (1998) and European Transport Policy 2010: Time to Decide (2001). According to the proposed approach transport users should be charged for all the additional costs they impose to society by their use of the transport system. Transport taxes and prices should thus vary according to infrastructure damage caused by vehicles, the degree of congestion, accident risk and environmental nuisances.

In the EU, the principles of such a charging policy are already implemented in the rail sector and a revision process is on-going for the road haulage sector: the so-called Euro-vignette directive[1] establishes the possibility to charge heavy road goods vehicles for the infrastructure (construction and maintenance) costs caused by their usage of the trans-European network. To manage congestion and environmental effects, charges may be differentiated to reflect the level of congestion and the sensitiveness of the environment (e.g. mountain passes).

Research activities under the Community Framework Programmes for research, technological development and demonstration activities (FP) have provided the analytical support to both the shaping of the pricing policy and its implementation. While the earlier FP projects have addressed more theoretical issues, such as operationalisation of the marginal cost principle, determination of optimal charging instruments and analysis of small-scale demonstrations, the focus has gradually shifted from theory to practical implementation of efficient pricing. More recently, demonstration projects in European cities have been carried out to test the acceptability of different pricing schemes and how to apply them.

The projects and expert networks supported by the Framework Programmes act as platforms for policy transfer. However, a clear analysis and understanding of this process is lacking so far. In this paper, we will first recall the principles of efficient pricing in transport and its socio-economic impacts. We will then briefly discuss problems of implementation and analyse the role of research networks for policy learning and knowledge transfer. The focus of the paper is on urban road pricing.

Principles of optimal pricing

External costs of transport

Good transport connections are essential for trade and economic development. Smooth, safe and secure transport systems reinforce sustainable economic growth and competitiveness and ensure smooth movement of passengers, goods and services. However, transport, and road transport in particular, causes several negative impacts to society as an unwanted and unintended side-effect of productive uses of the transport system. These so-called external costs include particularly the following:

- time wasted in congested traffic conditions,
- lives lost and injuries from accidents particularly on roads,
- health problems and premature deaths caused by air and noise pollution,
- lost biodiversity following the expansion of land use for new capacity to meet demand increases,

* Catharina Sikow-Magny (catharina.sikow@cec.eu.int) and Marcel Rommerts (marcel.rommerts@cec.eu.int) are both principal administrators at the Directorate General Energy and Transport of the European Commission.
The opinions expressed in this article are those of the authors and should not be attributed to the European Commission.
[1] The amendment of Directive 1999/62 is currently being discussed by the co-legislators.

– increasing contribution to climate change caused by "green house" gas emissions.

Several of the FP research projects have assessed the external costs of transport and developed methods for their monetary valuation. Following the state-of-the art methodology developed in the ExternE[2] series of projects, the UNITE[3] project has estimated the external costs of transport infrastructure use for all transport modes in different traffic and geographical contexts in Europe. The results indicate that the environmental cost of a lorry driving on the motorway from Basle to Karlsruhe is around 10 € cents/km whereas the cost doubles when driving in a city such as Berlin in day time and rises to almost 40 € cents/km at night because of higher noise annoyance. Regarding, for instance, the marginal emission cost of a ferry in the open sea, it is estimated to be 18 €/km or 1,622 € in total for a trip from Helsinki to Tallinn. For aviation, noise and air pollution costs amount to around 400 € for a Boeing 737–400 flight from Berlin, Tegel to London, Heathrow.

Towards a policy based on marginal cost pricing

The need to change the current pricing system stems from several factors:

– improved efficiency of the transport system,
– sustainable development and mobility,
– equity and cohesion issues.

In Europe the transport sector, in particularly road transport, is characterised by high fuel prices, mainly due to the various fuel taxes levied in most Member States of the European Union. On average these taxes represent around two thirds of the final consumer prices. The external costs do not, however, depend principally on the amount of fuel consumed but rather on the place and time of driving as well as the characteristics of the vehicle used. In addition heavy goods vehicles damage road surface and structures as a function of their axle load, and axle loading is not directly related to fuel consumption. Given its 'blunt' nature the fuel tax is an inefficient and therefore expensive means to achieve transport policy goals.

The policy of efficient charging aims to balance the level of the external costs and the cost of their mitigation. The European Commission's policy on transport pricing, as outlined in the various policy papers mentioned above, aims at a revenue neutral shift from taxation to charging transport at the point of use and according to the costs incurred.

The long term objective of efficient pricing in transport is to implement charging schemes that ensure that users of the transport infrastructure and services pay for the real costs they impose on society, such as costs related to infrastructure wear and tear, congestion, accidents and environmental damage. Marginal cost pricing, whereby charges reflect costs for actual use, is the preferred principle for such a scheme. The achievement of the long-term objective calls for an incremental approach, where a number of successive steps are taken. An application of the strategy has also to allow for differing speeds of implementation.

Socio-economic impacts of efficient pricing in transport

Economic theory has shown that marginal cost pricing leads to efficient allocation of society's resources. According to results from the EU Framework Programmes[4] and national research, implementing marginal cost pricing would be beneficial to the society. The overall welfare gain to the citizens ranges between 0.5–1.3 percent of the total regional income. Optimal pricing induces reductions in total transport volumes in the urban areas between 7–14 percent and in the inter-urban areas between 2–3 percent on average. In the peak, speeds increase by 32–70 percent in the urban and by 5–12 percent in the interurban areas. External costs of transport reduce by 13–35 percent in the urban and by 3–5 percent inter-urban areas.

The TRENEN[5] project analysed the impacts of efficient pricing in a number of European cities. For instance, in Amsterdam, with the current taxation and pricing regime, private cars account for 68 percent of all motorised trips[6] in the peak hours while the share of public transport is 32 percent. Because of severe congestion, the average speed on the roads is 13 km/h, which is only slightly higher than average speed by bike. Congestion pricing would enable rebalancing of the modes, private car's share would drop to 53 percent of trips and that of public transport would increase to 47 . The average speed on the roads would

[2] External Costs of Energy Use. For more information see http://www.jrc.es/pages/f-search.html.
[3] Unification of Accounts and Marginal Costs for Transport Efficiency, project to end early 2003, more details and intermediate results can be found at: www.its.leeds.ac.uk/research/index.html.

[4] See TRENEN, PETS, MC-ICAM, UNITE, IASON http://europa.eu.int/comm/transport/extra/web/index.cfm.
[5] Models for Transport Energy Environment, project completed, see final report at ports/strategic/trenen.pdf.
[6] Cycling is well developed in Amsterdam, however, non-motorised modes were not part of the modelling.

increase to 20 km/h. The benefits of efficient pricing more than outweigh its costs.

These modelling results have been recently confirmed by the London congestion pricing scheme.

However, cases of the practical implementation of marginal cost pricing has been rare in practice – except for the recent London congestion charging scheme and for some examples of heavy goods vehicle tolling systems in central Europe. The existing charging regimes typically rely on taxes that only remotely, if at all, reflect the underlying marginal costs, and revenue raising objectives.

Overcoming obstacles to implementation

A major requirement for a successful further development and implementation of efficient pricing is to pay greater attention to acceptance issues. At the EU level it is an issue of convincing citizens but also policy makers in the Member States to endorse or at least accept new charging schemes.

The EU, national governments and local authorities can use very different means when it comes to setting charges to transport and collecting general taxes but also when deciding upon regulations. Implementation of efficient pricing might imply changes in the respective roles and existing money flows. Also the ownership, degree of independence and charging principles for network infrastructure and terminals (ports, airports, etc.) differ considerably. Consequently, the necessary co-ordination and co-operation for different institutions at different decision-making levels might be difficult to achieve.

An interesting finding from the AFFORD[7] project is that politicians perceive the citizens to be more pro-car than they are in reality. The majority of citizens in the 6 cities[8] covered by the project could accept a well-structured package including pricing and other measures in cities where transport problems are perceived to be severe. The role of the media in influencing the discussion and public opinion on road pricing has not been sufficiently considered to date. For a pricing strategy to be accepted by citizens and by businesses it should comprise of the following elements:

- the objectives of the strategy have to meet main public concerns as to transport problems;
- the proposed strategies have to be perceived as effective solutions to the problems;
- revenues must be hypothecated and alternatives provided;
- fairness and distribution of costs has to be addressed meaningfully;
- people must have confidence in the effectiveness of the measures, use of revenues and anonymity of the system and
- there must be an "intelligent" marketing strategy.

Policy learning and transfer through EU-supported research projects and networks

Research projects and networks supported by the Framework Programmes may act as platforms for policy learning and policy transfer and may contribute to the emergence of a "pre-governance system" in the (urban) transport sector. However, there is no clear understanding of how these issues are related and of how they interact. One of the reasons is that the development over time of the project networks themselves, and the actors involved in them, has not been analysed. Academics with a background in science and technology assessment and political science have however discussed these issues in more general terms.

The process of European integration on S&T policy has been accompanied by the formation and organisation of interest groups at the European level. These groups represent the public and private sector and include scientific and economic interests (Grande and Peschke 1999). The practical collaboration takes place within, separate, "sub-networks". Luukkonen (1998) argues that EU research funding may act as a catalyst in the creation of new network configurations and considers the creation of networks a major impact of the EU framework programmes: she also suggests that there is an accumulation and continuity in many research networks.

Projects and networks can work where bureaucracies and/or markets do not yet do so. A policy network is a cluster or complex of organisations that are connected to one another by resource dependencies (Rhodes 1986). The Marsh and Rhodes (Rhodes 1997) approach to policy networks defines two types of policy networks: policy communities and issue networks. The actors involved in EU research projects on pric-

[7] Acceptability of Fiscal and Financial Measures and Organisational Requirements for Demand Management, project completed, see final report at http://europa.eu.int/comm/transport/extra/final_reports/urban/afford.pdf.
[8] Athens, Dresden, Edinburgh, Helsinki, Madrid and Oslo.

ing seem more a policy community, which is reflected by their shared values, the high quality of interaction and the continuity of the interaction over time.

The relationship between policy networks and policy transfer at the international level can be integrated through the notion of a policy transfer network (Evans and Davies 1999). Policy transfer is defined by Dolowitz (2003) as the process by which the policies and/or practices of one political system are fed into and utilised in the policy-making arena of another political system. Wolman and Pace (2002) suggest that policy transfer, since it is a form of policy learning that is done by governments, is a form of "organisational learning". A clear scheme for measuring the occurrence of policy transfer is lacking, as mentioned by Dolowitz and Marsh (1996) and by Evans and Davies (1999). In the view of Evans and Davies a policy transfer network is an action-oriented phenomenon set up with the specific intention of engineering policy change.

There is a possible linkage between networks, EU research funding and Aspinwall's concept of "pre-governance" in which the Commission informally gathers the participants of a possible future governance system (Aspinwall 1999). Aspinwall suggests that by engaging "industry and consumer groups" in research and discussions the Commission is actively establishing a constituency in the newer, undefined areas of transport policy, which are not yet regulated at the European level. This is the case for urban pricing. The networks help to define the nature of the problem to which the Commission later can offer potential solutions.

Formal network activities under the Framework Programmes

To ensure wide dissemination of results but also to allow for a debate among practitioners, a concerted action on pricing, CAPRI[9] was set up in 1998 under FP number 4. It aimed at co-ordinating research activities on pricing, disseminating and discussing the results of FP4 projects with the Member States, and identifying points of consensus and disagreement.

In FP5, CAPRI was followed by a thematic network project IMPRINT-Europe[10]. The overall aim of the network was to facilitate a two way process through a series of seminars, which enabled research results to be fed into pricing policy implementation, and to inform researchers about relevant policy questions. In this way it was intended to improve the quality of both policy making and research and to promote consensus by improving understanding of alternative points of view. Networking activities have continued also under FP6 where the focus is more on practical implementation and modal applications.

In parallel, support has been provided to a number of cities that intend to implement pricing in the near future, working together in the EuroPrice[11] network. The network brought together local government authorities and politicians in various countries to share information about the different strategies, technologies, information campaigns, and investment strategies employed in their efforts to introduce urban road pricing schemes. Three technical papers have been produced dealing with the full range of most relevant political questions about road user charging.

During FP5 the CUPID[12] thematic network developed and implemented a sound pan-European evaluation framework for urban pricing demonstrations. It worked closely together with the PROGRESS project described later. CUPID also had the role to facilitate the exchange useful information, to organise workshops and prepare papers on key issues and to run dissemination activities.

Research projects on urban pricing

In addition to the formally established networks described above, networking activities have also taken place through research projects on urban pricing. The community of researchers and cities involved in the different projects has remained rather stable over time. The research on urban pricing that was sponsored by the FPs during the second half of the 1990s specifically looked at real life impacts and acceptability. The projects have led to the conclusions described below.

Urban pricing measures do have an impact in real life

While the projects described before look at transport pricing on the aggregate, systemic level, several research projects have also looked at the implemen-

[9] http://www.its.leeds.ac.uk/ projects/capri/.
[10] http://www.imprint-eu.org/.

[11] The EuroPrice cities are: Amsterdam, Belfast, Bristol (coordinator), Edinburgh, Genoa, Leeds and Rome. For more information see http://www.europrice-network.org/.
[12] http://www.transport-pricing.net.

tation of specific pricing measures in practice. These projects use partial demonstrations or existing real life cases, sometimes in combination with modelling, to gain more detailed insights on user reactions and political barriers.

For example, CONCERT-P[13] assessed the efficiency and acceptability of integrated pricing and restraint measures, time-dependant and vehicle-based tolling and pollution-based pricing. The study results confirm the hypothesis that pricing measures are able to improve the usage of road space and can persuade motorists to switch to using public transport. In Trondheim, increased tolling reduced car use by 6.2 percent while increasing park and ride by 1.4 percent and public transport use by 4.2 percent. The Bristol case indicated that a daily toll combined with free public transport would reduce traffic flows up to 25 percent. Pricing measures also provide a new source of income. The tolling system in Trondheim has raised 125 million € in its first 12 years, all reinvested in transport and environmental measures.

The EUROTOLL[14] project validated the potential of pricing measures to combat congestion in 14 cases. The findings indicate that car users, in particular leisure travellers, shift travel times to off-peak or change routes if price differentials provide an incentive to so. Modal shift or a reduction in the overall number of trips was however found to be negligible. In Leicester, reactions to road pricing on an arterial road combined with P&R and a shuttle to the city centre were twofold: a number trips were shifted to the shuttle, while some of the car users preferred to avoid the charge by using secondary by-pass streets.

These projects demonstrated that pricing measures can change people's behaviour and travel patterns. The changes in behaviour don't need to be dramatic to have a noticeable effect on traffic conditions. Car users appear to change the timing, route or destination of their trip more easily than their travel mode.

Urban pricing schemes can be made acceptable

Acceptability by the public and at the political level is seen by many as the main obstacle for implementing marginal cost pricing in practice. The reasons behind opposition to efficient pricing are manifold: the distributional impacts are considerable, the benefits are not obvious to transport users, and behavioural changes might be induced affecting people's daily habits and even their economic or location choices. Several research projects have looked at acceptability and defined actions to increase acceptability.

The PRIMA[15] project has identified the criteria that foster acceptance or result in non-acceptance, of urban road pricing schemes by analysing successful cases as well as failures. The project showed that citizens consider road pricing as a means of last resort, i.e. traffic problems must be severe and no other way of solving them can be identified. In addition, road pricing should be a part of a policy package with clear, simple and transparent content and objectives. Other critical factors for acceptance are the level of the charges in the beginning, the distribution of the effects over different groups and the availability of travel alternatives. Adequate communication and transparency in the money streams are crucial.

The results of the PROGRESS[16] project show convincing financial, traffic and environmental impacts but at the same time demonstrate that the introduction of urban pricing is difficult. The PROGRESS demonstrations were partly of high profile and had obtained strong political support. The planned city-wide schemes in Bristol and Edinburgh did not take place. The area-wide electronic access control and parking pricing in Rome, an example of a hybrid pricing scheme, was successfully implemented. The tests of pricing concepts based upon satellite vehicle location systems with volunteers in Copenhagen and Gothenburg showed that the urban applications of this technology needs further development.

In the future research on urban pricing it will be important to begin to look beyond the current policies and marginal cost pricing in isolation and to assess the role transport pricing has in the overall transport policy development. The Commission's CIVITAS[17] Initiative is a major step in this direction and combines the objectives of the FP's transport and energy research programmes. The aim of CIVITAS is to assess the impacts on congestion, energy consumption, noise and air pollution of the introduction of ambitious integrated sustainable urban transport policy packages, involving pricing and other measures, technologies and infrastructures.

[13] http://btsa.es/concert/.
[14] http://www.cordis.lu/transport/src/eurotollrep.htm#7.
[15] ftp://ftp.cordis.lu/pub/transport/docs/summaries/road_prima_report.pdf.
[16] http://www.progress-project.org.
[17] http://www.civitas-initiative.org.

Conclusion

The European Commission has advocated the reform of transport pricing for over a decade. Research activities under the Community Framework Programmes for research, technological development and demonstration activities have provided the necessary analytical support to the development of the pricing policy as well as its implementation. They have helped to develop the principles of optimal pricing, supported formal network activities of experts and practitioners and supported projects testing the practical implementation of urban pricing at city level.

Research projects and networks supported by the Framework Programmes may also have acted as platforms for policy learning and policy transfer and contributed to the emergence of a "pre-governance system" in the urban transport sector. In this paper, we have attempted to explore relationships and interactions of these cases. However, more work is still needed to better understand these interlinkages. One of the reasons is that the development over time of the project networks themselves, and the actors involved in them, has not been analysed. This offers a clear opportunity for further analysis by academics with a background in science and technology assessment and political science.

References

Aspinwall, M. (1999), "Planes, Trains and Automobiles – Transport Governance in the European Union", in B. Kohler-Koch and R. Eising, eds., *The Transformation of Governance in the EU*, Routledge, London.

Dolowitz, D. P. (2003), "A Policy-Maker's Guide to Policy Transfer", *Political Quarterly* 74 (1), 101–8.

Dolowitz, D. and D. Marsh (1996), "Who Learns What from Whom: a Review of the Policy Transfer Literature", *Political Science* 10, 23–42.

European Commission (2001), *European Transport Policy 2010: Time to Decide.*

European Commission (1998), *Fair Payment for Infrastructure Use.*

European Commission (1996), *Towards Fair and Efficient Pricing in Transport.*

Evans, M. and J. Davies (1999), "Understanding Policy Transfer: A Multi-level, Multi-disciplinary Perspective", *Public Administration* 77 (2), 361–85.

Grande, E. and A. Peschke (1999), "Transnational Cooperation and Policy Networks in European Science Policy-making", *Research Policy* 28, 43–61.

Luukkonen, T. (1998), "The Difficulties in Assessing the Impact of EU Framework Programme in Research", *Policy* 27, 599–610

Rhodes, R. A. W. (1997), *Understanding Governance*, Open University Press, Maidenhead.

Rhodes, R. A. W. (1986), *The National World of Local Government*, London, Allen & Unwin, London.

Wolman, H. and E. Pace (2002), "Policy Transfer among Local Governments: An Information-Theory Approach", *Governance* 15 (4), 477–501.

Not Optimal, but Effective: the Multi-mode Ticket for Reducing Urban Traffic Congestion in Medium-sized Towns*

Ronnie Schöb*

Economists have long been aware of congestion externalities and proposed very sophisticated road pricing schemes to internalize the external costs of driving in urban regions. An optimal solution could be achieved by levying road user fees that exactly equal the additional costs that each car driver imposes on others. Since these costs depend among many other things, on location, time and type of car, an optimal policy requires, in principle, individual, time-dependent and locally differentiated road user charges (see Lindsey and Verhoef 2000 for a recent survey). Apart from Singapore, where such a differentiating scheme has been introduced and some far-reaching plans for Hong Kong that have been shelved yet again, not much further action has been observed.

Pragmatic second-best solutions such as cordon fee systems are more promising. Since London introduced its cordon toll system in the spring of 2003, many local administrations have renewed their interest in such road pricing schemes and are eagerly studying the first results. Nevertheless, such low-cost and easy-to-implement solutions are not very popular among scientists. Engineers promote the most sophisticated technology available, which normally requires huge set-up investments. The German highway toll system is a good example in this respect. Economists often do not dare to propose simple rule-of-thumb policies, which, although they may reap huge welfare gains, fail to meet the professional standard of providing optimal or almost ideal solutions to current economic problems.

This article introduces the idea of the "multi-mode ticket" that is described in more detail in the book *Alleviating Urban Traffic Congestion* (Arnott, Rave and Schöb 2005). The multi-mode ticket may be attractive for small and medium-sized towns that often face severe traffic jam problems and are interested in low-cost solutions. In the late 1980s, it was the Stockholm Urban Traffic Committee that came up with the idea of what we call the multi-mode ticket: each driver entering the inner city by car on workdays between 6.00 a.m. and 6.00 p.m. should provide a valid ticket for the Stockholm public transportation system. This could be a monthly ticket for commuters or a single return ticket for occasional travelers. When paying this "entrance fee", each car driver would receive a vignette that had to be put on the windshield for easy monitoring by traffic wardens in the inner city zone. Expectations were that this would reduce peak-period time traffic by 13 percent and total traffic by 9 percent and that the additional revenues would outweigh the administration costs by a quintuple. Although in the end it was not implemented the idea spread and for some time gained political and public appeal beyond Sweden. In Germany, the Green party suggested this scheme for medium-sized towns above 100,000 inhabitants. Since the implementation in both Sweden and Germany would have required a change in federal law, however, the idea completely disappeared from the political agenda. Apparently, modesty in political consulting became outdated and so did the Stockholm proposal. This was a pity, not only for Stockholm but for many other towns as well. Stockholm could have alleviated traffic congestion substantially for the last 15 years. Instead it is still waiting for more sophisticated road pricing systems.

The economic incentive structure of the multi-mode ticket is very simple and appealing for at least two reasons. First, even though the multi-mode ticket might not increase the marginal costs for a car trip downtown, it would reduce the opportunity costs of traveling downtown by bus or train thus inducing substitution of public transportation for private trans-

* This article is a condensed and non-technical version of Chapter 5 of Arnott, Rave and Schöb (2005).

** Ronnie Schöb is Professor for Economics at Otto-von-Guericke University Magdeburg (ronnie.schoeb@ww.uni-magdeburg.de).

portation. Second, since the policy replaces a ticket paid only by bus passengers with a ticket paid by all those entering the inner city by bus or car, revenues would go up dramatically. This would allow the local authorities to both reduce the deficit and increase the capacity of the public transportation system, which in turn would increase the public support for the scheme.

After a brief review of the external cost of urban traffic congestion, we will analyze the economic incentive structure of the multi-mode ticket in more detail and discuss ways to implement such a low-cost solution in medium-sized towns that suffer from severe traffic congestion.

Road congestion externalities[1]

There are several different externalities of urban traffic congestion. First, there are time-related marginal external congestion costs. Every car driver on a congested urban road imposes a cost on all other drivers by slowing them down. For a given road capacity, up to a certain traffic density speed is unaffected when one additional driver enters the road. But when a critical level is reached, each additional driver reduces the average speed and increases the travel time for all other drivers. Estimates show that the time-related marginal external congestion costs are on average between US$4 and US$6. Apart from time-related externalities, road congestion is responsible for several other externalities.

- Congestion also affects the probability and severity of traffic accidents. External costs arise if the total accident costs rise more than proportionately with respect to the traffic flow. Earlier studies pointed in this direction (Vickrey 1968; Newbery 1988), but some more recent studies have found that average accident costs (not including delay costs) may actually fall with traffic flow indicating a marginal external benefit of congestion (cf. Fridstrøm and Ingebrigsten 1991; Zho and Sisiopiku 1997). Speeding up traffic will reduce the frequency with which accidents occur, but at the same time will increase the severity of accidents. In heavy traffic, people drive more slowly and accidents cause fewer casualties.
- Congestion also leads to extra fuel consumption. The running costs on urban arterials are about 40 percent higher than on highways and rise quickly when congestion reduces speed to below 20 miles per hour on an urban arterial (Small 1992, p. 76). Thus, a car driver imposes not only time costs on other car drivers but also extra fuel costs.
- Closely related to additional fuel consumption is congestion-related pollution. Even though pollution depends in a very complex way on both the number of cars and the average speed on a congested road so that it is very difficult to find a significant statistical relationship (cf. Small and Gómez-Ibáñez 1999), pollution costs (including costs from noise) are normally assumed to increase with traffic flow. Small and Kazimi (1995) estimated that the total costs of pollution add up to 3.3 cents per mile in the Los Angeles metropolitan area.
- Finally, it should be noted that congestion fees can reap an additional second dividend that arises when tax revenues from congestion are used to cut other taxes that cause a deadweight loss. Given the huge inefficiency of the existing tax systems, the prospective welfare gains from using congestion fee revenues to reduce the welfare costs of taxation can be significant. Parry and Bento argue that "there is drastically more at stake in terms of economic welfare in what the government does with the congestion tax revenues than the entire welfare gains from internalizing the congestion externality" (Parry and Bento 2001, p. 662). Their simulations for linear demand and marginal cost curves indicate that the second dividend from reducing other distorting taxes is almost four times as high as the welfare gains from reducing congestion externalities.

By using the time-related marginal external congestion cost of US$4 to US$6 as a lower bound for the total marginal external cost, we can compare the cost of car commuting with the cost of commuting by bus. A survey of operating costs and the deficits of the public transportation authorities of some selected medium-sized towns in Europe and North America shows that all public transportation systems run a deficit that is normally covered by the municipality. Thus, municipalities have a genuine interest in raising additional revenues to cover these costs. The operation costs calculated per round trip are in the range of 100 to 160 percent of the price of a daily round-trip ticket which are in the range of US$2 to US$4 in most medium-sized towns. Since these average operating costs give us an upper bound for the marginal costs of bus commuting, this is also the range for the maximum marginal costs of a round trip not yet borne by the

[1] More detailed surveys can be found in Small (1992, pp. 78-85) and Small and Gómez-Ibáñez (1999).

passenger. Hence, some substitution of bus commuting for car commuting promises to improve welfare.

The theory of the multi-mode ticket

We focus our analysis on work-related traffic congestion. Workers who live in the suburbs and work downtown are free to choose whether to commute on a congested road or to take the bus. The number of trips by car is denoted by a (for automobile), the number of bus trips are denoted by b (for bus). Commuting is time consuming, but the time spent in the car or the bus may be more or less valuable for commuters, i.e. commuters consider the two modes as imperfect substitutes. Commuters can either buy daily tickets or a monthly ticket. The price of a daily bus ticket is given by τ_b. If someone decides to buy a monthly ticket, the price for daily tickets becomes zero.

The worker's decision problem

The marginal time costs of the individual commuter equal the average commuting time costs $APC(a)$ for all commuters. For car drivers, it is $APC(a)$, whereby we define these costs exclusive of any road user charges. Each commuter takes this time as given, but the average commuting time is rising with the total number of commuters. The individual commuting costs exclusive of the bus ticket are defined accordingly as $APC(b)$.

In Figure 1, the horizontal distance between the vertical axes measures the total number of commuting trips. The number of bus trips is measured from left to right and the number of car trips is measured from right to left. Average costs per bus trip are rising since commuters prefer some variety and find it increasingly boring to permanently commute by bus. For car trips, the curve is increasing because congestion increases the average commuting time. Without any regulation, commuters will split transportation modes such that their cost of traveling by car, $APC(a)$, equal their individual cost of commuting by bus, $APC(b) + \tau_b$. Hence, without any road pricing we observe an equilibrium with a_0 bus trips and $1 - a_0$ car trips.

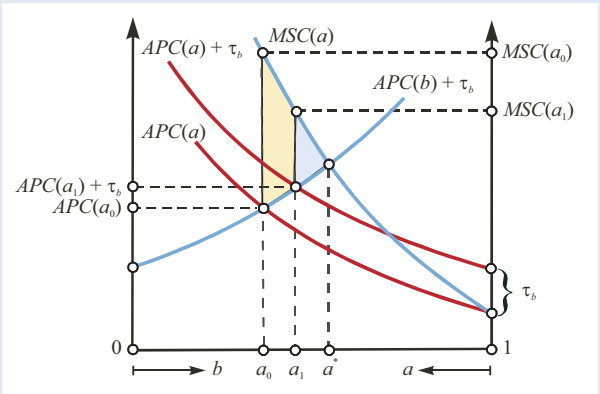
The multi-mode ticket and the day-by-day decision

If every commuter who drives downtown has to buy a daily ticket for the price of τ_b that allows him to travel either by car or by bus, the multi-mode ticket works like a congestion fee: it increases the price of commuting by car relative to the price of going by bus. Starting in the original equilibrium a_0 Figure 1, a commuter who – at the margin – decided to take the car before the multi-mode ticket was introduced, now faces higher car commuting costs and switches from car to bus. An additional $a_1 - a_0$ trip will now be made by bus rather than by car. Without the multi-mode ticket, the bus fare distorts the decision in favor of car commuting. This distortion is now eliminated irrespectively of the price of the multi-mode ticket. The multi-mode ticket thus unambiguously determines the allocation of a given number of commuting trips. Note however, that the introduction of the multi-mode ticket increases commuting cost, which in return may affect the total number of commuting trips.[2]

The marginal social cost of commuting by car is denoted by $MSC(a)$, the difference between $MSC(a)$ and $APC(a)$ indicates the marginal *external* costs of car commuting. In Figure 1, the trapezoid shaded in light gray then indicates the welfare gain from reduced traffic congestion that can be reaped when the multi-mode ticket is introduced. Thereby we consider the reasonable case we discussed in the last section that the marginal external congestion costs exceed the bus ticket price τ_b. The triangle in black indicates the remaining welfare loss with a^* being the optimal allocation. As long as the bus ticket price is low-

Figure 1

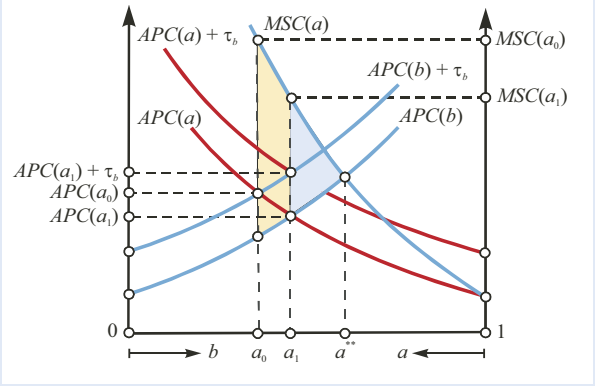
ALLOCATION OF ROAD TRIPS AND BUS TRIPS



[2] This effect is further elaborated in Arnott, Rave and Schöb (2005, chapter 5).

Figure 2

ALLOCATION OF ROAD TRIPS AND BUS TRIPS WHEN THE BUS TICKET IS A PURE MARK-UP



er than the marginal external costs, the multi-mode ticket fails to reach the optimal allocation, but the welfare effect of introducing the multi-mode ticket is unambiguously positive.

So far, we implicitly assumed that the bus fare equals the marginal cost of a bus trip. If the marginal costs of bus commuting were negligible, the bus fare would have to be considered as a pure mark-up. This would not affect the two allocations a_0 and a_1, but would affect the welfare analysis. This is shown in Figure 2.

The welfare gains become more significant when the bus ticket is a pure mark up over marginal cost of bus commuting. In this case, the $APC(b)$ curve represents the marginal social cost of bus commuting. (In Figure 1, by contrast, it is the $APC(b) + \tau_b$ curve which represents the marginal social costs of bus commuting.) In the initial equilibrium a_0, the allocation is the same as before but the welfare gain from introducing the multi-mode ticket increases by $\tau_b(a_1 - a_0)$. The area shaded in light gray in Figure 2 shows the welfare gain. Figure 2 also shows that the optimal allocation requires an even larger shift towards commuting by bus. In this case, the optimal allocation is defined by a^{**}.

The monthly multi-mode ticket

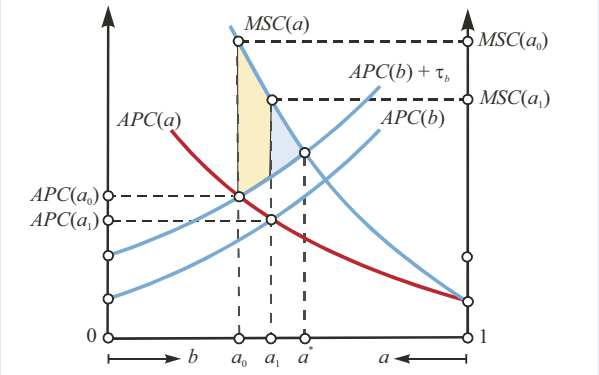
What happens if commuters decide to buy a monthly ticket rather than buying tickets on a daily basis? In contrast to the standard road pricing schemes, the possibility of buying a monthly (or annual) ticket does not affect the main property of the multi-mode ticket, namely the equivalence of the ticket price for commuting by car and bus. The monthly ticket does not affect the price of commuting by car. This remains at $APC(a)$. The costs for an additional trip by bus, however, fall to $APC(b)$ and the new equilibrium is at $APC(a_1) = APC(b)$. We therefore obtain the same condition for the optimal modal choice of the worker as before: for a given number of commuting trips, the equilibrium allocation for the monthly multi-mode ticket is the same as for the daily multi-mode ticket. Figure 3 shows both outcomes.

Compared to the daily ticket, however, commuting costs are now lower by the amount of τ_b This increases the incentive to commute compared to the case where only daily tickets are available. We may therefore expect an increase in total work-related trips, but workers will also substitute commuting by bus for commuting by car so that it will remain unclear *a priori* whether congestion increases or falls.

The properties of the monthly multi-mode ticket are very different from those of the cordon toll system in London that allows road users to pass the cordon when they have a monthly ticket. For a monthly ticket for passing the London cordon, the marginal costs of a trip by car do not change, but compared to the multi-mode ticket that allows for free bus trips as well, the marginal costs of bus commuting are higher by the bus fare τ_b. In Figure 3, a_0 indicates the cordon fee equilibrium for London for a given number of commuting trips. Compared to such a scheme, the

Figure 3

ALLOCATION OF ROAD TRIPS AND BUS TRIPS WITH A MONTHLY TICKET



monthly multi-mode ticket changes the relative prices for each trip decision. As monthly or annual tickets alleviate the administration significantly, without reducing the efficiency of the instrument, we have identified another advantage: the multi-mode ticket increases the price of car commuting relative to the price of the public transportation ticket irrespective of whether it is paid per trip or introduced as a periodical fee.

Two additional advantages of the multi-mode ticket are worth mentioning. First, in contrast to any cordon price scheme that does not change the price ratio between car trips and bus trips within the cordon, the multi-mode ticket changes the relative prices for trips within the cordon. To see this, we can reinterpret Figure 3 and consider the horizontal distance as the total number of trips within a cordon. The London cordon-fee equilibrium will be at a_0, irrespectively of whether the car driver paid a daily or monthly cordon fee. Although a cordon price scheme reduces trips to the center and through traffic, it does not reduce traffic within the cordon. The multi-mode ticket, by contrast, also creates an incentive to go by bus for another trip when the commuter has already passed the cordon. When car drivers have bought multi-mode tickets, they can use the bus for free in the whole cordon area. When deciding whether to undertake a trip within the cordon, the opportunity costs are therefore $APC(b)$ rather than $APC(b) + \tau_b$. The equilibrium for the inner-cordon mode allocation is at a_1 in Figure 3.

The second advantage is less apparent. Although the multi-mode ticket is not designed as a ticket that allows for price discrimination between peak times and off peak times, price discrimination may be achieved through variations in the public transportation capacity between peak and off peak times. If a bus runs every 10 minutes during peak time and only every 20 minutes during off peak times, the expected waiting time for the bus is 5 minutes lower during peak time. Thus the opportunity costs are lower in peak times than in off peak times. Of course, crowded buses and the risk of not being able to enter a crowded bus may work in the opposite direction. Another channel of price discrimination is offered by the existing price system of many local transportation agencies. They often offer special discounts for daily and monthly tickets if the bus is not used before 9.00 a.m. This price discrimination can be applied to car commuters as well. To guarantee effective monitoring, however, car commuters may not use this ticket before 10.00 a.m., otherwise cheating by those entering the inner city shortly before 9.00 a.m. may be too high.

From theory to policy

We will complement the analysis of the multi-mode ticket by briefly discussing how to implement a multi-mode ticket in practice. The first and most important question is how to determine the boundaries of the inner city. Initially, the cordon defined by the public transportation authority might be used to define the cordon of the multi-mode ticket. While the former is only defined by bus stops, the latter also requires borders for roads. As many towns apply a ring system for its public transportation system, one also has to clarify which rings should be included in the new scheme and which should not. Introducing a multi-mode ticket may lead to a rethinking of the public transportation cordon as it is important to include the most congested areas of the town. In this respect, designing an optimal cordon faces the same problems as designing an optimal cordon for any other road pricing scheme (see, for example, Newbery and Santos 2001).

Of particular interest when designing a cordon is the possibility for commuters to use park-and-ride facilities. The multi-mode ticket would certainly increase the value of park-and-ride systems many towns have installed or plan to build. For remote commuters, the multi-mode ticket makes it more attractive to drive to a park-and-ride facility where they can easily find a parking slot and from where a quick transfer to the center is ensured. These park-and-ride facilities will generate further time savings as they reduce the search for parking. As most park-and-ride schemes are available at no or low cost, they also allow the commuter to save the high parking fees downtown. To the extent that local authorities plan to enlarge existing park-and-ride facilities, they may earmark some of the additional revenues from the multi-mode ticket for this purpose.

The administration of the whole system can be based on the existing structures regarding the sale of bus tickets and the monitoring of parking in the center. In addition, fuel stations outside the cordon should be allowed to sell the multi-mode tickets. Commuters will learn quickly about the multi-mode ticket, but it might be a problem to provide all necessary information about the multi-mode ticket to out-of-towners. Here, a standardized symbol for "town with

multi-mode ticket" should be put on guide posts that are closer than, for example, 30 miles. This would give the occasional commuter the opportunity to buy a ticket in time.

Enforcement is easy with respect to commuter traffic. Local traffic wardens who watch for illegal parking can simply check whether all cars have a valid sticker on the windshield. More problematic, of course, is monitoring the through traffic, but occasional controls at lights within the cordon and at the cordon boundary, combined with high penalties for not providing a valid ticket, will reduce the incentives for free through traffic. Acceptability of these measures can be easily achieved with a tolerant policy in the initial phase where only *pro forma* tickets are issued to inform traffic offenders about the consequences non-adherence will have in the future.

Finally, it should be mentioned that marketing is important to gain political support. The Norwegian experience of selling the policy of introducing a price for a good that was free of charge before shows that the public must be informed in time about the intended introduction and its purposes. Information policy apparently must not simply highlight the advantages of reduced traffic congestion, it has to face people's expectation to be charged for things they wish to acquire, not for things they want to avoid (Jones 1998; Odeck and Bråthen 2002). Ison (2000) points out that earmarking revenue for investments in public transportation may be important for the public opinion. Hence, detailed ex ante plans on the improvement of park-and-ride facilities, the increased frequency in peak times and the modernization of the bus fleet may increase public acceptance.

The multi-mode ticket will not fully internalize the external costs of urban road congestion. Nevertheless, it offers towns with severe traffic problems a low-cost, easy-to-implement policy that promises significant welfare gains and is easy to administer and monitor. It may not be the final step towards more efficient urban transportation. A first modification may be to allow for a cheaper "bus-only" ticket alongside the multi-mode ticket. This allows for more price flexibility with respect to congestion pricing without changing the incentive structure. It may also be a promising start for many towns that, based on positive experiences with the multi-mode ticket, are considering whether to implement more sophisticated systems of road pricing.

References

Arnott, R., T. Rave and R. Schöb (2005), *Alleviating Urban Traffic Congestion*, MIT Press, Cambridge, MA.

Fridstrøm, L. and S. Ingebrigsten (1991) "An Aggregate Accident Model Based on Pooled, Regional Time-series Data", *Accident Analysis and Prevention* 23: 363–78.

Ison, S. (2000), Local Authority and Academic Attitudes to Urban Road Pricing, *Transport Policy* 7: 269–77.

Jones, P. (1998), "Urban Road Pricing – Public Acceptability and Barriers of Implementation", in K. Button and E. Verhoef, eds, *Road Pricing, Traffic Congestion and the Environment*, Edgar Elgar, Cheltenham, 263–84.

Lindsey, R. and E. Verhoef (2000), *Traffic Congestion and Congestion Pricing*, Tinbergen Institute Discussion Paper, no. 2000–103/3.

Newbery, D. (1988), "Road User Charges in Britain", *Economic Journal* 98, 161–76.

Newbery, D. and G. Santos (2001), *Urban Congestion Charging: Theory, Practice and Environmental Consequences*, CESifo Discussion Paper, no. 568.

Odeck, J. and S. Bråthen (2002), "Toll Financing in Norway: The Success, the Failures and Perspectives for the Future", *Transport Policy* 9:253–60.

Parry, I. and A. Bento (2001), "Revenue Recycling and the Welfare Effects of Road Pricing. *Scandinavian Journal of Economics* 103: 645–71.

Small, K. (1992), *Urban Transportation Economics*, Harwood Academic Publisher, Philadelphia.

Small, K. and J. Gómez-Ibáñez (1999), "Urban Transport", in P. Cheshire and E. Mills, eds., *Handbook of Regional and Urban Economics*, vol. 3,, *Applied Urban Economics*, North Holland, Amsterdam, 1937–1999.

Small, K. and C. Kazimi (1995), "On the Costs of Air Pollution from Motor Vehicles", *Journal of Transport Economics and Policy* 29, 7–32.

Vickrey, W. (1968), "Automobile Accidents, Tort Law, Externalities, and Insurance: An Economist's Critique", *Law and Contemporary Problems* 33: 464–87.

Zho, M. and V. Sisiopiku (1997), "Relationship between Volume-to-Capacity Ratios and Accident Rates", *Transportation Research Record* 1581: 47–52.

Land and Property Taxation in 25 Countries: A Comparative Review

Richard M. Bird and Enid Slack*

Every country has some form of tax on land and property. Such taxes have historically been local in most countries (although there are a few exceptions, such as Latvia and Chile, where they are mainly central taxes) and are often important sources of local revenue. One reason is that property is immovable – it is unable to shift location in response to the tax. Another reason is the connection between many of the services typically funded at the local level and the benefit to property values.

In a recent book (Bird and Slack 2004), we reviewed the main property taxes in the 25 countries listed in Table 1. Based on this comparative analysis, in this paper we discuss briefly the major policy alternatives with respect to taxing land and property – the choice of tax base, exemptions, methods of determining the tax base, tax rates, and the differential treatment of different classes of property (farms, residences, etc.).[1]

What is taxed?

Land vs. land and improvements

Table 2 summarizes the tax base in our 25 countries. In most countries, the property tax is levied on land and "improvements" (a term that includes structures, buildings, irrigation systems, and other man-made features). In a few, however, only the land portion of the property is taxed (e.g. Kenya). In Tanzania, unusually, only buildings are taxed. In countries where both land and improvements are taxed, the land portion is sometimes taxed more heavily than improvements.

In principle, a tax on land value only taxes location rents (the returns from a particular location regardless of the improvements to the site). Since improvements to land (such as structures) are not taxed, the owner has an incentive to develop the land to its most profitable use. Compared to a property tax on land and buildings that discourages investment in property, a site value tax thus encourages building and improvements. Assuming land is in fixed supply, a tax on land falls on landowners and cannot be shifted to others. Increased site value taxes will thus be capitalized into lower property values. Since the tax is borne proportionately more by owners of land and land ownership is unequally

Table 1

Property tax as a proportion of local revenues

	Property tax as % of local revenues
OECD:	
Australia	37.7[a]
Canada	53.3
Germany	15.5
Japan	25.5
United Kingdom	33.0[b]
Central & Eastern Europe:	
Hungary	13.6[c]
Latvia	18.2[d]
Poland	9.7
Russia	7.0
Ukraine	9.3
Latin America:	
Argentina	35.0[e]
Chile	35.1[f]
Colombia	35.0[g]
Mexico	58.7[h]
Nicaragua	6.4
Asia:	
China	4.9
India	7.0 to 41.0[i]
Indonesia	10.7
Philippines	13.4
Thailand	1.4
Africa:	
Guinea	32.0
Kenya	15.0
South Africa	21.0
Tanzania	4.0
Tunisia	32.4

For most countries, data are for 2000 or 2001.
[a] Includes only local taxation and not the state tax on land. – [b] Includes the local council tax and the local share of national non-domestic rates. – [c] Includes other local taxes such as a tourism tax. – [d] Percentage of local taxes. [e] This refers only to the municipal tax. There is also a property tax at the provincial level. – [f] The property tax is a national tax earmarked for local governments; 40 percent of revenues remain with municipalities where property is located. – [g] Property taxes as a percent of total Colombian local taxes. – [h] Percentage of municipal taxes. – [i] The range depends on the state.

Source: Bird and Slack (2004).

* The authors are, respectively, Professor Emeritus and Co-Director of the International Tax Program, Rotman School of Management, and Director of the Institute on Municipal Finance and Governance, Munk Centre for International Studies, both at the University of Toronto.

[1] Bird and Slack (2004) also discuss various aspects of tax administration (property identification, assessment appeals, and tax collection and arrears) as well as other taxes imposed on land, such as transfer taxes.

distributed, such a tax should be more progressive than a tax on land and improvements. Site value taxation thus scores well in terms of both equity and efficiency. Indeed, taxes on land are generally regarded as one of the least distortionary taxes, although more general taxes on property do of course distort decisions about improvements (investment) to property.

The valuation of land alone is difficult, however, because most urban real estate sales combine the value of land and improvements. The value of improvements thus needs to be subtracted to derive an assessed value for the land. While some consider such taxation unacceptably arbitrary, others argue that valuation of land alone is probably easier than valuation of property (Netzer 1998) and can often be estimated directly from sales and demolition records. The original arguments for site value taxation (George [1879] 1979) were made in a context in which cities such as San Francisco were growing rapidly. Land that was worthless one day was worth a fortune the next, owing largely to the rapid influx of population. Valuing land separately may be less of a problem in developing countries in which urban areas are growing rapidly (Bahl 1998). In many such countries, land and improvements are in practice assessed separately in any case, with land value being estimated on the basis of a land value map and building value in accordance with construction cost tables.

Another problem with taxing land only, however, is that, since the tax base is considerably smaller than the value of land and improvements combined, a higher and more distortionary rate is needed to generate comparable revenues.

Exemptions

In every country, some properties are excluded from the property tax base. Exemptions may be based on such factors as ownership (e.g. government), the use of the property (e.g. charitable purposes), or on characteristics of the owner or occupier (e.g. age or disability). In some countries, exemptions are granted by the central or state government; in other countries, exemptions are granted locally; in some, both levels can grant exemptions.

Common exemptions include government property,[2] universities, churches, cemeteries, public hospitals, charitable institutions, public roads, parks, schools, libraries, foreign embassies, and property owned by international organizations. In some countries, agricultural land and principal residences are also tax exempt.

Exemptions have been criticized on a number of grounds. First, to the extent that people working in exempt institutions use municipal services, they should be taxed. Second, the differential treatment between taxed and exempt properties has implications for economic competition among businesses and between businesses and government. Third, differential tax treatment affects location decisions, choices about what activities to undertake, and other economic decisions. Fourth, exemptions narrow the tax base and thereby increase taxes on the remaining taxpayers, reduce the level of local services that can be offered, or both. Finally, since the proportion of tax-exempt properties varies by municipality, disproportionate tax burdens are created across communities. This result is especially troublesome when higher-level governments determine what is exempt from local taxation.

If a case can be made for favoring certain property holders (such as churches and charitable organizations) to encourage their presence in the local community, these organizations should be rewarded directly with a grant rather than on the basis of their property holdings (Kitchen 1992). In the interest of transparency and accountability, all exempt property should still be assessed in the same way as other properties so that the value of the exemption is known. Only when this is done – which is unfortunately almost never the case in practice – will the full cost of land use for a particular purpose be taken into account in resource allocation decisions.

How is it taxed?

The next step is to determine the value to which the tax rate is to be applied. In general, two distinct assessment methodologies are used for property taxation: area-based assessment and value-based assessment, with the latter being divided into capital and rental value approaches (Youngman and Malme 1994). In addition, some countries use a system of self-assessment. Table 2 sets out the extent to which these approaches are used in the countries studied.

Area-based assessment

Under an area-based assessment system, a charge is levied per square meter of land area, per square

[2] In some instances governments make payments in lieu of taxes on their properties. Such payments are generally negotiated and are often much less than the property taxes would be.

meter of building (or sometimes "usable" space), or some combination of the two. Where measures of area are used for both land and buildings, the assessment of the property is the sum of an assessment rate per square meter multiplied by the size of the land parcel and an assessment rate per square meter multiplied by the size of the building. The assessment rates may be the same for land and buildings, or they may be different. For example, a lower unit value per square meter might be applied to buildings to encourage development.

A strict per unit assessment results in a tax liability that is directly related to the size of the land and buildings. With unit value assessment, the assessment rate per square metre is adjusted to reflect location, quality of the structure, or other factors. Market value has an indirect influence on the assessment base through the application of adjustment factors. For example, the assessment rate per square meter might be adjusted to reflect the location of the property within a particular zone in the city. Although the specific location of the property within the zone is not taken into account, properties in different zones will have different values. When the groups are defined narrowly enough, unit value begins to approximate market value. For example, a zone could be defined anywhere from an entire city to specific neighborhoods to properties on one side of a street.

As Table 2 shows, area-based assessments are commonly used in Central and Eastern Europe where the absence of developed property markets makes it difficult to determine market value. They are also used in parts of Germany (in the former GDR), China, Chile, Kenya, and Tunisia. A common example of unit-value assessment is in the assessment of agricultural land. In many countries, farm property is assessed per square meter, with the unit value varying with the location (region, accessibility to markets), fertility (irrigation, climatic conditions, soil conditions, hilliness), and sometimes with the crops grown.

Market value assessment

Market value (or capital value) assessment, used in all the OECD countries studied[3] and some others, estimates the value that the market places on individual properties. Market value is defined as the price that would be struck between a willing buyer and a willing seller in an arm's length transaction.

[3] The council tax in the United Kingdom uses a variation of the market value approach. See Bird and Slack (2004) for a description of how it works.

Three methods are commonly used to estimate market value:

- The *comparable sales approach* looks at valid sales of properties that are similar to the property being assessed. It is used when the market is active and similar properties are being sold.
- The *depreciated cost approach* values property by estimating the land value as if it were vacant and adding the cost of replacing the buildings and other improvements to that value. This approach is generally used when the property is relatively new, there are no comparable sales, and the improvements are relatively unique. The cost approach is also normally used to assess industrial properties.
- Under the *income approach*, the assessor estimates the potential gross rental income the property could produce and deducts operating expenditures. The resulting annual net operating income is converted to a capital value using a capitalization rate. This approach is used mainly for properties with actual rental income.

Rental value assessment

Under the rental value (or annual value) approach, property is assessed according to estimated (not actual) rental value or net rent. One rationale for using net rental value is that taxes are paid from income (a flow) rather than from wealth (a stock). In theory, however, there should be no difference between a tax on market value and a tax on rental value. When a property is put to its highest and best use and is expected to continue to do so, rental value will bear a predictable relationship to market value – the discounted net stream of net rental payments will be approximately equal to market value.

This relationship does not always hold, however. First, gross rents are often used rather than the economically relevant "net" rents that build in an allowance for maintenance expenditures, insurance costs and other expenses. Second, most countries tend to assess rental value on the basis of current use. There can thus be an important difference between market value and rental value. A property that is under-utilized would be assessed at a much lower value under the rental value approach than under the market value approach. From a land use perspective, a tax based on value in highest and best use is more efficient than a tax based on current use because it stimulates use to its highest potential by increasing the cost of holding unused or under-used land.

There are some problems with the use of rental value assessment. First, it is difficult to estimate rental value when there is rent control. Controlled or subsidized rents cannot be directly used to assess market rents unless the majority of properties are rent controlled. This has been an important problem in India. Second, because vacant land is not taxable under a tax based on rental value in current use (since there is no current use), an incentive is created for low return uses over high return uses. It may even become worthwhile to withhold rental properties from the market altogether.[4] If vacant properties are not taxed, the tax has to be higher on occupied properties to yield the same amount of revenue. These higher taxes further discourage investment.

In terms of tax administration, there are some additional difficulties with rental value (Netzer 1966). First, rental value is often difficult to estimate because there is little information on the annual rent of comparable properties for unique commercial and industrial properties. Second, net rents can be difficult to calculate because the distribution of expenses

[4] As noted above, if rental value were based on highest and best use, then vacant land would be taxable; the value would have to be estimated on the basis of other properties. Even if rental value were based on current use, it might be possible to assign a non-zero value to vacant land.

Table 2

Tax and assessment bases

	Tax base	Basis of assessment
OECD:		
Australia	Land or land and improvements	Market value or rental value or combination
Canada	Land and improvements (sometimes machinery included)	Market value
Germany	Land and improvements; farm properties also include machinery and livestock	Market value (rental income/construction costs); area in former GDR
Japan	Land, houses, buildings and tangible business assets	Market value
United Kingdom	Land and improvements; some plant and machinery	Market value for residential; rental value for non-residential
Central & Eastern Europe:		
Hungary	Unimproved value (plot tax); buildings (building tax)	Area or adjusted market value
Latvia	Land and buildings	Market value
Poland	Land, buildings, and structures	Area
Russia	Land for land tax; structures for property tax; assets for enterprise property tax	Area; inventory value of structures; value of assets
Ukraine	Land	Area
Latin America:		
Argentina	Land and buildings	Market value
Chile	Land and improvements	Area by location for land; construction value for buildings
Colombia	Land and buildings	Market value
Mexico	Land and buildings	Market value
Nicaragua	Land, buildings and permanent improvements	Cadastral value
Asia:		
China	Occupied land; land and improvements	Area; market value or rental value
India	Land and improvements	Mostly annual rental value; limited use of area and market value
Indonesia	Land and buildings	Market value
Philippines	Land, building, improvements and machinery	Market value
Thailand	Land and improvements (buildings and land tax); land (land development tax)	Rental value; market value
Africa:		
Guinea	Land and buildings	Rental value
Kenya	Land (but can use land and improvements)	Area; market value; or a combination
South Africa	Land and/or improvements	Market value
Tanzania	Buildings, structures or limited development[a)]	Market value (or replacement cost, if market value not available)
Tunisia	Land and improvements (rental housing tax); land only (tax on unbuilt land)	Area; rental value

[a)] Land belongs to the state and is not taxed; land rents are paid to the national government.

Source: Bird and Slack (2004).

between landlords and tenants differs for different properties. Third, assessors may not have access to rental income information because rental income is not always in the public domain in the same way as sales prices. Rental value assessment is used in a number of countries around the world, however, as Table 2 shows.

Area-based vs market-based assessment

Where it is possible to use market value, it is generally regarded as a better tax base. First, the benefits from services are more closely reflected in property values than in the size of the property. For example, properties close to transit systems or parks enjoy higher property values. Second, market value has the advantage of capturing the amenities of the neighbourhood, amenities that have often been created by government expenditures and policies. For example, two properties of identical size and age where one is located next to a park and the other is adjacent to a factory will pay the same tax under an area-based assessment system. A value-based assessment system would be fairer because the property next to the park would pay higher property taxes. Third, area-based assessment results in a relatively greater burden on low-income taxpayers than high-income taxpayers when compared to value-based assessment because it taxes all properties that are the same size the same amount, whether they are in high-income or low-income neighbourhoods. Similarly, older houses in a bad state of repair but with a large floor area will pay relatively high taxes. Furthermore, if a relatively poor neighbourhood becomes richer, there would be no tax change. A tax system that fails to take account of changes in relative values over time will result in inequities. If one value per square meter is chosen for all single-family homes, for example, and relative property values change as some locations become more desirable, then over a period of years inequities in the assessment system will result.

One advantage often claimed for unit value assessment is that property taxes tend to be less volatile than under market value assessment because they do not change when property values change. However, this "advantage" can also be a disadvantage, exacerbating inequities. It has also been argued that unit value assessment is easier to understand and cheaper to administer than value-based assessments particularly where the real estate market is not well developed. This is not true, however, for the multi-residential rental, residential condominium, commercial, and industrial properties that constitute the bulk of the tax base in most countries.

One problem with such properties, for instance, is what to include for tax purposes. Should atrium floors, servicing shafts, elevator spaces and so on be taxed even though they produce no direct revenue? Another problem is how to allocate shared facilities such as common entrances, halls, exits, aisles, atria or malls, among owners/tenants. Such common areas can be shared on the basis of the size of each unit relative to the total, the rent charged to each unit, or some other measure. A third problem in market economies has been the tendency towards the proliferation of multipliers that are applied to the area of improved property to reflect relative differences in value. In the Netherlands, for example, the system became so complex through such adjustments that it was finally abandoned (Youngman and Malme 1994).

At present, many transition countries employ some variant of area-based assessment. This choice no doubt reflects the nature of the available information on the physical area of building and land recorded in the old central planning records. As zones become more narrowly defined over time, however, it seems both likely and desirable that these systems will evolve into something closer to a market value system.

Self-assessment

Self-assessment requires property owners to place an assessed value on their own property. In Hungary, for example, the current local tax system is based on the principle of self-identification. Taxpayers are obliged to register and report their tax obligations to the local tax administration. In Thailand, self-declaration of property owners is made to local assessors who assess the self-declared value and identification in terms of how well it matches their data. Self-declaration of properties by landowners is also required in the Philippines, once every three years. The local assessor then prepares the assessment roll.

Where properties are assessed at market value and there is self-assessment, the taxing authority in some countries has the right to buy the property at the assessed value.[5] A system where the taxing authority can buy the property will only be credible if it actually can and will buy the property. In practice, this right

[5] Taiwan is an example (Youngman and Malme 1994, p. 12). This idea is an old one, used in Australia in the 19th century, for example, as noted by Bird (1974). It has seldom been effective.

seems to have been exercised only rarely, presumably because of the political and budgetary impossibility of large-scale property purchases. Tanzi (2001) has recently made a proposal along similar lines, that people should assess their own properties and then make the self-assessed values public. Anyone who wanted to buy their property at a price that exceeded the declared price, by some margin such as 40 percent, could make an offer. If the owner refused the offer, the bid plus a penalty would become the new assessment. Although appealing to economists, and frequently recommended in the past, such ideas on closer examination seem much less attractive on a number of grounds (Holland and Vaughan 1970) and have not proven acceptable in practice anywhere.[6]

Nonetheless, self-assessment is an appealing procedure to poor countries with little administrative capacity. It does not appear to require expert assessment staff, and it seems to be easy to implement. Indeed, in some cases, such as Bogotá, Colombia, self-assessment has at times appeared to be relatively successful in terms of increasing revenues from property taxes, albeit at a time of rapidly rising property prices. In general, however, self-assessment seems likely to lead to inaccurate estimates of property values, with a tendency toward under-estimation. It violates the principle of fairness on the basis of ability to pay because people with comparable properties will not necessarily pay comparable taxes. Generally, lower-valued properties have a lower rate of under-estimation than do higher-valued properties, making this assessment approach regressive (i.e. taxes are relatively higher on low-valued properties). Under-estimation also obviously erodes the size of the tax base, with the usual detrimental effects on tax rates and/or on service levels. In the end, there is no easy way to get people to tax themselves in the absence of a credible verification process. To minimize the obvious problems of under-statement associated with any self-assessment system, the government must be prepared to obtain (costly) expert assessments of individual properties in cases where it believes self-assessment is inaccurate.

At what rate is it taxed?

Tax liability is determined by multiplying the assessed value times the tax rate. Given the size of the tax base, the tax rate determines how much revenue the property tax will generate. Three major issues arise with respect to tax rates. Who sets them? Are they differentiated, and, if so, how? And, finally, how high are they?

Who determines the tax rate?

Tax rates are sometimes determined locally and sometimes by the central government. As shown in Table 3, there are very considerable differences between countries with respect to the extent to which local governments are free to determine tax rates. Sometimes rates are essentially set by the central government. Sometimes there is some local discretion, within centrally-set limits. Sometimes there is complete local discretion.

Where rates are determined locally, local governments first determine their expenditure requirements. They then subtract non-property tax revenues available (for example, intergovernmental transfers, user fees, and other revenues) from their expenditure requirements to determine how much they need to raise from property taxes. The resulting property tax requirements are divided by the taxable assessment to determine the property tax rate. Even where rates are locally determined, there are often limits placed on them by the central government. In Ontario, Canada, for example, tax rates imposed on non-residential property are effectively "capped" at present in many localities.

If a local government is to make efficient fiscal decisions, it needs to weigh the benefits of the proposed services against the costs of providing them. If local governments do not finance these services themselves, then the link between expenditures and revenues is lost and the choice of services will not be based on an accurate perception of their cost. Setting tax rates at the local level places accountability for tax decisions at the local level. Local determination of tax rates is particularly important in the many countries in which a senior level of government determines the tax base. Local tax rates may have to be set within limits, however, to avoid distortions. A minimum tax rate may be needed to avoid distorting tax competition. Richer local governments may choose to lower tax rates to attract business. With their larger tax bases, they can provide equivalent services at lower rates than poorer competing regions. The resulting location shifts are not always allocatively distorting, but they are generally politically unwelcome. In addition, a maximum rate may be

[6] For a brief review of the past history of this idea, and the problems with it, see Bird (1984).

needed to prevent distorting tax exporting, whereby local governments levy higher tax rates on industries in the belief that the ultimate tax burden will be borne by non-residents (Boadway and Kitchen 1999). Such tax exporting severs the connection between payers and beneficiaries and renders decentralized decision-making about taxing and spending inefficient.

Differentiated tax rates

Many local governments levy rates that differ by property class.[7] Different tax rates may be imposed for different classes of property (residential, commercial and industrial, for example). This system gives local governments the power to manage the distribution of the tax burden across various property classes within their jurisdiction in addition to determining the size of the overall tax burden on taxpayers.

Generally, where such variable tax rates are applied, properties are assessed at a uniform ratio (100 percent or some lesser percentage) of market value. Another and probably more common way to differentiate among property classes is through a classified assessment system, as in the Philippines. Under this system, classifications or types of property are differentiated according to ratios of assessed value, but a uniform tax rate is applied. In terms of accountability, variable tax rates would be more visible and easier to understand for taxpayers than a classified assessment system, which may, unfortunately, be one reason that differentiated rates are less commonly employed than differentiated assessment ratios. Indeed, even when assessment ratios differ substantially among classes of property, the differentiation is more often a matter of practice than of law and can only be determined by special study.

[7] Property tax rates can also vary according to the services received. In some jurisdictions, there is a general tax rate across the city and a special area rate or additional surcharge in those parts of the city that receive services only provided to them , for example, garbage collection, street lighting, transit etc.

Table 3

Characteristics of tax rate setting

	Different tax by property class	Local discretion over tax rates
OECD:		
Australia	Yes	Yes for local tax; limits on annual increase in revenues.
Canada	Yes	Yes (restrictions apply in some provinces)
Germany	Yes	Central base rates; locally determined leverage factors
Japan	No; assessment differentiation	Nationally set standard and maximum rates
United Kingdom	Two separate taxes	Residential tax only; tax ratios for bands set centrally
Central & Eastern Europe:		
Hungary	Yes	Yes, within legal limits
Latvia	No	No, but local governments can grant relief
Poland	Yes	Yes, subject to prescribed minimum and maximum rates
Russia	Yes	Yes, within narrow range set by senior governments
Ukraine	No	No
Latin America:		
Argentina	Yes	Yes
Chile	No	No
Colombia	Yes	Yes, subject to central government limits
Mexico	Yes	Yes
Nicaragua	No	No
Asia:		
China	No	No
India	Yes	Yes, subject to state restrictions
Indonesia	No	No, but can change valuation deduction
Philippines	No, assessment differentiation	Yes, subject to minimum and maximum rates
Thailand	Yes	No
Africa:		
Guinea	Yes	No
Kenya	Yes, but rarely differentiated	Yes
South Africa	No; relief mechanisms used	Yes
Tanzania	Yes	Yes
Tunisia	No	No

Source; Bird and Slack (2004).

Table 3 shows that in many countries tax rates are differentiated by property class, or there is assessment differentiation or tax relief for some classes of property. Variable tax rates (or other differentiation of property taxes among property classes) may be justified on a number of grounds:

- On the basis of fairness with respect to benefits-received, it can be argued that the benefits from local public services are different for different property classes. In particular, a case can be made on benefit grounds for taxing non-residential properties at a lower rate than residential properties (Kitchen and Slack 1993). Few examples of differentiation in this direction appear to exist, however.
- On efficiency grounds, it has been argued that property taxes should be heavier on those components of the tax base that are least elastic in supply. Since business capital tends to be more mobile than residential capital, efficiency arguments again lead to the conclusion that business property should be taxed more lightly than residential property. In reality, however, lower rates are generally applied to residential properties.
- Variable tax rates can also be used to achieve certain land use objectives. Since higher property taxes on buildings tend to slow development and lower taxes speed up development, a municipal policy to develop some neighbourhoods instead of others might support differential taxes in different locations as well as for different property classes.

An additional question about property tax rates is whether the tax is levied at a flat or graduated rate. In many countries, some graduation is in effect introduced by exempting low-value properties. In a few instances (for example, some provinces in Argentina) the tax rate increases with the value of the taxed property. In Thailand, the tax rate also increases, although in a way that results in rates being regressive. Many countries impose higher taxes on "idle lands" – though seldom with much effect (Bird and Slack 2004). Particularly in rural areas, some countries have occasionally attempted to use progressive land taxes as, in effect, proxy income taxes by attempting first to aggregate all land owned by a single person and then to impose a graduated tax. Such schemes have generally failed, however, owing both to the administrative difficulty of assembling the information – particularly when properties are located in different jurisdictions – as well as the political unreality of attempting to accomplish "land reform by stealth" in this way (Bird 1974).

The level of tax rates

One of the more striking features of land and property taxation in many developing countries is how low the tax rates are. Even in countries such as Argentina in which progressive rates are imposed, the top rate (on assessed value) seldom exceeds much more than 1 percent, and it is often lower. In Indonesia, for example, the centrally-set land tax rate is only 0.5 percent. Moreover, as a rule, the effective rate of property taxes is, owing to low assessment ratios and poor enforcement, much lower than the nominal or statutory rate. Other factors resulting in low effective tax rates in many countries are lags in reassessment and the inadequacy of adjustment for value changes. In the Philippines, for example, where the nominal rate is as high as 2 percent, the effective rate has been estimated at only 0.07 percent (Guevara, Gracia and Espano 1994).

Some special cases

Residential and non-residential property

In many countries, single-family residential properties are favored by deliberately under-assessing them compared to apartments or commercial and industrial property of comparable value; by legislating lower tax rates on these properties; or by granting special property tax relief measures in the form of tax credits, homeowner grants, or tax deferrals. The differential treatment of residences does not reflect the differential use of services by different property types. In many countries, single-family owner-occupied residential properties are presumably favored largely on political grounds: residential homeowners are much more likely to vote in local elections than are tenants.

There is little economic rationale for the usual higher taxation of non-residential property. Differentially higher taxation distorts land use decisions favouring residential use over commercial and industrial use. A similar rate on both uses would ensure that the choice is based on the highest and best use (Maurer and Paugam 2000). Special taxation of one factor of production (real property) may also distort productive efficiency by inducing a different choice of factor mix in producing goods and services.[8]

[8] As noted above, the central governments may need to establish minimum tax rates on non-residential properties to avoid distorting tax competition and maximum rates to avoid tax exporting.

Agricultural property

Finally, in most countries agricultural properties tend to be treated on explicitly favourable terms under most property tax systems.[9] In some countries, much agricultural land is simply not taxed. In others, rather than assessing farms at their market value (which reflects the highest and best use), farms are often assessed at their value in current use. The value of a farm for tax purposes is thus determined by its selling price if it were to continue to be used as a farm. Alternative uses of the farm (e.g. as a housing subdivision), or its speculative value, are not considered in the determination of value. Such favourable treatment of agricultural land is usually designed to preserve it from conversion to urban use. Basing the property tax on value in current use, however, is probably not sufficient to preserve farmland because the resulting tax differential is unlikely to be large enough to compensate for the much higher prices that would be paid if the land were converted to urban use (Maurer and Paugam 2000). Furthermore, favourable treatment of rural land can increase speculation at the urban fringe and hence end up increasing urban land prices.

References

Bahl R. (1998), "Land Taxes versus Property Taxes in Developing and Transition Countries," in Dick Netzer, ed., *Land Value Taxation: Can It and Will It Work Today?* Lincoln Institute of Land Policy, Cambridge, MA.

Bird, R. M. (1974), *Taxing Agricultural Land in Developing Countries*, Harvard University Press, Cambridge, MA.

Bird, R. M. (1984), "Put Up or Shut Up: Self Assessment and Asymmetric Information", *Journal of Policy Analysis and Management* 3, 618–20.

Bird, R. M. and E. Slack (2004), *International Handbook on Land and Property Taxation*, Edward, Elgar, Cheltenham, UK.

Boadway R. W. and Harry M. Kitchen, H. M. (1999), *Canadian Tax Policy, 3rd ed.*, Canadian Tax Foundation, Toronto.

Hernando de Soto (2000), *The Mystery of Capital, Basic Books*, New York.

George, H. ([1879] 1979), *Progress and Poverty*, Robert Schalkenbach Foundation, New York.

Guevara, M. M., J. P. Gracia, and M. V. C. Espano (1994), "A Study of the Performance and Cost Effectiveness of the Real Property Tax," Manila.

Holland, D. M. and W. Vaughan (1970), "Self-Assessment of Property Taxes", in A. P. Becker, ed., *Land and Property Taxation,* University of Wisconsin Press, Madison, WI.

Kitchen H. (1992), *Property Taxation in Canada,* Canadian Tax Foundation, Toronto.

Kitchen H. and E. Slack (1993), *Business Property Taxation*, Government and Competitiveness Project, School of Policy Studies, Queen's University, Kingston, Canada.

DeutKielMaurer R. and A. Paugam (2000), *Reform toward Ad Valorem Property Tax in Transition Economies: Fiscal and Land Use Benefits*, Land and Real Estate Initiative, Background Series 13, World Bank, Washington.

Netzer D. (1966), *Economics of the Property Tax*, The Brookings Institution, Washington DC.

Netzer D. (1998), "The Relevance and Feasibility of Land Value Taxation in the Rich Countries," in D. Netzer (ed.), *Land Value Taxation: Can It and Will It Work Today?* Lincoln Institute of Land Policy, Cambridge, MA.

Tanzi V. (2001), "Pitfalls on the Road to Fiscal Decentralization," *Working Paper* no. 19, Carnegie Endowment for International Peace, Washington, DC.

Youngman, J. M. and J. H. Malme (1994), *An International Survey of Taxes on Land and Buildings*, Kluwer Law and Taxation Publishers, Deventer.

[9] An interesting exception is the Philippines where, unusually, farm properties are taxed on a higher percentage of market value than residential properties.

An International Comparison of Selected Innovation Drivers

Lothar Funk and
Axel Plünnecke*

Contrary to the post-war period where growth and catching-up with the United States could largely be achieved through accumulation of production factors and from assimilating existing technologies, once European countries had moved closer to the technology frontier,[1] innovation has become the main engine of growth. In other words, the balance between imitation and innovation has shifted in favour of the second. At the heart of this ability to innovate lie all those factors that lead either to the introduction of new products (product innovation) or to the introduction of new production processes (process innovation). Additionally, a greater proportion of that innovation is radical rather than incremental (OECD 2004c, 82). According to the Sapir Report, "growth becomes driven by innovation at the frontier and fast adaptation to technical progress" (Sapir et al. 2004, 38).

The current study tries to shed some light on this hypothesis by analysing the 22 most important OECD countries within the last decade in order to answer the question, whether the economies at the technological frontier that had the highest rates of economic growth in recent years have indeed the best frameworks to innovate. Even if different international organizations – European Commission and OECD, for example – have conducted a range of very useful benchmarking exercises in the area of innovation[2] (European Commission 2004a; OECD 2004b), this question has not been analysed in depth by now to our knowledge.

Our study uses the United States as the benchmark economy.[3] The position of the US as the most technologically advanced country (European Commission 2004b, 174–177) results from the fact that it, in common with many European countries, has high hourly productivity rates, whilst it has, at the same time and in contrast to many European states, higher work volumes (Sapir et al. 2004, 34–35). Obviously, the trade-off between productivity and work volumes that is so apparent in much of Europe does not arise in the US.[4]

Our analysis demonstrates that, for countries at the technological frontier, the existing conditions for innovation have a significant effect on current growth rates. Human resources, financing possibilities as well as the institutional regulation of product and labour markets have a prominent position amongst those factors that influence innovation. At the technological frontier, those countries with high rates of growth also distinguish themselves from those with low rates of growth on important measures of innovation. Significant differences between the two groups of countries exist in the numbers of university graduates (including those in particularly important subjects, such as mathematics, sciences and technology – in short MST subjects), the availability of venture capital, the amount invested in information and communication technologies (ICT), the general conditions in which firms operate as well as in their demographic developments. Even if the other indicators of innovation are considered, it can be shown that the average ranking on such measures is correlated with economic growth rates after the end of the boom in the new economy.

Theoretical and methodological background

In contrast to traditional growth explanations, "modern theories emphasize research inputs and human

* Lothar Funk, University of Applied Sciences in Düsseldorf (lothar.funk@fh-duesseldorf.de); Axel Plünnecke, Cologne Institute for Economic Research IW (Pluennecke@iwkoeln.de).
This study is part of a research project by the Roman Herzog Institutes.

[1] Several European countries were able, above all in the immediate post-War decades, to pursue successfully a process of catching up. It was, thereby, possible to reach the technological frontier in important areas. However, the growth that is attributable to catching up came to an end, at the latest, at the beginning of the 1980s when the easy gains from imitating and incrementally improving existing technologies were exhausted and demand became saturated for the output of leading industries. A decrease in the importance both of imitating successful economies and of simple incremental innovations for economic growth are inextricably linked with the ending of catching up process (see Sapir et al. 2004, 35–37).

[2] Such benchmarks show the position of various national economies across a range of measures of innovation; they do not, however, provide an overall ranking based on those indicators that are most relevant for economic growth.

[3] Above all, we choose Germany as the base country, as it was the laggard among the countries analysed here. If growth rates over the last ten years are averaged out, then Germany, with, in real terms, an annualized growth rate in per capita GDP of only 1.2 percent, is last amongst 22 OECD countries.

[4] The US and other successful economies have managed to increase their productivity rates, despite increasing their rates of employment. This also applies to the low-skilled in the US. This implies that the US has obviously been more successful at transforming inventions in basic science into growth-enhancing innovations. High rates of hourly productivity alone are no indication of being a technological frontrunner. Several European countries, of which Germany is a good example, have high levels of labour productivity per hour; at the same time, however, they exhibit low levels of employment. Other countries that have higher employment levels have lower hourly productivity rates.

capital as the key drivers for long-run growth. They stress not only the importance of 'own' innovation but also the capacity to imitate and to absorb externally available know-how. Institutional factors and framework conditions are seen as an important part of the 'innovative system' in which innovative firms operate" (European Commission 2004, 175). There appears to be an emerging consensus that a narrow view of innovation ("science should somehow cause innovation") is inappropriate (Arnold and Thuriaux 2002, 1) and needs to be replaced by a broader view of innovation.

"Innovation is a fundamentally economic process. Schumpeter called innovation 'a new combination of factors of production'. This can be the result of an invention. But it can equally involve the exploitation of new natural resources, copying an idea from a distant market, or describing an old product in a new way. *Entrepreneurship – the act of making innovations – is not something related to science and research, but about changing the rules of the game in economic competition*. Exploiting an invention is, therefore, an important special case of innovation, but it is not the general case. Economists see innovative activity as a driver of economic development because it provokes imitation. Innovation forces competitors to react – often in creative ways involving improvement and 'innovating around' the first innovator's design to erode the 'supernormal' profit of the original innovator. Innovations give rise to changes in the economy, which may be several times larger than the effect of the original innovator. *The main driver of economic growth is therefore the process by which change diffuses through the economy"* (Arnold and Thuriaux 2002, 2).

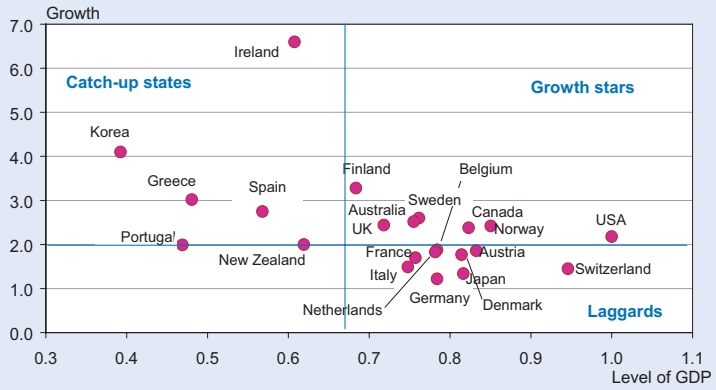
In other words, an economy's "national innovation capacity" is appropriately defined "as the ability of a nation to not only produce new ideas but also to commercialize a flow of innovative technologies over the longer term" (European Commission 2004, 175). It is vital that conditions for innovation and diffusion are approached as a whole (Arnold and Thuriaux 2003, 8). In the words of Jan Fagerberg (2003, 17): "Innovation processes are inter-temporal by nature. Current innovations depend on past innovations, and future innovations will depend on current innovations. This means that there may be a strong aspect of path dependency in innovation processes. Radical innovations open up new paths for future innovation activities and can dramatically influence what can be done profitably. Consequently: innovation and diffusion should be seen as an integrated process. It does not make sense to separate innovation and diffusion policy because a number of innovations occur in the wake of radical innovation and these only achieve economic impact as part of the diffusion process. What is important for innovation policy is to approach innovation as an integrated whole...; it is important to get on the bandwagon early, because as time goes by requirements become increasingly difficult to meet and unless one can jump on early it may be difficult to catch up at all because of the accumulated advantages associated with path dependency."

The arguments mentioned above have important repercussions for analysing the preconditions and effects of innovation empirically, as the incentive to engage in innovative investments, which involve risky experimentation and learning in particular at the technological frontier, is itself strongly affected by the economic environment (Sapir et al. 2004, 38). At least three implications result from such a broad perspective of innovation: firstly, it appears reasonable to link the rates of economic growth of countries to the conditions for innovation and diffusion, as is quite common nowadays (see e.g. Acemoglu et al. 2002; Fagerberg 2005; Fehn 2004). Secondly, although in each and every country there will be specific factors at work (see e.g. Boyer 2004), these will not be in focus here. Rather we will attempt to single out some general factors related to innovation capacity that may be of interest when debating the differences across countries in economic performance. These important innovation dimensions are human resources, financing conditions and more general framework conditions, for example in product and labour markets. The factors of innovation in this study are clearly multidimensional and, therefore, rather difficult to handle. Hence, we shall identify four reliable innovation input indicators for each of the three dimensions, express these in a comparable format and weigh them together, giving each indicator and dimension an equal weight in the calculation of the composite ranking indicator, which will be computed as an average rank of the single indicators in this article. Admittedly, there is an element of arbitrariness involved in such a calculation. This analysis may, for example, have a problem of omitted variable bias, which is, however, alleviated by using several different indicators for each dimension. Moreover, it would of course have been preferable to have prior knowledge about the true weights to use. Lacking such information, it appeared least ar-

bitrary to give each variable equal weight. Finally, it should be noted that the general approach we adopt here will – in line with the broad definition of innovation we follow – not only measures the direct effects of innovation activity (understood in a narrow sense as technological progress) on output but also indirect effects on output, such as reductions of technical inefficiency and improvements in allocative efficiency that cannot be separated empirically (see van Ark 2003, 10–13).

Figure 1

LEVEL OF ECONOMIC DEVELOPMENT AND ECONOMIC GROWTH[a)]



a) Average yearly growth rates in per capita gross domestic product over the period 1994 to 2004 in per cent and relative per capita gross domestic product in 1994 (USA = 1) in 22 OECD countries.

Source: Own calculations based on Eurostat (2005).

Innovation at the technological frontier: Catch-up states, growth stars and laggards

A recent study (Acemoglu et al. 2002) shows that the closer a country finds itself to the global cutting edge of technology, the more important radical innovations combined with fast adaptation to technical progress in all sectors become for economic growth. The basic conditions that promote the flow of radical innovations are of particular importance. By contrast, the importance of the ability to imitate others is severely reduced. According to this approach, the factors affecting innovation are of relatively little importance for economic growth in countries that are a long way from the technological frontier (catch-up countries); however, in national economies that are close to this technological frontier these factors are of decisive importance for growth. The following empirical analysis is based on this general empirical approach.

Per capita gross domestic product (GDP) will be a proxy for the technological advancement of a country. The US will be defined as the benchmark for technological advancement. Therefore, the US GDP per capita will be normalized to 1. In connection with their growth rates over the last ten years, the 22 OECD countries considered here can be grouped into three categories (Figure 1).

- **States that are catching up**: in the first group, countries can be found that have a relatively low level of per capita GDP and, at the same time, exhibit real GDP growth rates of more than 2 percent. These countries, such as Ireland, South Korea, Greece and Spain, can be described as states that are catching up. What also distinguishes these countries from most other countries is the fact that their rates of per capita GDP were less than two thirds of the US's level in 1994. This group of countries can profit significantly from imitating more successful economies and by incremental innovation.
- **Growth stars**: a second group of countries is characterized, firstly, by real GDP growth rates of over 2 per cent and, secondly, by already high levels of per capita GDP. This group can be described as growth stars. Over the last decade, Finland, Sweden, Australia, the UK, Norway, Canada and the US have belonged to this group. These states have, despite a relatively high technological level, high rates of economic growth.
- **Laggards**: a third group of countries, which are characterized by growth rates of less than 2 per cent, can be described as laggards. Belgium, Austria, the Netherlands, Denmark, France, Italy, Switzerland, Japan and Germany belong to this group. These countries are, because of their high levels of per capita GDP, undoubtedly highly developed economies. They, therefore, find themselves largely at the technological frontier. They have not, however, been able to take sufficient advantage of the opportunities for growth that have existed over the last decade.

Conditions for innovation

The growth stars as well as the laggards are, as groups of countries that are operating at the technological frontier, more dependent on radical innovations and fast adaptation to technical progress in all sectors for economic growth. This raises a couple of main questions. Firstly, what are the conditions that have the most favourable effects on growth? Secondly, did those economies that had the highest rates

of growth between 1994 and 2004 have the best frameworks for innovation and did countries which offer better conditions for innovation exhibit higher rates of growth in the later years of the period? By considering more recent years, influences will be excluded that, firstly, may have promoted innovation during the boom in the area of information and communication technologies (ICT) (the so-called new economy boom), but that may distort evaluations of the innovation factors in the medium term.

From a theoretical perspective, variations, in particular, in conditions for radical innovation play a central role in explaining the differences in growth rates between countries. These conditions cover:

Human resources. In the case of human capital, a paper by Aghion, Meghier and Vandenrusche (2003) demonstrates that, for 19 OECD countries between 1960 and 2000, the impact of the share of the population that is highly skilled on economic growth increases as the distance to technological frontier declines.

Financing possibilities. The role of finance and investment for innovation and growth is shown in papers by Levine (1997) and by Bassanini and Scarpetta (2002). The latter show that technological change – embodied by new ICT capital goods – has been a primary source of output and productivity growth in ICT-using sectors.

Table 1

An international comparison of human resources

Country	Persons with a tertiary education as a percentage of the population aged between 25 and 64		Persons with post-graduate research qualifications[a] as a percentage of the population aged between 25 and 64		Graduates in MST[b] per 100,000 employees aged between 25 and 34		Expenditure on education as a percentage of GDP	
	in %	Rank	in %	Rank	Value	Rank	in %	Rank
Growth stars								
Finland	33	4	1.9	4	1,785	2	5.8	11
Australia	31	6	1.3	8	1,659	4	6.0	8
Sweden	33	4	2.8	1	1,267	8	6.5	4
Norway	31	6	1.1	11	703	16	6.4	5
UK	27	11	1.6	6	1,727	3	5.5	13
Canada	43	1	n.a.	n.a.	855[c]	12	6.1	7
US	38	2	1.3	8	928	11	7.3	2
Laggards								
Denmark	27	10	0.9	14	799	14	7.1	3
Austria	14	20	1.7	5	528	20	5.8	11
Belgium	28	9	1.1	11	674	18	6.4	5
Netherlands	24	15	1.3	8	653	19	4.9	18
France	24	15	1.4	7	1,609	5	6.0	8
Italy	10	21	0.5	20	703	16	5.3	15
Germany	23	18	2.0	3	721	15	5.3	15
Japan	36	3	0.7	18	1,074	9	4.6	20
Switzerland	25	13	2.6	2	838[c]	13	5.3	15
Catch-up states								
Ireland	25	13	0.8	17	1,514	6	4.5	21
Korea	26	12	0.9	14	1,788	1	8.2	1
Greece	18	19	0.7	18	n.a.	n.a.	4.1	22
Spain	24	15	1.0	13	935	10	4.9	18
Portugal	9	22	n.a.	n.a.	n.a.	n.a.	5.9	10
New Zealand	30	8	0.9	14	1,497	7	5.5	13
Average values on the measures of human resources for the three groups of countries								
Growth stars	34	1	1.7	1	1,275	1	6.2	1
Laggards	23	2	1.4	2	844	2	5.6	2
Catch-up states	22	3	0.9	3	n.a.	n.a.	5.5	3

The data relate mostly to 2002.
[a] Those who have successfully completed a post-graduate research degree. – [b] Graduates of mathematics, engineering, the life sciences and technology. Values for Canada are for 2000 and for Switzerland are for 1998. – [c] 2000.

Source: OECD (12004e).

General framework conditions. The role of regulations for innovation and growth is shown in papers by Scarpetta and Tressel (2004) and by Nicoletti and Scarpetta (2003). There is evidence that high labour adjustment costs can have a strong negative impact on productivity. Such costs can, in particular, reduce incentives for innovation and the adoption of new technologies, and lead to lower productivity performance, when institutional settings do not allow wages or internal training to offset high hiring and firing costs. Reforms promoting private governance and competition tend to boost productivity. Both privatisation and market-entry liberalisation are estimated to have a positive impact on productivity.

As mentioned above, the performance of an innovation system cannot sensibly be explained by one solitary factor. Therefore, in order to evaluate the forces that promote innovation within a national economy, it is necessary to look at a bundle of factors as explained above. Due to the research design, it will not be possible, here, to cover the measures of the output of innovation that do not include the effects of research and development on the national economy. In other words, measures such as the number of new patents will not be discussed here. This paper aims to elaborate the link between basic input factors within an innovation system and economic growth per capita.

Human resources

The share of the population that is highly qualified is of particular importance for the ability to innovate, as mentioned above (Aghion et al. 2003). If the technological distance to the leading national economies is great, then the share of the population that is highly qualified has, empirically, hardly any role to play in influencing economic growth. This is because imitation and incremental innovation dominate. The share of the population that is highly qualified, the share of the population with a post-graduate research qualification, the numbers of successful candidates in mathematics, sciences and technology, and investments in education are important measures that can be used as indicators of a national economy's human resources.

If the three groups of countries are considered, it can be shown that the group of growth stars performs considerably better on these measures of human capital than the average for the group of laggards (Table 1). The share of the population aged between 25 and 64 that has a tertiary education is, at 34 percent in the former group of countries, nearly a third higher than the figure for the laggards. Similarly, the number of graduates in mathematics, sciences and technology per 100,000 employees aged between 25 and 34 is considerably higher amongst growth stars than it is amongst the laggards. Moreover, the former group of countries have a higher share of graduates who have completed post-graduate research degrees amongst those of typical graduating age. Indeed, the growth stars invest more, as a percentage of GDP, in education than the laggards.

The performance of the growth stars and the laggards differs significantly on the indicator measuring "persons with a tertiary education" and on the "MST" indicator. However, there is no significant difference between growth stars and laggards on investments in education and the share of the population with a doctorate.[5]

Financing possibilities

Technological advancement has been regarded, for a very long time, as a significant driver of growth; this is especially true in more recent growth theories that use two-sector models of the economy (Romer 1990; Grossman and Helpman 1991; Aghion and Howitt 1992). In such models, the goods sector of the economy manufactures products, and the factors of production (capital, labour and technological knowledge) are utilized. In the area of research and development (R&D), capital, labour and existing technological knowledge are similarly used; here, however, new technological knowledge is the output. By investing more in this sector, a higher rate of growth can be achieved; market imperfections can, however, lead to a sub-optimal supply of R&D (Romer 1996).

Expenditure on R&D as a percentage of GDP and the investment in ICT represent important indicators of investment and, therefore, the financing of innovations; this is particularly true for governments' promotion of research and knowledge in the area of basic science. This is, to a large extent, supported by governments and is carried out at universities and other research institutes (BMBF 2004, 167). One reason for the role of government in this area might be that free market mechanisms do not generate enough of such "products" because of their public-goods character. Moreover, state subsidies can fundamentally improve

[5] The p-value of the t-test that the means for the two groups are different is, for investments in education, 0.1 and, for the share of the population with a doctorate, 0.375.

Table 2

An international comparison of financing conditions

Country	Share of GDP spent on R&D[a] (2002)		Tax relief per R&D dollar spent in large firms[b] (2001)		Share of GDP that is available as venture capital (1998–2001)		Share of fixed-capital investments that is spent on ICT[c] (2000)	
	in %	Rank	in US$	Rank	in %	Rank	in %	Rank
Growth stars								
Finland	3.46[d]	2	–0.01	13	0.138	8	17.5	10
Australia	1.54	16	0.20	3	0.093	14	19.9	6
Sweden	4.27[e][g]	1	–0.01	13	0.207	5	21.6	3
Norway	1.67	15	–0.02	18	0.125	10	n.a.	n.a.
UK	1.88	14	0.10	8	0.219	4	22.8	2
Canada	1.91	12	0.17	4	0.251	2	21.4	4
US	2.67[h]	5	0.07	10	0.492	1	30.1	1
Laggards								
Denmark	2.52	7	0.11	7	0.082	16	19.1	7
Austria	1.93	11	0.12	6	0.044	21	12.8	16
Belgium	2.17[e]	10	–0.01	13	0.172	7	17.5	10
Netherlands	1.89[e]	13	0.10	8	0.241	3	20.9	5
France	2.20	9	0.06	11	0.119	11	12.6	17
Italy	1.11[e]	19	–0.03	21	0.076	17	16.1	12
Germany	2.52	7	–0.02	18	0.127	9	17.6	9
Japan	3.12	3	0.01	12	0.020	22	17.8	8
Switzerland	2.57[d]	6	–0.01	13	0.085	15	n.,a	n,a,
Catch-up states								
Ireland	1.15[e]	18	n.a.	n.a.	0.114	12	14.6	15
Korea	2.91[f]	4	0.13	5	0.202	6	n.a.	n.a.
Greece	0.65[e]	22	–0.01	13	0.059	20	15.7	13
Spain	1.03	20	0.44	1	0.095	13	15.5	14
Portugal	0.93	21	0.34	2	0.066	19	11.4	18
New Zealand	1.18[e]	17	–0.02	18	0.069	18	n.a.	n.a.
Average values of the financing measures for the three different groups								
Growth stars	2.49	1	0.07	2	0.218	1	22.2	1
Laggards	2.22	2	0.04	3	0.107	2	16.8	2
Catch-up states	1.31	3	0.18	1	0.101	3	n.a.	n.a.

[a] Research and development. – [b] Equivalent to the tax relief on every dollar spent on R&D in large companies. – [c] Information and Communication Technologies. – [d] 2000. – [e] 2001. – [f] R&D expenditures in the humanities and the social sciences have been excluded. – [g] Underestimated values. – [h] So-called capital expenditures have been excluded.

Source: OECD, 2004b; Cologne Institute for Economic Research IW.

the sub-optimal supply of innovations. However, for political-economic reasons and because of the risks of inefficient demands for subsidies so funding should be treated with care; this becomes increasingly important as the tasks of R&D become more application oriented (Farhauer and Henke 2002). A large amount of venture capital (even if it represents only a relatively small percentage of GDP) can, most readily, enable radical innovations which, as a result of the current structural changes in economies, obviously have a greater effect on growth than incremental innovations do.

Moreover, venture capital can increase the rate at which new companies are founded. Such companies, amongst other things, make a significant contribution to innovation as they are able to choose the most productive combination of factors of production. Unlike established companies, new firms do not have any costs associated with adaptation (for example, training, redundancies) when they introduce new technologies and production methods. The establishment of new companies also exerts competitive pressures on existing ones; the latter are then spurred on to create innovations (OECD 2004c, 88).

If an examination is made of the extent to which the growth stars and the laggards differ on the measures of financing and investment, the means across all four indicators are better for the growth stars than the laggards. The differences in average values for investments in ICT and venture capital are statistically significant.[6]

In the availability of venture capital in the early and expansionary phases the growth stars dominate.

[6] The p-value of the t-test on mean equivalence is, for tax relief on R&D, 0.4, and, for investments in R&D, just under 0.6; this indicates that there are no statistically significant differences between the two groups on these two measures.

Younger firms have there a far better climate in which to fund their expansions from external sources. For the growth stars, venture capital amounted to, on average, 0.218 per cent of GDP between 1998 and 2001. In the laggard countries, not even half that amount was available to establish new companies. These latter countries had to be satisfied with venture capital that amounted to approximately 0.1 per cent of GDP only. More recently, the group of growth stars have also invested in new ICT; these not only contribute towards higher productivity and growth, but also represent an important infrastructure for the wide diffusion of advanced technologies (OECD 2004c, 77). Whilst, for the growth stars, investment in ICT amounted to, on average, 22.2 percent of their fixed investment, this figure was a mere 16.8 per cent for the laggards.

Despite the fact that, on average, the growth stars perform better on indicators that measure investments in R&D and tax relief on R&D, there are states with low rates of growth that perform well on such indicators, and there are also states that have high rates of growth with relatively poor records in these areas. For instance, Switzerland, Japan, Germany, Denmark and France, which are all laggards, invest more in R&D than Canada, the UK, Norway and Australia, which are growth stars. Similarly, Austria, Denmark and the Netherlands, which are laggards, offer far greater tax relief on investments in R&D than Finland, Sweden and Norway, which belong to the group of growth stars.

General framework conditions

The general conditions for companies are of great importance for the potential success of new technologies as such investments are associated with great uncertainties. These risks can, in general, be managed better in a situation in which there are fewer labour-market regulations. Flexible markets, furthermore, create incentives for highly qualified young people to use their knowledge to create new knowledge; this means that they will not use their knowledge in socially unproductive rent seeking (Murphy, Shleifer and Vishny 1991).

Nonetheless, the growth stars Finland, Sweden and Norway have higher levels of labour-market regulation than many other growth stars. These latter countries are, however, typical small open economies with particularly high trade openness and labour demand elasticities that may be, due to lower insider power to set wages above market-clearing levels, not directly comparable in this respect with larger economies (Brandt et al. 2005, 65).

However, the other regulatory conditions for companies are better in the growth stars. A ranking of the conditions for firms (Matthes and Schröder 2005), which is based on World Bank data (World Bank 2004) and which does not take into consideration labour market flexibility, shows that the growth stars perform significantly better than the laggards and those states that are catching up. This measure covers data on the founding of companies, the registration of homes and property, the availability of credit, the disclosure requirements of public limited companies, the extent to which contracts are upheld, and the laws on insolvency.

Equally, the age structure of the population is of great importance for the willingness to innovate, as a younger population shifts the political majority in the direction of more innovative production methods (Gehrig 2000, 570–571). Of particular importance for the renewal of human capital is the ratio of pupils and students to the total number of employees. This indicator shows how strongly the basis of human capital as a share of total employment is being refreshed. It is, first and foremost, demographic developments that influence these factors. States whose populations are aging quickly have a paucity of younger people who, as carriers of newly created knowledge, are particularly important in driving radical innovations. A demographically more favourable age structure creates more impetus for innovation when the education system is of a high quality so that those starting work for the first time have a high level of competence. The results of the PISA studies for OECD countries in literacy, mathematics and the natural sciences are important measures of the quality of the education system (OECD 2004d).

To sum up, the growth stars perform on net better in their general conditions for innovation than both the laggards and those states that are catching up (Table 3). The growth stars exhibit significantly better averages for indicators that measure the general product, capital market and bureaucracy environments in which firms operate and that capture demographic developments.[7]

[7] The p-value of the t-test on mean equivalence is, for labour-market regulation, 0.122, and, for the quality of the education system as measured by PISA, 0.394. This means that the growth stars and laggards are not significantly different on these two measures.

Table 3

An international comparison of general framework conditions

Country	Labour market regulation (2003)		Conditions for companies[a] (2004)		Ratio of pupils and students to employees[b]		PISA results[c] (2003)	
	Index[d]	Rank	Index	Rank	in %	Rank	Value	Rank
Growth stars								
Finland	2.1	11	52.3	7	63	9	545	1
Australia	1.5	6	55.8	5	80	1	525	5
Sweden	2.6	17	52.6	6	64	7	510	10
Norway	2.6	17	59.2	2	59	13	493	16
UK	1.1	2	58.0	3	74	3	n.a.	n.a.
Canada	1.1	2	50.2	9	n.a.	n.a.	526	4
US	0.7	1	57.2	4	64	7	490	17
Laggards								
Denmark	1.8	8	45.9	14	56	14	494	15
Austria	2.2	12	44.6	17	52	19	496	14
Belgium	2.5	15	44.6	17	74	3	515	8
Netherlands	2.3	13	49.0	10	54	15	525	5
France	2.9	19	43.5	19	65	6	506	12
Italy	2.4	14	39.2	20	54	15	476	19
Germany	2.5	15	44.8	16	53	17	499	13
Japan	1.8	8	51.4	8	44	20	527	3
Switzerland	1.6	7	47.1	13	44	20	513	9
Catch-up states								
Ireland	1.3	4	48.4	11	70	5	508	11
Korea	2.0	10	44.9	15	61	10	538	2
Greece	2.9	19	31.3	22	60	11	466	21
Spain	3.1	21	47.4	12	60	11	484	18
Portugal	3.5	22	34.0	21	53	17	471	20
New Zealand	1.3	4	59.3	1	77	2	522	7
Average values for the measures of general conditions								
Growth stars	1.7	1	55.0	1	67	1	515	1
Laggards	2.2	2	45.6	2	55	3	506	2
Catch-up states	2.4	3	44.2	3	63	2	498	3

[a] Without labour market regulation. Range of the index: median of the individual measures = 50. – [b] Employees aged between 25 and 64. – [c] Results in literacy, mathematics and the natural sciences. OECD average = 500. – [d] Range of the index: from 0 (few regulations) to 6 (highly regulated).

Source: OECD, 2004a; 2004d; 2004e; Matthes/Schröder, 2004; Cologne Institute for Economic Research IW.

Overall evaluation of the drivers of innovation

If the average values for all the measures are considered, Canada performs the best. Then follow the US, the UK, Sweden, Finland and Australia (Table 4). It is only Norway from the group of growth stars that does not occupy a leading position. The national economies with a high per capita real income level and low levels of economic growth (the laggards) are positioned at the bottom of the list. From this group, Belgium performs the best. It is followed by Denmark, the Netherlands, Japan, Switzerland, France, Germany, Austria, and, well behind the other countries, Italy. The group of catch-up states is led by Korea, New Zealand and Ireland. These three countries perform better than the majority of the laggards. The catch-up states Spain, Portugal and Greece follow a long way behind, and are near the foot of the list.

Finally, the core hypothesis of this evaluation should be assessed. According to that hypothesis, the conditions for innovation should have a strong effect on the growth rates of those states that are close to the technological frontier. In other words, as explained in the basic theoretical approach above (Acemoglu et al. 2002), conditions for radical innovations are of significant importance for growth stars and laggards. Figure 2 supports this hypothesis. The average position for all of the measures of innovation used here (which mainly portray the conditions for innovation between 2000 and 2003) has a statistically significant effect on the growth rate in per capita GDP between 2000 and 2004.

The current study aims at selecting, according to theoretical considerations, indicators from a range of innovation benchmarks provided by the OECD; the selected measures can then be used in an analysis to

Table 4

The composite ranking of 22 OECD countries

Country	Average rank
Canada	5.7
US	5.8
UK	6.2
Sweden	6.6
Finland	6.8
Australia	6.8
Korea	7.3
New Zealand	9.9
Belgium	10.5
Denmark	10.8
Netherlands	11.0
Japan	11.2
Switzerland	11.5
France	11.6
Norway	11.7
Ireland	12.1
Germany	12.9
Spain	13.8
Austria	14.3
Portugal	17.2
Italy	17.4
Greece	18.2

The data relate to 2002.

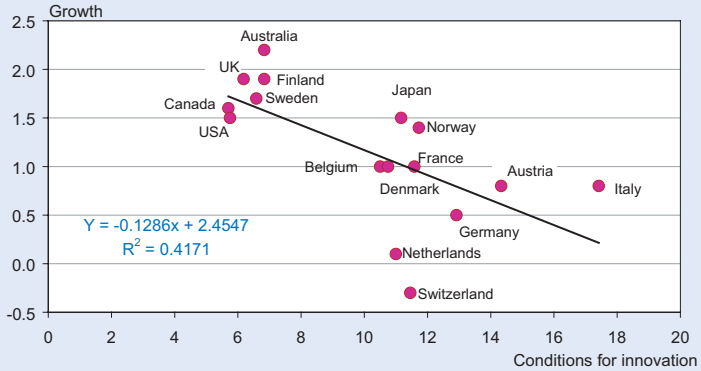
Source: Own calculation.

determine whether or not there is a statistical relationship between the country's average rank on such indicators and its macro-economic growth rate. If the measure of human capital alone were selected, the R2 measure would take an even higher value. It can, therefore, be tentatively concluded that human capital plays a particularly important role in innovation and growth.

It can, however, certainly be argued that those growth stars between 1994 and 2004 were also the fastest growing national economies between 2000 and 2004. The categorization of states according to their growth rates would appear, therefore, not to have been affected by the ebb and flow of the new economy. Those countries that grew strongly over the whole time period considered here had, at the start of this millennium, good conditions for innovation, and were also able between 2000 and 2004 to achieve a higher rate of growth than those states that, on the innovation measures, had worse values.

Figure 2

CONDITIONS FOR INNOVATION AND ECONOMIC GROWTH[a)]



a) Growth in per capita GDP between 2000 and 2004 in per cent and conditions for innovation of countries at the technological frontier.

Source: Own calculations based on Eurostat (2005).

Final remarks

The 22 OECD countries in the current study grew, on average, by 2.4 percent. Leading positions were taken by Ireland and South Korea, which because of their relatively low starting positions in 1994 were, to a large extent, still able to profit from a process of catching up. This catching up was characterized, first and foremost, by investments in already existing and predominantly labour-intensive technologies. Such a strategy is, however, no option for countries at the technological frontier. In highly developed industrial countries, the ability to innovate and a fast diffusion of these innovations as well as adaptation to technical progress plays, in comparison to imitation, an increasingly important role.

If Germany or, indeed, any other highly developed country with a poor growth record wants to increase its growth potential, it needs to improve the conditions for innovations at the technological frontier and the fast diffusion of these innovations. Improvements in human capital and the de-regulation of product markets should be of the highest priority. In particular, the government bureaucracy (including economically not sound subsidies) must be reduced considerably. In addition, more venture capital should be made available, and incentives to invest in ICT should be increased. A further level-headed de-regulation of the labour market and improved incentives in the welfare state so that innovators can find workers that match their new technologies more quickly could make an additional contribution to an improvement in performance on growth and innovation.

References

Acemoglu, D., P. Aghion and F. Zilibotti (2002), "Distance to Frontier, Selection and Economic Growth", *NBER Working Paper*, no. 9066, Cambridge MA.

Aghion, P and P. Howitt (1992), "A Model of Growth through Creative Destruction", *Econometrica* 60, 323–51.

Aghion, P., C. Meghier and J. Vandenrusche (2003), *Growth, Education and Distance to the Technological Frontier*, mimeo.

Arnold, E. and B. Thuriaux (2003), "Introduction", in European Commission, ed., *Future directions of innovation policy in Europe*, Brussels, 1–10.

Bassanini, A. and S. Scarpetta (2002), "Growth, Technological Change, and ICT Diffusion: Recent Evidence from OECD Countries", *Oxford Review of Economic Policy* 18, 324–44.

Baumol, W.J. (2004), *Education for Innovation: Entrepreneurial Breakthroughs vs. Corporate Incremental Improvements*, mimeo.

BMBF – Bundesministerium für Bildung und Forschung (2004), *Technologie und Qualifikation für neue Märkte*, Bonn and Berlin.

Boyer, R. (2004), "New Growth Regimes, But Still Institutional Diversity", *Socio-Economic Review* 2, 1–32.

Brandt, N., J.-M. Burniaux and R. Duval (2005), "Assessing the OECD Jobs Strategy: Past Developments and Reforms", *OECD Economics Department Working Paper*, no. 429, Paris.

European Commission (2004a), *Benchmarking Enterprise Policy*, Brussels.

European Commission (2004b), *The EU Economy: 2004 Review*, Brussels.

Eurostat, 2005, NewCronos, URL: http://europa.eu.int/comm/eurostat/newcronos/reference /display.do?screen= welcomeref&open=/&product=EU_MAIN_TREE&depth=1&language=de (accessed 20 January 2005).

Fagerberg,. J. (2003), "The Innovative Society", in European Commission, ed., *Future directions of innovation policy in Europe*, Brussels, 16–19.

Fagerberg, J. (2005), "The Dynamics of Technology, Growth and Trade: A Schumpeterian Perspective", in Hanusch, H. and A. Pyka, eds., *Elgar Companion to Neo-Schumpeterian Economics*, Edward Elgar, Cheltenham, in press.

Farhauer, O. and Henke, K.-D. (2002), "Wachstumstheoretische Konzeptionen und wirtschaftspolitische Folgerungen", *Das Wirtschaftsstudium* 31, 582–90.

Fehn, R. (2004), "Strukturwandel und europäische Wirtschaftsverfassung: Gibt es einen Zielkonflikt zwischen Effizienz und Sicherheit?", in Schäfer,W., ed., *Zukunftsprobleme der europäischen Wirtschaftsverfassung*, Duncker & Humblot, Berlin, 87–121.

Gehrig, T. P. (2000), "Zur Politischen Ökonomie des Technischen Fortschritts", in Külp, B. and V. Vanberg, eds., *Freiheit und wettbewerbliche Ordnung. Gedenkband zur Erinnerung an Walter Eucken*, Freiburg, 561–77.

Grossman, G. M. and E. Helpman (1991), *Innovation and Growth in the Global Economy*, Cambridge.

Levine, R. (1997), "Financial Development and Economic Growth: Views and Agenda", *Journal of Economic Literature* 35, 688–726.

Matthes, J. and C. Schröder (2004), "Business Regulation in International Comparison – Aggregating World Bank 'Doing Business' Data", *CESifo Forum* 6, no. 1, 42–49.

Murphy, K. M., A. Shleifer and R. W. Vishny (1991), "The Allocation of Talent: Implications for Growth", *Quarterly Journal of Economics* 106, 503–30.

Nicoletti, G. and S. Scarpetta (2003), "Regulation, Productivity and Growth: OECD evidence", *OECD Economics Department Working Papers*, no. 347, Paris.

OECD (2004a), *Employment Outlook*, Paris.

OECD (2004b), *OECD Science, Technology and Industry Scoreboard 2003*, Paris, http://www1.oecd.org/publications/e-book/92-2003-04-1-7294 [Download: 11-22-2004].

OECD (2004c), *Wirtschaftsbericht Deutschland*, Paris.

OECD (2004d), *Lernen für die Welt von morgen. Erste Ergebnisse von PISA 2003*, Paris.

OECD (2004e), *Education at a Glance*, Paris.

OECD (2004f), OECD-Education Database, URL: http://www1.oecd.org/scripts/cde/members/EDU_UOEAuthenticate.asp [Download: 1-20-2005].

Romer, P. M. (1990), Endogenous Technological Change, *Journal of Political Economy* 98, 71–102.

Romer, D. (1996), *Advanced Macroeconomics*, McGraw Hill, New York.

Sapir, A., P. Aghion, G. Bertola, M. Hellwig, J. Pisani-Ferry, D. Rosati, J. Vinals and H. Wallace (2004), *An Agenda for a Growing Europe. The Sapir Report*, Oxford University Press, Oxford.

Scarpetta, S. and T. Tressel (2004), "Boosting Productivity via Innovation and Adoption of New Technologies: Any Role for Labor Market Institutions?", *Policy Research Working Paper Series*, no. 3273, The World Bank.

Van Ark, B. (2003), "The Productivity Problem of the Dutch Economy: Implications for Economic and Social Policies and Business Strategy", *Research Memorandum GD-66*, Groningen Growth and Development Centre, Groningen.

World Bank (2004), *Doing Business in 2005 – Removing Obstacles to Growth*, Washington DC.

Tax Reform in the Slovak Republic

Ivan Mikloš, Marek Jakoby and Karol Morvay*

Previous tax systems in Slovakia and their imperfections

Economic policy in Slovakia has undergone a few principal qualitative turns in the course of a 15-year transition from a centrally-planned economy to a market economy. The need to adjust the tax system arose as an accompanying necessary features of such shifts. Besides dozens of parametrical changes, the tax system was subject to two cases of fundamental reconstruction. Both were linked to substantial changes in direction of economic policy. Firstly, the fall of the socialist economy required the creation of a tax system that could be comparable with that in standard market economies. Secondly, the overcoming of non-standard and somehow hybrid economic policy after 1998 and consequent upshifting of the reform process after 2002 called for further fundamental reform. An overview of the reform stages is provided in Box 1.

The tax system at the threshold of the transformation process corresponded to the needs of an administratively managed economy. It served as a tool for direct selection, subsidies and social policy. These functions were incorporated in the sales tax with numerous rates, which enabled an individual approach in accordance with the goals of the centrally planned economy. The sales tax system was marked by individual tax rates for each item and was, thus, in fundamental contradiction with the principles of modern market economy.

The system also included several items with negative sales tax (which in fact represented a subsidy) in order to maintain low prices and ensure affordability to all citizens. Understandably, the first fundamental change in the tax system included abolition of the negative sales tax in 1990. This step resulted in a first inflation shock (increase in prices of food which had benefited from the negative sales tax). This increase was compensated by a special social allowance. The abolition of the negative sales tax was followed by price liberalisation which came into force on 1 January 1991.

The timing of the launch of the first fundamental change of the tax system in January 1993 accidentally coincided with the birth of the Slovak Republic. The establishment of an independent state, however, did not have any direct impact on the shape of the system. The new system incorporated relevant components of the tax system in EU countries. Its basic pillar was a value added tax with two rates, which were later changed a few times. Personal income tax was progressive with six income brackets and rates ranging from 15 to 47 percent. However, the system also included five other tax rates applied to specific sources of income. Corporate income tax was set with linear rate of 45 percent.

Experience has consequently revealed numerous shortcomings of the system: dozens of legislation amendments increased its complexity, numerous exceptions and special regimes clashed with basic taxation principles. Furthermore, the progressivity of personal income taxation was increasing in the course of time: while the tax brackets were kept fixed, rising wages pushed taxpayers towards higher rates. Frequent parametrical adjustments to the system were adopted in response to the needs of public finances and macroeconomic development. Such adjustments were quite common mainly in the course of a macroeconomic stabilisation operation in 1999–2000 when it was necessary to correct the disequilibrium trends set forth by the non-standard economic policy between 1995 and 1998. Typical adjustments included shifting selected commodities from a reduced to a standard VAT rate and an increase in excise taxes. On the other hand, statutory corporate income tax rate was gradually reduced with the goal to promote economic activity.

The second and the most recent fundamental reconstruction of the tax system effective from January 2004 required much more political courage than the previous one because it was no longer copying standard tax systems of Western Europe but was rather based on a fundamental change of some taxation principles. A larger emphasis on competitiveness of the economy called for a more attractive and motivating tax system. The tax reform, with a flat income tax rate applying to all sources of income being its cornerstone, was part of a broader reform package,

* Ivan Mikoš – Minister of Finance of the Slovak Republic (imiklos@mfsr.sk); Marek Jakoby – Advisor to the Minister of Finance (mjakoby@mfsr.sk); Karol Morvay, senior analyst at M.E.S.A. 10 – Center for Economic and Social Analyses (morvay@mesa10.sk).

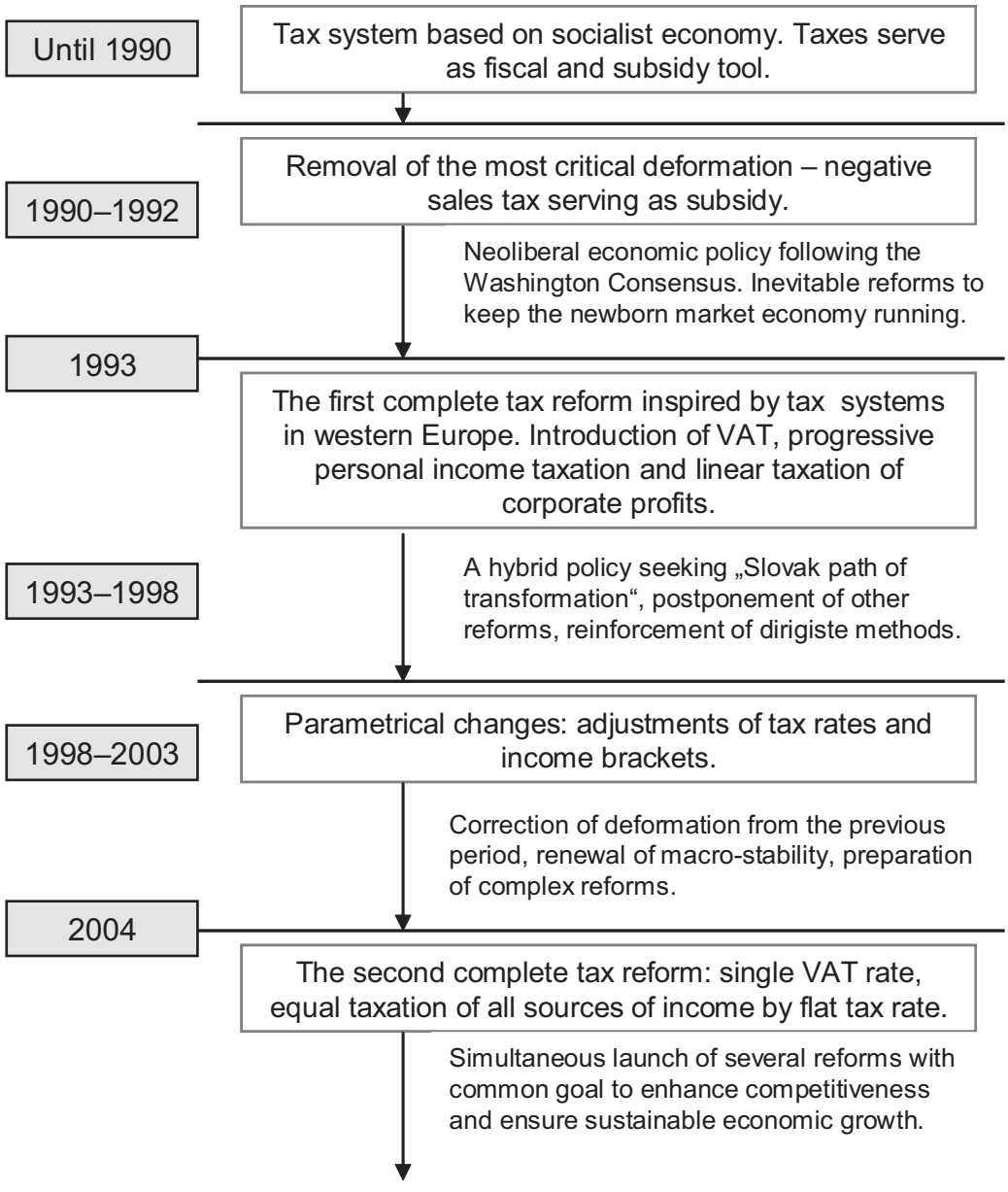
and one has to regard the tax reform as an integral part of the policy aimed at enhancing the competitiveness of the Slovak economy and ensuring its sustainable growth.

Political opposition towards tax reform

The Slovak tax reform was not challenged by opposition on the part of the public. In fact, people could assess the impact well in advance and a large group of winners profiting from the reform was quite clearly visible. The opposition was rather concentrated in those areas where reforms brought about a tangible increase in costs alongside possible future benefits (such as health care reform or an attempt to reform university education by introducing tuition fees). The social acceptance of the tax reform required minimising the group of losers who would not profit from the reform. A substantial increase in a tax-exempt allowance meant that people with low taxable income are not required to pay any income tax.

Box 1

Changes of Tax system and Economic Policy in Slovakia



However, the reform was opposed by left-oriented political entities (albeit determination of right and left is quite a problem on the changing Slovak political stage). Their opposition accented three points:

1) Abolition of progressivity allegedly eliminated solidarity.
2) Economic transformation brought sudden and frequently also a non-transparent, illegal or immoral distribution of wealth. Failure of other instruments to rectify such wrongful property acquisition had, in the opinion of the opposition, justified high income taxation of high-income earners. The opposition even put forward a few proposals for selective specific taxation regimes for some industries reporting huge profits (marked as "inadequate" by some politicians).
3) Abolition of reduced VAT which led to higher prices for elementary goods, such as food and medicines, thus negatively affecting mainly low-income households.

Although largely seen as a right-wing reform, it was not spared criticism from the right zone of the political spectrum. The right-wing critics questioned "too high tax rates" and the initial absence of the goal to reduce the overall income redistribution rate in the economy. A guiding principle of the process of setting up the reform parameters was the principle of fiscal neutrality, i.e. to collect as much in tax revenues as in a situation without reform. The critics, on the contrary, assumed that the reform would lead to a more radical decrease in state budget revenues that would consequently exert pressure to reduce expenditures.

However, the Ministry of Finance exercised caution in order to avoid any shocks in the course of implementation of the reform. Although the government set the reduction of the public expenditure/GDP ratio as one of its targets, the idea was not to achieve it by a sudden fall in the initial stage of the new system.

Basic philosophy of the tax reform

Avoiding the need to go into specific details, we can summarise framework principles of the Slovak tax reform (see also Box 2):

Shift from selectivity to universality. The new tax system subjects all kinds of income to equal taxation[1]. The system has thus reduced broad range of exceptions, special rates or special taxation regimes. Simplicity and transparency were thus significantly enhanced.

Enforced incentives to work. The elimination of progressive income taxation eliminated the disincentive to "earn more" and weakened the motivation to conceal one's income. A relatively low income tax rate combined with the absence of special dividend taxation stimulates business activity.

Shift from labour and capital taxation to consumption taxation. A reduced direct taxation leaves the key role to indirect taxes.

As we mentioned above, the tax reform was implemented alongside other reforms. It is therefore useful to point out the complementary effects of different reforms:

1) The tax reform is built on the same principles as the social system and labour market reforms. All these policies share incentives to activity as common feature: Tax policy through lower tax burden on labour income and child tax credit; labour market and social policies reach the same goal by more strict control of registered unemployed and by new social benefit system including work activation components aimed at long-time unemployed.
2) The tax reform supplements the public administration reform. The latter enforces role of self-governments, which enjoy increasing authority over generation of their revenues. The process of fiscal decentralisation included direct allocation of a given percentage of collected personal income tax to cities and municipalities[2]. The self-governments also gained power to set the rates of real estate taxes. Local authorities can thus actively shape the business environment in their respective areas. However, most self-governments used the new power to increase the tax rates dramatically in order to solve liquidity problems.

The government occasionally breaks with the ideological purity of the tax system by a policy aimed at attracting large foreign investments. Selective provision of investment incentives including tax abate-

[1] Except certain kinds of income which are not subject to income tax (inheritance, gifts, etc.) and sources of income explicitly tax exempted (pensions from public system, social benefits, income from the sale of real estate after five years of ownership, state lottery prizes, interest on Slovak eurobonds, etc.)
[2] Effective from January 2005, 70.3 percent of personal income tax is allocated directly to municipalities, 23.5 percent is transferred to regional self-governments and 6.2 percent remain in the state budget.

Box 2

Basic components of the tax reform

1. **Changes in indirect taxes:**
   - *Amendment to the Acts on Excise Taxes* – increase of certain excise tax rates: mineral oils, beer and tobacco products (August 2003)
   - *Amendment of VAT Act* – unification of VAT rates at 19 percent (before: 20 percent and 14 percent) as of January 2004.
2. **Changes in direct taxes:**
   - *New Income Tax Act* – introduction of 19 percent flat tax (January 2004)
     - Basic tax-exempted allowance increased from lump sum (SKK 38,760 in 2003) to 19.2 times the monthly living minimum (80,732 in 2004)
     - The allowance applicable also for non-working spouse with its proportional reduction linked to his/her own taxable income
     - Substitution of tax-exempt allowance per child with child tax credit
   - *New Real Estate Transfer Tax Act* – abolition of gift and inheritance tax and introduction of 3 percent flat rate for real estate transfer tax (January 2004); Abolition of real estate transfer tax (January 2005).
3. **Changes in indirect taxes in compliance with EU tax legislation**
   - *New VAT Act and new Excise Duties Acts* as of 1 May 2004.

ments creates exemptions of the principles of universality and equity. However, we consider this policy, principally motivated by efforts to reduce still high unemployment, to be of a temporary nature, without setting up long-term trends.

Initial effects of the reform

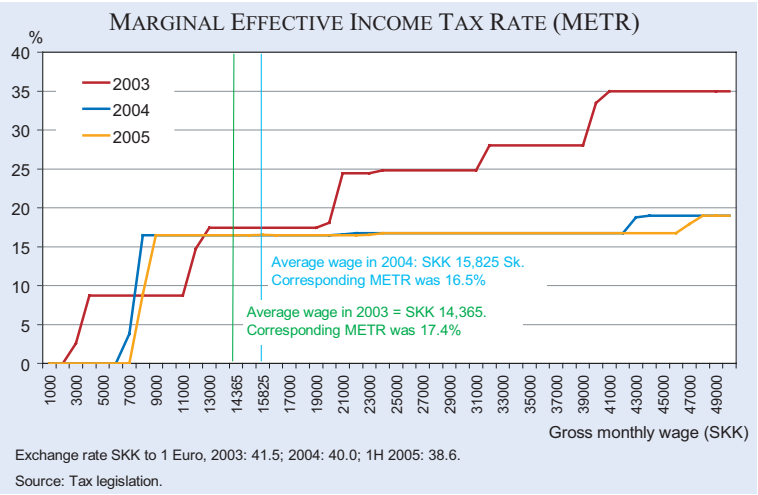
Personal labour income taxation

The tax reform brought a reduction in the effective taxation of personal income as well as of the tax wedge. This reduction, however, was minimal in the income range between 90 and 160 percent of the average wage (Figure 2). Figure 2 also shows that there is still some progressivity in the personal income taxation. Although the marginal income tax increased for people with taxable income in the range 60–90 percent of the average (Figure 1), the overall effect was favourable due to the introduction of a high basic tax-exempt allowance. We have also included 2005 in the comparison in order to compare the shifts resulting from systemic change (2004 vs. 2003) with those resulting from changes of built-in parameters, such as basic tax-exempt allowance (2005 vs. 2004). The taxation curves in 2005 (yellow lines in Figures 1 and 2) are shifted slightly to the right due to the automatically living-minimum-indexed tax-exempt allowance.

It is also worth to note the shift in total payroll tax burden in 2004 (Figure 3). Due to the reduction in social security contributions paid by employers the winners included workers in almost the whole spectrum of income. However, the introduction of an average wage-linked maximum assessment base negatively affected taxpayers who had previously paid contributions from lower assessment base that had been set as a lump-sum.

The introduction of a flat income tax and related features (tax-exempt allowance, tax credit, etc.) thus meant an increase in net income to the majority of

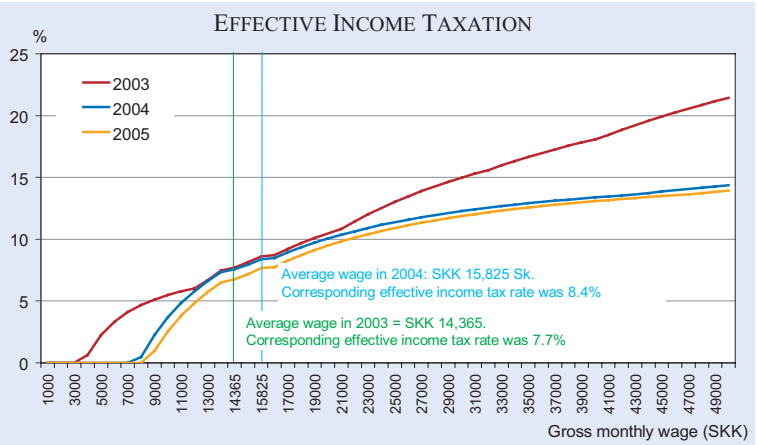
Figure 1



Exchange rate SKK to 1 Euro, 2003: 41.5; 2004: 40.0; 1H 2005: 38.6.

Source: Tax legislation.

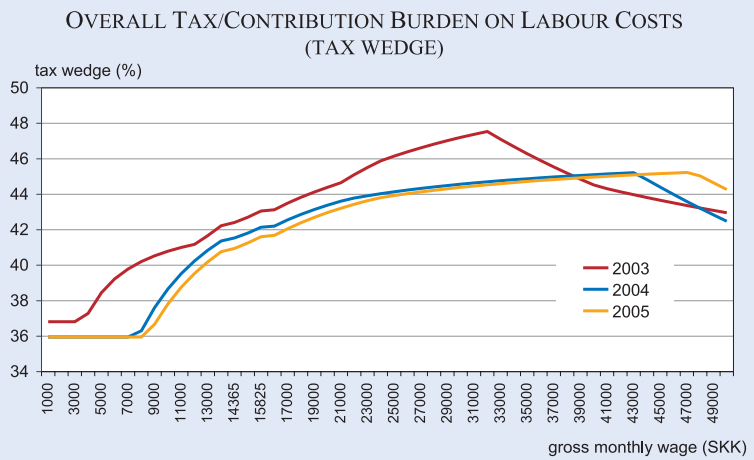
Figure 2



Exchange rate: see Figure1.

Source: Tax legislation.

Figure 3



Source: Tax legislation.

individual tax payers compared to 2003. As we indicated above, there was a group of taxpayers who experienced no change in their net income from a static point of view. This group consisted mainly of single individuals with monthly gross income ranging from SKK 12,000 to 20,000.

As the income taxation was only one component of the reform, it is necessary to consider indirect taxes to calculate the overall effect of the reform on net real income. The effect on two types of households is presented in Figure 4.

Figure 4 clearly indicates that winners of the tax reform from the point of view of impacts of tax changes on real income included families with children including households with non-working spouses taking care of children. It is also relevant, that any further increase in gross taxable income in the future will bring more in net income to all taxpayers than before the reform. The motivation to earn more in legal ways was thus substantially enhanced.

It is also worth noting that the real consumption of households grew by 3.5 percent in 2004. Although "distribution" of this growth by income segments might show some discrepancies, such robust growth dispelled the fears of the effect that the unification of VAT and increase in excise taxes might have on consumer prices.

While most of the working population was able to absorb the consequent increase in consumer prices with a higher net income, people without taxable income were not able cover higher living costs with tax allowances[3]. Pensioners were therefore compensated in the form of a lump-sum payment (rather non-systemic measure) paid in the autumn of 2004. Although we can say that pensioners were the short-time losers of tax reform, they are now benefiting from the pension reform, which has introduced a new indexation mechanism that provides space for an increase in the real purchasing power of pensions.

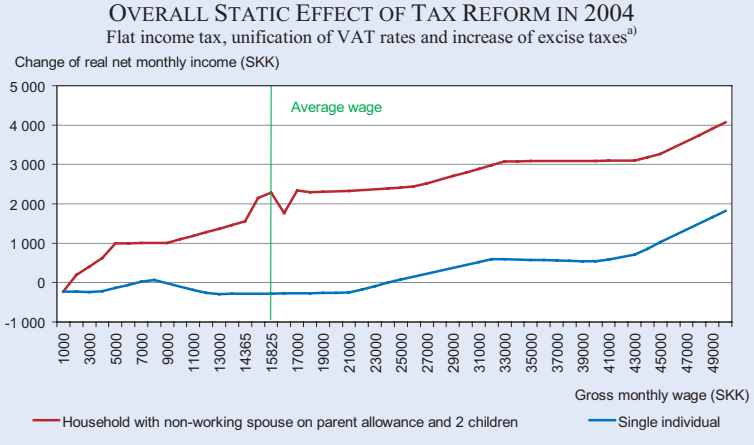
Capital income taxation

Unlike the case of personal income tax, there are no special indicators measuring the impact of the tax reform on the corporate sector. However, we can still use some (albeit simplified) parameters.

The Slovak tax reform was, beside other principles, based on the idea of significant reduction in capital income. The practical result of the abolition of dividend tax was an annual decrease in the effective taxation of business income from 36.3 to 19 percent (Table).

The reduction of taxation meant increased disposable income in profitable businesses and better values of return on sales. At the same time, more companies became potential bank clients.

Figure 4



a) Although the excise taxes on fuels and beer were increased already from August 1, 2003, their adjustment are in fact part of the 2004 reform.

Source: Tax legislation.

[3] Neither welfare payments nor pensions from the state-run pension scheme are taxed in Slovakia.

	2003	2004
	in %	
Statutory corporate income tax rate	25	19
Dividend tax	15	0
Effective taxation of distributed corporate profit	36,3	19

Source: Tax legislation; table: authors.

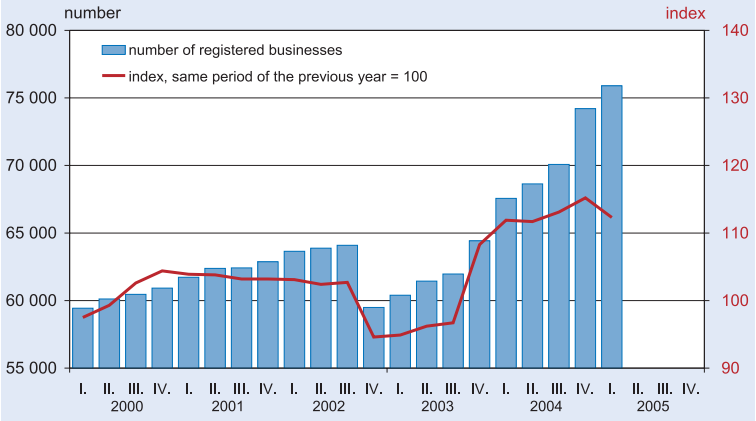
Higher disposable income has enabled more intensive investment activities and was probably one of the important factors behind the robust growth of gross fixed capital formation in 2005 (5.8 percent in constant prices), the highest growth rate since the last quarter of 2001.

Another useful indicator of perception of the tax reform by the corporate sector is the development in the number of businesses. It is significant that with the exception of agriculture (specific sector marked with long-term problematic issues) and the energy sector (dominated by a few large players), all industries reported a robust increase in the number of businesses at the end of 2003 (when the shape of reform was widely known) and in 2004 compared with the previous periods (Figure 5).

Beside lower effective taxation, the business community also welcomed a reduction in the depreciation period of buildings from 30 to 20 years, eliminated the restrictions of the depreciation rate in the first year of the depreciation period and liberalised treatment of loss carry-forward. Goliáš and Kičina (2005) note that private accountants in Slovakia were told by their clients that they were more concerned about their ability to write off legitimate losses than whether statutory corporate income tax rate was 15 or 25 percent. Contrary to the previous system valid until the end of 2003, loss can now be deducted from the tax base in the course of five years following the year when it incurred without even the obligation of write-offs and without the obligation to reinvest the same amount in capital goods.

Figure 5

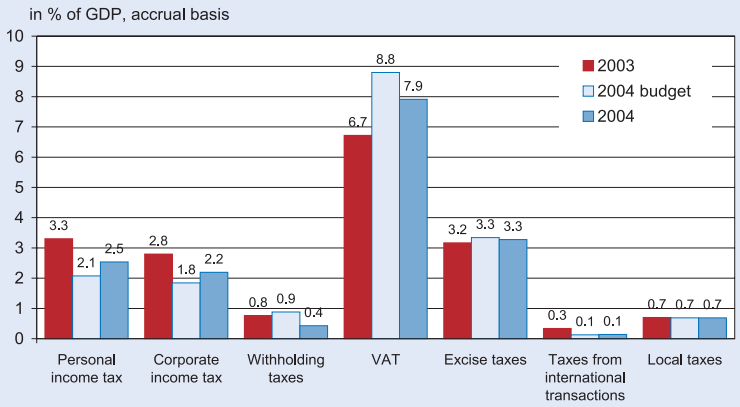
DEVELOPMENT OF NUMBER OF REGISTERED BUSINESS



Source: Statistical Office of the SR.

Figure 6

STRUCTURE OF TAX REVENUES IN 2003 AND 2004



Source: Ministry of Finance of the SR; calculations of the authors.

Fiscal implications

Total tax revenues in 2004, the first year of the new tax system, were in line with projections of the Ministry of Finance. The basic fiscal condition of the tax reform – fiscal neutrality – was thus met. At the same time, however, some shifts occurred in the structure of tax revenues. While the collection of personal as well as corporate income tax measured as a share of GDP exceeded the planned level (possibly as a result of higher motivation among taxpayers to report their income as well as postponement of some payments from 2003 to 2004), VAT collection failed to reach the projected value (Figure 6). However, the shortfall was probably at least partially affected by the entry of Slovakia into the EU and consequent problems with the collection of VAT from cross-border transactions.

A similar development of tax revenues is also visible in 2005,

with corporate tax being the "frontrunner". By the end of June, Slovak business paid SKK 10 billion (70 percent) more in corporate tax than projected.

Reactions from abroad

The Slovak tax reform has also become a frequently cited factor enhancing the attractiveness of Slovakia in the eyes of foreign investors. Although it is difficult to precisely specify the importance of the new tax system in the decision-making process of investors coming to Slovakia, it is quite obvious that the number of investors considering or already planning to invest in Slovakia increased in 2004. A hint in this context can be provided by the number of projects developed by the state investment agency SARIO. While the number was 22 in 2003 (with a total value of projected invested amount USD 1.44 billion), in 2004 the agency reported 47 accomplished projects worth USD 2.26 billion. However, we assume that the factors behind this increase are rather complex and include probably the whole spectrum of business-friendly reforms (such as more flexible labour legislation, more efficient business set up procedures) along with improving the infrastructure and a favourable labour cost/productivity ratio.

It is also significant, that Slovakia competes with other countries of the region for large investors offering generous state aid packages which in most cases also include partial or full corporate income tax holidays for up to ten years.

The Slovak tax reform provoked strong reactions from the foreign business community, and dozens of companies from the "old" EU member countries declared their intention to move headquarters or even independent units to "more favourable tax environments". While one has to understand such declarations as a highly useful and effective tool to exert pressure on domestic governments to carry out similar reforms, the highly competitive level of corporate profit taxation and relative simplicity and transparency resulted in a higher number of foreign businesses being registered in Slovakia.

Some top politicians in countries with traditionally high income taxation reacted in an irritated manner accusing Slovakia as well as other "flat-tax countries" of "tax dumping". Some of them consequently questioned the right of these countries to draw from the EU structural funds. Such argumentation is, however, built on completely wrong premises. As one of the strategic goals of similar tax reforms is to boost the economic growth by a lower tax burden on economic activity, their implementation in fact leads to higher GDP per capita and consequently to more limited space to offset the handicap of falling behind the EU average with structural funds.

It is worth noting that the reduction in income taxation along with the reduction in the tax burden measured both in relation to labour cost (tax wedge) as well as to GDP had been a long-term trend in most OECD countries already from the end of 1990s (Tax Foundation, 2004).

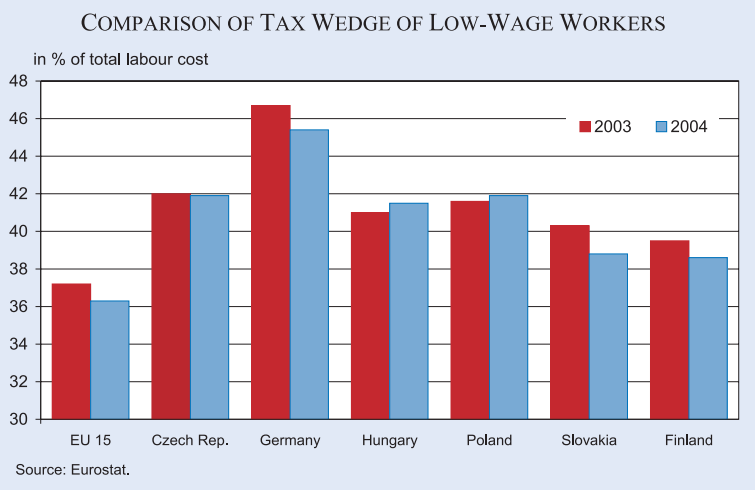
Still, however, the Slovak reform contributed to speeding up the process and fuelling the debates on radical tax reform in many other countries. Goliáš and Kičina (2005) note that the reform evoked fierce tax competition among central European countries, which continued to spread further to the west. They point at a reduction in the corporate tax rate in Austria from 34 to 25 percent (while the original plan was to "stop" at 31 percent), the reduction of the standard VAT rate in the Czech Republic from 22 to 19 percent in 2004, and the plan to decrease the corporate tax rate from 28 to 24 percent in 2006 and reductions in the statutory corporate tax rate in Poland and Hungary from 27 to 19 percent and from 18 to 16 percent, respectively.

Outlook on further reform steps

Increasing tax competition raises a simple question: what comes next? Fortunately, Slovak reformers do not consider the tax reform a one-off issue. It comes as a matter of course that adjustments are being made or planned. At the same time, the experts are considering modifications to some basic components of the new system. One of the most recent proposals of the Ministry of Finance is to replace the basic tax-exempt allowance with a tax credit. Such a step would lead to a higher net income for low-wage earners and consequently to higher incentives to create more jobs for people with lower qualifications. Such people make up most of the long-term unemployed in Slovakia and the goal to reduce substantially long-term unemployment presents one of the most important challenges to economic policy.

While the introduction of tax credit can contribute to some improvement, policymakers will have to look

Figure 7



closely at social contributions. Although total labour costs in Slovakia are still lowest from among the "V4" countries[4], it is obvious that the high contribution burden paid by both employers and employees from the first koruna of earned income (reaching 36 percent of total labour cost) effectively hampers the creation of low-wage jobs. While some improvement was achieved in 2004 (Figure 7) and a reduction in employer's contributions resulted in a moderate reduction of the tax wedge for low-wage earners (with monthly wages corresponding to 67 per cent of average wages in manufacturing), much has to be done in the entire system of social contributions, with due attention paid to the overall balance of individual components.

Despite the proposed adjustments, the modification of tax rates (income tax, VAT) is not the question of the day. If, however, public finances develop favourably (the basic requirement is to keep the general government deficit below 3 percent of the GDP from 2006), toughening tax competition in the region might exert pressure to further reduce the flat income tax. There is one negative scenario as well, though. The strongest opposition party, SMER, has officially declared a negative attitude to the flat income tax and declared ambitions to reintroduce income tax progressivity to the system. The next parliamentary election due in September 2006 will provide the answer to the future fate of the Slovak reform.

[4] V4 – four „Visegrád" countries – Czech Republic, Hungary, Poland, Slovakia.

References

Chren, M., *Fundamental Tax Reform in Slovakia – A Case Study on the Slovak Tax Reform*, (http://www.ig.wsiz.edu.pl/forum/paper/ MChren.pdf).

Tax Foundation (2004),"Fundamental Tax Reform: The Experience of OECD Countries", *Background Paper* no. 47.

Goliáš, M. and R. Kičina (2005) "Slovak Tax Reform: One Year After", in A. Blinov, S. Seheda and P. Udovenko, *Recent Economic Reform Experience from Central Europe: Inspirations and Suggestions for Ukraine*, ICPS, Kyiv.

Ministry of Finance of the Slovak Republic (2004), Možnosti znižovania odvodov na Slovensku (Possibilities to Reduce Contributions in Slovakia), Institute of Financial Policy (only available in Slovak).

IMF (2005), *Slovakia's Tax and Welfare Reforms*, IMF Working Paper.

Statistical Office of the SR, *Income, Expenditures and Consumption of Private Households in the SR 2003, 2004*, Statistical Report on Basic Development Tendencies in the Economy 2003, 2004.

Credit and Product Market Effects of Banking Deregulation: Evidence from the French Experience

Marianne Bertrand, Antoinette Schoar and David Thesmar*

Many economists believe that an efficient financial industry is central to economic development. While this conjecture dates back at least to Schumpeter, economists have only relatively recently taken up the systematic study of the impact of finance on economic growth. Starting with King and Levine's seminal paper (1993), a large number of studies have documented strong positive correlations between financial development, measured by the size of the financial sector, and growth and economic development. While these cross-country results are strikingly robust and consistent, they tell us little about the exact micro channels though which a more developed financial industry promotes growth.

In Bertrand, Schoar and Thesmar (2005), we undertook such a study of the micro channels by looking at the impact of financial market deregulation on the allocation of credit across firms, firms' behavior and product market dynamics. In particular, we analyzed the French banking deregulation in the mid-1980s. The French banking deregulation drastically reduced government interventions in banks' lending decisions, almost entirely abolished subsidized bank loans and allowed French banks to compete more freely in the credit market. In addition, several state-owned banks were privatized in the mid-1980s. According to most observers, the main effects of the reform were to move from a centrally-planned to market-based capital allocation, to decentralize the decision-making process on loan amounts and interest rates, and to introduce a stronger for-profit motive among banks.

* Marianne Bertrand is Professor of Economics at the University of Chicago, Graduate School of Business; she is affiliated with NBER, CEPR and IZA; mbertran@gsb.uchicago.edu.
Antoinette Schoar is Associate Professor of Finance at MIT-Sloan; she is affiliated with NBER and CEPR; aschoar@mit.edu.
David Thesmar is Professor of Economics at ENSAE-CREST, Paris; he is affiliated with CEPR; david.thesmar@ensae.fr.

While our analysis was restricted to a single country, the scope of regulations in place in France prior to the reform matches the experience of many other countries with regulated banking sectors. In this regard, the French reform is quite representative of the multiple changes many other countries would have to implement to liberalize their banking sector.

Our focus on France was also motivated by the availability of comprehensive and very detailed firm-level accounting data for this country. While most commonly used international firm-level data sets cover only publicly traded firms, the data used in this paper includes both private and publicly traded firms. The coverage of private firms was central to our analysis. First, since these firms typically have access to few other sources of external finance besides bank loans, they may be most affected by any changes in the banking sector. Second, and equally important, these firms represent a very large fraction of overall economic activity, making their coverage in the data necessary to any study on the impact of the banking reform on industry structure and dynamics.

While the French banking deregulation constituted an economy-wide shock, we isolated its effect on firm behavior and product market dynamics by studying differential changes post-reform across sectors that were more or less reliant on bank finance prior to the reform. The identifying assumption at the basis of this empirical strategy is that industries that were more financially dependent on banks prior to the reform should be more affected by the deregulation. In practice, we also assessed the robustness of our findings by using a US-based measure of external financing dependence (à la Rajan and Zingales 1998) as an alternative source of cross-sector variation in the strength of exposure to the banking reform.

Our findings, which are summarized below, are consistent with a model where distortions in bank lending create artificial barriers to entry in the real sectors of the economy. New entrants may be discouraged by the easy access to cheap credit for incumbent firms. Once banks become less willing to provide such (cheap) credit to poorly performing firms, prospective new entrants find it more attractive to come in and compete with incumbents. A more efficient banking sector therefore appears to play an important role in fostering a Schumpeterian "creative destruction" process that has been theoretically, and

increasingly empirically, linked to higher economic growth.

Such evidence of distortions in bank lending being associated with relative sclerosis and lower restructuring activity in the real sectors of the economy is reminiscent of Caballero, Hoshi and Kahyap's discussion (2003) of the role Japanese banks have played in the slowdown of the Japanese economy over the last decade. Our findings also complement recent work by Black and Strahan (2002) and Cetorelli and Strahan (2004) who study changes in industry-level entry rate, number and size distribution of firms in the context of the US interstate banking deregulation. One major difference between these papers and our study (in addition to the obvious focus on a different set of reforms) is our access to firm-level data. This data allows us not only to look at restructuring activities at the firm level but also to study the reallocation of capital across firms within industries.

The French banking reforms of the mid-1980s

After World War II, the French financial sector came under the centralized control of the Treasury, whose general aim was to channel savings and deposits into priority industries. To control the credit market, the Treasury set up a deposit network, consisting of savings banks, the postal checking system, the Bank of Foreign Trade, and four large cooperative banks. This network had privileged access to some deposits and the bond market, and a monopoly over the distribution of subsidized loans allocated by the Treasury. Increased governmental control over savings collection and use was also achieved through the nationalization of some of the biggest banks.

The economic turmoil after the 1974 oil shock further strengthened bureaucratization and state involvement in the banking sector. The government had to balance conflicting objectives: limiting money growth to stabilize the Franc's parity with the Deutsche mark while stimulating credit and investment. This was implemented through the "encadrement du crédit" program, which consisted in setting monthly ceilings on credit growth for each bank individually. A direct consequence of the "encadrement du credit" was to further strengthen the relative importance of subsidized loans and government control over lending decisions. By 1979, subsidized loans amounted to nearly half of all new loans granted to the private sector. In May 1981, a new socialist government was elected: fiscal policy became more expansionary, and a further nationalization of the banking sector was implemented. The Treasury also increased the pressure on state-owned banks to bail out failing industrial groups. The number of different loan subsidization programs increased dramatically, as the Treasury focused more and more on "job preservation" and preventing the shut down of poorly performing firms. As a result, the credit market became even more opaque, supporting more than 200 different interest rates for different loan subsidization programs.

The expected benefits from an increased centralization of the banking system did not pan out. In the fall of 1984, the socialist government announced a drastic reversal of policy. The goal was to transform the financial system into a decentralized credit market, where interest rates would be used to match the supply and demand of capital for each type of project. More specifically, three sets of reforms of the banking industry took place in 1985.

First, starting in 1985, most subsidized loans were eliminated. Also, the distribution of these remaining subsidized loans was no longer the monopoly of the Treasury-controlled deposit network, which improved transparency and competition on the lending market.

Second, the "encadrement du credit" was abolished in 1985 and capital flows in the economy became much more determined by market forces. Between 1985 and 1987, credit growth limits were gradually removed and replaced by a system of reserve requirement against deposits. Monetary policy was now conducted through interest rates on the money market and legal reserve requirements instead of quantity controls. Resources became much more available to expanding private banks. The money market was also reformed to stimulate inter-bank lending: private banks could borrow more funds from the Treasury network, which now had little use for them. In addition, the system of capital controls, strengthened in 1981 to defend the Franc, was progressively eliminated through a string of reforms ending in 1990 (Naouri (1986).

Third, market conditions became more transparent and conducive to fair competition. The 1985 Banking Act partially unified a myriad of banking regulations, and progressively also eliminated subsidized loans. Partial monopolies over deposits and lending enjoyed by some banks were progressively disman-

tled. Banks also faced more competition from other providers of external finance, as firms' access to the bond and equity market was facilitated.[1] Last, with the disappearance of subsidized credit, private banks no longer faced the unfair competition from the members of the Treasury Network.

Finally, a number of banks were privatized in the 1986–88 period (about 10 percent of the banks and 20 percent of the banking assets). Most industry observers believe, however, that the other regulatory changes we described above were more important in reforming the French banking industry in the mid-1980s than this partial privatization effort. Part of the rationale driving this belief was that about half of the bank assets that were privatized in the mid-1980s had just been nationalized in 1982.

The widely (at least anecdotally) discussed consequence of the reform was a change in banks' behavior. The reforms signalled that the Treasury was willing to let market forces shape the credit market landscape for the long run. These new conditions forced banks to change their lending practices and restructure internally, in part with the help of the diffusion of new technologies. A survey conducted in 1985 among French bankers showed drastic changes in attitudes about the internal management of banks (Rémy and Sergent 1986). According to the survey, the focus of bank managers was increasingly on reducing costs, controlling risks and introducing tighter performance monitoring. The greater competitive pressures were most intensely felt by banks in the Treasury network, as these banks had lost their privileged access to deposits and loan markets. The Treasury network's share in all deposits decreased by 28 percent between 1985 and 1990, and its share of loans went down by some 25 percent (Plihon 1995).

Data

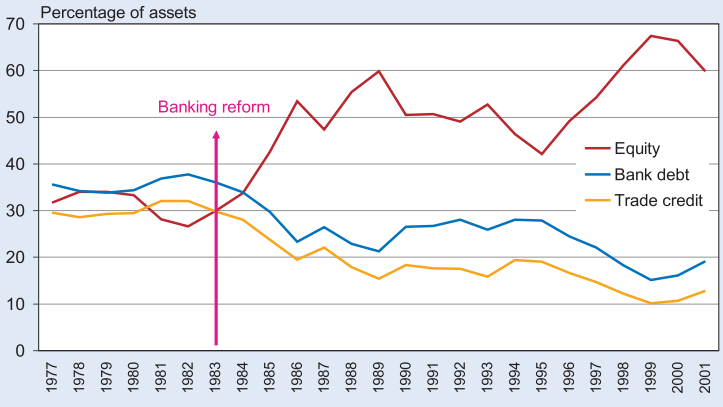
The firm- and industry-level data sets used in this study were based on accounting information available for all French firms, public or private, whose annual sales exceeded EUR 100,000 in the service sector and EUR 200,000 in other sectors. This accounting data was extracted from the tax files used by the Ministry of Finance for corporate tax collection purposes. French firms above these thresholds are required by tax authorities to fill in a detailed balance sheet and profit statement. Also included in the tax files is a four-digit industry classification code that is very similar to the SIC coding system in the US In addition, the data also contains reliable firm-level employment figures that have been cross-checked with information from employer labor tax reports. Individual firms can be tracked over time by the use of unique identifiers, which allows for the construction of a panel data set. We ended up with a data set of about 350,000 firm-year observations, which corresponds to about 15,000 firms per year. We excluded firms in the financial sectors from the sample (banking and insurance industries), since standard accounting measures are less meaningful in this industry.

Changes in financial structure and bank lending practices

In aggregate, the level of indebtedness of French firms went down dramatically in the mid-1980s. The ratio of total debt to assets was very high in the early 1980s, around 70 percent. Two years after the reforms, this ratio went down by 20 percentage points and remained stable around 50 percent over the 1986 to 1996 period (Figure).

These aggregate figures conceal a sharp contrast between bank dependent and non- bank dependent firms. As it turns out, firms in more bank-dependent sectors display larger changes in capital structure

AGGREGATE CAPITAL STRUCTURE OF FRENCH FIRMS



Source: Banque de France, Tableaux des opérations financières (Flow of funds), several years.

[1] Many reforms encouraging stock market finance and participation were undertaken in the mid-1980s. Among the prominent ones, in 1986, the monopoly of the Paris brokers was dismantled; between 1984 and 1990, capital controls were progressively removed; in 1986, the French stock market was among the first to become electronic. Tax breaks for stock market investment, simplifications of procedures for bond and equity issues were introduced at various points in the mid-1980s.

post-deregulation. They experience a larger drop in debt post-reform, but this drop in debt finance is only partly compensated by an increase in equity finance. The gap is filled with trade credit, which increases more for firms in bank dependent sectors. The effect of the banking reform is also reflected in an increase in the cost of capital in the more bank-dependent sectors. Furthermore, the largest changes in capital structure occur among the worst performing firms.

While consistent with the idea that banks are becoming more selective in their lending behaviour, these results could however also be driven by changes in the demand for bank capital. Due to the increase in the cost of capital (the mid-1980s where times of sharp monetary tightening and rise in real interest rates), firms might have been optimally restructuring their financing by relying less on bank loans.

To alleviate this concern and understand in more detail whether and how banks are changing their lending behavior after the reform, we looked at the correlation between new net bank loans and shocks to firm performance. The hypothesis we investigated is that banks were more willing to "bail out" poorly performing firms prior to the reform and that this behavior was dampened after the reform. Indeed, we found that in bank dependent industries, the ability of firms with a sudden drop in performance to raise debt was significantly weakened, in particular for structurally weaker firms. This is consistent with banks becoming less inclined to bail out poorly performing firms after the deregulation.

Last, we analyzed whether, conditional on getting new bank loans, firms were more likely to improve their performance after the reform. Reduced distortion in lending and subsequent improvement in banks' monitoring and screening abilities should reduce the provision of credit to firms that will subsequently perform poorly. Again, we found that getting more bank credit became more closely tied to subsequent good performance after the reform, in particular in more bank-dependent industries. This is consistent with banks making much more use of their soft information to grant loans.

Real effects of banking reforms

The previous section showed that the deregulation significantly changed the incentives of banks and led to stricter lending practices post-reform. We then analyzed whether this generated pressures on firms to engage in more cost-cutting and restructuring activities. Such a response would be expected if firms face stronger incentives to strengthen their credit rating. We found that, in the bank-dependent sectors, firms experience slower than average wage growth, invest less and outsource more post-reform. It is interesting to note that these evolutions are even more pronounced among worse performing firms.

Second, we investigated the hypothesis that the banking reform improved the dynamics and competitiveness of product markets. Our first approach was to look at entry and exit rates in various industries. We found that bank dependent industries experience a larger increase in the entry and exit rates of firms after the reform. This is consistent with the fact that changes in bank lending practices have increased the rate of turnover within industries. Both creations and destructions have increased.

If increased reallocation rates are indeed symptomatic of more dynamic and competitive industry structures, we might also expect market concentration to decrease after the banking reform, especially in the more bank dependent sectors. We found this to be the case in our data.

We then asked more directly whether credit allocation between firms improved during that period. We found two pieces of evidence that support this view. First, badly performing firms were more likely to exit their industry after the reforms than before. This was especially acute in bank dependent industries. This first point is consistent with capital being "withdrawn" sooner from worse performers. Secondly, we found that the best performing firms tend to access larger market shares after the deregulation, in particular in bank dependent industries. This last point is consistent with relatively more capital being allocated to better performers.

Finally, evidence of increased competition and improved allocative efficiency led us to look at several measures of efficiency and cost structure at the industry level. We found that bank dependent industries experience a sharper decline in labor costs after the deregulation, as well as stronger increases in employment and labor productivity.

Conclusion

Overall, our findings suggest that a well-functioning banking sector plays an important role in fostering a

Schumpeterian process of creative destruction. The distortions in the banking sector prior to the reforms may have created artificial barriers to entry for new firms by unduly protecting incumbents and thereby dampening the efficiency-inducing effects typically associated with a more competitive environment. Our analysis documents a novel multiplier effect that works through the impact of bank lending on product market dynamics.

In the context of the current policy debate, the French reforms provide a template for a successful multi-tiered deregulation of the banking industry. Contrary to many other episodes of bank deregulation, the French reforms relied on a combination of increased competition, abolishing interest rate targets and directed credit, simplification of regulations and elimination of bank subsidies. Ultimately, these reforms led to a systematic change in the structure and efficiency of the banking system. We conjecture that such thorough reforms may lead to more sustained change than many other recent episodes of banking reforms that relied mainly on bank privatizations without improving any of the other structural dimensions of the banking industry.

References

Bertrand, M., A. Schoar and D. Thesmar (2005), "Banking Deregulation and Industry Structure: Evidence From the 1985 French Banking Reforms", mimeo, University of Chicago.

Black, S. and P. Strahan (2002), "Entrepreneurship and Bank Credit Availability", *Journal of Finance* 57 (6); 2807–833.

Caballero, R., T. Hoshi and A. Kashyap (2003), "Zombie Lending and Depressed Restructuring in Japan", mimeo, University of Chicago.

Cetorelli, N. and P. Strahan (2003), "Finance as a Barrier to Entry: Bank Competition and Industry Structure in Local U.S. Markets", mimeo, Boston College.

King, R. and R. Levine (1993), "Finance and Growth: Schumpeter Might Be Right", *Quarterly Journal of Economics* 108: 717–38.

Naouri, J. C. (1986), "La Réforme du Financement de l'économie", *Revue Banque* 459: 211–21.

Plihon, D. (1995), "L'Evolution de l'Intermédiation Bancaire: 1950-1993", *Bulletin de la Banque de France* 21: 131–59.

Rajan, R. and L. Zingales (1998), "Financial Dependence and Growth", *American Economic Review* 88 (3): 559–86.

Rémy, C. and S. Béatrice (1986), "La Banque en Europe: Les Dix Prochaines Années", *Revue Banque* 457: 37–42.

Schumpeter, J.A. (1934), The Theory of Economic Development, Harvard University Press, Cambridge, MA.

Trends in Development Aid

During a "Millennium Development Summit" of the UN in September 2000 rich and poor countries had subscribed to contributing to an eradication of misery, poverty and underdevelopment on a worldwide scale within 15 years. These aims, the "Millennium Development Goals" (see Box) are so fundamental and self-evident that they can scarcely be disputed. What is disputable, however, is the question how the goals can be achieved and whether their achievement within 15 years, i.e. in 2015, is realistic or has ever been realistic. In any case, a considerable increase in development assistance is seen as a prerequisite for reaching the Millennium Goals.

It is against this background that a look at the factual trends of development aid, its structures and some of its issues could be worthwhile. We consider the official public development assistance of 17 EU countries, of the European Commission as well as of the US and Japan. This article deals with the amount of aid flows and with their regional and sectoral distribution. A further article – to be published in DICE Report 4/2005 – will focus on tied aid, debt relief, donor coordination and the structures of development administration.

The Millennium Development Goals

Goal 1: Eradicate extreme poverty and hunger
Goal 2: Achieve universal primary education
Goal 3: Achieve gender equality and empower women
Goal 4: Reduce child mortality
Goal 5: Improve maternal health
Goal 6: Combat HIV/AIDS, malaria and other diseases
Goal 7: Ensure environmental sustainability
Goal 8: Develop a global partnership for development
www.un.org/millenniumgoals

The size of development aid 1980–2003

It was already in 1970 that the UN General Assembly called upon the rich countries to provide 0.7 percent of their Gross National Income (GNI) as development assistance to poor countries. The discussion of the "Millennium Goals" has repeated and underlined that old demand. The protagonists of the Millennium Goals, including the economist Jeffrey Sachs, believe that the goals are attainable if all rich countries contribute 0.7 percent and also that 0.7 percent is a necessary precondition for realising the goals.

However, on average among donor countries, development aid has always been far from the 0.7 percent benchmark – and has practically not been increased since 1980 (Table 1). Between 1990 and 2003 official public aid as a percentage of GNI was reduced in several large donor countries, e.g. in France, Germany, Japan and the US. It is only in some small European countries, like Denmark, the Netherlands, Sweden and Norway, that public aid to developing countries has been kept at a high level, even surpassing the 0.7 percent level. In all other countries the percentages are far below that benchmark. However, since 2000 some countries have attained a slight increase in their aid, again measured in relation to GNI.

The regional and sectoral structure of development assistance

The majority of the very poor countries are located in Africa south of the Sahara. Twenty percent of the world's poor live in this region. Consequently, development aid is concentrated in this region (Table 2). Even within this group of countries there is a concentration. The Democratic Republic of Congo (DRC, formerly Zaire),is one of the top receivers of aid among 8 of the 20 donor countries considered here. These countries have also increased their assistance to the group of the very poor countries (least developed or LLDC) by about 30 percent since 1992. However, one must take into consideration that aid to some countries that do not belong to the very poor has been increased much more – e.g. aid to Iraq and Jordan.

The sectoral distribution of development aid has undergone a significant change between 1982/83 and 2002/03 (Table 3). From the 1960s to the 1980s development assistance was mainly used to foster the construction and improvement of the physical infrastructure, as well as for agricultural and industrial projects ("hard" infrastructure). By contrast the donor countries today have shifted their resources to support and improve what has been called "soft" infrastructure. The funds to promote social, economic and administrative competence in developing countries have been nearly doubled since the beginning of the 1980s (from 21.6 to 39.9 percent of total aid). Portugal and Greece have even dedicated three quarters of their total assistance for such purposes. For several other donor countries, aid for soft infrastructure is at about 50 percent of their total aid.

The remarkable change in the sectoral structure of aid is the result of a deeper understanding of the development process – on the part of the donor as well as on that of the receiving countries. Economic development is only superficially a problem of a lack of capital or of resources. Fundamentally, development depends, first, on economic institutions that give individually and socially reasonable incentives to farmers, traders, industrialists, workers and bureaucrats, and second, on good governance. Thus, donor countries also try to use development aid to support the struggle against corrupt elites and to promote reasonable forms of political participation – other aspects of improving the soft infrastructure.

Table 1

Development Aid, in percent of GNI

	1970	1980	1990	2000	2003
Austria	0,07	0.23	0.11	0.23	0.20
Belgium	0.46	0.50	0.46	0.36	0.60
Denmark	0.37	1.24	0.94	1.06	0.84
Finland	0.06	0.22	0.65	0.31	0.35
France	0.52	0.44	0.60	0.32	0.41
Germany	0.32	0.44	0.42	0.27	0.28
Greece	n.a.	n.a.	n.a.	0.20	0.21
Irland	n.a.	0.16	0.16	0.29	0.39
Italy	0.15	0.15	0.31	0.13	0.17
Luxemburg	n.a.	0.11	0.21	0.71	0.81
Netherlands	0.62	0.97	0.92	0.84	0.80
Portugal	n.a.	0.02	0.24	0.26	0.22
Spain	n.a.	0.08	0.20	0.22	0.23
Sweden	0.35	0.78	0.91	0.80	0.79
UK	0.39	0.35	0.27	0.32	0.34
EU	n.a.	n.a.	n.a.	n.a.	n.a.
Norway	0.33	0.87	1.17	0.76	0.92
Switzerland	0.14	0.24	0.32	0.34	0.39
Japan	0.23	0.32	0.31	0.28	0.20
USA	0.32	0.27	0.21	0.10	0.15

Note: Development aid is here official development assistance (ODA), net disbursements.

Source: OECD, DAC, Online (www.oecd.org/dac/stats/idsonline).

R. O. and A. R.*

* Anja Rücker holds an MA in Social and Economic Science and is a development advisor to the South African government; anja@ruecker@gtz.de.

Table 2

Regional distribution of official development assistance (ODA), in percent of net disbursements[a], 2003

Donor country	Top 3 recipients of net ODA 2003	Sub-Sahara Africa	South & Central Asia	Other Asia and Oceania	Middle East and North Africa	Europe	Latin America and Caribbean	Net ODA (disbursements) spent on LDC, 2003 (current million US$)	Net ODA spent on LDC as % of net ODA disbursements, 2003
Austria	Bosnia-Herzogowina, Ex-Yugoslavia unspecified, Turkey	39.7	8.9	5.2	12.4	26.9	7.0	64.29	12.73
Belgium	DRC, Ruanda, Cameroon	72.8	4.1	4.1	5.5	5,8	7.8	959.87	51.80
Denmark	Tanzania, Bangladesh, Uganda	52.8	12.7	11.1	7.1	5.7	10.5	448.02	25.63
Finland	Tanzania, Mozambique, Afghanistan	45.5	16.1	9.8	9.5	9.1	10.0	99.28	17.80
France	DRC, Cameroon, Serbia & Montenegro	58.1	4.8	8.2	14.7	8.5	5.8	2,247.59	30.99
Germany	DRC, Cameroon, China	44.1	12.1	8.9	10.6	12.8	11.6	1,551.36	22.87
Greece	Albania, Georgia, Iraq	17.5	10.3	2.4	11.3	54.4	4.2	10.79	2.98
Ireland	Uganda, Mozambique, Ethiopia	73.4	7.1	3.5	6.1	4.2	5.7	226.05	44.85
Italy	DRC, Nicaragua, Palestinean administr. Areas	62.8	7.3	1.4	11.9	9.4	7.1	722.09	29.68
Luxemburg	Vietnam, Cape Verde, Laos	43.4	7.5	13.0	10.7	9.2	16.2	55.88	28.80
Netherlands	DRC, Tanzania, Indonesia	49.4	15.2	10.2	7.3	7.6	10.3	979.49	24.60
Portugal	Timor-Leste, Cape Verde, Angola	51.2	8.9	24.9	6.1	5.7	3.2	150.84	47.14
Spain	Nicaragua. Honduras, Bolivia	26,3	6,2	6.4	12.8	12.8	35.6	160.36	8,.8
Sweden	DRC, Tanzania, Mozambique	50.8	14.8	8.7	6.8	8.2	10.7	608.76	25.37
UK	India, Tanzania, Bangladesh	43.8	24.3	5.6	8.1	10.3	7.8	1,348.57	21.47
Norway	Tanzania, Afghanistan, Iraq	46.9	18.0	6.2	10.7	11.1	7.2	577.61	28.29
Switzerland	Serbia & Montenegro, DRC, Tansania	39.2	21.8	8.5	5.6	12.3	12.6	255.21	19.65
European Communities	Serbia & Montenegro, Sudan, Afghanistan	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	2,309.70	32.19
Japan	Indonesia, China, Philippines	17.7	28.2	38.4	4.6	1.9	9.3	1,077.95	12.14
USA	Iraq, DRC, Jordan	35.9	14.4	7.4	21.4	7.7	13.2	4,238.56	26.08

[a] Including imputed multilateral flows, e.g. making allowance for contributions through multilateral organisations, calculated using the geographical distribution of multilateral disbursements for the year of reference.
DRC: Democratic Republic of Congo. LDC: Least developed countries.

Source: Compilation by the authors from OECD, DAC, Statistical Annex of the 2004 Development Co-operation Report (www.oecd.org/dac/stats/dac/dcrannex) ; OECD, DAC, Online (www.oecd.org/dac/stats/idsonline).

Table 3

Sectoral distribution by donor country, in percent of total bilateral commitments, 1982/83 and 2002/03

Donor country	Social and administrative infrastructure		Economic infrastructure		Agriculture		Industry and other production		Commodity Aid and programme Assistance		Emergency Aid		Other		Sum	
	1982/83	2002/03	1982/83	2002/03	1982/83	2002/03	1982/83	2002/03	1982/83	2002/03	1982/83	2002/03	1982/83	2002/03	1982/83	2002/03
Austria	12.7	39.3	62.3	7.7	1.2	1.4	13.5	2.2	1.3	0.4	0.7	9.1	8.2	40.0	100.0	100.0
Belgium	11.1	25.7	4.2	4.7	3.5	4.3	34.3	0.6	1.9	2.7	0.4	6.2	44.6	55.8	100.0	100.0
Denmark	20.4	38.5	22.6	19.2	19.4	4.6	32.1	3.1	n.a.	0.6	-	8.4	5.4	25.6	100.0	100.0
Finland	15.9	49.3	21.1	5.4	16.9	4.3	35.1	2.2	0.2	2.6	3.7	11.9	7.2	24.4	100.0	100.0
France	51.3	30.1	14.8	4.4	9.5	2.4	6.9	0.9	6.7	3.4	0.2	6.3	10.6	52.5	100.0	100.0
Germany	29.7	36.2	28.7	11.6	10.3	3.1	13.8	1.1	2.9	0.9	0.7	3.8	14.0	43.3	100.0	100.0
Greece	n.a.	83.1	n.a.	1.6	n.a.	0.7	n.a.	0.7	n.a.	0.0	..	5.1	n.a.	8.7	_	100.0
Ireland	n.a.	65.6	n.a.	2.1	n.a.	5.1	n.a.	0.4	n.a.	5.7	-	7.0	100.0	14.1	100.0	100.0
Italy	19.3	13.1	21.6	1.2	18.5	1.5	10.8	0.8	9.8	5.0	1.7	6.8	18.3	71.6	100.0	100.0
Luxemburg	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	..	..	n.a.	n.a.	_	-
Netherlands	32.0	24.6	14.9	6.8	17.7	2.5	9.1	0.4	4.3	3.2	1.9	3.3	20.0	59.2	100.0	100.0
Portugal	n.a.	75.5	n.a.	3.6	n.a.	2.0	n.a.	0.9	n.a.	0.8	..	0.8	n.a.	16.4	_	100.0
Spain	n.a.	41.0	n.a.	17.4	n.a.	4.5	n.a.	2.4	n.a.	0.7	..	4.7	n.a.	29.3	_	100.0
Sweden	21.4	33.2	5.8	8.0	12.9	2.1	12.2	0.8	0.6	3.1	12.7	22.8	34.4	30.0	100.0	100.0
UK	18.7	37.6	26.9	7.5	7.8	5.3	25.3	2.0	3.7	0.3	0.4	12.6	17.2	34.7	100.0	100.0
Norway	19.6	48.6	24.9	7.2	17.5	4.5	12.8	0.8	0.1	3.6	7.0	18.7	18.1	16.6	100.0	100.0
Switzerland	20.4	19.6	2.1	9.4	23.0	5.3	31.4	4.0	8.9	5.1	11.2	17.7	3.1	38.8	100.0	100.0
European Communities	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.		
Japan	13.1	20.4	43.8	27.8	10.4	6.8	15.4	1.6	4.0	0.4	0.1	0.7	13.2	42.4	100.0	100.0
USA	17.3	37.1	4.2	3.3	14.2	1.8	13.4	4.3	26.1	13.8	1.9	12.7	22.9	27.0	100.0	100.0
Average	21.6	39.9	21.3	8.3	13.0	3.5	19.0	1.6	5.4	2.9	2.2	8.3	22.5	35.0		

Note: Commitments and disbursements may differ in a given year. 1982/83 means the average for 1982 and 1983.

Source: Compilation by the authors from OECD, DAC, Statistical Annex of the 2004 Development Co-operation Report (www.oecd.org/dac/stats/dac/dcrannex) ; OECD. DAC Online (www.oecd.org/dac/stats/idsonline).

Foreign Students' Access to Labour Markets upon Graduation

As recent policy changes demonstrate, OECD countries are increasingly seeking to retain their foreign university students upon graduation and facilitate their access to the labour market. Starting in 2001, Australia allows foreign students to remain in the country up to six months following graduation while they apply for permanent residence. Germany enacted new laws as of 2005 making it possible for all foreign students to remain in the country for one year and to find work. Canada, as recently as May 2005, extended the post-graduation work permit from one year to two for graduates of certain universities. There appear to be several rationales for such developments. For one, foreign graduates can be easily integrated into the labour market of the host country due to their linguistic mastery and their familiarity with the local social manners. Second, university graduates represent a highly skilled portion of the labour force, adding to a country's potential "brain gain".

In absolute terms, the United States, the United Kingdom, Germany, France and Australia are the leading host countries for foreign students. In relation to their total student population, Australia, Switzerland, Austria, Belgium, Germany, the United Kingdom and France have the highest intake of foreign students. In 2002, all of these seven countries had at least 100 foreign students for every 1000 enrolled (Table 1).

The possibilities for students to change their resident status upon graduation and obtain access to the labour market differ among OECD countries. Out of nine countries examined, five of the countries have provisions that allow a student to remain in the foreign country unemployed (between 6 and 36 months depending on the country) following graduation. In the case of Germany, New Zealand, the United Kingdom and the United States, the explanation for this allowance is that the student is assumed to be seeking work. The United Kingdom only makes this allowance for science and engineering, and select business graduates. Australia allows all graduates six months to reside in Australia while and if they apply for permanent residence. In contrast, Ireland, Poland and Sweden have no provisions allowing students to remain upon graduation in their country, even if they have a job offer. Students studying in these countries must often leave the country when their student visa expires (Table 2). In Canada, remaining in the country is contingent upon obtaining a job offer at graduation. Graduates of Canadian universities have the advantage over other foreign workers of not needing to obtain confirmation from Human Resources Canada saying that their job does not displace others in the Canadian labour market.

All of the nine countries examined have temporary work schemes that allow foreign workers, upon finding employment, to stay in the host country for two or more years. However, explicit privileges for graduates of host-country universities are normally not granted, unless the work permit is exclusively provided for graduates, as is the case with Canada.

In obtaining permanent residency only Australia and New Zealand facilitate their foreign students. If a candidate received his degree from an Australian institution, Australia grants both extra points and lowers the pass-mark on the assessment test for "Skilled Migration". In New Zealand, graduates applying under the "Skilled Migrant" category also obtain extra points on a points-assessed test if they hold a New Zealand qualification (Table 2).

From the countries examined, we see that the special residence laws for university graduates can facilitate both a foreign graduate's entrance into the labour market of the country in which he studied and long term settlement and immigration to the country. Unfortunately, it is difficult to know exactly how many

Table 1

Foreign students from throughout the world as a percentage of all foreign and domestic students, 2002

Australia	17.7	Hungary	3.3	Norway	4.8
Austria	12.7	Iceland	4.1	Poland	0.4
Belgium	11.0	Ireland	5.2	Portugal	n.a.
Canada	n.a.	Italy	1.5	Slovak Republic	1.1
Czech Republic	3.4	Japan	1.9	Spain	2.4
Denmark	7.4	Korea	0.2	Sweden	7.5
Finland	2.4	Luxembourg	n.a.	Switzerland	17.2
France	10.0	Mexico	0.1	Turkey	1.0
Germany	10.1	Netherlands	3.7	United Kingdom	10.1
Greece	1.6	New Zealand	9.5	United States	3.7

Source: OECD, Education at a Glance – OECD Indicators 2004, Paris 2004, p. 306.

foreign students stay in the foreign country after graduation, be it temporarily or permanently. "There has still been no measurement of this at international level, and it is often necessary to make do with incomplete data, particularly due to the lack of information on change from the status of student to holder of a work permit or permanent immigrant status" (Tremblay 2004). Nevertheless, some statistics are available in specific countries.

Tremblay compared the number of permanent residence permits granted to foreign students by Australia under the Skilled Migrant Programme with OECD data on the number of students studying in Australia. She found that in 2002, out of 41,148 foreign graduates in Australia, 6,274, or 15.2 percent, were granted permanent residence permits. In 2003, out of 47,638 graduates, 8,890 graduates, or 18.7 percent were granted permanent residence permits. These numbers, however, are merely low range estimates of foreign graduate stay-rates, since several possibilities exist for students to stay in Australia. For example, a student can stay in Australia under a temporary permit and later apply for permanent residence, or obtain permanent residence through marriage.

Finn (2003) calculated stay rates of foreign doctorate recipients in the United States by tracing tax records. Finn found that of the 5,557 foreign doctorate recipients in the class of 1987, 49 percent were paying taxes two years later in 1989. More recent short-term stay rates are dramatically higher. Of the 8,770 foreign doctoral recipients who graduated in 1999, 71 percent were still paying taxes two years later. Finn also looked at five and ten year stay-rates, concluding that 65 percent of the foreign doctorate class of 1996 were still paying US taxes five years later in 2001, while 58 percent of the foreign doctorate class of 1991 were still paying taxes ten years after graduation in 2001.

M. L.* and W. O.

References

Finn, M. (2003), *Stay Rates of Foreign Doctorate Recipients from U.S. Universities*, Oak Ridge.

Tremblay, K. (2002), "Student Mobility Between and Towards OECD Countries: A Comparative Analysis", in OECD (ed.), *International Mobility of the Highly Skilled*, Paris, 39–67.

Tremblay, K. (2004), "Links Between Academic Mobility and Immigration", paper presented at the Symposium on International Labour and Academic Mobility: Emerging Trends and Implications for Public Policy, Toronto, http://www.wes.org/ewenr/symp/KarineTremblayPaper.pdf.

* Marcus Lampert is an undergraduate from Princeton University and did an internship at the Ifo Institute in summer 2005.

Table 2

Regulations regarding possibilities for student visa holders to change residence status, 2005

	Temporary residence opportunities		Advantages and paths for graduates to obtain permanent residence
	Length of time a student can remain in the country, following graduation, without a job offer	Length of time a student can remain in the country, following graduation, with a job offer	
Australia	The student may remain in Australia without a job offer until the expiration of the student visa. If the graduate wishes to apply for permanent residence he or she can apply to remain in Australia for an additional six months after completing his or her studies.	It depends on the type of work permit applied for. There is no temporary work scheme specifically designed for graduates of Australian institutions.	Permanent residence is granted in the Skilled Migration Category according to a points-assessed test. Applicants receive extra points and are given a lower pass mark if they hold an Australian qualification.
Canada	Without a job offer, students must leave when their student visa expires.	One year. Two years for those who studied and have found work outside of certain metropolitan areas. In contrast to other foreigners, graduates do not need to obtain confirmation from Human Resources Canada saying that their job does not displace others in the Canadian labour market.	None specifically for graduates of Canadian Educational Institutions. Graduates with at least one year of work experience may apply for permanent residence, which is granted according to a points-assessed test. While educational qualifications do earn additional points, having studied in Canada does not earn one extra points.
Germany	Graduates can remain in Germany for up to one year after graduation without a job offer for the purpose of finding employment.	If the student has found a job offer or employment, it is up to the local authority to grant a work permit. There is no national law limiting the length of a work permit.	Foreign graduates of German universities do not specifically have advantages for permanent residence. Germany does however immediately grant permanent residence to those defined as “highly skilled”. The “highly skilled” includes “researchers and scientists with specialized training and expertise”.
Ireland	Without a job offer, students must leave when their student visa expires.	Even with a job offer, a work permit can only be applied from outside of Ireland, usually in the country of the person’s residence. A work permit is usually valid for two years and can be extended while remaining in Ireland.	None specifically designed for foreign graduates of Irish universities. Ireland has no permanent residence status. One can remain in Ireland on a work permit and obtain habitual residence rights or become a citizen.
New Zealand	Since the academic year ends in December and student visas are usually valid until the following March, graduates have a three-month time during which they can stay in New Zealand without a job offer. If the student still does not have a job offer, he or she can apply for a “Graduate Job Search Work Permit”, which allows the graduate to remain in New Zealand for six months to look for employment.	A regular work permit is valid for a maximum of two years but is renewable.	Graduates applying under the Skilled Migrant category for permanent residence obtain extra points on the points-assessed test if they hold a New Zealand qualification. Applying for permanent residence is not grounds for remaining in the country after studying.

(Table 2 continued)

	Temporary residence opportunities		Advantages and paths for graduates to obtain permanent residence
	Length of time a student can remain in the country, following graduation, without a job offer	Length of time a student can remain in the country, following graduation, with a job offer	
Poland	Without a job offer, students must leave when their student visa expires.	All residence permits are issued for a maximum of two years, but can be renewed. Obtaining a work permit is handled by the local authorities, and it is possible that the graduate might have to go to a consulate outside of Poland to change from a student to a work visa. This can usually be done at a consulate near Poland, like Berlin or Prague, for example.	No specific advantages for foreign graduates of Polish universities. A foreigner can obtain permanent residence after having lived in Poland for five continuous years.
Sweden	Requests are handled on a case by case basis, but generally graduates must return to their home country upon completion of their studies.	Requests to stay and work are handled on a case by case basis.	No specific advantages for foreign graduates of Swedish universities.
United Kingdom	Without a job offer, most students must leave when their student visa expires. However, science and engineering graduates and graduates of the world's top 50 business schools can remain for one year following graduation without a job offer for the purpose of finding employment.	It depends on the type of work permit applied for. There is no temporary work scheme specifically designed for graduates of United Kingdom institutions.	None specifically for graduates of UK Educational Institutions. The Highly Skilled Migrant Programme grants permanent residence according to a points-assessed test. While educational qualifications do earn additional points, having studied in the UK does not earn one extra points.
United States	Students can remain in the country without a job offer for 12–36 months after graduation, depending on their visa and degree, to find employment. The specific time periods are: 12 months on an F-1 Student Visa 18 Months on a J-1 Exchange Visitor Visa 36 Months for PhD graduates on a J-1 Exchange Visitor Visa.	If, after the time granted by F-1 and J-1 to find work, the student has found an employer, the employee can be granted an H1-B visa. The visa lasts for three years and can be renewed for an additional three years.	None specifically for graduates of US Educational Institutions. Those already in the job market have the possibility of obtaining permanent residency (a "Green Card") through employment-based immigration.
Note: Many countries have agreements with other countries that make it significantly easier to study and subsequently join the workforce as a foreigner. For example, all members of the European Economic Area (EEA), which includes Germany, Ireland, Poland, Sweden and the United Kingdom, are supposed to allow free movement of labour and residence. This means that it is significantly easier for a student from an EEA country to study and work in another EEA country than described above. Likewise, Australian and New Zealand have more lenient visa and immigration laws with one another. Mexican and Canadian workers can obtain a special work permit for the United States			

Sources: Tremblay (2004). "Extending Your Stay", "Studying in Australia", "Working in Australia", Australian Department of Immigration and Multicultural and Indigenous Affairs (http://www.immi.gov.au/). "Immigrating to Canada as a Skilled Worker", "Studying in Canada", "Working Temporarily in Canada", Citizenship and Immigration Canada, http://www.cic.gc.ca/english. "Zuwanderung A-Z", Bundesministerium des Innern, http://www.zuwanderung.de/. "Working in Germany", Auswärtiges Amt, http://www.auswaertiges-amt.de/. "Information on Work Permits and Working Visas/Authorisation", Embassy of Ireland in Berlin, http://www.botschaft-irland.de/. "Work Visas/Authorisations Information Leaflet", The Department of Enterprise, Trade and Employment, http://www.entemp.ie. Immigration New Zealand, http://www.immigration.govt.nz/. "Permit for Residence", Polish Ministry of Foreign Affairs, http://poland.gov.pl/?document=1618. The Home Office of the United Kingdom, http://www.workingintheuk.gov.uk, http://www.ukvisas.gov.uk. "J-1 Student Status", http://www.oiss.yale.edu/visa/j1status.htm. "Information about F-1 Status", http://www.oiss.yale.edu/visa/f1status.htm. "Introduction to H1-B Visa", http://www.workpermit.com/ us/ employer_h-1b.htm. "H1-B Visa", http://www.internationalstudent.net/H1-B_Visa.htm. Embassies and consulates.

Export Promotion through Embassies?

Most governments regard the promotion of their nation's exports as a major task. One instrument of export promotion used by practically all countries is to assign an export promotion role to their – already existing – foreign embassies as well as to set up additional consulates abroad. But is there really a link between exports and the number of official foreign missions (embassy plus, possibly, consulate[s])? This question has been addressed in a recent paper by Andrew K. Rose (2005).

First we look at the factual number of embassies and consulates (E&C) held by 22 major export countries (henceforth: club countries or club). The Table shows that the 22 club countries hold many more E&C in countries outside the club than within the club (2597 vs. 1560). Within the club, the Netherlands holds by far the highest number of E&C (206), followed by Switzerland (136). Most club countries hold about 50 or 60 E&C in other club countries. With respect to the number of E&C held in a club country by other club countries, the picture is similarly diverse. In the US club countries hold 247 E&C, followed by Germany where 134 E&C are held. In most other club countries much fewer E&C are held.

Moreover, it is interesting to compare the totals of column 23 with the totals of line 23. For most countries the number of E&C "received" and "provided", so to say, do not differ substantially. However, for some countries the similarity does not hold. In the US and Germany the number of E&C received (247 and 134, respectively) is much larger than the number of E&C provided (80 and 68, respectively). In the Netherlands and Switzerland it is quite the other way round: There are many more E&C provided than received.

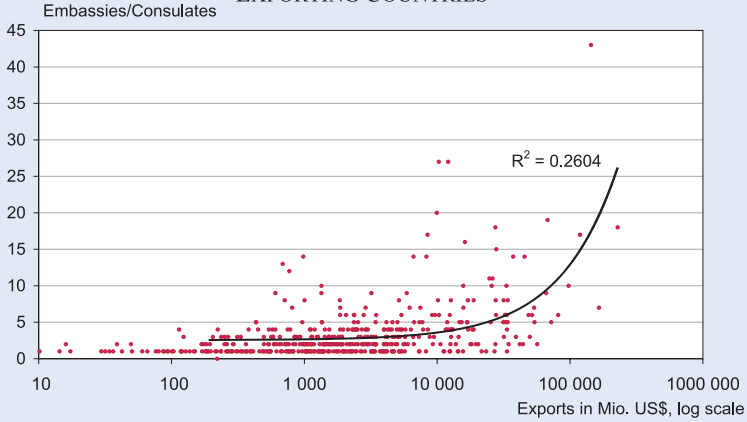
Now let us relate exports (figures not shown) to E&C for the club countries. 22 countries export and provide E&C to 22 destination countries (a single country only to 21 destination countries). Thus, we have 22 x 21 datapoints for exports and E&C. In the Figure each point stands for the export value and the number of E&C provided from one club country to another club country. From the outset, one cannot expect a very strong correlation because exports are influenced by many more (and more important) factors than by E&C. And E&C are influenced by many more (and more important) considerations than by exports. Nevertheless, the correlation coefficient (0.26) is not quite low.

The question whether there is a link between exports and E&C, possibly a causal link running from E&C to exports, is treated by Rose by econometric means. His estimation model of exports is basically rooted in the well-known (and not new) gravity approach, which usually employs distance, GDP, GDP per capita, population, area and the like as factors determining exports of country A to country B. Rose adds further factors, not always employed in gravity approaches, like whether countries A and B have a common language or currency, whether one of the two countries is landlocked or an island, whether they belong to a regional trade agreement or whether A has been a colony of B and so forth. This long list of variables is now further enhanced by A's number of E&C in B (and, of course, vice versa).

The dataset Rose employs is much larger than our Table and Figure suggest. He considers the exports (and E&C) of the 22 major exporters, but not only to 22 but to 200 destination countries. The problem of a possible reverse causality (running from exports to E&C) is tackled by the introduction of instrumental variables which measure the geopolitical importance of a country as well as possible preferences of the diplomatic corps for certain countries.

The general result is that more than 75 percent of the variation in bilateral exports is explained by the mo-

Exports and Embassies/Consulates of Major Exporting Countries



Note: The export data relate to 2002 and 2003 (average).
Source: Own compilation from a dataset provided by Andrew K. Rose (see Rose, 2005).

Embassies and Consulates of 22 major exporting countries

Embassies and Consulates of	in																						Total E&C in club countries	Total E&C in non-club countries	World
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)
Australia (1)		1	2	2	3	1	2	2	2	2	6	1	1	1	1	1	1	1	1	2	1	7	41	56	**97**
Belgium (2)	2		3	3	2	3	2	1	0	1	1	1	1	1	1	1	3	1	2	2	1	5	37	71	**108**
Brazil (3)	2	1		3	2	1	3	1	1	2	2	1	1	2	1	1	2	1	2	1	1	8	39	78	**117**
Canada (4)	2	1	4		4	1	4	5	1	2	6	1	3	1	1	2	2	1	1	1	1	18	62	85	**147**
China (5)	4	1	3	4		3	3	2	1	3	5	2	2	1	2	3	2	2	2	2	3	6	56	160	**216**
France (6)	2	3	4	6	4		7	3	1	5	2	1	1	2	2	2	6	1	4	2	2	11	71	163	**234**
Germany (7)	3	1	5	4	4	5		4	1	3	1	1	1	2	5	4	6	1	2	4	2	9	68	141	**209**
India (8)	4	3	2	3	2	1	4		2	4	2	1	1	1	1	3	3	1	3	3	3	5	52	134	**186**
Indonesia (9)	6	3	1	4	1	2	10	3		4	4	2	3	1	1	1	2	1	1	2	1	7	60	72	**132**
Italy (10)	6	5	7	5	3	10	14	4	1		2	1	1	2	1	4	2	1	8	3	4	11	95	134	**229**
Japan (11)	6	1	8	5	6	3	5	4	5	2		3	1	1	1	5	3	1	2	2	2	17	83	122	**205**
Korea (12)	2	1	2	4	5	1	3	2	1	1	10		1	1	1	2	1	1	2	1	1	8	51	76	**127**
Mexico (13)	1	1	1	4	2	1	2	1	1	2	1	1		1	1	1	2	1	1	1	1	43	70	50	**120**
Netherlands (14)	7	5	13	14	3	18	19	3	5	16	6	1	9		4	2	20	8	5	6	15	27	206	255	**461**
Poland (15)	2	1	4	4	3	4	5	2	1	2	1	1	1	1		4	2	2	1	2	2	4	49	86	**135**
Russia (16)	2	2	3	3	3	3	6	4	1	4	4	2	1	1	4		2	2	2	4	2	4	59	169	**228**
Spain (17)	3	1	5	3	2	10	7	1	1	4	1	1	2	2	1	1		1	3	2	3	9	63	102	**165**
Sweden (18)	1	1	1	1	3	1	2	1	1	1	1	1	1	1	2	3	1		1	2	1	2	29	63	**92**
Switzerland (19)	8	2	12	8	2	17	8	4	2	14	2	2	4	3	1	2	7	2		5	4	27	136	165	**301**
Turkey (20)	3	3	1	1	2	5	14	1	1	2	1	1	1	3	1	2	1	1	3		2	5	54	94	**148**
UK (21)	6	1	9	5	4	8	8	9	1	4	4	1	4	2	1	3	8	1	2	4		14	99	162	**261**
US (22)	4	1	4	7	5	4	6	4	3	6	6	1	10	2	2	4	2	1	1	3	4		80	159	**239**
Sum (23)	76	39	94	93	65	102	134	61	33	84	68	27	50	32	35	51	78	32	49	54	56	247	1,560	2,597	**4,157**

Note: E&C means: Embassies and Consulates; 1 means: there is one embassy and no consulate; 3 means there are one embassy and 2 consulates; China is China, P.R. (Mainland). The figures mainly relate to 2004.

Source: own compilation from a dataset provided by Andrew K. Rose (see Rose, 2005).

del and that the coefficients of the variables are plausible in sign and size. But the estimation also shows that there is still a role for E&C to explain export differentials. One consulate more is on average associated with a 6 to 10 percent higher export level. This effect is strongest for the first "consulate" (which is an embassy) and decreasing in the number of consulates. The effect is significant, but relatively small compared with that of most of the usual gravity model variables.

R. O.

References

Rose, A. K. (2005), The Foreign Service and Foreign Trade: Embassies as Export Promotion, NBER Working Paper 11111, February 2005.

Global Competitiveness

Competitiveness has become a fixation for political leaders, the popular press, corporations, and national and international institutions. Even average citizens worry about the "competitiveness" of a nation when, puzzled, they observe how outsourcing or manufacturing relocation takes jobs from their home country.

The World Economic Forum presents a new index of competitiveness: the Global Competitiveness Index. This new index is designed with the goal of unifying the two indexes currently produced by the World Economic Forum (the Growth Competitiveness Index and the Business Competitiveness Index). The new index is described here without being assessed.

Competitiveness is defined as the set of institutions, policies and factors that determine the level of productivity of a country. The level of productivity, in turn, sets the sustainable level of prosperity that can be earned by an economy. In other words, more competitive economies tend to be able to produce higher levels of income for their citizens. The productivity level also determines the rates of return obtained by investments in an economy. Given that the rates of return are the fundamental determinants of the aggregate growth rates of the economy, a more competitive economy is one that is likely to grow at larger rates over the medium to long run.

The new index is based on three principles:

– The first principal is that the determinants of competitiveness are complex. This complexity is captured by 12 pillars of economic competitiveness which are allocated to three subindexes:

  Subindex 1: Basic Requirements:
  - Institutions
  - Infrastructures
  - Macroeconomic Stability
  - Security
  - Basic Human Capital

  Subindex 2: Efficiency Enhancers:
  - Advanced Human Capital
  - Goods Market Efficiency
  - Labour Market Efficiency
  - Financial Market Efficiency
  - Technological Readiness
  - Openness and Market Size

  Subindex 3: Innovation Factors:
  - Business Sophistication
  - Innovation

– The second principle on which the Global Competitiveness Index is founded is that economic development is a dynamic process of successive improvement, in which economies find increasingly sophisticated ways of producing and competing. In other words, the process of economic development evolves in stages. The countries are allocated to one of the (three) stages of development (or the transitional stages).

– The third principle on which the new index is founded is that, as economies develop, they move from one stage to the next in a smooth fashion rather than in abrupt jumps. Thus, the weights of each of the subindexes change as a country develops. The theory of stages of development suggests that, for less advanced economies, the basic requirements are more important; for intermediate economies, efficiency enhancers are the key; and for advanced countries, innovation factors are central.

The Table displays the results of computing the Global Competitiveness Index for 2004. The countries are ranked in decreasing order. The second column contains the value of the index and the third column the rank number. The following six columns contain the information (value and rank) for each of the three subindexes: the basic requirements subindex, the efficiency enhancers subindex, and the innovation factors subindex, respectively.

The most competitive country in the world is the United States, with an overall score of 5.21, followed by Finland (5.04), and Denmark (4.95), Switzerland (4.93), and Sweden (4.92), which constitute the top five. The United States does not score particularly well in basic requirements (rank 18: the main cause is its dismal macroeconomic stability – rank 83 – and its not-so-great performance in security – rank 37). But it is the world's leader in both efficiency enhancers and innovation factors. Since the United States is in the third stage of development (the innovation stage), the weight of the basic requirements is relatively minor, and so the other two subindexes put this country in the leading position. Finland leads the world in basic requirements, but it only ranks 6th in efficiency enhancers (it is interesting to notice its 18th place in labour market efficiency) and 4th in innovation factors. Denmark is 2nd, 5th, and 8th, respectively.

The next five are Germany (4.86), Singapore (4.85), Hong Kong (4.81), United Kingdom (4.80), and Japan (4.79). Of particular interest is the extremely

The global competitiveness index

Country	Overall index		Three main components					
			Basic requirements		Efficiency enhancers		Innovation factors	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank
United States	5.21	1	5.50	18	5.02	1	5.18	1
Finland	5.04	2	6.03	1	4.54	6	4.70	4
Denmark	4.95	3	5.99	2	4.55	5	4.42	8
Switzerland	4.93	4	5.88	6	4.45	8	4.61	6
Sweden	4.92	5	5.76	8	4.45	7	4.69	5
Germany	4.86	6	5.75	10	4.27	14	4.74	3
Singapore	4.85	7	5.89	5	4.60	4	4.15	14
Hong Kong SAR	4.81	8	5.95	4	4.65	2	3.90	20
United Kingdom	4.80	9	5.48	20	4.61	3	4.38	9
Japan	4.79	10	5.35	22	4.26	15	4.94	2
Taiwan	4.72	11	5.31	23	4.38	10	4.44	7
Netherlands	4.72	12	5.59	14	4.36	11	4.32	10
Iceland	4.70	13	5.80	7	4.42	9	3.97	19
Norway	4.69	14	5.96	3	4.25	16	4.01	18
Canada	4.66	15	5.58	15	4.30	12	4.23	13
Australia	4.63	16	5.70	11	4.28	13	4.04	17
France	4.60	17	5.49	19	4.15	20	4.30	11
Austria	4.57	18	5.61	12	4.13	21	4.11	16
Belgium	4.54	19	5.51	17	4.12	22	4.14	15
New Zealand	4.54	20	5.60	13	4.25	17	3.87	23
Ireland	4.38	27	5.23	27	4.11	23	3.88	21
Spain	4.10	34	5.08	33	3.79	28	3.53	29
Slovenia	4.09	35	5.13	32	3.63	32	3.41	32
Czech Republic	4.06	38	4.83	42	3.60	36	3.35	34
Lithuania	4.06	39	4.80	43	3.61	34	3.30	37
Portugal	4.05	40	5.21	28	3.64	31	3.20	41
Slovak Republic	4.03	41	4.78	45	3.61	35	3.16	44
Latvia	3.97	44	4.86	40	3.47	47	2.93	63
Hungary	3.96	46	4.75	47	3.51	44	3.08	48
Greece	3.84	51	4.94	37	3.47	48	3.13	47
Italy	3.80	56	4.56	54	3.48	46	3.48	30
Poland	3.57	72	4.25	69	3.13	63	2.99	58

low score of Germany in labour market efficiency (82nd in the world). Sweden is ranked quite low in that same pillar (25th), which is led by Hong Kong (1st) and Singapore (2nd).

W. O.

Reference

World Economic Forum, *The Global Competitiveness Report 2004–2005*, Basingstoke 2004, pp. 64–5.

End-of-Life Vehicles: Free Take-back and Other Instruments

End-of-life vehicles (ELV) can exert three major externalities on the environment: by dumping the ELV in the environment, by releasing pollutants in the course of ELV treatment operations (recycling, recovery, reuse) and by land filling of automobile-shredding residue. Although the EU figures on the annual amount of ELV seem to be relatively uncertain, the problem is not on a minor scale. For the EU15 and Norway it is estimated that in 1998 there were 11.5 million cars deregistered, of which 7.7 million were ELV (the bulk of the difference being exported for ELV treatment or reuse abroad).

Before the adoption of an EU directive on ELV (Directive 2000/53) there were (and still are) specific regulations and/or voluntary industry agreements in the member countries. Most of the regulations and agreements stipulate technical and administrative procedures but do not rely on the use of economic instruments. It is only in the Netherlands and in Sweden that economic instruments for addressing the mentioned externalities of ELV are used.

In the Netherlands a "disposal fee" (or recycling fee) was introduced during the 1990s. The instrument consists of three elements: The first element is a target (or several targets) to increase the recycling/recovery/reuse (RRR) rate. The second is a fee that has to be paid by the first owner of a new car registered in the country. The third part is a "recycling premium" paid to dismantlers, transporters and recyclers for the extra-cost they incur for meeting the RRR target rates. The system is managed by a private company.

In Sweden a "deposit-refund" system is used. It was introduced already in the 1970s. The producer or importer of a car has to pay a "recycling fee" to the Vehicle Disposal Fund. The fee, which is fixed by the Government, raises – of course – the price of the car. The Vehicle Disposal Fund pays a "scrapping premium" to the final car owner if the car is correctly deregistered for scrapping. Beside this premium, there might be a positive or negative value of the ELV. This is to be negotiated between the final owner and the dismantler. If the dismantler buys the car before deregistration, he is entitled to the scrapping premium.

The EU Directive 2000/53 introduced a third economic instrument: the "free take-back". Free take-back implies that the final car owner can deliver the car to a (registered) dismantler free of charge. The latter receives financial resources from the car-making industry. The amount of the resources going to all dismantlers has to be determined in light of the RRR targets set likewise by the EU directive.

Mazzanti and Zoboli (2005) consider each of the three economic instruments capable of dealing with the three mentioned ELV externalities, whereas the free take-back system – as they cautiously formulate – "could be, ex ante, a good candidate as the best economic instrument." The main reason is that car makers have to meet the RRR targets and want to do this at low costs. Thus, they have an incentive to innovate upstream in a way that makes recycling cheaper.

However, the authors regard it as uncertain how far free take-back or other economic instruments are capable of exerting optimal dynamic incentive effects for innovation in complex and inter-linked industrial subsystems and to induce optimal innovation paths. The necessity for innovation to meet the RRR targets of the directive is seen by them not only upstream but also downstream. Moreover, not only product and process innovations are required, also new markets must be created or arise. The authors see a possible solution in an integrated policy approach which should lead to a "shared inter-industry interest in innovation". A cornerstone of such a solution may be enforceable voluntary agreements within the car sector and between car sector and government.

Presently the EU directive is under implementation in most EU countries. It leaves some room for countries to keep main elements of their existing system for dealing with ELV. Table 1 informs about the status of implementation of the directive in EU-15 and Norway, while Table 2 describes the current financial conditions related to ELV.

R. O.

Reference

Mazzanti, M. and R. Zoboli (2005), "Economic Instruments and Induced Innovation: The Case of End-of-Life Vehicles European Policies", *FEEM Working Paper*, no. 80.

Table 1

ELV legislation & EU ELV Directive implementation

Country	Existing Legislation	Directive Implementation Status	Requirements deviating from Directive
Austria	Ordinances issued 11/2002 based on Waste Management Act (2002); Voluntary Agreement (1992)	Implemented	Economic operators not specified; main tasks placed exclusively on manufacturers/importers; extreme complex report duties
Belgium	Voluntary Agreement 1999; Renewal Voluntary Agreement in negotiation; regional legislation: Flanders 1999, Brussels 2002 and Wallonia 2003	Implemented	Free take-back by 1st January 2006 on the condition of a progressive management plan approved by different administrations before 1st July 2004
Denmark	Consolidation Acts 373 & 860 (1999); Executive Order 141 (2000)	Modification of Executive Order published 19th June 2002, covering the period until end of 2006	As defined in "car package" (treatment, requirements to companies, compensation)
Finland	No	In preparation. Adoption expected Q3, 2004	Free take-back as of 1st January 2005
France	Voluntary Agreement (1993)	The French ELV transcription decree has been sent to the council the 13th March for adoption. (Decree N° 10)	Research: Increase recoverability (2002 90%) and recyclates use; Parts marking; Information
Germany	Legislation (Altauto VO)	Implemented since 1st July 2002	Last owners deliver ELVs to certified dismantlers. Recovery quota. Cost-free take-back. Technical requirem. for dismantlers
Greece	General Waste Law – Presidential Decree regarding the implementation of ELV Directive (05/03/2004)	Presidential Decree regarding the implementation of ELV Directive was published on the 5th March 2004	Manufacturers only are liable to set up the system. Athens, Thessalonica, Patra, Volos & Heraklion: Cost-free take-back, up to end 2006. Rest of Greece: Cost-free take-back, up to end 2009. AMVH has been approved as a collective system by law published on the 17th of June 2004
Ireland	Voluntary Agree, in discussion commencing 09/01	Primary legislation finalized 3rd July 2003	None
Italy	National and regional laws	Law published on 7th August 2003	ELV owners deliver ELVs to authorised centres – Carmakers are free to create a free take-back network. If they do so they are allowed to guarantee free take-back only if their vehicles are delivered to one of these authorised centre.
Luxembourg	General waste law only	Law implemented 1st April 2003	None, however, producers are responsible for Monitoring
Netherlands	Special body created by 5 organisations: ARN (1993)	Implemented since 4th June 2002	Dutch Government has brought forward the EU 2015 targets to 2007 and 2006 to 2003
Norway	Government System (19789	Implemented since 16th July 2002	As per ELV Directive
Portugal	Voluntary Agreement (1999) and Waste Management Law (1998)	Implemented since 24th August 2003	Management integrated system is to be organised, transport and treatment costs must be borne by producers if negative market values appear
Spain	Legislation (RD 1383/2002)	Issued 3rd January 2003	Owners deliver ELVs to authorized centres. Producers shall take back ELVs. Negative market values from 1st January 2007 if deficit in treatment chain audited by third party. Recovery quotas.
Sweden	Car scrapping legislation (1975), producer responsibility ordinance (1998), Regulation on vehicle dismantling operation (April 2002), Government bill (May 2003) Ban of certain metals (July 2003)	Remaining issues: Cost-free take-back 2007, Article 9.2	Car industry responsible for targets & reports for all cars and for free take-back (cars registered Jan 1998)
UK	Statutory instrument 2003 N° 2635	In effect 3rd November 2003	None
Note: The information relates to July 2004.			

Source: European Automobile Manufacturer Association (ACEA), ACEA Country Report, July 2004 (www.acea.be).

Table 2

Current financial ELV take-back conditions

Country	Market Conditions	Cost-Free Take-Back	Fund System
Austria	x	Free take back when new or used car is purchased	n.a.
Belgium	x	From 01 January 2006	no
Denmark	Last owner is negotiating with the dismantler	Indirectly: Last owner pays for treatment (approx. DKK 800–1,000), receives fund compensation (DKK 1,750) when presenting CoD	Annual Payment of DKK 90 per car owner. This amount is added to the liability insurance premium + minimal producer contribution for new vehicles (0.5 €)
Finland	x	As of enacting of legislation	n.a.
France	x	Only at shredders	no
Germany	Market conditions for ELVs until 31st December 2006	From 1st January 2007	n.a.
Greece	x	Athens, Thessalonica, Patra, Volos & Heraklion: Cost-free take-back, up to end 2006, Rest of Greece: Cost-free take-back, up to end 2009. Common system of all manufacturers	no
Ireland	Under Discussion	Mid 2005	no
Italy	x	New vehicles from 2002, old car park as of 2007	No fund system is foreseen
Luxembourg	x	Producer liable	no
Netherlands	n.a.	Levy System, paid by importer, possibility to charge it to the first registered owner	n.a.
Norway	n.a.	Incentive to last owner. Treatment cost covered by Government	A car scrapping fee paid on new vehicle sales
Portugal	LOIs for whole car park are signed with 5 dismantlers	„Scrapping bonus“ for old cars when a new car is purchased. New vehicles from 2002, old car park as of 2007. Producer borne costs for transport and depollution (according to the dates above) if negative market values appear	Contribution to system provider (Valorcar) based on car share
Spain	x	Cost-free take-back: from 1st January 2005	n.a.
Sweden	n.a.	Cost free for vehicles on the market from January 1998 by producer responsibility	Premium paid to last owner, higher for cars registered before 1998. A car-scrapping fee paid to the state-owned fund on each new car registration, higher for personal imports. The fund system is under review, starting fall of 2003 and may become changed
UK	x	From 1st January 2007	n.a.
Note: The information relates to July 2004.			

Source: European Automobile Manufacturer Association (ACEA), ACEA Country Report, July 2004 (www.acea.be).

Childcare Services

Childcare governance

In Finland and Sweden local (municipal) authorities are responsible for providing childcare. The central government identifies goals, formulates guidelines and sets financing rules. In both countries, municipalities have to provide pre-school services and out-of-school-hours care as required by parental work commitments. In Sweden, this obligation concerns all children of one year and older, but unlike Finland, parents do not have recourse to the judicial system to exercise this right. In Finland, the childcare guarantee concerns all children who are not yet old enough to go to primary school (age seven in both countries). Throughout the year, Swedish municipalities have to provide a childcare place "without unreasonable delay", i.e. within three to four months upon application. Legislation on local service provision may set general rules on basic standards, but otherwise does not provide an excessive number of detailed provisions on the nature and scope of service provision. Hence, the systems allow for variation in service provision across municipalities.

In England, each local authority is now expected to develop, plan and coordinate childcare and pre-school services including nurseries, children's centres, day-care services, playschools and out-of-school-hours care. Working with their partners through children's trusts (partnerships including local community representatives, Jobcentre Plus, schools, health agencies, NGOs and commercial private childcare providers), local authorities are responsible for local childcare facilities to serve local needs.

Types of childcare service

The provision of childcare services in Finland differs across municipalities, but broadly speaking there are three types of childcare: centre-based day-care facilities, group family day-care centres, and home-based family day care (Table 1). Most centres provide full-time care, i.e. services that start at 7.00 a.m. until the late afternoon (5.00 or 6.00 p.m.). Finnish municipalities also organize supervised play activities and open day-care centres, providing social networks for parents (and their children) who would otherwise provide personal care at home on a full-time basis. In 2000, Finland launched its "pre-school programme" for six-year-olds and municipalities have to offer this pre-school education for free and comply with the new core preschool curriculum. In general this type of pre-school education is provided for five hours per day starting at 9.00 a.m. There are a few private day-care centres in Finland, covering 2 percent of children under age three and 4 percent of three to six-year-olds. Often such private centres provide specific services (e.g., an emphasis on music classes or foreign languages), and they are generally more expensive than municipal day care.

In Sweden, municipalities operate the pre-school facilities, which are largely centre-based day-care facilities. The obligation for municipalities to provide day care in line work requirements means that in general childcare centres are open from about 6.30 a.m. or 7.00 a.m. to 6.30 p.m. Since 2003, municipalities also have to provide free access to pre-school centres to all four to five-year-old children and pre-school classes for six-year-olds. To complete the scope of service provision, there are open pre-school services for parents who would otherwise provide full-time parental care, but who wish their children to interact with their peers for a few hours per day.

UK childcare policy has traditionally relied on private sector provision of day-care facilities, and as a result there is great variety in the type of childcare services that is available (Table 1). The multitude of services includes day care centres as, for example, day nurseries, children centres and family centres, and family day care services through childminders all registered with the Office for Standards in Education (Ofsted). "Sessional" day care, for example, in playgroups is typically available for children from two years of age, for about four or five hours per day. In addition, there is the free early education offer for all three- and four-year-olds of 55 free two and one-half hour sessions per term for six terms before reaching statutory school age – the first term following their fifth birthday. In the United Kingdom, the prevalence of providers varies with the type of service. Full day-care services are largely commercial (75 percent), while play groups are often organised by NGOs and churches. Out-of-school-hours care is likely to be organised by either NGOs or local authorities, while the free early education offer is predominantly organised by local authorities.

Participation in childcare

The use of childcare services by very young children is highest in Sweden (Table 2, Panel A). In 2002, about 370,000 children or 80 percent of all one to

five- year-olds in Sweden made use of a formal childcare service. Already 45 percent of children not yet two years of age use childcare, and childcare attendance grows rapidly when children get older: about 85 percent of the two-year-olds use formal childcare. In the other countries, the use of childcare among children not yet three years of age is much more limited. In Finland and Scotland about 45 percent of two-year-olds use a formal childcare service, while in these two countries, the use of childcare by children before their fist birthday is highest in Scotland at 12 percent. In both England and Scotland just over a quarter of all children under age three uses formal childcare. By contrast childcare use among older children generally exceeds 90 percent in the year directly prior to entering school (Table 2).

Table 2, Panel B shows that in Sweden almost 90 percent of children age one to five are in centre-based care while only 10 percent use a family-based setting (use of family day care is particularly low among those not yet two years of age). Formal care in family-based environment is also uncommon in England and Scotland. By contrast, formal family-day care is common in Finland.

W. O.

Reference

OECD, *Babies and Bosses, Reconciling Work and Family Life*, vol. 4, Canada, Finland, Sweden and the United Kingdom, Paris 2005.

Table 1

Diverse childcare services

Main characteristics of care facilities for children

	Age group	Type of providers	Overview of the service
FINLAND			
Children aged 0–6			
Day-care centres	4 months to 6 years old	Mainly run by municipalities	
Family day care		Operated by municipalities or private	
Three-family day care			Two or three families alternate to take care of children in a home-based environment. Generally on a full-time basis
Group family day care			Two or three childminders provide care in a facility generally provided by local authority. Generally full-time care
Playground activities		Operated by municipalities or private	
Open day-care centre		Operated by municipalities or private	Care provided up to ten continuous hours
Preschool classes	6 years old	Public	
Children aged 7–12		Mainly based in schools	Free "morning and afternoon" sessions provided in comprehensive schools for pupils in grade 1 and 2 (7 and 8 years old). Three hours per day arranged between 7.00 a.m. and 5.00 p.m.
SWEDEN			
Children aged 0–5			
Preschool centres		Centre-based facilities mainly run by municipalities	Full-day care provided from 7.00 a.m. to 6.30 p.m.
Family-based		Family-based care provided by childminders	
Open pre-school		Services providing collective activities for children who otherwise do not use pre-school facilities	Services provided for a couple of hours per day
School-aged children (6–12)			
Out-of-school care		Provided mainly in schools but also in childcare centres ("leisure time activities" centres)	

Table 1 (continued)

	Age group	Type of providers	Overview of the service
UNITED KINGDOM			
Children aged 0–5			
Full day-care centre	0–5 years old	Includes services provided outside the home. Most frequent are day nurseries (for-profit or not), Children's centres and family centres. Nursery schools can also provide care for children between three and five	Childcare provided for at least four hours per day
Sessional care	0–5 years old	Includes playgroups or crèches that offer short term childcare while parents are unable to look after children	Childcare is organised by session of less than four continuous hours per day. Children can attend two session per day but not more than five sessions per week
Nursery schools	3–4 years old	Provided by public, private sector and voluntary sector (in Scotland). Free of charge for parents when provided by the public sector and often at only a residual charge in the private sector	Educate preschool aged children. Generally open school hours (9.00 a.m. to 3.30 p.m.) during term time, sometimes morning only. Children usually attend for half a day
Childminders	Usually up to 8 years of age	Private service of a self-employed person who provides day care for more than two hours per day; usually based in the childminder's own home, with flexible hours	Childcare is provided according to parents needs
School-aged children (5–12)			
Out-of-school care	5–12 years old	Includes summer camps, holiday play schemes, breakfast clubs, after-school clubs. A range of activities are offered including sports, drama, arts, and crafts and music	

Source: OECD (2005).

Table 2

Use of childcare

Panel A. Participation of children in formal childcare is highest in Sweden Participation rate in education and care facilities, % of the child population concerned			
	Age 0– <3	Age 3 to age of compulsory school[a)]	Children under school age
Finland	25% 1% for <1–year-old children 28% for 1– <2–year-old children 44% for 2– <3–year-old children	68% 61% for 3– <4–year-old children 68% for 4– <5–year-old children 72% for 5– <6–year-old children 94% for 6– <7–year-old children are enrolled in free pre-school[b)]	50%
Sweden	65% 45% for children under 2 85% of children at age 2– <3	 89% for 3– <4–year-old children 91% for 4– and 5–year-old children	81%
England	26%[c)]	82% 97% for 3– to 4–year-old children[d)] 93% for 3– <4–year-old children 100% for 4– <5–year-old children	
Scotland	27% 12% for <1–year-old children 23% for 1– <2–year-old children 47% for 2– <3–year-old children	80% 94% for 3– <4–year-old children 100% for 4– <5–year-old children 40% for 5– <6–year-old children	

Panel B. Children of all age-groups use centre-based care Proportion of children in each type of care		
	Age 0– <3	Age 3 to age of compulsory school[a)]
Finland		
Family-based care	48%	30%
Centre-based care	43%	62%
Other	9%	7%
Sweden		
Family-based care	10%	10%
Centre-based care	90%	92%
England		
Family-based care	9%	10%
Centre-based care	26%	40%
Pre-school	6%	48%
Other (nannies, au pair)	10%	14%
Scotland		
Family-based care	18%	6%
Centre-based care	82%	16%
Preschool facilities		78%

[a)] Mandatory school age is 7 in Finland and Sweden, and 5 in England and Scotland. – [b)] Preschool is provided at age 6 in Finland and Sweden, 3 in the United Kingdom. – [c)] Estimates for year 2001 from Woodland, S., M. Miller and S. Tipping (2002), "Repeat Study of Parents' Demand for Childcare", Department for Education and Skills Research Report, No. RR348 London. – [d)] Office for National Statistics, Provision for children under 5 years of age in England, January 2004 (provisional). Available data for England is not directly comparable with information for the other countries, as data include non-regular care (e.g. baby-sitting services not included elsewhere), and because data have not been adjusted for multiple use of formal childcare services by one and the same child.

Source: OECD (2005).

New at DICE Database

"Natural Environment" and "Experts' Assessments"

In the third quarter of 2005 two major additions have been made to the DICE Database (free access: www.cesifo.de/DICE). Both are new topics. One is "Natural Environment", the other is "Experts' Assessments of Governance Characteristics". "Natural Environment" describes in a country-comparative way the environment-related regulations under which economic agents have to make their decisions. In addition to such regulations, important characteristics of the environment itself are also described.

The new topic "Experts' Assessments" characterises – as do other topics of DICE – the framework conditions of an economy. However, the new topic differs from others in that it does not describe facts about the framework (in words or figures) but provides assessments of it. All of the assessments are in quantitative form; most variables are received on an annual basis and cover a longer range of time. This should facilitate econometric research, focusing on the relation between framework conditions and economic reforms on the one hand and economic performance on the other. For researchers interested in such questions DICE should increasingly become a "one-stop centre", providing easy access to information from many sources.

Forthcoming Conferences

Global Economic Imbalances: Prospects and Remedies (1)
Munich, 11 – 12 November 2005

The issues to be treated on this third CESifo-Delphi Conference (first part) include the role of the Euro in a changing international environment, the viability of the US current account deficit, multilateral exchange rate regimes and policy coordination, financial and exchange rate crises, the role of currency and trading blocs.

Scientific organisers: Helge Berger, Thomas Moutos and Sarantis Kalyvitis

Guns and Butter: The Economic Causes and Consequences of Conflict
Munich, 9 – 10 December 2005

The conference will address topics such as the effects of conflict on economic activity, the potential role of globalisation in fostering a more peaceful or a more violent world, the relationship between development and conflict, how to measure the costs of conflict, and war finance.

Scientific organiser: Gregory D. Hess

Money, Macro and International Finance
Munich, 24 – 25 February 2006

The objective of this CESifo Area Conference is to bring together scholars of international macroeconomic questions in order to stimulate interaction and cooperation between them. All CESifo research network members are invited to submit their papers.

Scientific organiser: Paul de Grauwe

Applied Microeconomics
Munich, 10 – 11 March 2006

This CESifo Area Conference will provide a forum for scholars dealing with questions of applied microeconomics. The aim of the conference is to stimulate interaction and cooperation between the participants. All CESifo research network members are invited to submit their papers.

Scientific organiser: Christian Gollier

Global Economy
Munich, 7 – 8 April 2006

This is the third area meeting for CESifo's Global Economy group. Topical questions are how the gains from globalisation differ from those from trade, causes and effects of marginalisation, the role of culture and local identity, new forms of global institutions and arrangements.

Scientific organiser: John Whalley

Public Sector Economics
Munich, 21 – 23 April 2006

This CESifo Area Conference provides an overview of the current research undertaken by members of the Public Sector Economics area of the CESifo network. Its objective is to stimulate interaction and

cooperation between them. All CESifo research network members are invited to submit papers.

Scientific organiser: Frederick van der Ploog

Employment and Social Protection
Munich, 26 – 27 May 2006

This CESifo Area Conference gives CESifo members the opportunity to present and discuss their ongoing research. All CESifo research network members are invited to submit papers.

Scientific organiser: Kai Konrad

Global Economic Imbalances: Prospects and Remedies (2)
Munich, 2 – 3 June 2006

The issues to be treated at the third CESifo-Delphi Conference (second part) include the role of the Euro in a changing international environment, the viability of the US current account deficit, multilateral exchange rate regimes and policy coordination, financial and exchange rate crises, the role of currency and trading blocs.

Scientific organisers: Helge Berger, Thomas Moutos

New Books

Doing Business in 2006: Creating Jobs
World Bank and International Finance Corporation, forthcoming in 2006

Doing Business in 2005: Removing Obstacles to Growth
World Bank and International Finance Corporation, 2005

International Institutions in the New Global Economy
Lisa L. Martin (ed.)
Edgar Elgar, 2005

Ageing and Pension Reform Around the World
Giuliano Bonoli (ed.)
Edgar Elgar, 2005

Banking Regulation and World Trade Law: GATS, EU and Prudential Institution-building
Lazaros E. Panourgias
Hart Publishing, 2005

Innovation and Institutions
Steven Casper (ed.)
Edgar Elgar, 2005

Competing with the Government: Anti-competitive Behaviour and Public Enterprises
R. Richard Geddes, David E. M. Sappington, J. Gregory Sidak (eds.)
Hoover Institution, 2005

Institutions, Sustainability, and Natural Resources:
Shashi Kant and R. Albert Berry (eds.)
Springer, 2005

Growth, Trade and Economic Institutions
Tapio Palokangas (ed.)
Springer, 2005

New Ideas in Contracting and Organizational Economics Research
Harvey S. James, Jr. (ed.)
Nova Science Publishers, 2005

New Institutional Economics Abstracts (network journal)
Michael Sykuta, (ed.)
Social Science Research Network, 2005

A New Country-comparative and Institutional Database

Deposit Insurance Around the World

The database, existing since 1999, has been recently updated and enlarged. It is available as an Excel workbook online at the Finance Research website of the Development Economics Research Group, World Bank. For a summary, see World Bank Policy Research Working Paper 3628, June 2005.

CESifo, a joint initiative of the University of Munich's Center for Economic Studies and the Ifo Institute for Economic Research

Economic Studies

Formerly *ifo Studien*

Vol. 51, No. 2–3/2005

www.cesifo-economic-studies.de

DICE

Database for Institutional Comparisons in Europe

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
or
ochel@ifo.de