

Public sector reform and evaluation.

Approaches and practice in international perspective

Hellmut Wollmann
Humboldt-Universität
Institut of Social Science
Unter den Linden 6
D 10099 Berlin
hellmut.wollmann@rz.hu-berlin.de

in.: Hellmut Wollmann (ed.) 2003,
Evaluation in Public Sector Reform. Concepts and Practice in International Perspective,
Cheltenham/Northampton: Edward Elgar, pp. 231-258

In this concluding chapter an attempt will be made, in a comparative perspective, to identify some salient patterns and trends in the ‘twinned’ development of public-sector reforms and evaluation therein and thereof. For this purpose, cursory accounts on the development in some of the countries under inspection in this volume will be given before some summarising remarks are made.

I. Some country profiles

Analytical framework

Empirically the following cursory accounts will draw on the preceding chapters of this book as well as on other related sources (particularly the country reports in Pollitt and Bouckaert 2000, pp. 129 ff.).

Conceptually they will be guided by an analytical scheme in which, drawing on the ‘neo-institutionalist’ debate¹ and on the institutional transformation literature², the following factors are hypothesised to have particular (explanatory) relevance:³

- 1) Starting conditions. Probably more than in other areas of institution building and institutional choice, the institutional choice in public sector reforms is influenced by the very starting conditions, that is, the current format and profile of the public sector (in its various dimensions) from which the modernization process is bound to take off (see also Pollitt and Bouckaert 2000, pp. 62 ff.). The leverage of the starting conditions on the modernization course is readily apparent, since in the perception of the relevant actors they define the country’s modernization deficit or reform need as measured against the imperatives of the dominant modernization discourse—or,

conversely, the starting conditions may, in the interpretation of the actors, signal no need or a minimal need for modernization and may suggest even a modernization 'lead' over their respective country on some crucial scores (Wollmann 1996, pp. 15ff.);

- 2) Socio-economic and 'external' factors (budgetary crises, 'external' influences [for instance, of the European Union or World Bank]);
- 3) Institutional and cultural traditions (such as legacies or path-dependencies⁴). Public-sector modernization decisions are likely to be strongly influenced also by the institutional and cultural traditions and givens of the country's institutional world. The impact of these factors is probably the stronger, the more firmly (to the point of eliciting a path-dependency) such institutional and cultural givens are empirically and normatively rooted in the country's history and tradition. The different strands of the Rechtsstaat (in Continental European countries) and of the 'civil culture'/'public interest' traditions (in the Anglo-Saxon world) are exemplary (see Pollitt and Bouckaert 2000, pp. 52 ff, Wollmann 2000b, pp. 4 ff.);
- 4) Institutional (polity) setting (unitary/centralised versus federal/decentralised, majoritarian versus consociational/consensual, multi-actor versus single-actor⁵ countries);
- 5) Actor constellations, intentions, interests, 'will and skill'. As the decisions on public sector modernization are prepared and taken in discourse and decision-making arenas made up of the relevant political, administrative and socio-economic actors (as well as of academics and consultants), the actors' constellations, intentions, interests and 'will and skill' are likely to have significant influence on the institutional choice finally embarked upon;

- 6) (national as well as international) discourses and discourse coalitions⁶, including the increasing interpenetration of national and international discourse arenas and networks with ensuing exchange, learning and adaptation processes.⁷

For an (explanatory) account of the specific profile of evaluation in and on public sector reforms in the different countries, particularly the distinction should be recalled between internal evaluation in terms of (self-) evaluative tools and procedures within the operating administrative units, on the one hand, and external evaluation particularly of the social science-guided evaluation research type, on the other hand. In explaining the former, the emergence and state of public sector reforms appears to be a strong predictor. With regard to the latter, the existence of institutional actors (such as parliament, court of audit) that advocate external evaluation and the degree of ‘maturity’ of the country’s evaluation culture⁸ probably have a strong impact.

As a caveat it must be mentioned, however, that the following sketches—including the tentative ‘causal interpretation’—are liable, due to brevity and selectivity, to be imperfect and patchy.

Great Britain

Since the early 1980s Great Britain has been a frontrunner and pacesetter in what, under the now familiar label of New Public Management (see Hood 1990), has become the main current in the international debate on and practice of public-sector modernisation guided by neo-liberal and managerialist beliefs. The decisive shift is explained by a constellation of factors. When elected into office in 1979, the new Conservative government under Margaret Thatcher was determined to initiate large-scale neo-liberal and managerial reforms. The country's 'starting conditions' (among other things, over-centralised government and the quasi-monopoly of public personnel in the delivery of services) called for major changes. Great Britain's unitary ('Westminster') system provided the institutional levers to effect and enforce such changes from central to local levels nationwide.

The Financial Management Initiative of 1982, the Next Steps Initiative of 1988 (which led to creation of some 140 executive agencies), and the 1991 Compulsory Competitive Tendering (CCT) and Citizen Charters were crucial steps. The Financial Management Initiative of 1982 embraced the whole of central government with its philosophy of decentralised management, decentralised budgets, more targets and more professionalism (see Pollitt and Bouckaert 2000: 273, Gray and Jenkins 1992: 64ff). The National Audit Office and Audit Commission were created with a mandate that stressed the '3 Es' (economy, efficiency and effectiveness).

From the outset the introduction and employment of performance measurement systems was advocated as a central management and evaluation tool (for details and references see Pollitt and Bouckaert 2000, pp. 273ff.). The setting of performance goals

and indicators (by contracts or legal prescriptions) and their monitoring, measuring and reporting (via 'internal' evaluation) has been applied to the agencies as well as to the local authorities. The annual publication of national 'league tables' for schools and hospitals, which receives considerable media attention, was just another expression of the 'performance indicator culture' that has emerged in the UK (see also Jann and Reichard, chapter 3 of this volume).

The performance indicator (PI) movement has still gained more momentum another push under the New Labour Government which, in repealing CCT and replacing it with its Best Value (BV) regime, statutorily introduced in the Local Government Act 1999, turned to use an expanding set of performance indicators to particularly put the local authorities and their service-related activities under a rigorous top-down surveillance (see Wilson and Game 2002: 337 ff.). A crucial role came to be played by the Audit Commission which was established in 1982 in order to monitor the financial and managerial competence of local government (and a number of specified quangos, including the health service). In the Local Government Act 2000 the responsibility of the Audit Commission was extended to ascertain whether the local authorities are providing "value for money". The 2001 White Paper (DTLR 2001) took the BV regime a conspicuous step further by stipulating an elaborate evaluative scheme, revolving around the Audit Commission, in which the local authorities were to be classified as "high performers", "strivers", "coasters" or "poor performers" whereby, as a remarkably centralist and interventionist feature (see Wilson and Game 2002: 338), the "good performers" are to be rewarded by central government by getting more money and "more freedoms", while the "poor performers" can be sanctioned to the point, in the last resort, to place the failing council into the hands of Government-

appointed administrators. The first round of such stringent top-down performance measurement and assessment was recently carried out by the Audit Commission on the local authorities of the counties. The report (assessing all counties under the aforementioned four categories) was published in early December 2002, arousing considerable public attention and controversy (particularly from those county authorities that came out with a poor rating).

By contrast, full-fledged external evaluations have seldom been undertaken during the Conservative era.⁹ Under the New Labour government external evaluation has been on the rise. In line with the new emphasis on ‘evidence-based decision-making’ (Sanderson 2000, p. 433), a Centre for Management and Policy Studies has been created in the Cabinet Office; the Centre is intended to provide a ‘window in the heart of government’ for research and evaluation evidence (Cabinet Office 2000). Mention should be made, for instance, of the evaluation, conducted by a team of independent consultants (see DETR 2001), on an extensive pilot programme that preceded the introduction of the full BV regime (see Wilson and Game 2002: 338).

New Zealand

New Zealand, also a frontrunner in the new modernisation wave, has generated the ‘purest’ theoretical formulation of the New Public Management doctrine (see Halligan, chapter 5 in this volume) and also has gone further by simultaneously realising the privatisation, agencification and marketisation elements of NPM (see Pollitt and Bouckaert, chapter 2 in this volume). A similar constellation as that in the UK (overcentralist government, quasi-monopoly of public personnel in the delivery of

services, and a majoritarian Westminster government setting) has probably had some bearing on the sudden shift in 1984, except for the fact that it was effected by a Labour government, which in the face of a pressing economic crisis turned (almost overnight) from its traditional social-democratic to a neo-liberal persuasion. This policy change was conceived and engineered by a small (and almost secret) circle of Labour Party leaders, business leaders, neo-liberal university economists and like-minded Treasury officials (see Halligan 2001, p. 85). The pronounced management orientation of the reform was evidenced by the emphasis placed on guiding the newly created agencies by performance and reporting requirements—that is, by variants of internal evaluation. In fact, this performance orientation has even been characterised by an ‘obsessive concern with outputs and accountability’ (Halligan, chapter 5 in this volume).

Notwithstanding this heavy stress on performance management, New Zealand has also initiated major evaluative reviews, first by mandating the Steering Group for Review of State Sector Reforms in 1991 and then by commissioning Allan Schick, an internationally renowned expert, to conduct a comprehensive review in 1994¹⁰ (see Halligan, chapter 5 in this volume).

Australia

Australia’s 1983 Financial Management Improvement Program reforms included strong elements of management and programme budgeting as well as mandatory evaluation to ‘close the loop’ for a new system of results-oriented management (see Halligan, chapter 5 in this volume and Pollitt and Bouckaert 2000, pp. 202 ff). Australia’s modernisation approach did without the theoretical NPM-related stringency that was characteristic of

New Zealand's path. The constellation of factors that triggered Australia's modernisation move was quite different from that of New Zealand. The main motive of the new Labour government's push for large-scale public-sector reforms was the political concern 'to re-establish ministerial control and greater responsiveness to government policies and priorities' (see Halligan, chapter 5 in this volume and Pollitt and Bouckaert 2000, p. 202). Vis-à-vis the federal structure of the country (as a crucial starting condition), agencification—being one of the pivotal concepts of NPM—was understandably of little relevance.

Reflecting the specific orientation of Australia's reform drive, due to the absence of agencies, the introduction of indicator-based performance management and measurement systems was given less importance, whereas the employment of evaluation has been writ large. First, this applies to the employment of policy and program evaluation at large. In fact, the frequency and volume of policy evaluations has lead the (Australian) State Services Committee to gibe that the Australian approach can be 'characterised as evaluating everything that moves' and as a 'picture of evaluation overkill' (State Services Committee 1999, quoted from Halligan, chapter 5 in this volume). Both the country's court of audit and its parliament have proven to be advocates of wide-spread evaluation.¹¹

Regarding the evaluating of public-sector reforms, Australia has probably gone further than any other country. In 1991, after a decade of intensive reforms, the first (and probably the most extensive yet) evaluation was undertaken by the Task Force on Management Improvement, a quasi-independent group of public servants (see Halligan, chapter 5 in this volume). In addition, for instance, a prominent business leader was commissioned to do further (more sectoral) reviews. Small wonder that Australia has

been ranked, in an internationally comparative assessment of the 'evaluative culture' at large in 21 countries, among the 'top five' (see Furubo and Sandahl 2002a, p. 11).¹²

USA

In the USA, a significant move towards reforming the federal government was undertaken under the (Democrat) Jimmy Carter by the Civil Service Reform Act of 1978, which introduced performance appraisal and merit pay. In 1992 the Bill Clinton administration launched a major reform project. The centrepiece of the programme, entrusted to vice-president Al Gore, was the National Performance Review (NPR), which produced a report, subtitled 'Creating a Government that Works Better and Costs Less'. Borrowing directly from the title of Osborne and Gaebler's bestseller, 'reinventing government' (known as REGO) became the trademark and battle-cry of reform (see Pollitt and Bouckaert 2000, p. 282; Rockman 2001, p. 8). For one thing, the 'better and less government' shibboleth fell well in line with the country's long-standing managerialist tradition (which reaches back to the progressive movement of the late 19th century) as well as with American anti-bureaucratic and anti-big government beliefs.¹³ Moreover, the politically shrewd Clinton/Gore reform policy, in calling for 'less government and less government spending', embraced a traditionally Republican position and was meant to bring the Clinton administration political gains (see Rockman 2001, p. 8ff.). Furthermore, the new administration was faced with a huge federal deficit inherited from the Reagan and Bush administrations.

Conceptually the NPR aimed at turning the agencies of federal government into performance-based organisations (PBOs) or more flexible decentralised management

structures that would focus on results (see Rist and Paliokas 2002, p. 230). The US Congress ratified elements of the NPR by passing the Government Performance and Results Act (GPRA) of 1993, which required executive agencies to periodically report on their achievements with regard to their agency and programmatic goals. For the first time, Congress created a legislative structure whereby data on government performance would be fed systematically into the budget process, thus linking performance results with funding decisions (see Rist and Pakiolas 2002, p. 230).

The reform legislation significantly revolved around the idea of institutionalising and stepping up agency-based (internal) evaluation procedures. At the same time, the (internal) evaluation function has strong advocates and watchdogs both in the General Accounting Office (GAO), which reports to the US Congress, and in the Office of Management and Budget (OMB), which reports to the president. Since the Civil Service Reform Act (CSRA) of 1978, which was enacted under the Carter administration, within each agency the Office of Inspector General (OIG) has had the responsibility of auditing operations to advance efficiency and effectiveness, and it may execute special audits and evaluation studies (see Christensen, Lægreid and Wise, chapter 4 in this volume). Furthermore, the Government Performance and Results Act of 1993 requires performance auditing in federal agencies. Since 1998 outputs have had to be quantified and measured by indicators in the entire federal administration. Moreover, in 2001, under the Bush administration, OMB put the Executive Branch Management Scorecard System into effect. This ‘can be seen as a form of oversight and ranks agency performance on five areas of management against stated criteria of success’ (Christensen, Lægreid and Wise, chapter 4 in this volume). In sum, the federal

government (internally) produces a large volume of evidence about program performance.

Regarding external evaluation, it must be noted that this rise of management-oriented and result-based internal evaluative procedures has been accompanied, since the 1980s, by a decline in external evaluation (see Rist and Pakiolas 2002, pp. 230 ff.), in which the US federal government has been a ‘world leader’ since the 1960s.

However, it should be borne in mind that, in line with a legislative tradition which goes back to the upsurge of evaluation in the mid-1960s, evaluation requirements are still included in legislation on administrative reforms such as in the Civil Service Reform Act of 1978. So, while making allowances for reservations (such as by Rist and Pakiolas 2002), one may still speak, particularly in comparative terms, of a ‘flourishing culture of evaluation’ (Pollitt and Bouckaert 2000, p. 284).¹⁴

Finally, it should be added that this account on the USA addresses only the federal level. In order to get a more complete and more adequate picture, developments at the state and local government levels would have to be included. ‘In fact, state and local government can be seen as leaders and initiators of the movement to reinvent government’ (Christensen, Lægreid and Wise, chapter 4 in this volume, referring to Brudney et al. 1999).

Sweden

Since the early 1980s, when it faced a mounting budgetary crisis and, consequently, a challenge to its traditional welfare state model, Sweden has embarked upon a ‘double track’ reform trajectory. On the one hand, the Swedish government turned in 1985 to

further decentralisation of public functions to the counties and municipalities in what, based on the country's traditional multi-function local government model, already has represented an unusually decentralised government system (see Premfors 1991, 1998). On the other hand, a series of strong financial management reforms were implemented between 1988 to 1993, including result-oriented budgeting and accrual accounting. Results-oriented management was officially adopted for all state organisations from 1988 (see Pollitt and Bouckaert 2000, pp. 264ff.). The reason for this 'mixed' strategy, combining 'traditional' reforms with NPM-inspired components, may be seen in the particularities of Sweden's 'starting conditions': on the one hand, the existence of the country's politically and functionally strong local government levels invited further decentralisation as means to further pare down the (already comparatively spare) central agencies. On the other hand, some 300 agencies, which operate as single-purpose bodies with a high degree of autonomy from central government guidance, lend themselves to result-based operation and, hence, performance management (see Furubo and Sandahl 2002b, p. 119).

In its evaluation profile Sweden is characterised, first, by widespread adoption and practice of result-oriented steering, monitoring and reporting tools in the agencies as well as in the local authorities. In addition to purely economic information, the annual reports of the agencies contain primarily information about what has actually been produced, and what this output has cost, and not normally evaluation. Many agencies are often given special assignments by the government to evaluate certain matters and to report back the results in their annual report (see Furubo and Sandahl 2002b, p. 120).

Second, a number of strong public authorities play a significant role in the (external) evaluation of government activities. The National Audit Board (RRV) conducts about 20 audits of agency and ministry-level performance per year (see Christensen, Lægreid and Wise., chapter 4 in this volume). The Performance Audit Department within RRV investigates and promotes efficiency and effectiveness in government activities. Another important player is the Swedish Financial Management Authority (ekonomistyrningsverket [ESV])¹⁵ which, among other things, makes comparisons among national organisations, and between Sweden and other countries, and benchmarks performance.

Third, the typical ‘commission system’ bears mentioning (see Vedung 1992). Including representatives of political parties and stakeholders, such as trade unions, businesses, and academics, these commissions are appointed for the preparation of practically all legislative projects and drafts. They are seen as a conduit of evaluative knowledge (see Furubo and Sandahl 2002b, pp. 116 f.). The parliament can establish commissions to evaluate specific aspects of public management reform.

Finally, there is evidence of a growing interest in the academic research community in conducting social science-based evaluations of public management reforms (see Vedung 1997).

In sum, the variants of internal and external evaluation add up to a extensive and dense evaluation network and potential. These are embedded in an evaluation culture and tradition which dates back to the 1960s (when Sweden was among the few European frontrunners)¹⁶ and earlier.¹⁷

Netherlands

In the early 1980s, under Prime Minister Lubbers, the Netherlands government initiated an administrative reform package, in which the decentralisation of executive functions to lower levels of government was writ large. In the 1980s new autonomous public bodies (ZBO) were created. Some were long-established (such as the state universities), but almost half of them were established after 1980 (see Pollitt and Bouckaert 2000, p. 247). In the 1990s the decentralisation drive was accelerated by the setting up of departmental agencies as the Dutch variant of agencification. Between 1991 and 1998 more than 20 agencies were established (see Pollitt and Bouckaert 2000, p. 247) with the intention to increase the efficiency within the central government by means of result-oriented management. The (time-honoured) inspectorates within the ministries have become increasingly involved in evaluation, quality assessments and impact studies, in addition to their traditional legality control. Moreover, in each ministry a central financial and economic affairs department has been set up in order to stimulate and coordinate evaluation activities within the ministry (see Leeuw, chapter 6 in this volume). As a management concept and tool the ‘new steering model’ was formulated and put in practice (see Kickert and In’t Veld 1995).

The evaluation system got a decisive push in 1990 when the court of audit published a government-wide study of evaluation practices in the executive branch of the central government. This triggered activities of government and parliament in which systematic and periodic evaluation research was envisaged as a crucial instrument for reviewing the effectiveness and efficiency of policy programs. In practical terms it was concluded that the evaluation function needed to be installed in the existing

departmental frameworks and structures in order to optimise the linkage to existing policy and budgeting processes (Bemelsman-Videc 2002, p. 98). There has been a trend toward developing and refining performance indicators for a widening range of public services. Autonomous public bodies (ZBOs) as well as agencies are obligated to present data on costs and benefits annually (see Leeuw, chapter 6 in this volume).

The National Court of Audit (Algemene Rekenkamer) has an important advocate and actor in external evaluation. Over the years it has repeatedly carried out government-wide investigations as well as performance audits and occasional meta-analyses of evaluations ('meta-evaluation'; see Leeuw, chapter 6 in this volume).

Germany

Germany has been a latecomer to the international NPM-inspired modernisation movement. The reason for this time lag can be seen in the country's specific 'starting conditions'. First, the traditional federal-decentralised constitutional fabric, characterised by a politically and functionally strong local government level, as well as the traditional principle of subsidiarity (leaving the delivery of social services largely to non-public welfare organisations), has made key NPM concepts (such as agencification and outsourcing) appear less pertinent and less necessary. Furthermore, in the Rechtsstaat tradition the salience of legal regulation and judicial review of public administration have been institutional, cultural and normative impediments to an easy access and adoption of private-sector-derived managerialism. Last but not least, Germany's multi-level administrative system has had a good record of administrative reforms over time, including the reform push of the 1960s and 1970s (see Wollmann

2000a, pp. 920 ff. and in chapter 7 of this volume). The dramatic shift and overture to NPM-inspired concepts in the early 1990s was triggered largely by the mounting budgetary problems arising from the financial costs of German unification and from the need to meet the Maastricht parameters.

Since the early 1990s local governments have taken the conceptual and practical lead under the guidance of the so called ‘New Steering Model’ which, drawing on NPM managerialism and, in part, on pertinent Dutch experience, was formulated and propagated by a municipality-financed consulting institute (KGSt). A growing number of municipalities and counties have introduced budgeting, cost-to-achievement accounting, and in some cases accrual accounting and controlling—with a focus on (internal) performance management tools. It is only recently that outsourcing has gained momentum.

Somewhat later the Länder have entered the reform trail, exhibiting the typical variances among them.. In some Länder traditional schemes of decentralisation and of reorganizing the regional (meso-) level have been adopted. In most Länder variants of internal performance management have been addressed.

Finally, in late 1999 the federal government turned to administrative reforms, revolving also largely around performance management concepts.

Concerning the speed and direction of public-sector reforms, two features of Germany’s politico-administrative system should be considered. First, at each of the three government levels the actors operate quite independently on administrative reform issues—vertically as well as horizontally—which fosters the traditional ‘incrementalist’ reform style. Second, administrative reforms have been an ‘executive’ issue, while the

parliaments have shown little interest—perhaps regarding and respecting administrative reforms as an executive prerogative.

Following and mirroring the adoption of managerialist concepts and components, variants of management-oriented and output(‘product’)-based internal monitoring and (information feedback-related) controlling procedures have been extensively introduced. Furthermore, intermunicipal benchmarking, through inter-municipal cooperation, has been put in place.

External evaluation so far has hardly been undertaken at any level. This may be explained by the short time that has passed since the introduction of the NPM-inspired reform measures. But it also reflects that neither the federal and Länder courts of audit nor their parliaments have so far shown interest in initiating and conducting evaluation of the public-sector reforms underway. The scant attention that the evaluation of public-sector reforms has so far elicited in Germany contrasts with the observation that the evaluation of ‘substantive’ policies has played a significant role in Germany since the ‘first wave’ of the 1960s (see Wagner and Wollmann 1986; Derlien 1990). Germany has continued to have policy evaluation at a fairly high and stable level (see Wollmann 1989), especially since the early 1990s, because of EU structuring funding in East Germany and of ensuing evaluation (see Derlien 2002, pp. 84 ff.).¹⁸

Mention should finally be made, however, of the (applied) social science and public policy research that academic (mostly university-based) research groups and institutions have conducted on public-sector reforms measures—with a focus on the local government level and on the institution-building dimension of reforms (in the implementation research stance) rather than on outputs (see Wollmann, chapter 7 in this volume).

Italy

Italy is another example of a Continental European latecomer to NPM-inspired public-sector reforms. Throughout the 1990s, in several legislative waves, a host of reform laws and decrees were passed in an attempt to reform Italy's political and administrative arenas (including privatisation, organisational and personnel reforms, and introduction of management tools) in an intentionally radical manner (see Lippi, chapter 8 in this volume). Well into the 1990s public-sector reforms had been largely impeded by the Italian legalist tradition, with its body of administrative laws and its priority of legal review over economic efficiency (see Stame 2002). Moreover, reforms were blocked by party competition. In the early 1990s the politico-economic context changed dramatically. The corruption scandals which broke after 1992 led to crisis for the country's political elites. The mounting budgetary problems were aggravated by the perceived need to meet the Maastricht parameters as a precondition of becoming a founding member of the European monetary union. Thus, there was mounting pressure for stringent public-sector reforms. Since the early 1990s, NPM-guided reforms have been on Italy's political agenda.

The first major modernisation push came in 1993 with the newly formed (post-scandal) government of 'technocrats' under Prime Minister Ciampi and under Minister of Public Function Sabino Cassese, a reformist law professor. A 1993 decree, with explicit NPM references,¹⁹ introduced performance management (controllo di gestione) and prescribed the formation of evaluation units (nuclei di valutazione) in each public agency. In 1997 and 1998, a set of legislative acts, the 'Bassanini laws', were passed by

the centre-left Ulivo (Olive tree) coalition (under Franco Bassanini, again a reformist law professor, heading the Ministry of Public Function). These laws pushed for the streamlining of the administration and for the realisation of administrative federalism (see Stame 2002). So far the modernisation of political and administrative structures appears to have advanced furthest at the local government level (see Lippi, chapter 8 in this volume).

Following the legislative introduction in 1993, the prescribed performance control (controllo di gestione) has been realised and put in place by various public agencies,²⁰ with performance indicators varying widely for each sector. Many local authorities have changed their operations by adopting management tools, such as result-oriented management control (see Lippi, chapter 8 in this volume). Although the court of audit (Corte di Conti) was instructed, by legislation, to take up performance auditing, it has largely retained its traditional judiciary orientation and profile.

Policy evaluation at large saw a real upsurge in Italy during the 1990s, particularly resulting from the spree of external evaluation conducted on the European Structural Funds (Italy is a prime beneficiary of European funding; see Stame 2002). In the field of public-sector reforms, however, external evaluation so far has seldom been undertaken.²¹ ‘Neither the scientific community nor the institutions have sought to evaluate public-sector reform policy’ (Lippi, chapter 8 in this volume).

Japan

While Japan is another latecomer to NPM-guided public-sector reforms the country saw some noticeable changes during the major public-sector reforms of the 1980s (see

Chapter 9 in this volume). In reaction to the oil price shocks of the 1970s, in 1981 the (Second) Provisional Administrative Reform Commission (SPARC)²² was set up; SPARC put forth a number of reform recommendations, particularly the controversial privatisation of the Japan National Railways. Subsequently a series of Administrative Reform Commissions (ARCs) was established in order to monitor the reform process. In early 1993 the reform movement picked up momentum. In 1995 the Decentralisation Promotion Act emphasised decentralisation by strengthening local governments.

Following the national elections of 1996 and the formation of a new government under Ryutaro Hashimoto, public-sector reform policy shifted conspicuously into high gear. A reform commission was set up which, unusually enough, was chaired by the prime minister himself. The commission's recommendations related to a broad spectrum of reforms, including strengthening the cabinet (*vis-à-vis* the previous 'autonomy' of the ministerial bureaucracy), reorganising the central government, reforming the civil service and policy evaluation system as well as pronouncedly NPM-inspired components, such as agencification and performance management—with explicit borrowing and 'importing' from US and UK experience (see Muramatsu and Matsunami, chapter 9 in this volume).²³ Various factors led to the dramatic shift in public-sector reforms under the Hashimoto commission. First of all, it was a reaction to the mounting budgetary crisis which followed the burst of the bubble economy (see Chapter 9 in this volume). Second, mounting criticism of the inflexibility and reform-resistant inertia of the government's bureaucracy (almost forming a 'government within the government') finally came to a point.

As a major consequence, some 60 Independent Administrative Institutions (IAIs) were created to operate under a 'sunset'—that is, a pre-set 'termination'—

formula and to monitor and report on their performance on the basis of performance criteria.

The Basic Law on the Administrative Reform of the Central Government of 1998²⁴ was intended to reorganise central ministries and agencies—particularly by institutionalising and employing evaluation—by the year 2001. Conceptually and especially instrumentally, two stages of evaluation are stipulated. First, inside each ministry and agency the establishment of an evaluation unit was prescribed; furthermore, it was mandated that every government agency should self-evaluate the policies and programs under its jurisdiction in terms of effectiveness, efficiency, and cost-benefit analyses both prior to and following the enactment of the programmes. Second, within the Ministry of General Affairs—which was to play a crucial role in the development of a evaluation system—a ‘Committee on Policy Evaluation and Evaluation of Independent Administrative Institutes’ was designed to fulfil a government-wide, interorganisational coordinating and ‘meta-evaluating’ function (see Yamaya 2002, p. 344).

While the already existing Management and Coordination Agency continues to monitor the efficiency and economy of government, the Board of Audit (that is, the court of audit) is still largely concerned with checking the budgetary correctness of government activities.

Although Japan has recently moved conspicuously towards institutionalising internal evaluation (implementation still is in an incipient stage; see Yamaya 2002, p. 344), external evaluation is still lacking. ‘Contractual research’ still has a mostly prospective (ex ante) consultancy character—the lion’s share is carried out by the commercial consulting sector and research firms (so-called ‘think tanks’; see Wollmann

2002a, p. 11577 with references). This also reflects the reserve and distance Japan's university-based social science has traditionally exhibited with regard to applied, or evaluative, research (see Wollmann 1983; Watanuki 1991).

In sum, notwithstanding the rapid moves which Japan has recently made towards evaluation in and of public-sector reforms, she has much room for development.²⁵

Brazil

A caveat should be made when considering Brazil as a Latin American case in point. On the one hand, it shows the most interesting (and politically and geographically, the most relevant) case in the region. On the other hand, being more advanced than the other Latin American countries (with the exception of Chile) in terms of the matters under discussion here, Brazil is not typical or representative (for more comprehensive overviews, see Monteiro, Chapter 10 in this volume and Bresser-Pereira 2001).

Since colonial times Latin Americans have been subject to a kind of state-society relationship heavily based on bureaucratisation, centralism, formalism, fiscalism and authoritarianism (see Monteiro, Chapter 10 in this volume)

Leaving aside the historical peculiarities of administrative development in Latin America which lie in the colonial past, reference should be made to the military regimes of the 1960s that brought forth 'bureaucratic-military states' whose policies were characterised by an internationalisation of the economy and by centralisation, 'statisation' ('étatatisation') and authoritarian control of the public sector. Since the 1980s Brazil's development has, like that of other Latin American countries, been marked by two features: to wit, the return of the country to democracy and a deepening budgetary

crisis (triggered by the peso crisis in Mexico). Reflecting the reintroduction of democracy, Brazil's 1988 constitution aimed at a far-reaching decentralisation of the country by devolving political and administrative responsibilities to the regional and local levels (see Bresser-Pereira 2001, p. 152). At the same time, in reaction to the 'first generation' reform concepts and demands of the World Bank, the country undertook the deregulation and privatisation of public enterprises (see Monteiro, Chapter 10 in this volume). NPM-specific modernisation concepts were not considered at that point, apparently also due to the influence of the World Bank which, at that time, embraced a 'sequencing' strategy according to which developing countries should first complete civil service and (Weberian, as it were) bureaucratic reforms before engaging in NPM-inspired management reforms (see Bresser-Pereira 2001, p. 147 with references).

Since 1995 Brazil has seen a second wave of public-sector reforms in which NPM components have been adopted. Key concepts were laid down in the Plano Diretor da Reforma do Estado (Guidelines to State Reform) under the Cardoso presidency. The main elements were a broad spectrum of goals, including institutional adjustments of the public sector, checks and balances among state powers, strengthening of citizen participation, decentralisation and introduction of NPM components: control by contracted outcomes, managed competition and social control (Bresser Pereira 2001, p. 158). These concepts were turned into constitutional language by the constitutional amendment of 1998.²⁶

Under the programmatic heading 'Entrepreneurial Management' evaluation has made its entry to Brazil's administrative system (Bresser-Pereira 2001, p. 161). The process of evaluation has come to be directed by the technical staff of the Ministry of Planning, Budget and Management—the key player in the evolving evaluation system.

Divided in three steps, evaluation begins with self-evaluation (conducted by the program managers themselves) of expected or ‘contracted’ results. Then it goes on to a second phase, in which programme performance is to be set against the President’s Strategic Directives and to Sector Strategic Directives. The final step is an overall analysis by the Ministry on the programme’s performance. In this institutional and instrumental design, evaluation has been observed to be a tool with which the central government imposes and enforces institutional reforms. ‘The generation of goals, objectives, indexes and indicators typically follows a top-down logic according to which managers have to adopt (and adapt themselves to) externally-defined standards’ (Monteiro, Chapter 10 in this volume). Hence these evaluation tools have been used largely in internal (in-house) evaluation; external evaluation seldom has been commissioned or conducted.

II. Patterns and trends in sector reforms and evaluation

Public sector reforms: convergence or divergence?

In taking up the much-debated question as to whether the trajectories of public sector modernisation pursued in the various countries show convergence or divergence, the empirical evidence is somewhat ambivalent.

On the one hand, the sample of countries considered in this volume (and beyond) suggests that NPM-inspired public-sector modernisation and ‘twinned’ evaluation are on a ‘victorious march’ throughout the OECD world. This suggests convergence.²⁷

On the other hand, significant differences between the countries exist with respect to the timing, packaging and mix of reform concepts and components—whether NPM-related or ‘traditional’—hinting at divergence.²⁸

The divergence hypothesis becomes even more compelling when it is assumed that a country’s trajectory of institutional development is greatly shaped by factors which are deeply rooted—even in a ‘path-dependent’ manner—in the tradition and history of the country. Without further unfolding and substantiating this argument, it should suffice to single out two factors that plausibly make a difference in shaping the divergence of trajectories between countries and cultural groups (see, for example, Wollmann 2000b, 2000c with references).

One set of factors is evident in the juxtaposition of the Rechtsstaat and public interest tradition. The Continental European countries are still based on the Rechtsstaat tradition, in which the activities of public administration are for the most part

regulated by legal provisions and subject to juridical review. This puts a structural limit to the relevance of management tools, such as contracting and outsourcing. By contrast, in the Anglo-Saxon countries public administration is seen as being guided by ‘public interest’ principles as being rooted in the Common Law tradition in which—unlike the Roman Law tradition of the Continental European countries—no divide between state and society or between public and private law is recognised. Hence, the concepts of contractualism, marketisation and managerialism have a much easier cognitive and normative entry and acceptance (see Pollitt and Bouckaert 2000, pp. 52 ff, Wollmann 2000b, pp. 4 ff).

Another pivotal factor is the type of local government, particularly in its political and administrative functions in the country’s government system. In countries that have a politically and functionally strong local government system (such as Germany and the Scandinavian countries), the multi-function local government model is likely to be maintained, if not strengthened, thus limiting extensive use of (NPM-inspired) outsourcing of public functions. By contrast, in countries with politically and functionally weak local governments (such as the UK or New Zealand), the outsourcing and ‘quangoisation’ of public functions may be widespread.

Referring to the ideal-type exaggeration in which Johan Olsen has juxtaposed the ‘Supermarket State’ and the ‘Sovereign State’ models of government (1988), two scenarios can be depicted.

First, one might discern a type of public-sector trajectory in which the main function of the (programmatically ‘lean’) public sector is simply to facilitate public services, and the delivery of the services should essentially rely on marketisation, outsourcing and on ‘one function’ agencies—marginalising the traditional

democratically-elected and politically accountable local self-government. Some of these (ideal-type) features of Olsen's 'Supermarket State' can be identified, in real terms, in Anglo-Saxon countries.

By contrast, another pattern may be seen (again with some ideal-type exaggeration) in a public-sector trajectory in which, despite its absorption of significant neo-liberal and NPM impulses to increase the cost-efficiency, the state's main function continues to be legal regulation and application; and in which the bulk of public tasks continue to be carried out by (a still Weberian) public administration, primarily on the decentralised level of multi-function local self-government. These (ideal-type) features of Olsen's 'Sovereign State' can be said to characterise the politico-administrative system of Germany and, perhaps of Sweden (and the other Scandinavian countries)—with an even stronger emphasis on decentralised multi-functional local government.

Evaluation in and of public sector reforms

The introduction, institutionalisation and employment of the evaluation function (in the broad sense) has, on the one hand, shown significant commonalities among the countries. On the other hand, considerable differences in the timing, the profile and packaging of the evaluative variants and approaches are apparent.

Internal evaluation in and of public sector reforms—setting off a 'third wave' of evaluation

Perhaps the most striking and salient common trend can be seen in the new conceptual focus, strategic emphasis and institutional gravity of the goal-oriented, performance-

related and result-based conduct of the administrative activities and operations. During the 'planning phase' of the 1960s and 1970s, it is true, the evaluation function was conceptually guided by the idea of a 'policy cycle' (goal setting, implementation, evaluation) in which it played a crucial role in the policy-making process. But evaluation was functionally and operationally aimed at identifying the goal attainment of public policies rather than at a permanent monitoring and feedback reporting as an integral part of the 'public management package' (see Furubo and Sandahl 2002a, pp. 20 ff.). It is the strategic importance of performance management and of the permanent monitoring and feedback reporting as a primarily internal and self-evaluative operation that gives a new dimension, quality and gravity to NPM-inspired evaluation: the 'third wave' in the development of evaluation.

Reflecting this intrinsic 'twinning' with performance management and performance indicators (PI), approaches to internal evaluation have progressed furthest in countries that have been frontrunners in public management. This holds particularly true for the Anglo-Saxon countries - whether based of an almost orthodox conception of NPM, such as in New Zealand, but also in the U.K. (with the recent upsurge of performance measurement under New Labour's Best Value regime) or on a long-time managerialist tradition, such as in the USA. But it applies also to countries with a comparatively long-standing auditing tradition (such as the Netherlands) or those like Sweden, where the time-honoured agencies lend themselves to performance-oriented thinking and practice. In countries with a pronounced Continental European Rechtsstaat and legal regulation tradition that have been latecomers to adopting NPM, it may be this very Rechtsstaat tradition (and the significant role the application of law still plays in the day-to-day activities of public administration) that puts an institutional, cultural and

normative limit to performance management, and hence to concomitant forms of internal evaluation.

External evaluation in and of public sector reforms—still lagging behind, but gaining momentum

The development and state of external evaluation on public-sector reforms (that is, of evaluation initiated or conducted by institutions outside the operating unit) exhibits considerable cross-country variation that appears to be influenced by different sets of actors and factors.

External evaluation of public sector reforms in terms of methodologically rigorous (social science-based) evaluation research is still more or less a rarity throughout all countries, particularly with a comprehensive mandate.

Yet evaluative efforts which are directed, in a “meta-evaluation” stance, at externally assessing (and comparing) the performance reports of public agencies based on internal evaluation (monitoring) procedures are increasingly being undertaken, particularly by institutions of the auditing type and provenance. The performance assessment which has recently been put in place in the U.K. under the Best Value regime (revolving around the Audit Commission which reviews the performance of the local authorities on the basis of manifold indicators and criteria) is the most conceptually and instrumentally most advanced (and, it should be recalled, most centralist) example yet of a combined application of external and internal evaluation. On the top of it, the Best Value evaluation machinery has been given teeth by allowing central government to “reward” or “punish” the local authorities according to their high

or low performance quality. Another telling example is offered by Japan where central agencies have recently been created which operate under a “sunset” formula according to which they will be continued or terminated depending on the results of internal as well as external evaluation.

Institutions within the executive

Whereas on the ‘demand side’, the operating administrative units (agencies, divisions and sections with sectoral ministries) often show little interest in inviting and allowing external evaluation and outside evaluators to ‘penetrate’ their internal operations, executive government actors with cross-cutting (coordination, planning, budgetary, planning etc.) interests are more readily disposed and eager to initiate and conduct evaluation of the activities of sectoral ministries and agencies, including their administrative reform measures. In some countries (such as Australia and New Zealand) the Minister of Finance has been charged with promoting evaluation, including ‘meta-evaluating’ of the (internal) evaluative activities of sectoral ministries; in others a ministry or agency explicitly put in charge of engineering the administrative reform (such as the Ministry of Planning and Budget in Brazil or a newly created unit within the Ministry of General Affairs in Japan) performs this function.

Parliament

The United States Congress has traditionally promoted the (external) evaluation of policies, as a crucial instrument of its scrutinising function, by passing new legislation along with evaluation requirements. But in some parliamentary (Westminster) systems, particularly Australia's, parliament has taken an active part in external evaluation. In other countries, such as in Germany, parliaments have, in the past, called for the evaluation of 'substantive' policies (environmental protection, economic promotion etc. policies), but so far have paid little attention to the evaluation of administrative reform measures, probably because they traditionally have regarded administrative reforms as an executive prerogative.

Court of audit

In most countries the court of audit has moved in the direction of policy evaluation, including administrative reform measures. In the USA, the General Accounting Office (GAO) is designed to support Congress, but in 'parliamentary' countries that have moved towards] the GAO model, such as Sweden or the Netherlands, the courts of audit have been explicitly assigned, or have assumed, an advocacy role in initiating, conducting or monitoring ('meta-evaluating') evaluation of public-sector reforms. In some countries, such as Japan, the function of the court of audit has only recently been redefined accordingly. In other countries, such as Germany and Italy, the courts of audit have largely retained their traditional focus on scrutinising the budgetary compliance and bookkeeping of administrative activities, and have refrained from reviewing performance and policies.

International actors and discourses

In most countries the incipient moves towards external evaluation came with the advent of the NPM modernisation discourse, which in turn was strongly influenced by the ongoing international debate and practice. In some countries, the introduction of policy evaluation and of evaluation of public-sector reforms was clearly promoted by international or supra-national organisations (such as the European Union structural funds in the case of Italy, see Stame 2002; and Lippi, chapter 8 in this volume) and transnational conceptual borrowing and learning (as in the case of Japan, see Muramatsu and Matsunami, chapter 9 in this volume).

Scientific/research community

On the supply side, the extent of external evaluation was also influenced by the degree to which researchers, whether university-based or commercial, have embarked upon forming an active and assertive evaluation research community advocating the need for methodologically sound and intellectually independent policy evaluation. The recent expansion and activities of national as well as international evaluation societies express the growing interest and readiness of researchers and analysts world-wide to engage in professional evaluation and ‘contractual research’ (see Wollmann 2002a).

It needs to be pointed out, however, that particularly with regard to external evaluation, there is still a glaring discrepancy between the claim of transparency of political and administrative processes and activities and the still low profile of (external)

evaluation of public-sector reforms (see also Pollitt and Bouckaert, Chapter 2 in this volume).

1. For its ('historical', cultural, actor-centred, etc.) variants see, for example, Peters 1995.
2. See, for example, Wollmann 1997 (on institutional transformation of local government structures in post-socialist countries). In this context, mention should be made particularly of the concept of institutional 'legacies' and 'path-dependencies' which was first applied to the (institutional) transition and transformation in Latin American and the to the former socialist Central East European countries, see Karl and Schmitter 1991.
3. For similar conceptual schemes see Pollitt and Bouckaert 2000, Chapter 2 in this volume, and Christensen and Lægreid 2001 for an explication of their 'transformative' approach.
4. See Note 2, above.
5. See Schmidt 2000, pp. 232 ff.
6. On the concept of discourse and of discourse analysis, see Wittrock, Wagner and Wollmann 1991, pp. 43 ff; Hall 1993; Schmidt 2000, pp. 229 ff. On the application of discourse analysis on public sector reform, see Wollmann 1996, pp. 21 ff; and Wollmann 2002c.
7. On the concepts of (policy) learning, see Rose 1993; and Dolowitz and Marsh 1996.
8. On the concept of 'maturity' of evaluation culture and an attempt to grasp it by means of a nine-indicator scale, see Furubo and Sandahl 2002a, pp. 7 ff.
9. See Pollitt and Bouckaert 2000, p. 274: 'Conservative ministers tended to take the line that reform was essential, and self-evidently desirable, and that formal, public evaluation might prove a delay and distraction'.

10. Allan Schick concluded that although the reforms were ‘more comprehensive and rigorous than those introduced in other countries, they have been neither complete nor perfect’ (quoted in Halligan, chapter 5 of this volume).
11. See Halligan, chapter 5 in this volume: ‘[Parliamentary] committees have produced over 3000 reports covering the past three decades [that] have come to provide an integral component of the evaluation system’.
12. This ranking was made by Furubo and Sandahl by formulating nine indicators and by filling them (qualitatively) on the basis of available information (from the 21 country reports in their volume). See Furubo and Sandahl 2002a, pp. 7 ff.
13. See Dwight Waldo 1980: ‘We did not want a European-style state, we did not need a European style state, and we did not develop a European-style state’ (quoted from Stillman 1998, p. 172).
14. This is corroborated by the Furubo/Sandahl ranking according to which the USA is still ‘number one’ on the ‘evaluative culture’ score (see Furubo and Sandahl 2002a, p. 10).
15. Before 1998, ESV was a part of RRV, but became a governmental agency in its own right in mid-1998.
16. For an early assessment, see Levine 1981: ‘Evaluation is ... endemic throughout the Swedish system’.
17. In Furubo/Sandahl’s ranking of 21 countries Sweden is among the ‘top five’ on the evaluative culture scale (see Furubo and Sandahl 2002a, p. 10).
18. It should not come as a surprise that on the Furubo/Sandahl scale on ‘evaluative culture’ in 21 countries (see previous note), Germany is in the upper third, that is, in seventh place (see Furubo and Sandahl 2002a, p. 10).

19. See Stame 2002, note 5: ‘Most of these reforms... are under the direct inspiration, sometimes even wording, of their English prototypes. The famous “efficiency, effectiveness, economy” (3Es) principle became a refrain of public managers’.
20. In research on the implementation of the 1993 decree Rebora found (see Rebora 1999) ‘that, in general, Ministries have created a system of internal control through consulting ... local governments have preferred a mix of internal control and consulting’ (quoted from Stame 2002).
21. It should be remembered, however, that it was a comprehensive (highly critical) evaluation-type report (named ‘Giannini report’ after its main author) on the state of public administration, commissioned by Parliament and published in 1979, that broke the ground for radical administrative reforms.
22. SPARC was composed of business, labor and academic representatives as well as former high-ranking civil servants.
23. See Muramatsu and Matsunami, Chapter 9 in this volume: ‘The secretariat of this commission studied [US vice president] Al Gore’s Government Performance Program, the British program of evaluation by Public Service Agreements (PSAs) and later the United States Government Performance and Result Act (GRPA)’.
24. Law No. 103 of 1998. For details see Yamaha 2002, pp. 341 ff.
25. Hence, in Furubo/Sandahl’s aforementioned ranking of 21 countries in which, on the basis of 6 indicators, the ‘advancedness’ (‘maturity’) of the country’s ‘evaluative culture’ is (tentatively) assessed (see note 18), Japan comes out at the very bottom.
26. ‘The constitutional amendment played an important part in public management reform, because, besides allowing for the reform, it roused a national debate that changed traditional views of public administration’ (Bresser-Pereira 2001, p. 158). At

the same time the responsibility for the reforms was transferred from the (small) Ministry of Federal Administration and Reform of State to the larger (and more powerful) Ministry of Planning, Budget and Management.

27. See Thoenig, Chapter 11 in this volume: 'At least in OECD countries, the current disparity between a few pioneering countries and other more cautious countries can be expected to narrow'.

28. See Pollitt and Bouckaert 2000, p. 96, and Chapter 2 of this volume; Christensen and Lægreid 2001, pp. 20 ff; Wollmann 2000b (on the British and the German government systems); and Wollmann 2000c (on local government systems in the UK, France and Germany).

References

- Barzelay, M. (1997), 'Central Audit Institutions and Performance Auditing: A Comparative Analysis of Organizational Strategies in the OECD, in *Goverance*, vol. 10, no.3, pp. 235- 260
- Bemelmans-Videc, M. L. (2002), 'Evaluation in The Netherlands 1990-2000: Consolidation and Expansion', in Jan-Eric Furubo, Ray C. Rist and Rolf Sandahl (eds), International Atlas of Evaluation, New Brunswick and London: Transaction, pp. 115-128.
- Bresser Pereira, Luiz Carlos (2001), 'New Public Management Reform: now on the Latin America reform agenda, and yet...', in Revista Internacional de Estudos Políticos, Special issue[September], 143-165.
- Cabinet Office (2000), Adding it Up: Improving Analysis and Modelling in Central Government, London: Central Office of Information.
- Christensen, Tom and Per Lægreid (2001), 'A Transformative Perspective on Administrative Reforms', in Tom Christensen and Per Lægreid (eds), New Public Management, Aldershot: Ashgate, pp. 13-39.
- Department of the Environment, Transport and the Regions (DETR) (2001), Improving Local Public Services: Final Evaluation of the Best Value Pilot Programme (London: DETR)
- Department of Transport, Local Government and the Regions (DTLR) (2001), Strong Local Leadership – Quality Public Services (London: DTLR)

- Derlien, Hans-Ulrich (1990), 'Genesis and Structure of Evaluation Efforts', in Ray Rist (ed.), Program Evaluation and the Management of Government, New Brunswick and London: Transaction, pp. 147-176.
- Derlien, Hans-Ulrich (2002), 'Policy Evaluation in Germany: Institutional Continuation and Sectoral Activation', in Jan-Eric Furubo, Ray C. Rist and Rolf Sandahl (eds), International Atlas of Evaluation, New Brunswick and London: Transaction, pp. 74-91.
- Derlien, Hans-Ulrich and Ray C. Rist, (2002), 'Policy Evaluation in International Comparison', in Jan-Eric Furubo, Ray C. Rist and Rolf Sandahl (eds), International Atlas of Evaluation, New Brunswick and London: Transaction, pp. 439-455.
- Dolowitz, D. and D. Marsh (1996), 'Who learns from whom? A review of policy transfer literature', Political Studies, 343-355.
- Furubo, Jan-Eric, Ray C. Rist and Rolf Sandahl (eds) (2002), International Atlas of Evaluation, New Brunswick and London: Transaction.
- Furubo, Jan-Eric and Rolf Sandahl (2002a), 'A Diffusion-Perspective on Global Developments in Evaluation', in Jan-Eric Furubo, Ray C. Rist and Rolf Sandahl (eds), International Atlas of Evaluation, New Brunswick and London: Transaction, pp. 1-26.
- Furubo, Jan-Eric and Rolf Sandahl (2002b), 'Coordinated Pluralism—The Swedish Case', in Jan-Eric Furubo, Ray C. Rist and Rolf Sandahl (eds), International Atlas of Evaluation, New Brunswick and London: Transaction, pp. 115-127.

- Gray, Andrew and Bill Jenkins (1992), 'Implementing Evaluation: Lessons from the UK', in J. Mayne, et al. (eds), Advancing Public Policy Evaluation, Amsterdam: Elsevier, pp. 59-68.
- Hall, Peter A. (1993), 'Policy Paradigms, Social Learning and the State. The Case of Economic Making in Britain', Comparative Politics, 275-290.
- Halligan, John (2001), 'The process of reform in the era of public sector transformation: Directive and retrospective', in Tom Christensen and Per Lægreid (eds), New Public Management: The Transformation of Ideas and Practice, Ashgate: Aldershot, pp. 73-89.
- Hood, Christopher (1991), 'A public management for all seasons?', Public Administration, 69[1], 3-19.
- Karl, T. L. and P. C. Schmitter (1991), 'Modes of transition in Latin America, Southern and Eastern Europe', International Social Science Journal, 43, 269-284.
- Kickert, Walter and R. In't Veld (1995), 'National government, governance, and administration', in Walter Kickert and F. van Vught (eds), Public Policy and Administration Science in the Netherlands, London: Harvester Wheatsheaf, pp. 45-62.
- Levine, Robert A. (1981), 'Program Evaluation and Policy Analysis in Western Nations', in Robert A. Levine, Marian A. Solomon, Gerd-Michael Hellstern, and Hellmut Wollmann (eds), Evaluation Research and Practice, Beverly Hills and London: Sage, pp. 12-27.
- Lippi, A. (2000), 'One theory, many practices. Institutional allomorphism in Italian reorganization of local government', in S. Gherardi and B. Jacobsson (eds),

- ‘Managerialism as the Latin of our times: Reforming Italian public Sector Organisations’, Scandinavian Journal of Management, 4.
- Olsen, J. P. (1988), ‘Administrative Reform and Theories of Organization’, in C. Campbell and Guy Peters (eds), Organizing Governance: Governing Organizations, Pittsburgh, US: University of Pittsburgh Press, pp. 233-254.
- Osborne, D. and T. Gaebler (1992), Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector, Reading, US: Addison Wesley.
- Pollitt, Christopher (1995), ‘Justification by works or by faith? Evaluating the New Public Management’, Evaluation, 1[2 (October)], 133-154.
- Pollitt, Christopher and Geert Bouckaert (2000), Public Management Reform. A Comparative Analysis, Oxford: Oxford University Press.
- Premfors, R. (1998), ‘Reshaping the democratic state: Swedish experiences in a comparative perspective’, Public Administration, 76[1], 141-159.
- Rebora, G. (1999), La valutazione dei risultati nelle amministrazioni pubbliche, Milan: Guerini e associati.
- Rist, Ray C. and Kathleen L. Paliokas (2002), ‘The Rise and Fall (and Rise Again?) of the Evaluation Function in the US. Government’, in Jan-Eric Furubo, Ray C. Rist and Rolf Sandahl (eds), International Atlas of Evaluation, New Brunswick and London: Transaction, pp. 225-245.
- Rockman, Bert (2001), ‘Politics by other means: Administrative reform in the United States’, International Review of Public Administration, 6[2], 1-13.
- Rose, Richard (1993), Lesson-Drawing in Public Policy, Chatham, US: Chatham House.

- Sabatier, Paul (1987), 'Knowledge, Policy-Oriented Learning and Policy Change. An Advocacy Coalition Approach', in Knowledge: Creation, Diffusion, Utilization, pp. 649-692.
- Sandahl, Rolf (1992), 'Evaluation at the Swedish National Audit Bureau', in J. Mayne et al. (eds), Advancing Public Policy Evaluation, Amsterdam: Elsevier, pp. 115-122.
- Schmidt, Vivien A. (2000), 'Values and Discourse in the Politics of Adjustment', in Fritz W. Scharpf and Vivien Schmidt (eds), Welfare and Work in the Open Economy, vol. 1, Oxford, UK: Oxford University Press, pp. 229 ff.
- Stame, Nicoletta (2001), 'Evaluation in Italy: an inverted sequence from performance management to program evaluation?', in Jan-Eric Furubo, Ray C. Rist and Rolf Sandahl (eds), International Atlas of Evaluation, New Brunswick and London: Transaction, pp. 273-290.
- Stillman, Richard J. (1998), Creating the American State, Tuscaloosa and London: University of Alabama Press.
- Vedung, Evert (1992), 'Five Observations on Evaluation in Sweden', in J. Mayne, M. L. Bemelsmans-Videc, J. Hudson and R. Conner (eds), Advancing Public Policy Evaluation, Amsterdam: North-Holland, pp. 71-87.
- Vedung, Evert (1997), Public Policy and Program Evaluation, New Brunswick: Transaction.
- Wagner, Peter and Hellmut Wollmann (1986), 'Fluctuations in the development of evaluation research: Do "regime shifts" matter?' International Social Science Journal, 108, 205-218.
- Watanuki, J. (1991), 'The impact of social sciences on the process of development in Japan', in Peter Wagner, Carol Weiss, Björn Wittrock and Hellmut Wollmann,

- (eds), Social Science and Modern States, Cambridge: Cambridge University Press, pp. 221-229.
- Wilson, D. and Chris Came (2002), 'Local Government in the United Kingdom', 3rd edition, Houndmills: Palgrave
- Wittrock, Björn, Peter Wagner and Hellmut Wollmann (1991), 'Social science and modern state', in Peter Wagner, Carol Weiss, Björn Wittrock and Hellmut Wollmann, (eds), Social Science and Modern States, Cambridge: Cambridge University Press, pp. 28 ff.
- Wollmann, Hellmut (1983), Development and State of Policy Research in Japan, Tokyo: National Institute for the Advancement of Research (NIRA).
- Wollmann, Hellmut (1989), 'Policy analysis in West Germany's federal government: A case of unfinished governmental and administrative modernisation', Governance, 2[3], 233-266.
- Wollmann, Hellmut (1996), 'Verwaltungsmodernisierung: Ausgangsbedingungen, Reformanläufe und aktuelle Modernisierungsdiskurse', in Christoph Reichard, and Hellmut Wollmann (eds), Kommunalverwaltung im Modernisierungsschub?, Basel: Birkhäuser, pp. 1 ff.
- Wollmann, Hellmut (2000a), 'Local Government Modernization in Germany: Between Incrementalism and Reform Waves', Public Administration, 78[4], 915-936.
- Wollmann, Hellmut (2000b), 'Comparing Institutional Development in Britain and Germany: (Persistent) Divergence or (Progressing) Convergence?', in Hellmut Wollmann and Eckhard Schröter (eds), Comparing Public Sector Reform in Britain and Germany, Aldershot: Ashgate, pp. 1-26.

- Wollmann, Hellmut (2000c), 'The development and present state of local government in England and Germany', in Hellmut Wollmann and Eckhard Schröter (eds), Comparing Public Sector Reform in Britain and Germany, Aldershot: Ashgate, pp. 107-131.
- Wollmann, Hellmut (2000d), 'Local government systems. From historic divergence towards convergence? Great Britain, France, and Germany as comparative cases in point', Government and Policy, 18, 33-55.
- Wollmann, H. (2001a), 'Germany's trajectory of public sector modernization: continuities and discontinuities', Policy & Politics, 29[2], 151-169.
- Wollmann, Hellmut (ed.) (2001b), 'Evaluating Public Sector Reforms', special issue of Revista Internacional de Estudios Politicos, 127-143.
- Wollmann, Hellmut (2002a), 'Contractual Research and Policy Knowledge', International Encyclopedia of the Social and Behavioral Sciences, 5, 11574-11578.
- Wollmann, Hellmut (2002b), 'Verwaltungspolitik und Evaluierung: Ansätze, Phasen und Beispiele im Ausland und in Deutschland, Evaluation und New Public Management', Zeitschrift für Evaluation, 1, 75-101.
- Wollmann, Hellmut (2002c), 'Verwaltungspolitische Reformdiskurse und -verläufe im internationalen Vergleich', in Klaus König (ed.), Deutsche Verwaltung an der Wende zum 21. Jahrhundert, Baden-Baden: Nomos, pp. 489-525.
- Yamaya, Kiyoshi (2002), 'The Art of Policy Evaluation in Japan', in Jan-Eric Furubo, Ray C. Rist and Rolf Sandahl (eds), International Atlas of Evaluation, New Brunswick and London: Transaction, pp. 337-355.