

The Demand for Evaluation from a Public Choice Perspective

By Stefan Mann*

Summary

From the perspective of Public Choice Theory it is highlighted which motives exist for the persons in charge to commission the evaluation of political programs. It is questioned in what differences to non-commissioned evaluations these motives could result. The hypotheses are presented that commissioned evaluations are less theoretically ambitious and that the result will be more friendly on the program than evaluations that are carried out by scientists due to own interests. These hypotheses are tested for evaluations of structural policy in Germany between 1990 and 2000. It shows that commissioned evaluations are not less ambitious, but, on average, are much more in favour of the evaluated program than non-commissioned evaluations. As a conclusion it is recommended to put evaluations in different hands from that of the office in charge, for example to the courts of audit.

1. Introduction

The discipline of policy evaluation does not suffer from a lack of standards (Bussmann, 1997; Wottawa and Thierau, 1998; Vedung, 1999). The existing standards, however, provide a scientific ideal. They do not answer the question about the impact of the evaluation's context. The context of the evaluation of public programs or organizations can, from an economic point of view, be easily put into two categories: Evaluations that are commissioned and paid for by the organization in charge and evaluations that are not. The term evaluation is mostly used for commissioned studies only. Though, non-commissioned research papers that deal with the results of a political program or the success of an organization usually fulfill the criteria for evaluations and are therefore also grouped under this name henceforth.

The term non-commissioned, in our context, does include evaluations that are commissioned by independent institutions as research financing foundations or governmental bodies that are not responsible for the program evaluated. We therefore draw the distinction line between evaluations that are commissioned by the office that was directly involved in the respective program and evaluations that are carried out for any other purpose.

This paper is concerned with the question what difference for the methodology and for the result of the

evaluation the matter of economic dependency from the responsible office could have. It starts to answer the question by highlighting the motive for the office or the politician in charge (henceforth called the principal) to ask for evaluation of a public program.

One possible hypothesis could be that the principal shares the evaluating scientist's (henceforth called the agent) interest in the ultimate truth about the outcome of the program. If this hypothesis proved true, the principal-agent relationship would not have a direct impact on methodology and result of the evaluation except that the stimulus of the principal to provide full information to the agent would perhaps be stronger. However, as an alternative hypothesis it could be assumed that utility maximizing motivations of principals as suggested by Public Choice Theory could be an important force.

Section 2 approaches the question what impact the assumptions of Public Choice Theory would have for the demand for evaluations. In Section 3, commissioned evaluations of structural policy in Germany from the last ten years are compared with evaluations without this principal-agent relation. The comparison considers the applied methodology and the study's results. Conclusions

* Universität Rostock, Institut für Agrarökonomie und Verfahrenstechnik, Justus-von-Liebig-Weg 7, 18059 Rostock; e-mail: stefan.mann@agrarfak.uni-rostock.de

from this comparison and policy recommendations are drawn in Section 4.

2. Implications of Public Choice Theory

Public Choice Theory arose from the notion that not only individuals act as such, but they also form groups with (more or less) common interests which act collectively. Dunleavy (1991; 3) names four premises with which Public Choice Theory operates in predicting the behavior of groups:

- people have sets of well-formed preferences
- their preference orderings are consistent
- people are “maximizers” who seek biggest possible benefits and least costs in their decisions
- people are basically egoistic, self-regarding and instrumental in their behavior

Public Choice Theory — with this set of assumptions — has most intensely been used to predict the behavior of politicians and non-governmental interest groups. In some applications, however, it has also been analyzed for the group of bureaucrats which implicates the notion of utility maximization would have. This group is of special interest for our purpose as it is usually the group of bureaucrats in charge of structural policy that is commissioning evaluations of particular programs.

Reviewing Public Choice literature on the rationale of bureaucratic behavior, an interesting start of the debate is provided by Downs (1967). He acknowledges the heterogeneity of people working in administration and is therefore only able to present a catalogue of self-interest motives (e.g. prestige) and broader motivations (e.g. pride). In effect, however, one of the results of all instrumental motivations is that officials always distort information communicated upwards to superiors or politicians so as to present their own or their section's activity in the most favorable light.

The most famous model of bureaucratic behavior is provided by Niskanen (1971). He introduces the assumption that bureau heads desire as large a budget as possible. This is justified on the ground that income, prestige, power and amenities are a positive monotonic function of budget size; moreover, pressure from subordinates for larger budgets on the one hand and from executive and legislative committees on the other ensures the survival of the budget-maximizing bureaucrat in the same way that competition produces the survival of the profit-maximizing firm. However, as a difference to firms, Niskanen shows that the output of the bureau with such goals will be larger than optimum.

This model has been broadly discussed and several modifications have been suggested. One modification which was accepted by Niskanen (1989) himself is that

bureaucrats do not aim to maximize the total budget but their discretionary budget (or “slack-budget”), i.e. the budget they can spend in a way that is useful for them (Migué and Bélanger, 1974). Dunleavy (1991) goes even further in showing the disincentives for pure budget maximization and therefore develops the Bureau-shaping model. This model implies that rational bureaucrats concentrate on bringing their agency into line with an configuration conferring high status and agreeable work tasks. Their goal is to transform the agency into a central control, transfer or contracts agency. More recently, another aspect was added by Moe (1997) who emphasizes the aspect of political uncertainty and the bureaucrat's subsequent need to apply ex ante control mechanisms. Such control mechanisms may include decision procedures, civil service rules, independent forms of organization or timetables.

Apart from the undisputed correction that a rational bureaucrat would maximize his discretionary budget rather than his total budget, the approaches cited above are not really competing (even if they sometimes try to appear so) but rather complementary. They draw a rational bureaucrat's picture as somebody who tries to present his work in the brightest light in order to get more slack-budget in a bureau with almost unlimited options and guarantees for options in the future.

Why would such a bureaucrat demand an evaluation of a program under his supervision? The brief answer is: Either because he has to or because an evaluation may increase his utility.

Let us consider the latter case first: How can an evaluation increase the bureau's utility? What conclusions would an evaluation have to have in order to increase the discretionary budget for a program? It is most likely that the essence of the evaluation should sound like “The program had remarkable success, but in order to solve the problem in its full size, more flexibly usable funds are necessary.” Such a statement which emphasizes the importance of the problem at hand and the ability of the applied program to solve the problem will encourage the bureau's development towards a central agency. It will as well increase available funds for the bureau's relatively free allocation. Statements that would be maximally disliked, on the other hand, would rather be “The program is a waste of money. Terminate it immediately!”, which would be read as a threat even for the existence of the bureau. Hence, we would assume that voluntary demand for evaluation will exist only if the office can ensure that the evaluation's result will be beneficiary for the program and therefore for the bureau.

Some political programs may have a “built-in” evaluation, such as the Structural Funds of the European Commission that require ex-ante, interim and ex-post evaluation. In this case, a rational bureaucrat won't have the option to not demand an evaluation. He will, however,

try to push the evaluation into a positive direction, into a direction that again meets his demands.

What options does the principal have to influence the agent's result? As many questions between principals and agents, an exact answer depends on the specific characteristics of the principal, the agent and the environment. Generally speaking, there are three options that can be applied. It starts with the selection of the agent. With his given set of information, the agent is likely to choose an agent where a friendly evaluation appears most likely. The second option is to supply information about the program only selectively, for example to withhold unsuccessful projects. Eventually, the last option is to put pressure onto the agent. We have to consider that the agent will usually be a private firm that, for financial reasons, is interested to be appointed for more evaluations in the future. Getting contracts in the future is an option which the principal can tie on the good result of a current evaluation (see first option). This, however, would be a case of moral hazard which should not become obvious to the public.

Hence, arguing from a Public Choice perspective, it is likely that the results of commissioned evaluations are biased towards a positive outcome due to the, admittedly hidden, strategy of the bureaucrat in charge. As a first hypothesis, we therefore suggest that commissioned evaluations draw a more positive picture about the evaluated program than non-commissioned evaluations.

Another question that arises is the influence of the described principal-agent relation on the method of the evaluation. Assuming that the principal can gain positive utility from an evaluation, it would be rational for him to maximize his cost-benefit ratio by getting as many evaluations as possible for a given budget. On the other hand he would have the incentive to minimize costs for a given amount of evaluations in order to have enough "slack-budget" left. But how can average costs of evaluations be minimized?

Now consider the of the supply curve for evaluations, that is the relation between evaluations offered and the price received for them. The scope of such a supply curve is very much dependent on the methodology applied as there are rather simple and very sophisticated ways of carrying out an evaluation. Simple ways of evaluating are obviously less cost-intensive than thorough in-depth analysis. The most simple form of evaluation is just describing the very direct impacts of the program, i.e. in the case of structural policy how many funds have been spent for which projects and possibly how many new jobs were created within this project. Very sophisticated analyses, however, would try to model macroeconomic effects of the political program or would apply broader economic theory for arguments in favor or against the efficiency of the program.

Our second hypothesis from a Public Choice perspective would therefore be that commissioned studies tend to be merely descriptive evaluations compared with more rigorous and focused non-commissioned studies.

3. Testing for Systematic Differences Between Evaluations

3.1 Normative Framework

Before testing for differences between commissioned and non-commissioned evaluations, the comparability of criteria has to be discussed for the field of structural policy. Can be argued that commissioned studies just shared values with the office in charge, while non-commissioned studies dismissed the whole approach from the beginning?

This aspect could be a serious threat for programs in, say, educational policy, as there is large dissent among scientists about the ultimate target of educational policy. In structural policy, however, three primary targets can be identified that are largely consensual within the scientific community:

- a) Economic Growth. Structural policy is meant to increase overall productivity, usually measured in terms of GNP (Riedel and Scharr, 1997a).
- b) Regional Convergence. Within states or state unions, it is attempted to reduce productivity and income gaps between regions (Sala-i-Martin, 1996)
- c) Labor Generation. The creation of jobs is another central target of structural policy (Deitmer, 1993).

Structural Policy measures in Germany mainly take place in disadvantaged regions, so there is no competition between targets a) and b). And as we know from our macroeconomics lessons, there is a clearly positive relation between economic growth and labor generation (Okun's Law), so that a), b) and c) are complementary targets rather than substitutive.

As a result, the reference manual of primary goals of structural policy is uncritical and borne by broad consent. Evidence can be taken from the fact that after official evaluation guidelines for structural policy had been published by the European Commission (1999), no rigorous critique has been brought up. If differences between commissioned and non-commissioned evaluations exist, the possibility of different normative frameworks that have been in use can therefore largely be excluded.

3.2 Method

In order to verify the hypotheses, a sample of commissioned and non-commissioned studies of one

field of evaluations within a certain period of time and within a specified area was inspected. It was grouped according to methodology and result. By ordered probit and probit analysis, it was analyzed whether systematic differences between the two groups occurred.

Namely, the field of structural policy was chosen because the sample size of this category is sufficiently large. For all objective 1-areas (in Germany the whole East) and for most objective 5b-areas (some parts of Western Germany; since 2000 partly objective 2-areas), an external ex-ante, interim and ex-post evaluation has to be carried out according to Community regulations. It may be interesting to note, by the way, that not all commissioned evaluations are accessible for the public. On the other hand, the field of structural policy has been of some scientific interest for economists during the last years. Therefore, there do not only exist a large number of commissioned papers, but also of non-commissioned papers published as articles, books or PhD-theses.

In order to attain comparability between evaluations involved, only studies from the last ten years (1991–2000) evaluating programs that were carried out within Germany were included in the analysis. Furthermore, only studies were included that contained, at least in one sentence, an attempt to evaluate the efficiency of structural policy, not merely describing what has happened. Still, the heterogeneity of all papers included was enormous and ranged from the evaluation of structural policy in one district to structural policy in the whole of Germany.

Categorizing the different evaluations is not an easy task as every evaluation stands for itself with its individual motivation, approach and structure. It becomes therefore necessary to fix some criteria that can be applied for putting all available papers into the right “boxes”.

Commissioned studies are for our purpose defined as studies that are commissioned by an organization that plays a central role in the program that is evaluated. This appears to be easy in the first place, but led, for instance, to the exclusion of the work by Franzmeyer et al. (1993) which was commissioned by the Federal Ministry for Commerce in order to evaluate the reform of EC-Structural Funds. It could, in this case, not really objectively be decided whether the role of the Federal Ministry within this reform was significant enough to group the paper as being commissioned.

The next relevant categorization concerns the depth of the analysis. What we call “Theoretical Analysis” is, in effect, a remarkably heterogeneous group of papers which only share their reliance on some sort of economic method like

- Econometric models like HERMIN (Riedel and Scharr, 1997)
- Economic scenarios (Breitenacher et al., 1995)

- Welfare Economics (Loy et al., 1996)
- Sectoral classifications (Schultz et al., 1995)

Papers in which neither of those methods were applied but which were largely restricted to a description about the funds that were spent and the projects that were realized, were put onto the category “Descriptive Analysis”. A pure survey did not qualify for being adjoined to the theoretical category.

The most difficult task was to find a reasonable (not to mention an objective) system to group evaluations according to their result. Eventually, it was chosen to group evaluations in five categories which were defined as follows

Very positive was adjoined to papers that fully stated the success of the program without mentioning serious shortcomings.

Rather positive was adjoined to papers that mentioned some criticism about the program but came to conclusion that it had an overall positive effect.

Neutral was not claiming that the paper didn’t make any judgement; this category was rather adjoined to papers in which it did not become clear whether negative or positive aspects prevailed.

Rather negative was adjoined to papers that mentioned some positive effect but mainly recalled arguments for not supporting the program.

Very negative was adjoined to papers that did only see disadvantages and no benefit in the program.

Hence, the following two functions were tested:

$$M = f(C)$$

$$R = f(C)$$

whereas M was the methodology of the study (0 — descriptive analysis, 1 — theoretical analysis), R was the result of the study (1 — very positive; 5 — very negative) and C was describing the character of the study (1 — commissioned, 0 — non-commissioned)

21 commissioned evaluations and 17 non-commissioned evaluations were screened. Albeit the total of 38 appears as a small sample in social science research, it is nevertheless representative. The “Population”, i.e. published evaluations about structural policy in Germany during the last ten years, is not much larger than that. All regional ministries and major research institutes were approached for the process of sampling.

3.3 Results

Outcomes regarding methodology and result of all structural policy evaluations that have been screened are reflected in Tables 1 and 2.

Table 1:

Comparison of applied methods evaluating structural policy

	Descriptive Analysis	Theoretical Analysis
Commissioned Studies	17 ¹	4 ²
Non-commissioned Studies	10 ³	7 ⁴

Table 2:

Comparison of obtained results evaluating structural policy

	Very positive	Rather positive	Neutral	Rather negative	Very negative
Commissioned Studies	2 ⁵	14 ⁶	5 ⁷	0	0
Non-commissioned Studies	1 ⁸	3 ⁹	6 ¹⁰	3 ¹¹	4 ¹²

Probit analysis that was applied to Table 1 in order to reveal differences in the methodology between the two sample groups, showed that the hypothesis of bureaucrats commissioning mainly simple, descriptive analyses had to be rejected as probability of an influence of the variable was 84 % only (Pseudo-R²: 4.4 %). Given the sometimes difficult decision whether still to consider an evaluation as descriptive or already as theoretical, it should be made clear that there is no more reason to assume that commissioned studies apply a methodology less ambitious than non-commissioned studies.

Ordered probit analysis that was carried out to test whether there is a dependency of the outcome on being a commissioned evaluation showed that, on our five-step scale, being commissioned improved the likely outcome by 1.4 steps. With a likelihood of 99.9 % (Pseudo-R²: 12.8 %), the evaluations' results were influenced by being a commissioned study. Even by considering the difficulty to group all evaluations appropriately on the scale, and even if there were sometimes misjudgements, the dependency would most probably endure. In this case, the hypothesis as suggested by Public Choice Theory had to be accepted.

4. Conclusions

The outcome of the analysis suggests that commissioned evaluations are biased. There is no reason to assume that evaluations on structural policy that are written to pursue a PhD, to do some publishing in referred journals or to follow own scientific interests are systematically biased in one direction. If there exists a difference in the outcome between commissioned and

non-commissioned papers, as it was proven, the only plausible explanation is that commissioning bureaucrats tend to influence the results. They may do so consciously. Particularly in cases where principal and agent have been working together for a couple of years, it is, however, more likely that the agent delivers what he knows will suit the principal's interests. He may, in a way, carry out self-censorship.

This is not at all to say that commissioned evaluations are worthless. They tend, in the field studied here, to reveal a lot of achievements of structural policy measures and the do usually mention a few shortcomings. They are methodologically not less sophisticated than non-commissioned evaluations. The only critique that remains is that their conclusions let the policy appear in a brighter light than the results of a purely scientifically guided evaluation.

A conclusion close at hand from these results would be that the administration should not commission evaluations any more. This, however, would be most counter-productive as the efficiency of public funds keeps to be an issue of uttermost importance, and it would be too dangerous to rely on the self-interest of scientists who may concern themselves with the subject or they may not. Thus, the question has to be raised how the institutional background could be altered in order to commission evaluations without the disadvantageous principal-agent relationship.

¹ Blume et al. (1999); Böhm and Volkert (1998, 1999); Büstro (1997a, 1997b); Friedrich et al. (2000); Geißendörfer (1998); Geißendörfer et al. (1998, 1999); Hagen and Toepel (1997); Kment and Borsch (1998); Riedel et al. (1996); Riedel and Scharr (1997b); Schwab et al. (1998); Seibert et al. (1997); Toepel et al. (2000); Toepel and Weise (1995).

² Breitenacher et al. (1994); Riedel and Scharr (1997a, 1999); Schultz et al. (1995).

³ Barth and Karrasch (1995); Baumann (1997); Bursig (1991); Fock (2000); Forstner and Clemens (1998); Geldermann et al. (1996); Henkel (2000); Hummel (1997); Ochel (1997); Riedel and Wiesner (1997).

⁴ Beckmann (1995); Deitmer (1993); Hoffmann et al. (1997); Kroker (1998); Loy et al. (1996); Schrader (1995); Striwe et al. (1996).

⁵ Geißendörfer et al. (1999); Kment and Borsch (1998).

⁶ Blume et al. (1999); Böhm and Volkert (1998); Büstro (1997b); Breitenacher et al. (1994); Friedrich et al. (2000); Geißendörfer (1998); Geißendörfer et al. (1998); Hagen and Toepel (1997); Riedel and Scharr (1997a, 1997b, 1999); Schultz et al. (1995); Schwab et al. (1998); Seibert et al. (1998).

⁷ Böhm and Volkert (1999); Büstro (1997a); Riedel et al. (1996); Toepel et al. (2000); Toepel and Weise (1995).

⁸ Henkel (2000).

⁹ Deitmer (1993); Hummel (1997); Ochel (1997).

¹⁰ Bursig (1991); Forstner and Clemens (1998); Geldermann et al. (1996); Hoffmann et al. (1997); Riedel and Wiesner (1997); Schrader (1995).

¹¹ Baumann (1997); Fock (2000); Kroker (1998).

¹² Barth and Karrasch (1995); Beckmann (1995); Loy et al. (1996); Striwe et al. (1996).

A possible answer would be to found a public organization with the only task to commission evaluations of public programs. A necessary precondition for the success of such an organization would be that links between this organization and ministries in charge of the programs that have to be evaluated are as weak as possible. This organization would therefore have no incentive whatsoever to push the result of the evaluation into any direction, so there would be no risk of influencing the agent except for a proper and attentive accomplishment of his work. Sticking at the German example, there would be good reasons to place such an organization on the Federal level. The two fields with the

most intense ongoing evaluation activities, educational and structural policy, are mainly in the hands of the *Länder* and the European Commission, so that a federal agency would provide the greatest degree of independence.

It is a widespread consensus that the amount and importance of evaluation will rather increase than decrease in the future. This is another reason why the creation of an relatively independent governmental Evaluation Organization or the integration of such tasks in an existing organization (eg. the Federal Court of Audit) would be a worthwhile challenge, contributing to the quality and objectivity of evaluation.

References

- Barth, U., P. Karrasch* (1995): Wirkungen der regionalen Wirtschaftsförderung im Regierungsbezirk Leipzig. Berlin: KSPW.
- Baumann, S.* (1997): Strukturpolitik in den neuen Bundesländern. Bayreuth: Verlag P.C.O.
- Beckmann, K.* (1995): Probleme der Regionalpolitik im Zuge der Vollendung des Europäischen Binnenmarktes. Frankfurt: Peter Lang.
- Blume, L., K.-P. Gaulke, J. Rother* (1999): Ökonomische Wirkung der Städtebauförderung in Mecklenburg-Vorpommern. Berlin: Duncker & Humblot.
- Böhm, D., B. Volkert* (1998): Bericht zur Ex-ante-Evaluation zum operationellen Programm für die EU-Gemeinschaftsinitiative KONVER II in Baden-Württemberg. Stuttgart.
- Böhm, D., B. Volkert* (1999): Finanzielle Existenzgründungsförderung des Landes Baden-Württemberg — eine Erfolgskontrolle. Stuttgart.
- Breitenacher, M., G. Illing, E. Langmantel, A. Städtler, L. Uhlmann*: Gesamtwirtschaftliche Wirkungen der Existenzgründungspolitik sowie Entwicklungen der mit öffentlichen Mitteln — insbesondere Eigenkapitalhilfe — geförderten Unternehmensgründungen. München 1994.
- Büstro* (1997a): Zwischenbewertung des Einsatzes der Europäischen Strukturfonds für die Periode 1994 bis 1996 im Bundesland Mecklenburg-Vorpommern. Rostock.
- Büstro* (1997b): Zwischenbilanz für die Interventionen des EFRE im Rahmen des Gemeinschaftlichen Förderkonzeptes und des Operationellen Programms der Jahre 1994 bis 1999 in Mecklenburg-Vorpommern. Rostock.
- Bursig, B.* (1991): Die Regionalpolitik der Europäischen Gemeinschaft. Frankfurt: Peter Lang.
- Bussmann, W.* (1997): Einführung in die Politikevaluation. Basel: Helbling & Lichtenhain.
- Deitmer, I.*: Effekte der regionalen Strukturpolitik auf Investitionen, Beschäftigung und Wachstum. Münster 1993.
- Downs, A.* (1967): Inside Bureaucracy. Boston: Little, Brown.
- Dunleavy, P.* (1991): Democracy, Bureaucracy & Public Choice. New York: Harvester.
- European Commission* (1999): Evaluating socio-economic programmes. Luxembourg.
- Fock, T.* (2000): Beitrag der Wirtschafts- und Strukturpolitik zur Lösung von Beschäftigungsproblemen. In: Johann Heinrich von Thünen — Gesellschaftspolitische Aspekte seines Werkes und ihre Bedeutung für die Entwicklung des ländlichen Raumes. Berichte über Landwirtschaft; 213. Sonderheft. Münster-Hiltrup: Landwirtschaftsverlag.
- Forstner, B., Clemens, D.* (1998): Einzelbetriebliche Investitionsförderung in Schleswig-Holstein: effizient und nachhaltig? In: bauernblatt 52 (51/52), 74–78.
- Franzmeyer, F., B. Seidel, C. Weise* (1993): Die Reform der EG-Strukturfonds von 1988. Berlin: Duncker & Humblot.
- Friedrich, W., S. Gehle, B. Lagemann, G. Machalowski, H. Schruppf* (2000): Die Gemeinschaftsinitiative "KMU" in Ostdeutschland. Essen: RWI.
- Geißendörfer, M.*: Zwischenbewertung der Gemeinschaftsinitiative LEADER II Bayern. Triesdorf 1998.
- Geißendörfer, M., O. Seibert und H. von Meyer* (1998): Ex-post-Evaluierung der Gemeinschaftsinitiative LEADER I in Deutschland. In: Berichte über Landwirtschaft 76 (4) 540–579.
- Geißendörfer, M., M. Thomas, O. Seibert* (1999): Weitere wissenschaftliche Begleitung der Ziel 5b II-Strukturförderung in Bayern. Triesdorf.

- Geldermann, A., C.-G. Eucken, M. Odening (1996): Landwirtschaftliche Investitionsförderung in den neuen Bundesländern – eine kritische Analyse dargestellt am Beispiel Brandenburg. In: Landwirtschaftliche Investitionsförderung: Bisherige Entwicklung, aktueller Stand, Alternativen für die Zukunft. Frankfurt.
- Hagen, K., K. Toepel (1997): Europäische Strukturfonds in Sachsen: Zwischenevaluierung für die Jahre 1994 bis 1996. Berlin.
- Henkel, G. (2000): Dorferneuerung in Deutschland — Bilanzen und Perspektiven eines erfolgreichen Programms. In: Johann Heinrich von Thünen — Gesellschaftspolitische Aspekte seines Werkes und ihre Bedeutung für die Entwicklung des ländlichen Raumes. Berichte über Landwirtschaft; 213. Sonderheft. Münster-Hiltrup: Landwirtschaftsverlag.
- Hoffmann, H., Jahnke, D. und Kögl, H. (1997): Effektivität der Investitionsförderung. In: Rostocker Agrar- und Umweltwissenschaftliche Beiträge, Heft 7. Rostock, 179–202.
- Hummel, M. (1997): Subventionen: Motor für Wachstum und Beschäftigung? In: Oppenländer, K.-H.: Wiedervereinigung nach sechs Jahren: Erfolge, Defizite, Zukunftsperspektiven im Transformationsprozess. Berlin.
- Kment, E., R. Borsch (1998): Endbericht über die Zwischenevaluierung Interreg II für den Österreichisch-Bayerischen Grenzraum. München.
- Kroker, R. (1998): Impacts of Regional Policy on Regional Growth and Employment. Europäische Zeitschrift für Regionalentwicklung 7/98, 11–12.
- Loy, J. P., Koester, U., Striwe, L. (1996): Struktur- und Verteilungswirkungen der landwirtschaftlichen Investitionsförderung am Beispiel Schleswig-Holsteins. In: Landwirtschaftliche Investitionsförderung: Bisherige Entwicklung, aktueller Stand, Alternativen für die Zukunft. Frankfurt.
- Migué, J. L., G. Bélanger (1974): Toward a General Theory of Managerial Discretion. Public Choice 17: 23–43.
- Moe, T.M. (1997): Positive Theory of Public Bureaucracy. In: Mueller, D. C.: Perspectives on Public Choice. Cambridge: University Press.
- Niskanen, W.A. (1971): Bureaucracy and Representative Government. New York: Aldine-Etherton.
- Niskanen, W.A. (1989): A Reflection on Bureaucracy and Representative Government. In: Blair, A., S. Dion: The Budget-Maximizing Bureaucrat. Pittsburgh.
- Ochel, W. (1997): Die Förderung der neuen Bundesländer und Ost-Berlins durch die Regionalpolitik der Europäischen Union. In: Oppenländer, K.-H.: Wiedervereinigung nach sechs Jahren: Erfolge, Defizite, Zukunftsperspektiven im Transformationsprozeß. Berlin: Duncker & Humblot.
- Riedel, J., F. Scharr (1997a): Die Europäischen Strukturfonds in den neuen Bundesländern. München: ifo.
- Riedel, J., F. Scharr (1997b): Der Europäische Fonds für Regionale Entwicklung (EFRE) in Thüringen. München: ifo.
- Riedel, J., F. Scharr, (1999): Wirtschaftsförderung im Transformationsprozess. München: ifo.
- Riedel, J., G. Wiesner (1997): Aufschwung Ost: Konjunktur für Wirtschaftsförderprogramme. In: Oppenländer, K.-H.: Wiedervereinigung nach sechs Jahren: Erfolge, Defizite, Zukunftsperspektiven im Transformationsprozess. Berlin: Duncker & Humblot.
- Riedel, J., G. Wiesner, G. Bratzke, G. Wagner (1996): Wirkungen der Wirtschaftsförderung auf unternehmerische Innovationen in der Region Halle–Leipzig–Dessau. München: ifo.
- Sala-i-Martin, X.X. (1996): Regional cohesion: evidence and theories of regional growth and convergence. European Economic Review 40 (6), 1325–1352.
- Schrader, H. (1995): Impact Assessment of the EU Structural Funds to Support Regional Economic Development in Rural Areas of Germany. In: Journal of Rural Studies, Vol. 10 No.4, 357–365.
- Schultz, B., B. Middeldorf, M. Holst, B. Neumayer (1995): Evaluierung des Mitteleinsatzes des EFRE im Rahmen des Operationellen Programms Mecklenburg-Vorpommern 1991 bis 1993. Berlin.
- Schwab, O., J. Lang, W. Jaedicke (1998): Evaluierungsstudie über den Einsatz der Strukturfondsmittel der Europäischen Union im Rahmen der Gemeinschaftsinitiative INTERREG II in den Freistaaten Sachsen und Bayern. Berlin.
- Seibert, O., M. Geißendörfer, M. Thomas, S. Gaerdes (1997): Fondsübergreifende Zwischenbewertung des Ziel 5b II-Programmes Bayern (1994–96). Triesdorf.
- Striwe, L., J. P. Loy, U. Koester (1996): Analyse und Beurteilung der einzelbetrieblichen Investitionsförderung in Schleswig-Holstein. Agrarwirtschaft 45 (12) 423–434.
- Toepel, K., R. Sander, W.-C. Strauss (2000): Europäische Strukturpolitik für die Stadterneuerung in Ostdeutschland. Frankfurt: Peter Lang.
- Toepel, K., C. Weise (1995): Die Förderung von gewerblichen Investitionen in Ostdeutschland durch den Europäischen Fonds für Regionale Entwicklung. Berlin: DIW.
- Vedung, E. (1999): Evaluation im öffentlichen Sektor. Wien: Boehlau.
- Wottawa, H., H. Thierau (1998): Lehrbuch Evaluation. Göttingen: Huber.

Die Nachfrage nach Evaluationen aus Public Choice-Perspektive

Zusammenfassung

Aus Perspektive der Public Choice Theorie wird die Fragestellung beleuchtet, welche Motive es für die Verantwortlichen geben kann, politische Programme evaluieren zu lassen, und welche Unterschiede zu nicht beauftragten Evaluationen daraus resultieren könnten. Es werden die Hypothesen aufgestellt, dass vom Verantwortlichen in Auftrag gegebene Evaluationen erstens theoretisch weniger anspruchsvoll sind und zweitens vom Ergebnis besser ausfallen als Evaluationen, die Wissenschaftler aus eigenem Antrieb ausführen. Am Beispiel der Evaluationen von Strukturpolitik in Deutschland im Zeitraum 1990–2000 werden diese Hypothesen überprüft. Es zeigt sich, dass in Auftrag gegebene Evaluationen zwar nicht weniger anspruchsvoll sind, aber im Durchschnitt ein deutlich besseres Ergebnis des zu evaluierenden Programmes konstatieren als andere Evaluationen. Als Fazit wird empfohlen, die Evaluation politischer Programme in andere Hände zu legen als das Programm selbst, z. B. in die der Rechnungshöfe.