

# MPRA

Munich Personal RePEc Archive

## **Evaluating election platforms: a task for fiscal councils? Scope and rules of the game in view of 25 years of Dutch practice**

Bos, Frits and Teulings, Coen

CPB Netherlands Bureau for Economic Policy Analysis

June 2011

Online at <http://mpra.ub.uni-muenchen.de/31536/>

MPRA Paper No. 31536, posted 14. June 2011 / 10:46

*Paper to be published in the Proceedings of the Banca d'Italia Fiscal Policy Workshop on Rules and Institutions for Sound Fiscal Policy after the Crisis Perugia, 31 March - 2 April 2011*

**Evaluating election platforms: a task for fiscal councils?  
Scope and rules of the game in view of 25 years of Dutch practice**

Frits Bos and Coen Teulings<sup>1</sup>

**Abstract**

In some countries - the Netherlands, UK and USA - the expected economic implications of election platforms of political parties are evaluated by independent economic institutions prior to the election. This paper analyzes the merits and limitations of this process, taking 25 years of Dutch experience as a point of reference. In particular in times of financial crisis and unsustainable public finance, evaluation of election platforms can serve as a disciplining device for unrealistic or (time) inconsistent promises by politicians. More in general, it can help political parties to credibly inform voters about the implications of their platforms, to design more efficient policies and to reach consensus on them. It can also create a level playing field for political parties not represented in the government, in particular those with limited resources for economic information and expertise. However, there may be adverse effects, in particular when trade-offs are presented in an unbalanced way or when the rules of the evaluation provide too much room for gaming and free lunches.

*Key words: Evaluation of election platforms, Fiscal watchdogs*

*JEL code: A11, D7, E02, E61, E62, H0*

---

<sup>1</sup> The authors would like to thank Peter Dekker, Bas Jacobs and Wim Suyker for comments on an earlier draft.

## 1. Introduction

In some countries - the Netherlands, UK and USA - the expected economic implications of election platforms of political parties are evaluated by independent economic institutions prior to the election. We analyze the merits and limitations of this process, taking the Dutch experience as a point of reference<sup>2</sup>.

In the Netherlands, some months before the elections, on request of the political parties, CPB Netherlands Bureau for Economic Policy Analysis (CPB) publishes an economic evaluation of their election platforms. For example, what are the consequences of the platforms for the government budget, economic growth, employment, the purchasing power of various types of households and the environment? In March 2010, CPB compared the election plans of nine Dutch political parties (see CPB, 2010). This comparison and analysis was the seventh evaluation of election platforms in twenty-five years.

In the UK, since the election of 1997, the Institute for Fiscal Studies (IFS)<sup>3</sup> publishes policy briefings during election time. These policy briefings review the policies advocated by the three main UK political parties in their ‘manifestos’. They also discuss the track record of current and previous government and the sustainability of public finance without any change in policy. A wide range of policy issues is covered, like ‘pensions and retirement’, ‘environment’, ‘living standards, inequality and poverty’ and ‘families and children’. The latter includes policy proposals about education, parental leave, child care and relevant taxes and benefits. According to one of the major UK-newspapers, these IFS policy briefings are a great service to voters: “Many voters will have asked themselves why the main parties have been so vague about what they intend to do about the crisis in public finances. But it took the IFS to launch a blunt, impartial and authoritatively damning attack on all three parties, accusing them of not being straight with voters” (The Independent, 6 May 2010).

---

<sup>2</sup> Previous assessments of the Dutch experience with the economic evaluation of election platforms are Haffner and Van Bergeijk (1994), Don (2003) and the papers in Graafland and Ros (2003).

<sup>3</sup> The IFS is an independent economic research institute funded by public and private grants. It produces academic and policy related findings on UK taxation and public policy. It was launched in 1971 in order to improve UK tax policy and to make the UK tax debate more informed and rational by bringing together political, legal, accounting and economic arguments. Its first major success was the Meade report published in 1978. This report by a committee chaired by Professor James Meade, winner of the Nobel Prize in economics in 1977, evaluated the UK tax system as a whole, what are the objectives, to what extent are these objectives met by the current UK tax system and what reforms should be undertaken?

In the USA, the estimates by the Congressional Budget Office (CBO) of budgetary costs and savings of some major policy proposals, e.g. on health care reform, play a major role in elections for president, the house and the senate<sup>4</sup>. CBO traditionally focuses on static scoring of such policy proposals, i.e. including the direct effects and their behavioral responses but without macroeconomic effects. However, since 2003, also the consequences of different macro-economic assumptions are shown ('dynamic scoring').

Though evaluation of election platforms by independent economic institutes is quite rare, calculating the budgetary costs of major policy proposals or analyzing their macro-economic consequences are very common all over the world. Such analyses are used in political decision-making in many countries. Also international organizations, like OECD, IMF, the European Commission and the World Bank, make similar analyses. So, introducing independent evaluation of election platforms will generally imply two types of changes. Firstly, such methods currently often used for analyzing official government policy proposals are to be applied to the election platforms of government parties and opposition parties. Secondly, these analyses should be conducted by independent institutions instead of government units subject to political interference, e.g. a Ministry of Finance or research institutes with a clear political signature. This second type of change may turn out to be the most difficult. The abrupt end of the new Hungarian fiscal council shows that independent fiscal watchdogs are most useful when their advice bites most, but then they are also politically most vulnerable.

Economic theory on decision making in a democracy provides a rationale for independent evaluation of election platforms (see Swank, 2003). Decision making under direct democracy is likely based on too little information, because information is an (impure) public good. Collecting such information is costly and when collected also many others can benefit from this information. This raises a free rider problem. An alternative solution to ensure sufficient information for voters is to delegate the making of policy to a number of agents, i.e. introduce representative democracy instead of direct democracy. However, parties are inclined to provide incorrect information. They will adopt a view of the economy that is

---

<sup>4</sup> The CBO was established in 1974 to serve the American Congress and plays a major role in measuring the budgetary impact of new legislation and other policy proposals: "CBO's score can doom legislation or smooth the way to passage; it can compel committees and members to modify pending legislation, even after political deals have been negotiated; and it can complicate or thwart the president's legislative ambitions" (Schick, 2007, p. 133).

meant to increase their chances of winning the elections. Voters may also distrust this information for being politically biased, even when it is actually objective. Financing independent institutions for analyzing the economy and policy proposals may help to solve this information problem.

The first ideas for an economic analysis of election platforms in the Netherlands date back to 1972<sup>5</sup>. A national economic journal asked all major political parties to specify their proposals for five policy issues: wage and price policy, redistribution of income, employment, economic growth and spatial planning. Each issue was illustrated with questions, e.g.:

- How should the government fight inflation? Should the wages of civil servants be frozen or should all price increases be forbidden?
- Are you in favour of a maximum income and how high should this be?
- How to avoid that young people do not find a job in line with their education? Should foreign employees continue to work in the Dutch economy?
- Should economic growth be reduced in order to limit pollution and exhaustion of natural resources?
- Should economic growth be reduced in the urban agglomeration in the Western part (the Randstad) and be stimulated in the Northern and Southern part of the Netherlands?

The seven major political parties wrote an essay on these issues and these essays were published three weeks before the elections. The editorial motivated this interrogation of political parties: “Dutch political parties fail continuously in being frank and transparent about their political preferences and policy proposals. For problems like the environment, transport, spatial planning and inflation only ad hoc measures are taken that avoid hurting the voters. ... Political parties try to seduce voters by vague promises instead of by being honest and clear about how to fight inflation and reduce the negative effects of economic growth’ (p. 1057).

In this editorial, the idea of an independent evaluation of election platforms is related to the credibility and commitment problem of politics. Since voters have a hard time to evaluate the costs of policy proposals, politicians are tempted to make more promises than is financially viable. Furthermore, politicians are tempted to spend money on specific interest groups today, as to achieve their electoral support, leaving the distribution of the tax burden

---

<sup>5</sup> See ESB, Dutch Journal on Economics and Statistics, No 2874, pp. 1057-1079.

to finance this spending open for future decision making. All kind of institutions have been designed to resolve these problems<sup>6</sup>, fiscal councils being a rather recent new branch on this tree. Fiscal councils are supposed to provide an independent analysis of the governments' fiscal policy on a regular basis (see Hagemann, 2010, Calmfors, 2011 and Calmfors and Wren-Lewis, 2011 for an overview). The evaluation of election platforms can be interpreted as another branch of this tree or just a new task for fiscal councils.<sup>7</sup>

A major merit of evaluating election platforms is the timing. For a fiscal watchdog to be effective, providing information and policy advice at a very early stage of decision-making is important. This maximizes the time for interaction with policy makers, reduces conflicts with policy makers (in particular when such information comes as an unhappy surprise) and reduces the likelihood that policy advice is ignored simply because it conflicts with earlier political statements. According to the IFS "With public attention more focused on policy debates than usual, the election campaign is one of the best opportunities we have to ensure that the sort of rigorous, evidence-based social science research that we undertake has a high impact on policy development and discussion" (IFS Policy Briefings for the General Election, 2010).

The structure of this paper is as follows. Section 2 describes the history of the evaluation of election platforms in the Netherlands. Section 3 provides an overview of various approaches to the evaluation of election platforms, varying in scope from just an overview of direct budgetary effects to full-fledged welfare analysis including their general equilibrium effects. Section 4 discusses in greater detail the rules of the 'game' in the Netherlands. For example: How do you ensure the quality and neutrality of the evaluation? And how to avoid political parties will manipulate the game? Section 5 concludes.

---

<sup>6</sup> Calmfors and Wren-Lewis (2011) distinguish six reasons for deficit bias: informational problems (e.g. due to over-optimism among voters or politicians about future economic growth), impatience, exploitation of future generations, electoral competition, common-pool problems and time inconsistency.

<sup>7</sup> Calmfors (2011) distinguishes six tasks for fiscal councils: 1. Provision of 'objective' macro economic forecasts on which government budget proposals can be based. 2. Costing of various government policy initiatives. 3. Ex ante evaluation of whether fiscal policy is likely to meet its medium-term targets. 4. Ex post evaluation of whether fiscal policy has met its targets. 5. Analysis of the long-run sustainability of fiscal policy and 6. Normative recommendations on fiscal policy. . The evaluation of election platforms should be added as a seventh task.

## **2. The Dutch history of evaluating election platforms**

The evaluation of election platforms in the Netherlands started in 1986. The CPB<sup>8</sup>, the government's economic expert institute, had just published the new official macro-economic forecasts for the new period of government. The three major political parties asked then to investigate the economic consequences of their election platforms. To what extent would these proposals help to increase economic growth and improve public finance? And what would be the consequences for unemployment and the purchasing power of various groups of households? CPB prepared for each party a separate paper. These three papers were published jointly after the election.

Three years later, four political parties requested an evaluation of their election platforms and the results were jointly published a month before the election. As a consequence, since 1989 the results of the evaluation of election platforms have played an explicit role in Dutch elections.

---

<sup>8</sup> CPB Netherlands Bureau for Economic Policy Analysis (CPB) is the oldest fiscal council in the world. It was founded in 1945. CPB is fully independent as far as the contents of its work are concerned. It is publicly funded and part of the Ministry of Economic Affairs, Agriculture and Innovation. Research is carried out on CPB's own initiative, or at the request of the government, parliament, trade unions or employers' federations. Forecasts and analyses by CPB play a major role in the official decision-making process of Dutch economic and fiscal policy (see Bos and Teulings, 2010).

**Table 1. Evaluation of election platforms in the Netherlands (1986-2010)**

	1986	1989	1994	1998	2002	2006	2010
Number of political parties	3	4	5	5	8	8	9
Seats in parliament prior to elections	133	137	131	137	123	139	148
A. Direct budgetary effects during the next election cycle							
- Overview of budget cuts, expenditure and taxes	x	x	x	x	x	x	x
- Detailed overview of policy measures			x	x	x	x	x
- Overview of policy measures by function					x	x	x
- Effect on government deficit and debt	x	x	x	x	x	x	x
- Effect on employment in the government and healthcare sector					x	x	x
B. Economic feedbacks on budget, labour market and bbp during the next cycle	x	x	x	x	x	x	(a)
C. Long term effects on labour market and bbp		x	x	x	x	x	x
D. Effects on purchasing power of various groups of households	x	x	x	x	x	x	(a)
E. Long term effects on budget					x	x	x
F. An analysis of the impact on specific sectors							
- Environment and mobility			x	x	x	(a)	x
- Education and innovation						x	x
- Health Care		x			x		x
G. An analysis of the impact on specific sectors using a welfare criterion							x
(a) Absent due to time restrictions following the fall of the government							

Table 1 provides an overview of the seven evaluations of election platforms in the Netherlands. We report the number of political parties that participate and the number of seats in parliament these parties cover. Furthermore, we report the scope of the evaluation, using exactly the same categories as will be applied in our theoretical discussion in the next section.

Three conclusions can be drawn straight away. First, the number of parties that participate in the evaluation has steadily increased from three to nine. During the past 25 years, the political landscape in the Netherlands has become fragmented. More and more parties participate in the election and more and more parties gain access to the parliament. In



1968, the three major parties had 90% of the seats in Parliament; at present, this has been reduced to 55%. Despite this fragmentation, the number of seats in parliament covered in the evaluation has been rather stable and covers now even 99% of all seats. This is remarkable, since CPB only evaluates a program on request of a political party. Hence, parties can refuse to participate in the evaluation. Nevertheless, almost all parties choose to do so.

Apparently, the revelation principle is at work here. Given that the institution of the evaluation of election platforms has obtained a vested position in the Dutch political system, new parties find it indispensable to participate, since non-participation would inevitably send a signal to the voters that the party's election platform is economically unsound. Though the great majority of the voters will not spend a minute in reading the evaluation, many journalists do, and they report extensively on the evaluation in the newspapers and politicians refer repeatedly to the evaluation during their television performances. Moreover, the evaluation plays a major role in the negotiations on a new government that start immediately after the election. Not all parties participate in these negotiations, but a party of which the platform has not been evaluated would become less attractive as a coalition partner and will start the negotiation process with a backlog.

This revelation principle is well illustrated by the participation of the GreenLeft Party since 1994, the Socialist Party and SGP since 2002 and the PVV since 2010. The decision to participate by the GreenLeft Party and the Socialist Party was part of an official change in strategy. Their new strategy was to become regarded as a serious opposition party with a high quality program and to increase their chances for participating in a new coalition government. Following this strategy, it does not suffice to say no to policy proposals by the government. Each time, a realistic and financially sound alternative should be presented. Independent evaluation of their election platform fitted well in this new strategy. The SGP (a small right-wing Christian party) was very reluctant to participate, but felt forced to do it because a direct political opponent (ChristenUnie, another small Christian party) decided to join (see van Liedekerke, 2003, p. 142). Since 2005, there is a major new political party: PVV (an anti-Islam party). Last year, it also decided to request for evaluation of its election platform. This party has now become the official supporter of the current minority government. Its success in the elections and its important role in current Dutch government would have been unlikely without such evaluation of its election platform.

A second conclusion is that over time the description of policy proposals and their direct budgetary effects became much more detailed, extended and comparable. In the first decade, only an overview was provided of the major budget cuts, extra government

expenditure and tax measures. Since mid 1990s, also a more detailed description of policy measures is provided. Since 2002, changes in government expenditure or taxes are broken down by function, e.g. public administration, defense, education, health care, social security and transfers to corporations. As a consequence, the policy measures of political parties can now be compared by function.

A third conclusion that can be drawn from Table 2 is that the scope of the analysis has increased gradually. The core part of the Dutch evaluation of election platforms consist of four components, all pertaining to the next period of government<sup>9</sup>:

- Description of policy measures.
- Direct budgetary effects of the policy proposals, i.e. what are direct effects on the government deficit and debt in billion euro and as percentage of GDP without taking into account interactions with the national economy. In practice, this may include some behavioural effects, e.g. evasion of taxes or effects on the labour supply not incorporated in the macro-economic model.
- Analysis with a macro-model showing feedback effects for a core set of macro-variables, like economic growth, inflation, employment and the government budget. The direct budgetary effects of the policy proposals are used as inputs for this analysis.
- Analysis of the change in purchasing power of various groups of households. This includes the direct effects of policy proposals on household income plus the macro-economic effects on purchasing power, e.g. via changes in prices and wage rates.

Over the years, the core set of four components has been supplemented with other information. Since 2002, the analysis of direct budgetary effects has been extended with information on the employment effects for the government sector (broken down into e.g. central government, local government, safety, defense and education) and health care. A major purpose of many policy proposals is to establish long term effects that become only

---

<sup>9</sup> In 2010, due to the unexpected fall of government, time constraints implied that of this core-part of the evaluation only the description of proposals and their direct budgetary effects could be presented. Most of the political parties were not very happy with the absence of effects on purchasing power and medium-term macro-economic effects. For two parties, next time such absence of a politically very important part of effects of might be a reason not to participate.

visible after the next period of government. Long run labor market effects have been discussed since 1989. The scope of such long analyses has been extended each time. The analysis of the effects on environment, congestion and mobility was included since 1994.<sup>10</sup> Long term effects on government finance have been included since 2002. Since 2006, also the long run effects of proposals with respect to education and innovation are being assessed and last year an analysis of the housing market has been added. Last year, an innovation of the analysis of education and innovation was the link with long term economic growth. A specific feature of the housing market analysis and the analysis of mobility was that an estimate of welfare effects was included. Reform of health care was often a major issue in Dutch policy. Analyses of the major issues involved have therefore been included in 1989, 2002 and 2010. As a consequence of all these extensions, the evaluation provides a very broad overview of the consequences and trade-offs of the policy platforms of Dutch political parties.

**Table 2**      **The consequences of the election platforms of 2 Dutch political parties in 2010: Summary overview (changes relative to basis, unless otherwise mentioned; Dutch GDP in 2010 about 600 bln euro)**

	SP	VVD
Improvement EMU-balance 2015 (euro bn; ex ante)	10	20
Sustainability of public finances (euro bn)	16	39
Purchasing power of households 2015 (euro bn, ex ante)	+ 1¼	- 1½
Profits of companies 2015 (including housing corporations) (euro bn, ex ante)	- 4¼	- 1¼
Structural employment (% point)	- 1	5%
Accessibility by public or private transport (welfare gain, bn euro)	- ¼	¼
Car usage (2020, in %)	0	0
Public transport usage (2020, in %)	+ 5	0
Reduction of greenhouse gases (Mton Co2)	21	2
Quality of nature (2020)	+	--
Quantity of nature (2020)	0/-	--
GDP effect education (structural, %)	¼	4

<sup>10</sup> In 2006 this was not possible due to time constraints resulting from an unexpected fall of government. In 2002, three of the eight participating parties, i.e. 71 seats in parliament, opted out for such analysis. In the evaluation of 2010, such opting out was not allowed anymore: political parties were given the choice either to participate on all issues or not to participate (see section 4).

Science/innovation 2015 (budget, mld euro)	- 0.06	- 0.1
Housing market (welfare gains % gdp)	0.4	0.3
Change in house prices (2015, %)	- 6	- 2
Change in net rent (2015, %)	- 3	10
Health care, employment, 2015, thsnd )	+15	- 50
Own risk health insurance (2015, level, euro)	0	300
More (+) or less (-) market forces cure	--	++

This is illustrated by table 3, which presents a summary of the outcome of the evaluation in 2010 for the two parties with the most special election platforms, the Socialist Party (SP) and the Liberal Conservatives (VVD). The differences in the platforms are clearly visible. The Socialist Party has a much more lenient position regarding the reduction of the budget deficit than the Liberal Conservatives, both by the end of the election cycle in 2015 and in the long run sustainability gap. Obviously, cutting public expenditure more heavily as done by the Liberal Conservatives implies that purchasing power is reduced more strongly. The Socialist Party cares less about profits and more about the environment and Health Care than the Liberal Conservatives. Liberal Conservatives want to raise housing rents and private contributions in health care and want to introduce market forces in the organization of healthcare. The evaluation of the election platforms offers therefore a clear overview of two entirely different policy views.

There are several mechanisms underlying the gradual increase in scope (see also section 3). First, a partial evaluation has loopholes. Parties seek the weak points in the evaluation by making proposals of which the benefits are communicated clearly in the evaluation, but of which the cost fall outside the scope of the evaluation, in particular for budgetary cost beyond the end of the next election cycle. An attempt to cover these loopholes leads to a gradual trend towards an increasing scope of the evaluation. Second, political parties have diverging preferences regarding the topics on which they want to focus the election campaign. Green parties want to focus on environmental problems, conservative parties stress issues of law and order, liberal parties put most attention on education. Each party feels set at a backlog by not including (or: treating less extensively) their favorite topic. This yields a strong pressure to cover ever more special topics in the evaluation.

The evaluation of election platforms is fully embedded in the official decision-making process for the next period of government, i.e. in deciding on the new policy plans and new fiscal framework (see Bos, 2008 and Bos and Teulings, 2010). After the election, the

evaluation plays a major role during the formation of a new coalition agreement<sup>11</sup>. The evaluation offers an initial overview of the economic and financial implications of the parties' proposals. It is therefore a good starting point for negotiating the terms of a coalition agreement<sup>12</sup>. This applies not only to the proposals of parties involved in the coalition agreement. In practice, the evaluation serves as a data base on all kinds of policy measures that could be considered during the negotiations; in particular the budget cuts and extra revenue generating measures by other parties are a popular source of inspiration. On request of the political parties involved, CPB commonly provides also analyses of provisional and final coalition agreements.

Measuring the actual impact in the Netherlands of evaluating election platforms is difficult. Its impact should be assessed by comparison to a non-observable counterfactual, i.e. what would have happened without such evaluation? Would the political platforms have been different, would voters have voted differently and would coalition-agreements and actual policy practice have been different? At the start of the 1980s, already before the start of evaluating election platforms, major political parties agreed on drastic fiscal consolidation and restructuring of the Dutch economy. So, what would have been at that time the value added of evaluating election platforms, e.g. when CPB would have decided unilaterally to start such evaluation and publish the results before the elections? Similarly, what would have been the most recent political platforms and coalition agreement without any evaluation of the election platforms?

However, Dutch politicians agree that evaluating election platforms makes a substantial difference. For example, after the unexpected fall of the previous government, the most recent election date was set in such a way to leave just sufficient time for an evaluation of election platforms. According to Dutch politicians, the direct benefits for the general public at large must not be exaggerated (see Liedekerke, 2003). Few voters will read the evaluation report, but fall back upon media analysis of the report. But press reports magnify certain results, sometimes completely forget others that might be just as interesting and could

---

<sup>11</sup> In the Netherlands, parties usually form governments on the basis of wide-ranging coalition agreements. The coalition agreement plays an exceedingly important role during the government's term in office. It sets out the result of the give and take among the coalition partners on many policy issues.

<sup>12</sup> So, for this purpose, the overview of the policy proposals by political parties in the evaluation is preferred to the officially platforms published by the political parties themselves. The latter may differ in content and detail substantially from that used in the evaluation, see section 4.

even be seriously flawed. Political parties can play a role in this by framing and spinning the results.

Dutch political parties seem to agree that the major and direct beneficiaries of the evaluation are the political parties themselves. “The calculation effort pushes parties to be clear about their political programme and final options; it confronts them with hard choices that might be left unnoticed if not for the CPB involvement. In short: the calculation disciplines parties and precludes that wishful thinking turns into party politics. ... the content of the political programme cannot be made up of elusive promises.. it pushes the discussion within the political party itself to a higher level, in which people have to think twice before introducing a proposal. Wild politics is weeded out.... cross-party discussion will become smoother, because everybody is constrained by the same analytical framework and used to the same types of ever returning arguments (what will this proposal cost, how effective can it be, et cetera). Political discussion civilizes through the exposure to the calculation effort. ... [It] brings political programmes closer together and therefore simplifies the cabinet formation process” (Liedekerke, 2003, p. 138).

### **3. What is the proper scope of the evaluation?**

From a pure economic theory perspective, the platform of a political party only has to make statements about the preferred welfare distribution within and between generations, since these statements are normative. Conditional on this preferred distribution of income/welfare, the Pareto criterion allows ranking all alternatives according to their efficiency (since income is just one aspect of welfare, we refer to the distribution of welfare rather than income in what follows). This ranking belongs to the domain of positive analysis and can therefore in principle be done by scientists. Since society can be classified in large number of subgroups, each with different attributes and interests, even this relatively simple objective of a preferred distribution of welfare is a multidimensional problem. Since any change of institutions always affects both efficiency and the distribution of welfare and since instruments to redress the distributional effects of institutional change are usually missing, the debate on the appropriate distribution of welfare inevitably spills over into the debate on the efficiency. Hence, even when politicians focus strictly on the welfare distribution, the decision problem remains highly complicated.

However, reducing politics to the problem of a preferred welfare distribution constrains the domain of politics too much. Politicians have views on a much wider set of issues than just the welfare distribution. For example, whether or not healthcare should be privatized invokes a political debate that cannot be reduced to distributional issues alone. Ranking both sides of the debate on a welfare criterion would probably be a misperception of the political content of that debate. Alternatively, one could sketch the effects of various policy options and the tradeoffs that it involves which go beyond the traditional fundamental trade off, that between equity and efficiency. These observations raise the question how to set up an meaningful evaluation of election platforms. If constraining politics to a statement on the proper welfare distribution does not work, what else can be a defensible position about the scope of an economic/scientific evaluation of election platforms without the evaluation itself becoming a political statement?

Table 3 provides an overview of various topics that can be included in the evaluation. The topics are ranked in the order of an increasing scope: an evaluation should at least include topic A and can be gradually extended by including further options, starting from option B. The table provides a summary of the main arguments pro and contra the extension of the evaluation with that topic. Below, each option will be discussed in greater detail. In our discussion of the pros and cons of very alternatives we draw upon the Dutch experience. The

first evaluation in 1986 covered the topics A, B, and D. The last evaluation in 2010 covered all topics A till G, except for topic B and D, which were omitted for lack of time (the elections were held prematurely due to the collapse of the previous government).

**Table 3. A summary of pros and cons of various approaches to the evaluation of election platforms**

Topic	Advantage	Drawback
A. Direct budgetary effects during the next election cycle	<ol style="list-style-type: none"> <li>1. simplicity</li> <li>2. close alignment to the role of fiscal watchdog</li> </ol>	<ol style="list-style-type: none"> <li>1. no trade offs shown</li> <li>2. implications for deficit in later cycles ignored</li> </ol>
B. Economic feedbacks on budget, labour market and bbb during the next cycle	<ol style="list-style-type: none"> <li>1. shows medium run economic effects of policies</li> </ol>	<ol style="list-style-type: none"> <li>1. overrates expansionary policies and underrates structural reform</li> </ol>
C. Long term effects on labour market and bbb	<ol style="list-style-type: none"> <li>1. shows implications for a snapshot of the welfare distribution</li> <li>2. provides insight in the tradeoff political parties face</li> </ol>	<ol style="list-style-type: none"> <li>1. ignores changes in socio economic status</li> <li>2. ignores lifecycle effects</li> </ol>
D. Effects on purchasing power of various groups of households	<ol style="list-style-type: none"> <li>1. avoids the trap of overrating the short run effect on effective demand</li> </ol>	<ol style="list-style-type: none"> <li>1. puts a high burden of proof on the assessment of long run effects</li> </ol>
E. Long term effects on budget	<ol style="list-style-type: none"> <li>1. avoids burden shifting to future governments</li> </ol>	<ol style="list-style-type: none"> <li>1. raises a commitment issue: what is the value of early announcement of future policies?</li> <li>2. how to deal with issues where nobody has a serious plan regarding the future?</li> <li>3. if combined with alternative D: how to deal with variations in future growth rates?</li> <li>4. trade off between long term government finance and long term household income/profits is not shown</li> </ol>
F. An analysis of the impact on specific sectors, e.g. education	<ol style="list-style-type: none"> <li>1. provides a broader overview of the effect of platforms</li> <li>2. helps creating consensus on the economic impact of policies</li> </ol>	<ol style="list-style-type: none"> <li>1. choice of sectors to include is arbitrary</li> <li>2. requires detailed knowledge of these sectors</li> <li>3. not only of first order, but also of second order effects</li> <li>4. prior communication with political parties required</li> <li>5. risk of gaming</li> </ol>
G. An analysis of the impact on specific sectors using a welfare criterion	<ol style="list-style-type: none"> <li>1. allows an integral evaluation of costs and benefits of e.g. market distortions of taxes and subsidies</li> </ol>	<ol style="list-style-type: none"> <li>1. limits the role of political preference beyond what politicians view as their area of competence</li> <li>2. marginal utilities (prices) to tradeoff various inputs are not always available</li> </ol>



### **Topic A. Direct budgetary effect for the next election cycle**

The simplest evaluation is just an analysis of the effects of the proposed election platform on the government budget by the end of the next election cycle (each election cycle coincides with the term of a cabinet, which lasts 4 years, unless the cabinet loses confidence in parliament during the cycle). This position links directly to the role of fiscal watchdogs. Parties are evaluated by the effect of their program on the government budget. Anything else is left over to the marketing skills of the political parties and to the imagination of the voters. The advantage of this position is that its modesty is easily defensible. However, the evaluator does not provide the voter any help in grasping the implications of the proposed policies. More seriously, by focusing on the consequences for the budget by the end of the next election cycle, one runs the risk that political parties seek proposals that provide means for the next government and shift the burden for the budget to future governments. For example, parties have proposed to change the fiscal treatment of pension premiums. Currently, pension contributions are tax deductible, while benefits are taxed. Reversing that rule reduces the budget deficit at the end of the current election cycle, but raises it in the future.

### **Topic B. Economic feedbacks on budget, (un)employment, and GDP**

One further step is to include the indirect effects of the election platforms for the budget, the labour market, and GDP by the end of the election cycle. This provides information on the impact of an election platform on the economy. The caveat of including these economic feedbacks at the end of the election cycle is that they are likely to be dominated by medium run effects of policies on effective demand. For example, suppose that the budget deficit and unemployment have gone up during the past election cycle due to a recession. Suppose one party wants to counter these adverse effects by reducing the replacement rates. In the short run this reduces effective demand due to the reduction in purchasing power for the unemployed. The positive effect of lower benefits on the budget deficit might therefore be offset by lower consumption taxes. In the long run the lower benefits and the effect of a lower replacement rate lead to a reduction of the deficit. Similarly, expenditure cuts might reduce the deficit, but raise unemployment due to their effect on effective demand. Since these cuts have to be made anyway to keep the budget balanced, this is merely an issue of timing. As long as the evaluation does not provide the voter a shadow price of a lower deficit by the end of the election in terms of future GDP and future (un)employment, the voter will find it hard to weight short run benefits against long run cost. Establishing this shadow price is therefore a major challenge for economic theory, which has not been solved till today.

Summarizing, structural policies proposals are underrated by including only an evaluation of the economic impact during the next election cycle, since the evaluation is dominated by medium run effects on effective demand, while the structural effects tend to take longer before they are fully realized. For example, according to the evaluation of the election platforms of 2006, the greatest difference in unemployment rate between two parties was 0.4 %-point. One party proposed a reduction of the replacement rate by 2.5% point. The main positive effect of this measure on labor supply and GDP is realized only after the end of the next election cycle. For this reason, some observers/economists prefer leaving out the medium run economic effects and concentrating on the structural effects.

### **Topic C. Long term effects on (un)employment and GDP**

An obvious solution to the problem of overrating the medium term effects of policies on effective demand is to include an analysis of their structural or long terms effects in the evaluation. However, this puts a large burden of proof on the estimates of long run equilibrium effects of policies. Moreover, the public find it a hard to believe these long run effects, while the short run effects on effective demand have larger credibility in the eyes of the public. From an insider point of view, these judgments are highly debatable. For example, recent studies of the CPB on the effects tax reforms on labour supply reveal that the estimates of their effects are highly reliable, while the effects on effective demand are might be far more debatable. These conclusions are in line with similar evidence of the Institute of Fiscal Studies for the United Kingdom.

### **Topic D. Effects on purchasing power of various groups of households**

One step further is to include the effect of policy proposals on the purchasing power of various subgroups in society during the next period of government. From the point of view that politics most important role is to decide on the welfare distribution, this is obviously meaningful information for the voter. The combination of this topic and topic D. gives a handle on the tradeoff between equity and efficiency.

However, there is a major caveat here. The tradeoff between equity and efficiency stems from policy makers' inability to distinguish between effort and ability. Hence, redistributive taxation undermines the incentives for providing effort, or equivalently, it reduces labour supply. A positive evaluation of this trade off requires a broadly shared view on the elasticity of labor supply and the effect of the replacement rate on unemployment. Such an agreement might be hard to achieve.

As argued by Piketty (1995), political parties might be deeply divided on this issue for perfectly understandable reasons. Effort, social background, and income are positively correlated. Left wing parties might explain these positive correlations by arguing that income is determined social background determines income, while right wing parties might stress the role of effort. Since left wing people usually have a lower social background than right wing people, they provide less effort since they think it does not matter anyway, while right wing people think the opposite and therefore provide a lot of effort. These beliefs and the behaviour they induce generate exactly the correlations that we find in the data. Therefore, an econometrician might have a hard time to provide convincing evidence on the benefit of either side. This shows why the classical distinction between positive statements on efficiency and normative statement on the distribution of welfare is not that clear cut in practice.

Remarkably, a generally accepted view on this issue has emerged in the Netherlands. The CPB has modeled the labor market in its MIMIC model (see Gelauff and Graafland, 1994, Donders and Graafland, 2000 and Folmer, 2009). This exercise has generated a view on the relevant elasticities that is broadly accepted by all political parties.

Focusing on the direct effect of policies for purchasing power keeping constant the socio economic status of a person ignores an important aspect of the tradeoff between equity and efficiency. Reducing the replacement rate lowers the purchasing power of an unemployed, but raises the probability for that person to get reemployed. By keeping constant the socio economic status, this aspect is ignored. From the point of view of individual households (see Di Tella et.al. 2001 and 2003), losing one's job has much larger implications for well being than slight policy changes regarding tax brackets or replacement rates. Ignoring these implications overstates the negative effects of this type of policies.

A further drawback of this analysis of purchasing power is that it focuses on the current status and ignores future effects. This is a serious limitation when analyzing for example the intergenerational impact on lifetime welfare of increasing the retirement age.

### **Topic E. Long term effects on budget**

When an economy is approximately in a steady state, there is little need for a separate analysis of the effect of election platforms on future government budgets. As long as a policy leads to a balanced budget today, it will also do so in the future. However, the economies of all OECD countries are not at all in steady state. The ageing of society is a major risk for the long run sustainability of the public finances. Any policy increasing public pensions or health

care for elderly is currently affordable, but might be a nightmare in the future. Hence, it is useful to include a long term perspective in the evaluation of election platforms. This approach is particularly useful for an evaluation of the level of public debt at the end of the next election cycle. As noted before, the economic discipline has not agreed on a proper shadow price for public debt. In a long term framework, the intertemporal budget constraint solves this issue. The requirement not to let public debt explode provides a meaningful constraint.

Inspired by generational accounting developed by Auerbach, Gokhale and Kotlikoff (1991), CPB has worked out a set of rules for the long term evaluation of public finance. We summarize the main lines below<sup>13</sup>. The growth of labour productivity is assumed to be exogenous and a series of policy parameters are fixed at their current level, like the ratio of the public elderly pension to wages, health care expenditure per person of a particular age divided by the real wage, and tax revenues and public consumption, the latter two as a share of GDP. Using the expected evolution of the demographic composition of the population (accounting for the expected increase in life expectancy), we can calculate the evolution of the primary surplus and public debt. It would be accidental if the debt ratio would converge to a stable path. The expenditure cut required to let the debt ratio converge is called the sustainability gap. The platforms of political parties are evaluated by their effect on the sustainability gap and the distribution of purchasing power across generations. When a party submits a platform that does not close the gap, the CPB arbitrarily closes the gap, e.g. in 2040, by raising the tax rate. In that case, future generations will bear a disproportional share of the burden of implicit public debt. This provides an objective shadow price for public debt: the purchasing power of future generations<sup>14</sup>.

Although this approach is quite effective in dealing with the long term budget constraint of the government, this approach faces a number of hairy problems. In the simple set up of scenario A that started the whole exercise, the CPB allowed parties to submit proposals that could be implemented during the next election cycle and evaluated these proposals solely by their effect on the deficit by the end of that cycle. Parties were not allowed to submit proposals that had an effect only after the end of the next election. The

---

<sup>13</sup> For a more detailed description, see Draper and Armstrong (2009) and Horst et al. (2010).

<sup>14</sup> The interest rate could also be regarded as a shadow price of public debt. However, business cycle effects, specific market circumstances and central bank policies to keep interest rates close to zero in order to stimulate economic growth can invalidate the interest rate as a shadow price of public debt.

sustainability analysis opened the way for proposals that were relevant only for their effects on the long run. But what is the credibility of a proposal that is only going to have real effects 20 years from now? Is such a proposal credible in the first place? Fundamentally, politics is not able to commit itself, because there is no outside power that can force politics to live up to its promises, see Acemoglu (2003). Moreover, the politicians that rule 20 years from now are different from those who rule today. Why would future politicians consider themselves to be committed to the proposals done by their predecessors? Hence, the CPB has imposed a number of constraints for a long term policy proposal to be included in the evaluation of election platforms for the current election cycle. We return to this issue in section 4 when discussing the rules of the game.

Further complications arise when these long term projections for the budget are combined with topic D., the long term effects of policies on (un)employment and GDP. Suppose a party wants to raise future GDP by investing in education. Since many policy parameters are expressed as a share of GDP (or: real wages), an increase in productivity does not have as beneficial an effect on sustainability, since an increase in productivity raises expenditure one-for-one. Here, the stylized representation of policies as a fixed share of GDP, without taking into account the fact that a higher GDP implies more of the policy being available, disrupts the analysis. When these issues arise special attention is required to avoid bizarre conclusions.

### **Topic F. Including extended analysis of specific sectors**

A further option is to include the analysis of the economic effects of policy proposals on specific sectors, and whenever possible also the spillover of these sectors to GDP and (un)employment. The problem is that the choice of what sectors to include is arbitrary. Sectors with a large public involvement are obviously the first candidates. However, the last evaluation of the CPB included education, highways/road pricing, and environment, but did not include the effects of changes in the police force and confined the analysis of health care mostly to summarizing the direct budgetary effects. Policy proposals with respect to education were classified into three groups: promising, not promising or neutral. A proposal is promising if its social benefits exceed the social costs. Proposals for which not sufficient empirical studies are available are put in the group ‘Effects not known’. Also a link was made with long term economic growth. Policy proposals regarded as promising or not promising are then quantified in terms of their effect on long term GDP volume growth.

In practice, three criteria determine whether or not or how a sector is included:

- An intensive public debate on a sector raises the likelihood of it being included.
- Following up on the first criterion, when many political parties have announced proposals for that sector, a sector is likely to be included.
- Finally, a sector can only be included when economic science has a well established body of knowledge about it.

The latter criterion is quite important in practice. For example, the economic discipline has spent enormous resources to the analysis of education. This research has provided in great deal of broadly shared insights into the impact of education on the economy. Health care plays a larger role in the current political debate than education, not the least due to the ageing of the society. However, there is no generally accepted body of knowledge on the effect of competition policies for the health care system. Hence, a serious analysis of this issue in an evaluation of the election programs is hazardous. The evaluator can hardly avoid taking a political stance.

Here, the second and the third criterion contradict. Political parties demand the CPB to include healthcare in its evaluation of election platforms, because only by including health care, parties can credibly communicate their policy proposals to the voters. Since health care is at the political front line, it is quite understandable that political parties express this preference. At the same time, the CPB must restrain its judgments on this issue since there is no shared body of knowledge.

A serious complication in the analysis of specific sectors is that one needs to know not only the first order effect of a particular proposal, but also the second order effect. The reason for this necessity is simple. Suppose that a particular proposal has a positive net discounted value per euro spend and can be scaled up to any level. A typical example is raising the skill level of the workforce. Calculations by Hanushek and Woessman (2010) have shown that the net discounted value of investments in human capital is enormous. How should a proposal by a political party to double the expenditure on education be evaluated? Or to triple it, for that matter? Clearly, there must be a limit beyond which further investments no longer have a positive discounted value, the standard economic concept of diminishing returns. But what is that limit? As another example, most research suggests that reducing class size has a negative net discounted value, since it is costly and not very effective (see e.g. Dobbelsteen et al. 2002 and Woessman and West, 2006). Reversing the argument, increasing class size must have a positive discounted value. What to do when a political party proposes to raise class size to a 1000 pupils? In most cases, the discipline

spends a decade to decide on the magnitude of first order effect. Establishing the second order effect is likely to take at least another decade, see Teulings and Van Rens (2008) for a first attempt for investing in human capital.

There is a further reason for worry. If the argument regarding the importance of second order effects is correct, then using cross country evidence on the first order effect is inappropriate for the evaluation of the first order effect in particular country. The magnitude of the first order effect of an investment in education depends crucially on the initial state of a country's education system. The effect of an improvement is likely to be much higher if the system is in a state of disarray than if the system is already functioning perfectly.

Ultimately, the issue is about the proper role of economists. Is their role to give a best judgment on what is a fruitful direction for policy makers to go, without claiming to have an idea about the optimum, or is their role to classify political parties on how close their proposals correspond to a hypothetical optimum? Given uncertainty on the first question, some restraint in answering the second might be appropriate. Remarkably, political parties press for answering the second, as this is a way for them to communicate the sincerity of their proposals to the voters. For example, the unwillingness of the CPB to assess the effect of a more market oriented organization of health care lead to an uproar among some political parties who made proposals in this area.

The positive way to view this process is that it is a reflection of the gradual emergence of a political consensus on what is an efficient policy. Similar to the way all political parties use the results of the Mimic model as a "true" representation of the tradeoff between equity and efficiency with regard to marginal tax rates, the pressure to include an economic analysis of the education system in the evaluation of election platforms can be interpreted as a process towards a shared view on the effect of education on the economy. This makes clear that the evaluation of election platforms along these lines is only feasible if there is public/scientific debate on the evaluation standards to be applied.

A final issue regarding the evaluation of specific sectors is the question whether political parties should be informed a priori about the evaluation-methods that will be applied. Not informing parties a priori makes the evaluation like a gambling game, where political parties have to guess about the methodology and hence the type of proposals that "score" best. In this way, the evaluation process is unlikely to contribute to sound economic policy. Moreover, the evaluation is allegedly based on scientific and hence reproducible knowledge. It is hard to square the presumption of reproducibility with not informing parties a priori about the methodology that is going to be applied.

However, prior information will induce parties to try to game the system, finding loopholes in the announced methodology that yield an artificially positive evaluation. An escape clause for this type of practices is a minimum requirement for the credibility of the evaluation. However, there is a further complication. Making available the evaluation methodology implies that all parties have equal access to the effects attributed by the evaluator to all kind of policy proposal. What to do if a particular party comes up with a unique proposal not considered by other parties. Should other parties be informed about this proposal and the way in which it is evaluated? Or should this proposal be excluded from the evaluation on the grounds that the methodology for the evaluation of this proposal was not common knowledge among all parties? Stated differently: is a politician only involved in making normative choices on the proper distribution of welfare distribution, or is he also an entrepreneur who gets rewarded on the electoral market for coming up with Pareto improvements? In practice, the CPB has chosen to provide prior information on the methodology, but to allow political parties an advantage who come up with Pareto improvements themselves by not revealing these proposals to other parties.

### **Topic G. Extended with a welfare criterion**

The evaluation under topic F. is restricted to some sector specific outcome, like health status (for health care), congestion (for infrastructure), or skill measures (for the education system) and effects on GDP and (un)employment. One can generalize the evaluation by using welfare criterion instead. The Hicks-Kaldor criterion -just adding up the monetary value of all relevant aspects, whether traded or not- is most convenient. Moreover, it is the only criterion that has an objective legitimate provided offsetting transfers can be implemented or that there are no a priori reasons that the existing welfare distribution is better justifiable than the alternative distribution. The latter might apply for institutional reforms in small sectors, where small groups of insiders capture large rents. The advantage of using a welfare criterion is that it provides an easy way of aggregating various aspects of people's well being into a single statistic, in particular aspects that tend to be ignored when taking GDP as a criterion. Leisure is an obvious example. An evaluation of proposals according to their effect on GDP implicitly sets the value of a change in leisure equal to zero. A welfare criterion uses the net wage rate as the valuation/price of a change leisure. Using welfare allows a positive integral evaluation of all aspects that are shown to be relevant by applying people's revealed preference for each of these aspects.



However, the logic of the valuation of different aspects of well being by market prices is not easily communicated to the wider public. The paradox of the public perception of economics is that it blames economics for focusing on monetary GDP only and ignoring other aspects of well being. However, at the same time it views the standard approach of economics for including these aspects as an inappropriate intervention in the domain of other social sciences. Economists are imprisoned in the cave of the concept of GDP while at the same time being accused of not willing to leave the cave. It is reflected in the conviction of politicians that increasing hours worked is economically beneficial, irrespective of the value of a reduction of leisure. Though unsatisfactory from an economist' point of view, this is the way it is. One interpretation is that we allow consumers to decide in the tradeoff between for example butter and milk on the basis of market prices, but that we want politicians to decide in the tradeoff between leisure and other consumption, based on politicians' rather than consumers' relative valuations of leisure versus other consumption. One might wonder why, but that being the case, the only option for an evaluation of election platforms is to provide separate information on the effects on GDP and leisure, and leave aggregation of both aspects to the voters/politicians.

In practice, using the welfare concept is therefore appropriate only in small number of special cases. The CPB has applied the concept when evaluating proposals to the reform of the housing market. This market is heavily distorted, both its rental and its private ownership segment. The CPB designed a method to evaluate the cost of these distortions, using the concept of a Harberger triangle (see Hines jr., 1999). In this way, reforms could be evaluated both on their effect on "aggregate" welfare as on the welfare distribution. The welfare concept is the only way to give voters a handle on the size of the distortions implied by prevailing institutions.

#### 4. Which rules of the game?

The evaluation of elections platforms can be regarded as a game in which political parties compete for maximizing the number of votes for their party while respecting their fundamental political preferences. The benefits of evaluating election platforms depend critically on the rules of this game. For example, it is important that the rules give the right incentives to political parties: they should not try to manipulate the game, they should reveal their real preferences and should not adjust their policy proposals for the wrong reasons, e.g. due to an unbalanced or inaccurate presentation of trade offs and effects.

Rules used for evaluating election platforms serve different purposes. Three different purposes of such rules can be distinguished:

- Independence of the political process;
- Good communication between the political parties and the evaluator;
- Good quality of the evaluation and limited room for gaming and free lunches.

#### **Table 4. Rules for independence of the political process.**

---

1. Election platforms are only evaluated on request of the political party involved.
  2. All political parties (likely to be) represented in the parliament can participate.
  3. All political parties are treated as equally important.
  4. Election platforms are evaluated as if the party is the government and has all seats in Parliament.
  5. The evaluator informs the parties in advance about the time schedule, rules, topics and presentation.
  6. Political parties do not get information about the policy proposals made by other political parties.
  7. The evaluator does not communicate with the press about first results.
  8. Policy proposals and results are presented in a uniform way; the text is descriptive and neutral.
  9. In text and tables, parties are ranked by the current number of seats in Parliament.
- 

Tables 4, 5 and 6 provide for each of these purposes an overview of the rules used in the Netherlands.

An election platform is only evaluated on request of political parties (rule 1 in Table 4). What are the advantages and drawback of this rule in comparison to the alternative of obligatory participation? The rule of voluntary participation has two drawbacks. Firstly, evaluation will only occur on requested by a political party, even when political and

economic circumstances indicate major benefits of such analysis. Secondly, when not all political parties request an evaluation, voters will get an incomplete set of information.

Nevertheless, this Dutch rule has some clear merits. It avoids a conflict with a political party that does not want its political platform being analyzed. Ensuring the cooperation of the political party is also important for a proper interpretation of the policy proposals. Furthermore, the number of parties participating in the analysis of election platforms has gradually increased to nearly all political parties. As a consequence, the evaluation now compares the election platforms of nearly all political parties.

A major reason for this seems to be the revelation principle (see also Lecq, 2003). The parties with the 'best' and most solid and honest economic plans have an incentive to participate. But then not participating provides a negative signal to the voters. The revelation principle will also apply in case of obligatory participation, i.e. when an expert institute decides unilaterally to start evaluating election platforms irrespective of whether the political party consents or not. Knowing that their election platform will be analyzed anyhow, political parties may decide to cooperate and provide extra details on their plans in order to avoid negative signaling and publicity.

A major drawback of obligatory participation is that parties do not want to cooperate and do not want to clarify and specify their policy proposals. This limits the quality, detail and scope of the evaluation. However, also for obligatory participation, the revelation principle may work in the longer run and induce political parties to cooperate and provide more information.

During the evaluation process, political parties are not allowed to see the policy plans of the other political parties (rule 6 in table 4). Similarly, they are informed about the draft and final results of the effects of their policy plans, but they are not informed about those of the other parties. This information is only disclosed after official publication of the evaluation of election platforms. Such information may then contain surprises. For example, some party may have innovative proposals that would have also suited their party. It may also turn out that they cut/spend less on a specific policy theme than another party. This may conflict with their political profile, e.g. being the party that is most environmental friendly, champion for education, best for the poor income, best for realizing a smaller government or solid public finance.

**Table 5. Rules for good communication between political parties and the evaluator**

---

1. Policy proposals send to the evaluator are regarded as the election platform.
  2. Statements by political parties in the press are not the responsibility of the evaluator.
  3. The evaluator is transparent about the methodology to be used.
  4. The evaluator publishes the baseline projection before the analysis of election platforms.
  5. Political parties can change their policy proposals during the game.
  6. Political parties can ask the evaluator for advice, e.g. how to meet their targets in alternative ways.
  7. Political parties can put forward text proposals for the description of their policy proposals.
- 

The proposals send to the evaluator are evaluated and not the election platforms officially published by political parties (rule 1 in table 5). This is done for two reasons.

The first reason is that official election platforms do not contain sufficiently clear and well-specified information about the policy proposals. Official election platforms are mostly qualitative, focused on convincing potential voters and without much specification of policy measures proposed. A frequent annex of such official election platforms is a simple budgetary overview. What is needed for the evaluation can be regarded as extended and more detailed version of such a budgetary overview. For example, the simple budgetary overview may show that subsidies are to be cut by 1 billion euro. But in order to assess whether this is practically and legally feasible and to be able to say something about its consequences more specification is needed: which subsidies are to be cut by which amount?

The second reason is the interactive nature of the evaluation (see rules 4 and 5 in table 5). When the evaluation is published some months after official election platform, this interaction is likely to have resulted in some changes in a party's policy proposals. The analysis of election platforms has therefore much similarities with a mix of topdown- and bottom-up budgeting (see Ljungman, 2009): some general targets in terms of government deficit and debt or other variables like the purchasing power of various groups of households are usually defined at the start and these are then made consistent with the initial set of specific policy proposals after one or more rounds of negotiations and deliberations. This may result in changing or deleting such specific policy proposals or adding new ones, but it may also imply that the level of ambition in terms of general targets is adjusted.

Advantages of this interaction between political parties and the evaluator are that the policy proposals become more realistic, detailed and effective in reflecting and meeting the party's economic and political preferences. Drawbacks are that the evaluation takes more

time and resources and may also allow more room for strategic and misleading behaviour by political parties.

More in general, evaluation of election platforms could be regarded as two way interaction between policy makers and economists<sup>15</sup>. First, it gives economists an opportunity to inform politicians. This is well appreciated by Dutch politicians: “The evaluation of election platforms is one of the most pleasant of our jobs, a real highlight. We have a special team put together for this. You learn a lot, for example how to translate general policy ideas into specific policy proposals. You get a good notion of the major policy tools for the various policy targets.” “We often use the booklet. It is well written. The overview comparing the policy platforms does not always provide good news, but gives a fair view of the choices made by the different political parties.”<sup>16</sup>

Secondly, the evaluation of election platforms gives politicians also an opportunity to inform economists, e.g. about their political preferences and the issues and trade-offs they are interested in. The evaluation of election platforms can therefore provide important feedback for the economic expert institute involved: by analyzing a broad range of policy proposals from (many) different political parties, the relevance and quality of models, knowledge and skills are tested. Economists involved in the evaluation are also generally very positive: “It is one of the most exciting and interesting jobs for a young economist at CPB and well worth the many extra working hours.” “Good for CPB and good for the country.”

The evaluator is transparent about the methodology to be used (rule 3 in table 4.2). This may include separate publications on the models used, literature surveys on the effectiveness of various type of policy measures (e.g. with respect to education) or studies on the efficiency of national institutions with respect a policy area (e.g. housing market, education, health care or social security).

The evaluator publishes the baseline projection before the evaluation of election platforms (rule 4 in table 5). Publications on the baseline may also include statements on the consequences of specific policy proposals. For example, a study on the sustainability of public finance may include also an overview of major policy proposals that could be considered to improve sustainability.

---

<sup>15</sup> A similar conclusion but about empirical models and policy making was drawn by Butter and Morgan (2000).

<sup>16</sup> Statement during evaluation of last year’s evaluation of election platforms.

Political parties can put forward text proposals for the description of their policy plans (rule 6 in table 5). However, the text should be neutral and descriptive and should not contain all kinds of unwarranted marketing statements.

Political parties are not allowed to opt out for one or more issues (rule 1 in table 6); they could only decide not to participate at all or participate on all issues. One political party did not want an analysis of the consequences for the environment, but decided nevertheless to participate. This rule of not allowing opting out for some issues seems to contradict the rule that political parties are free to decide whether to participate.

Analyzing election platforms is quite different from forecasting. Macro-economic forecasts by CPB and other institutes are indeed not very accurate. The uncertainty of the baseline projection as such is not a problem provided it is not (politically) biased and people are sufficiently aware of the uncertainty of the projection. Furthermore, for analyzing and comparing election platforms the same baseline projection is used for all political parties, e.g. on the macro-economic development and the sustainability of public finance.

---

**Table 6. Rules for quality and objectivity of the evaluation**

---

1. Political parties cannot opt out for one or more topics.
  2. The baseline projection is a neutral extrapolation based on the assumption of unchanged policy.
  3. Only new policy proposals are included; this excludes policy in the baseline projection.
  4. The same methodology is used for evaluating the election platforms of all parties.
  5. Only policy proposals are included that are sufficiently clear and well-specified.
  6. Only policy proposals that can be made (unilaterally) by central government are included.
  7. Policy proposals should be legally and practically feasible during the next period of government.
  8. Policy measures of which the effects cannot be assessed sufficiently reliably are not included.
  9. Policy proposals should have real effects during the next period of government.
- 

However, the baseline is very important for the framing of policy proposals, e.g. because politicians and voters are myopic and loss-averse (see Tversky and Kahneman, 1986 and Kahneman, 2003). The baseline used is a neutral extrapolation based on unchanged policy. But different macro-economic assumptions influence the perceptions of the sustainability of public finance and the development of real income of households.

Also the definition of unchanged policy is very important for such perceptions. What is unchanged policy, e.g. for taxes, social benefits and expenditure on education,

infrastructure and health care? Should a strict legalistic approach be taken or should e.g. the developments during last 5 or 10 years corrected for policy changes and demographic changes be extrapolated? Policy included in the baseline is by definition not included in the evaluation of election platforms. As a consequence, depending on the baseline, a policy proposal can be included in the evaluation or not. Promises by political parties in their election platforms about not raising taxes or guaranteeing the real income of poor households will have a different meaning depending on the baseline. Different assumptions about unchanged policy can also sketch a rather different picture of the problems to be solved by the next government. For example, according to a strict legalistic approach, public finance may be sustainable, but according to a more economic and plausible approach there may be serious budgetary problems to be solved.

The baseline by the evaluator used for the next period of government (see CPB, 2010c) is to a substantial extent legalistic, but is in several respects also quite different from a purely legalistic approach. A major example is health care. The total public and private expenditure on health care are expected to increase by 4% per year in real terms, i.e. corrected for the general price change of GDP. This is more than could be expected due to only economic growth (1 3/4%) and ageing. It is assumed that public expenditure on health care for each age cohort increase in line with economic growth and that the remaining 1% increase per year is financed privately. This would mean a drastic increase of private expenditure on health care that would affect the real income of households substantially. In their election platforms, politicians can decide whether to agree with the assumption of a drastic increase in private health care expenditure, or whether to take supplementary measures, e.g. further increasing social security contributions for health care or find ways to reduce the rise of total health care expenditure, e.g. by efficiency gains.

Similarly, the baseline for the long term calculations on sustainable public finance assumes constant arrangements, i.e. the same quality of social benefits and public services for the same level of tax-rates (see van der Horst et al, 2010, p. 15-17). This assumption is also used by similar studies by the European Commission and OECD. What does this assumption mean? Individual public expenditure, like old age benefits, unemployment benefits and expenditure on health care and education, the expenditure per person/pupil (by age cohort) are linked to the general increase in wage rates. Collective public expenditure, like that for defense, infrastructure, subsidies to corporations and public administration, are linked to the development of GDP and are therefore assumed to stay constant as a percentage of GDP. Taxes are assumed to remain constant as a percentage of the tax base, e.g. income for the

income tax and private consumption expenditure for VAT and excise duties. This is clearly not a legalistic approach. For example, according to Dutch law the major tax bracket for income tax should increase in line with price change and not in line with wage rates. A legalistic approach would mean that in the long run all households become subject to the highest tax rate of 52%. This would solve all problems of the sustainability of Dutch public finance, but would not provide a realistic picture of the future.

Widening the scope of analysis is often important to provide a more balanced picture of the effects and trade-offs of policy proposals (see section 3). It also helps to avoid free lunches, i.e. policy proposals which seem to have only benefits and no drawbacks. However, in order to avoid free lunches, specific rules are needed to decide which policy proposals should not be included in the evaluation at all. The latter implies to serve no lunch at all for the political party.

Some policy measures amount simply to double-counting, e.g. policy measures already included in the baseline scenario. For example, the future revenues of natural gas and the financial assets of the social security funds are already included in the baseline projection on the government budget. As a consequence, proposals to use future revenue of natural gas or the assets of social security funds to improve the government budget, to reduce tax rates or to finance extra expenditure are not accepted.

Some policy measures are no policy measures, as they are just an alternative estimate of some revenue or expenditure in the baseline.

Other policy measures only amount to a re-arrangement of the financial portfolio of the government or a re-arrangement of revenue and expenditure between various parts of the general government should also be ignored. For example, the sale of offices and leasing them back. Introducing transfers or financial transactions between various parts of general government without changing the overall budget deficit and net worth will also be ignored, e.g. transferring the substantial financial assets of Dutch provinces to the Dutch central government.

Some policy measures are presented as ‘magical solutions’ for improving government finance. Tanzi’s chronicle of the bankruptcy of Argentina (2007) gives some beautiful examples, e.g. tax revenues would be boosted by privatization of tax collection, more sophisticated computers for tax collection or the introduction of a single tax on all transactions while abolishing all other taxes. In general, it is wise to be very skeptic to such magical solutions, to ignore them in scoring and to motivate why it is not only uncertain but also very unlikely that such proposals will solve any of the budgetary problems.



Several specific rules serve also as a filter for accepting policy proposals. A first rule is that policy proposals should be specified sufficiently. For example, a proposal to improve the labour market position of the young, the elderly or the low-educated should be specified further. Not only the amount of money available should be known, but also the design of the specific policy proposals, e.g. via more schooling, less social benefit or a tax credit. Without specification, the efficiency of these proposals and the distributional consequences cannot be assessed.

A second rule is that policy proposals should be (unilaterally) subject to decision-making by the central government. For example, the central government cannot decide how local government, the European Commission or private social housing corporations should spend their money. The government has also a limited influence on agreements between employers and trade unions, e.g. on wage moderation and pension contracts. Only decisions that can be made by the central government are included in the evaluation, e.g. cutting general or specific transfers to local government or changing the tax treatment of pensions. Contributions to the EU cannot be reduced unilaterally by the government, as this is the subject of negotiations at European level. Similarly, wage rates of civil servants cannot be reduced unilaterally, as this is the subject of negotiations with the trade unions. The outcome of such negotiations depends critically on the labour market.

A third rule is that the policy proposal should be legally and practically feasible during the next period of government. For example, abolishing provinces requires a change in the constitution and can therefore not be realized in one period of government. Similarly, European laws and international agreements on human rights drastically limit the possibilities to further tighten asylum policy.

A fourth rule is that policy proposals whose consequences cannot be assessed reliably are ignored. For example, the economic effects of major reform of the institutions in health care –introducing a free market for hospitals: allowing the free entry of privately funded hospitals and allowing loss making hospitals to go bankrupt- are hard to assess.

The sustainability analysis opened the way for proposals that were relevant only for their effects on the long run. Hence, the evaluator has imposed a number of constraints for the long term effects of a policy to be included in the sustainability analysis: First, the proposal must have real effects during the current election cycle. Second, the proposal must be logically defensible. E.g. a proposal to raise the retirement age by 1 day by the end of the election cycle, and by 5 years in some 20 years from now, is not viewed as logical proposal. Obviously, the 1 day increase is only included to meet the first requirement. Third, we cut off

the effect of gradual changes by 2040, to avoid proposals with effects that are quantitatively important only after 2040. One party proposed to cap mortgage deductibility at 1 million euro and not to index this cap forever. Practically nobody has a house above this cap today, but without indexing that will be totally different 40 years from now. Hence, the big revenues come in the far future. Finally, the evaluator is very reluctant to include proposals that affect tax rates and the like, because tax rates are typically decided upon in a yearly policy cycle. Claiming that you raise the tax rate in 10 years from now is therefore non-credible. To the contrary, raising the retirement age by one month a year over the next 24 years is credible. Societies do not decide on the rules for retirement every year. Hence, such a proposal is credible.

The latter constraint introduces a distinction between institutions with and without commitment value. True as this distinction between credible and non-credible proposals may be, it introduces a large degree of discretion on the side of the evaluator. This is undesirable, since the evaluator can easily be accused of being a politician instead of evaluating political platforms. However, this type of judgments is unavoidable if one wants to include an analysis of the long term effects of election platforms in their evaluation. Indeed, experience shows that political parties seek the boundaries of the rules outline before. They seek proposals that minimize the impact on purchasing power next election cycle, but that maximize the impact on the sustainability gap. Clearly, these rules are a binding constraint.

Most policy proposals are not free lunches but involve trade-offs or effects that may be difficult to quantify. Two examples can illustrate how the evaluator then nevertheless tries to come up with pragmatic solutions. This can be done quantitatively, but could also be solved by a qualitative analysis or restatement of the proposal or in the description of the proposal in the evaluation. Sometimes, it may also be necessary to introduce an additional rule, e.g. a maximum on the budget cut for civil servants.

A very common proposal is to reduce fraud with taxes and social security benefits by 'better inspection and detection methods'. Without any specification of the difference with current methods to detect and reduce fraud, no savings are recorded. Also, such new or more intensive methods generally first cost extra money. A practical compromise often used is that the political party 'invests' some money (say 300 mln euro) for improving or extending detection methods and that this leads to a saving of exactly the same amount of money by reducing tax and social benefit fraud.

Our second example refers to proposals for reducing the number of civil servants. This often seems to be a free lunch, as possible negative effects of such budget cuts on

quality and quantity of the services provided are difficult to assess in advance. In case of substantial cuts on the tax office, reduction of tax revenue may even seem to be likely, but how to estimate then by how much? However, when political parties have to specify their proposal, these proposals become much less a free lunch. For example, reducing the number of civil servants of the central government by 20% would imply that all major units are cut by this percentage. However, 25% of the number of civil servants of the central government is the tax office, 30% consist of police, prison and administration of justice and 8% work on road maintenance or planning new infrastructure. Political parties are generally not willing to make major cuts in these units of central government; such cuts become also visible in overview table on budget cuts, e.g. in the functions public administration, safety and infrastructure and in the table on employment effects for the government sector. Many of the proposals are often also overlapping, e.g. proposals for a general efficiency cut, a reduction of overhead, a reduction of the purchase of consultancy services, some years no compensation for inflation and many additional cuts for specific units of government. As a consequence, by having to specify their plans they usually also substantially modify and reduce their proposed budget cuts. In the description of such proposals, the evaluator makes also explicit that these budget cuts are generally not or only to a limited extent increases in efficiency.

A new situation occurred last year: in the baseline scenario already substantial budget cuts were included for the central and local government and nearly all political parties wanted on top of that very substantial extra budget cuts. Making the consequences of their proposals explicit via tables and text was not sufficient anymore. In order to keep the proposals realistic for only one period of government, maxima for budget cuts -on top of what was already in the baseline scenario- had to be set by the evaluator. These maxima were partly inspired by just released government reports discussing various alternatives for fundamental reform of the government budget. For central government, the maximum was 9% for some parts and 6% for the rest. For local government, 20% reduction of the general transfer to provinces was accepted and 10% of those to municipalities. On some other parts, more budget cuts were allowed, e.g. for defense a larger percentage is plausible in one period of government considering the high share of short term contracts, the importance of purchases of military goods and investments and the possibility to obtain substantial revenue by selling military and non-military assets.

Such rules for filtering policy proposals are intended to make the evaluation more reliable and realistic. However, they could also serve as a filter biased against innovative policy proposals. Skepticism of the evaluator regarding the existence of free lunches might in

fact favor small groups of insiders who collect large rents that could potentially be extracted to the benefit of the wider public. The skepticism regarding the feasibility of such reforms acts as a conservative force. Changing the rules during the game and inventing rather ad hoc rules, e.g. a maximum on specific budget cuts, introduces a substantial amount of arbitrariness. When this occurs, this should be well motivated by the evaluator. The quality of these arguments in combination with the general reputation of the evaluator is then crucial for retaining the credibility of the evaluation.

## 5. Conclusions

Summing up, there is a wide variety of approaches to the evaluation of election platforms, each with their own pros and cons. A comprehensive and long term analysis allows a more balanced presentation of all relevant tradeoffs and implications that are otherwise easily swept under the carpet. However, showing these tradeoffs and implications in a sufficiently reliable and impartial way may be hard and will demand substantial resources and economic skills. Moreover, the presentation of these tradeoffs by the evaluator must be perceived by an overwhelming majority of the political parties and the voters as being fact based and scientifically justified. If such resources and skills are not available or if a broad agreement on the relevant tradeoffs is absent, then it is preferable to constrain the scope of the analysis. In general, the evaluator should take great care not to become part of the political game. That requires that he constrains himself to positive statements, and that he is aware of the fact that constraining political decision making to a choice of the appropriate welfare distribution is trying to lock politicians up in a far too small domain.

Evaluating election platforms is in many respects not high tech-economic analysis, e.g. based on one very big econometric model with thousands of equations. It is a mixed bag of analyses, assumptions and facts: simple and sophisticated analysis, bookkeeping and behavioural analysis, macro and micro, quantitative and qualitative, assumptions about unchanged policy and use of all relevant information, in particular about the government budget and the national economy, institutions and laws.

Evaluating election platforms could be regarded as a game for competing political parties. The rules of this game should ensure the objectivity and quality of the evaluation, give the right incentives to political parties and limit the room for gaming and free lunches.

Voluntary participation by political parties seems to give voters an incomplete set of information. However, in the Netherlands nearly all political parties request for participation, as they do not want to give voters the negative signal that they have something to hide or fear from such evaluation. This information revelation principle may also apply in case of obligatory participation. Knowing that their election platform will be analyzed anyhow, political parties may decide to cooperate with the evaluator in order to avoid negative signaling and publicity.

Over time the scope of analysis is likely to increase in order to cover the loopholes of more partial evaluation or to better incorporate the major different political preferences and

issues. A more encompassing scope of analysis will then also increase the willingness of political parties to participate.

Dutch practice shows that evaluating election platforms can help to reach consensus on policy issues. For example, to what extent are budget cuts needed to improve the health of public finance, to what extent are the budget cuts proposed sufficient for this, what are the consequences of policy measures for the real income of poor households and how effective are the various tools to reduce unemployment?

The credibility and commitment of election platforms depends critically on their link with actual policy, i.e. to what extent will the promises made in the election platforms be reflected in coalition agreements and policy practice? It is therefore very important that the evaluation of election platforms is embedded in the political calendar and decision-making process.

## References

Acemoglu, D., 2003, Why Not A Political Coase Theorem? Social Conflict, Commitment, And Politics, *Journal of Comparative Economics*, Volume 31 (4, Dec), pp. 620-652.

Bos, F., 2008, The Dutch Fiscal Framework; History, Current Practice and the Role of CPB, *OECD Journal on Budgeting*, pp. 1-48. Also published as CPB Document 150.

Bos, F. and C. Teulings, 2010, CPB and Dutch Fiscal Policy in View of the Financial Crisis and Ageing, CPB Document 218; an earlier version was presented at the Conference on Independent Fiscal Institutions, Budapest, March 18-19, 2010.

Butter, F.A.G., den, 2003, An Economic Perspective [on Analysing Election platforms], in Graafland and Ros, pp. 33-44.

Butter, F.A.G. den and M.S. Morgan (eds.), 2000, *Empirical Models and Policy Making: Interaction and Institutions*, Routledge, London.

Calmfors, L., 2011, The role of independent fiscal policy institutions, CESifo Working paper No 3367. Also published as Report of the Swedish Fiscal Policy Council 2010/9.

Calmfors, L. and S. Wren-Lewis, 2011, What should fiscal councils do?, CESifo Working paper No 3382.

CPB, 2006, Charting choices 2008-2011; economic effects of eight election platforms, CPB Document 139.

CPB, 2010a, Focusing on quality; report from the CPB Review Committee 2010 (M.F. Hellwig, T.M. Andersen, R.W. Boadway, J. Elmeskov, R. Griffith and S. van Wijnbergen).

CPB, 2010b, *Keuzes in kaart 2011-2015*; effecten van negen verkiezingsprogramma's op economie en milieu ('Charting choices 2011-2015; nine election plans and their effects for the national economy and the environment'), joint publication with Planning bureau for the Environment.

CPB, 2010c, Economische Verkenning 2011-2015, CPB Document 203.

Di Tella, R., R.J. MacCulloch and A.J. Oswald, 2001, Preferences over inflation and unemployment: evidence from surveys of happiness, *The American Economic Review*, Vol 91 (1), pp. 335-341.

Di Tella, R., R.J. MacCulloch and A.J. Oswald, 2003, The macro economics of happiness, *The Review of Economics and Statistics*, vol 85 (4), pp. 809-827.

Dobbelsteen, S., Levin, J. and H. Oosterbeek, 2002, The causal effect of class size on scholastic achievement: Distinguishing the pure class size effect from the effect of changes in class composition, *Oxford Bulletin of Economics and Statistics*, 64, pp. 17-38.

Don, F.J.H., 2003, Economic Analysis of Election Programmes: What, How, and Why?, in Graafland and Ros, pp. 21-30.

Don, F.J.H. and J. Verbruggen, 2006, Models and methods for economic policy; 60 years of evolution at CPB, *Statistica Neerlandica*. Also published as CPB Document.

Donders, J.H.M. and J.J. Graafland, 2000, CPB models and employment policy in the Netherlands, Chapter 2, pp. 10-25, in Butter and Morgan (eds.), 2000.

Draper, D.A.G. and A.M. Armstrong (eds.), 2007, GAMMA, a Simulation Model for Ageing, Pensions and Public Finances, CPB Document 147.

Folmer, K., 2009, Why do macro wage elasticities diverge; a meta-analysis, CPB Memorandum 224.

Gelauff G.M.M. en J.J. Graafland, 1994, Modeling welfare state reform, MIMICing tax policies and the labour market, Amsterdam: Elsevier.

Graafland, J.J. and A.P. Ros (eds.), 2003, *Economic Assessment of Election Programmes: Does it Make Sense?*, Kluwer Academic Publishers, Dordrecht/Boston.



Haffner, R.C.G. and P.A.G. van Bergeijk, 1994, The economic consequences of Dutch politics, *De Economist* Vol. 142, nr. 4, pp. 497-505.

Hagemann, R., 2010, Improving Fiscal Performance Through Fiscal Councils, OECD Economics Department, Working Papers no. 829.

Hanushek, E.A, and L. Woessmann, 2010, How much do educational outcomes matter in OECD countries?, NBER Working Papers 16515.

Hines jr, J.R., 1999, Three sides of Harberger triangles, *Journal of Economic Perspectives*, Vol. 3, issue 2, pp. 167-188.

Horst, A. van der, L. Bettendorf, N. Draper, C. van Ewijk, R. de Mooij en H. ter Rele, 2010, *Vergrijzing verdeeld; toekomst van de Nederlandse overheidsfinanciën* [Ageing and distribution; the future of Dutch public finance], CPB Special Publication 86.

Jonung, L. and M. Larch, 2004, Improving fiscal policy in the EU: the case for independent forecasts, European Commission, *Economic Papers* no. 210.

Kahneman, D., 2003, A psychological perspective on economics, *American Economic Review*, Vol. 93, No. 2, pp. 162-168.

Lecq, F. van der, 2003, Calculated choices: Who profits?, in Graafland and Ros (2003), pp. 83-96.

Liedekerke, L. van, 2003, The views of the political parties, in Graafland and Ros (2003), pp. 137-144.

Ljungman, G., 2009, Top-Down Budgeting—An Instrument to Strengthen Budget Management, IMF Working Paper 09/243.

Piketty, T., 1995, Social mobility and redistributive politics, *The Quarterly Journal of Economics*, vol. 110, no 3., pp. 551-584.

Schick, A., 2007, *The federal budget; politics, policy and process*, Brookings Institutions Press, Washington DC.

Swank, O.H., 2003, *A Rationale for Assessments of Election Programmes*, in Graafland and Ros (2003), pp. 97-106.

Tanzi, V., 2007, *Argentina: An economic chronicle; How one of the richest countries in the world lost its wealth*, Jorge Pinto Books Inc., New York.

Teulings, C. and T. van Rens, 2008, *Education, growth and income inequality*, *The Review of Economics and Statistics*, MIT Press, vol. 90(1), pp. 89-104.

Tversky, A. and D. Kahneman, 1986, *Rational choice and the framing of decisions*, *Journal of Business*, 59(4), Part 2, pp. S251-278.

Woessmann, L. and M. West, 2006, *Class-size effects in school systems around the world; evidence from between-grade variation in TIMSS*, *European Economic Review*, 50, p. 695-736.