

Measuring central bank independence: a tale of subjectivity and of its consequences

By Gabriel Mangano

Centre Walras-Pareto, University of Lausanne, BFSH 1, 1015 Lausanne, Switzerland, and London School of Economics

The paper first scrutinises the two indices of central bank independence (CBI) most commonly used in the empirical literature. It defines and discovers an impressive interpretation spread, a major criteria spread but a negligible weighting spread in those indices. Second, it examines the robustness of the empirical 'common knowledge' on the benefits of CBI. It finds that, when rankings produced by various CBI indices are regressed with, among others, average inflation, 87.5% of the regression coefficients are not statistically significant. Third, following recent theoretical developments, it suggests an alternative approach to the measurement of a central bank's operational status.

1. Introduction

The idea that granting formal independence to a Central Bank (CB) is a way to improve its country's (or region's) inflationary performance is currently enjoying considerable popularity. The statutes of the European Central Bank, for instance, formally prohibit the governments of EMU member countries from attempting to influence the conduct of monetary policy (European Commission, 1991a and 1991b). In the UK, the newly-elected Chancellor (finance minister) has recently won many praises for having introduced fresh legislation conferring operational independence on the Bank of England. Central Bank Independence (CBI) has become a fashionable notion.

The theoretical underpinnings for such claims are now quite well known: the credibility of monetary policy is supposed to be enhanced when it is implemented by an independent CB, as government pressures for a more expansionary stance can be more easily resisted; the eradication of this so-called time-inconsistency problem of monetary policy (and/or the weakening of the political business cycle) should then lead workers to lower their inflationary expectations, and thus to moderate their wage claims; as a consequence of lower wage settlements, average inflation (as well as its variability) should be reduced.¹

On the empirical front, the number and consistency of studies examining the inverse CBI-inflation relationship also seem to build, at least at first sight, quite a

¹ For a clear and concise review of the complete theoretical argument for CBI, see ch. 3 of Cukierman (1992), and the references therein.

strong presumption in its favor: Eijffinger and de Haan (1996) present an impressive list of papers which, across different samples and different periods, all offer evidence supporting that relationship.²

A number of authors, however, have revealed some significant weaknesses in these theoretical and empirical conclusions. Posen (1994), whose results have been indirectly confirmed by DeBelle and Fischer (1994), deals a serious blow to the theoretical justification for CBI. When testing for a number of predictions which should be verified if higher CBI really leads to higher credibility, Posen finds that none of these predictions is supported by the data, and therefore that even if inflation is affected by CBI, there is no sign that it is so through the standard credibility channel.

Similarly, other studies have questioned the various empirical findings on CBI. Cargill (1995) argues that because of major flaws in the measurement of CBI in certain countries, its negative link with inflation is weaker than originally thought. Walsh (1993) and Forder (1996) stress that as long as measures of CBI are considered valid only when they correlate satisfactorily with inflation, they cannot logically be used to test the hypothesis on which they are constructed.

This paper deals explicitly with these empirical concerns, and tries to gauge the seriousness of the measurement problems affecting most (if not all) CBI indices. It shows that it may be premature to take the conclusions of the numerous empirical studies on CBI at face value, given that the measures of CBI on which they rely do not appear to offer a fully satisfactory representation of a CB's statutes.

The majority of these studies base their investigations and conclusions on one of two widely-respected CBI indices: those computed and presented by Grilli, Masciandaro, and Tabellini (1991), and by Cukierman (1992) in his ch. 19.³ In this paper, I start by dissecting each of these indices and examining their consistency. I then bring some other published indices into the picture, and use them to test the robustness of the well-established common knowledge on the benefits of CBI. Finally, I motivate and suggest alternative ways in which the CB's actual status could be captured.

I show, first, that, although the two indices mentioned above share some common features, they both suffer from a rather large 'subjectivity spread': any empirical result based on either of them therefore appears questionable. My results reinforce (and allow for an explicit quantification of) similar suspicions expressed by other studies (see Eijffinger and de Haan, 1996; Cargill, 1995).

I then regress the country rankings produced by these and other CBI indices, rather than the value of the indices, on inflation and growth, and find that, although regression coefficients are usually of the expected sign, their statistical

² It is worth noticing, however, that the evidence on the link between CBI and the variability of inflation is much more mixed.

³ Twenty out of the 26 empirical studies on CBI listed in Table 7 of Eijffinger and de Haan (1996) use at least one of these measures.

significance is generally poor (few of the *t*-statistics are above their null-hypothesis rejection value at a 5% level of significance). This tends to confirm the impression gained earlier, namely that the conclusions reached in the empirical literature about the supposed effect of CBI on various other economic variables (notably inflation) are on shakier foundations than is usually realised.

I finally argue that it is the whole of a CB's institutional structure (including features such as its accountability), more than its mere legal independence, which should be the centre of future measurement attempts. In particular, the conclusions of recent studies which introduce agency theory into the theoretical analysis of the interaction between CBs and political authorities (Fratianni *et al.* 1993; Persson & Tabellini, 1993; Walsh, 1995) are well worth being applied to future empirical investigations.

This study therefore differs from former research on CBI in at least four respects:

- (i) it offers an in-depth examination of the subjectivity and consistency of the two most frequently used indices of CBI, and through this detailed analysis, introduces four supplementary measures;
- (ii) it places emphasis on the country-rankings produced by various CBI indices, rather than on index-specific values;
- (iii) it considers most indices of legal CBI, rather than just one or two, in its investigation of the inverse CBI-inflation relationship;
- (iv) it suggests a different empirical approach, motivated by recent theoretical findings on the structural characteristics of CBs.

By questioning the reliability of existing CBI indices, it draws attention to one of the fundamental weaknesses of the existing empirical literature on inflation and CBI, and suggests that greater caution should be exerted when invoking its results.

2. A review of existing indices

Attributing a unique value to a whole set of legal characteristics which influence the way a CB operates is not an easy task; it involves a lot of subjectivity. Even the authors of the two most thorough attempts at such a task acknowledge this fact. Grilli *et al.* (1991) realise the arbitrariness of their aggregation method (p. 367), but stress that they prefer to emphasise simplicity. Cukierman (1992) admits openly (p.379) that his weighting scheme is simply designed to minimise data-availability problems.

It would be somewhat narrow-minded, however, to reject all measures of CBI constructed so far only on the grounds that they are subjective. What is more interesting is to examine the practical consequences of such unavoidable subjectivity. This paper concentrates on that crucial but neglected issue. It examines in detail the consistency of the two most widely-used CBI indices, compares the country rankings they produce with those derived from other published indices, and tries to

verify whether the much-acclaimed inverse relationship between CBI and average inflation is 'subjectivity-proof'.⁴

In this section, the two CBI indices proposed by Grilli *et al.* (1991) and Cukierman (1992) (henceforth, GMT and CUK, respectively) are dissected.⁵ A directly-comparable version of each of these indices is constructed by analysing their individual characteristics. With the help of these alternative indices, a method of quantifying the subjectivity spread of the original indices is discussed, and a measure of the components of such a spread is offered. Finally, the country rankings produced by the six indices considered, as well as their respective values, are compared.

2.1 Reconsidering the data

It should be noted from the outset that, while other indices of legal CBI have arguably received at least as much attention as CUK and GMT,⁶ the initial focus here is only on these two indices mainly because they are built upon an extensive range of detailed criteria (15 in the case of GMT, 16 for CUK). This allows for a closer scrutiny of their consistency and comparability.⁷

Tables 1 and 2 show the values taken, in each of the 17 countries considered, by each of the characteristics constituting (respectively) GMT and CUK, after the following procedure has been applied to the original data.⁸ (i) refining the definition of some characteristics, in order for all of them to be expressible on a binary scale (Yes/No) and in order for those which are common to both indices to be directly comparable; (ii) normalising on a [0; 1] scale the value attributed by the original studies to each variable,⁹ and (iii) where necessary, correcting values given by GMT and completing values missing in CUK using the data set available in the

⁴ i.e., Whether the conclusions of the statistical tests linking CBI and average inflation are robust with respect to the subjectivity of the CBI indices used in those tests.

⁵ Unless specified otherwise, CUK always refers to Cukierman's unweighted index LVAU relevant for the 1980–89 period (see Footnote 7 for a justification of that choice of period).

⁶ Bade and Parkin (1982) and its subsequent extension in Alesina (1988) are two other oft-quoted studies. There have since been other attempts at measuring legal or actual CBI, and some of them are introduced and briefly described in Section 3.

⁷ The quest for greater comparability also guided the choice of the countries to be included in the sample, as well as the selection of the relevant period. On the first count, only the countries present in both GMT's and CUK's sample are considered in this section, which yields a total of 17 countries. On the second count, the values retained for CUK's characteristics are the ones attributed to them by Cukierman (1992) in the 1980s, given that GMT is claimed by its authors to cover the 'flexible exchange-rate period' (hence presumably post-1972), and that in the countries studied here, there is no major difference anyway between the values shown in Cukierman (1992) for the 1970–79 or the 1980–89 period.

⁸ The original data itself is given, for the sake of completeness, in Tables 1a and 2a.

⁹ In CUK's case, where most characteristics take several values between 0 and 1, the threshold for the choice between a 0 or a 1 value—i.e. indirectly the refined definition of the variable—was chosen, when possible, to yield maximum similarity with GMT for that characteristic.

Table 1 Grilli, Masciandaro and Tabellini (1991) index: refined and normalised criteria

	AU	B	CH	CN	D	DK	FR	GR	IR	IT	J	NL	NZ	Ö	SP	UK	US
CBGA	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
CBGT	1	0	1	1	1	0	0	0	1	0	0	1	0	0	0	0	0
MPFA	0	0	1	0	1	1	0	1	1	0	0	1	0	1	0	0	1
PSSO	1	0	1	1	1	1	0	0	1	0	1	1	0	1	1	0	1
LDCR	0	0	1	0	1	0	0	1	0	0	0	1	0	1	0	0	1
CLGI	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	1
CLGM	0	0	1	0	1	0	0	0	1	0	0	0	1	1	1	1	1
CLGL	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
PGDP	1	1	1	0	1	1	1	0	0	0	1	1	0	1	0	0	1
GMT9	0.56	0.22	0.89	0.44	0.78	0.33	0.22	0.33	0.67	0.22	0.33	0.67	0.22	0.67	0.33	0.33	0.78
CBBA	0	0	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0
CBBT	0	1	0	0	1	0	1	0	0	0	0	1	0	0	1	0	1
CBBG	0	0	1	0	1	0	0	0	0	1	0	1	0	0	1	1	1
CLGA	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	1	1
DRSA	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1
CBSR	0	1	1	1	0	1	1	0	0	0	0	0	0	1	0	0	0
GMTN	0.47	0.33	0.73	0.47	0.73	0.40	0.33	0.33	0.47	0.27	0.33	0.60	0.20	0.60	0.33	0.40	0.73

Notes: Original data refined, when possible and suitable, on the basis of Appendix C in GMT (1991), and normalised to be expressed on binary scale.

GMT9 is a simple average over the normalised values of the nine refined criteria common to both GMT and CUK indices, as measured by Grilli *et al.* (1991).

GMTN is a simple average over the normalised values of the nine refined and six remaining criteria used in GMT (1991).

CBGA: CB Governor Appointment (1 if *not only* appointed by Government, 0 otherwise).

CBGT: CB Governor, Terms of Office (1 if *explicitly* appointed for more than 5 years, 0 otherwise).

MPFA: Monetary Policy Formulation Authority (1 if no Government approval required, 0 otherwise).

PSSO: Price Stability as Statutory Objective (1 if at least among statutory objectives, 0 otherwise).

LDCR: Legal Directives for Conflict Resolution (1 if strengthen CB position, 0 otherwise).

CLGI: CB Lending to Government, Interest-rate (1 if at market interest rate or if *lending not available*, 0 otherwise).

CLGM: CB Lending to Government, Maturity (1 if *explicitly* limited to 1 year or less or if *lending not available*, 0 otherwise).

CLGL: CB lending to Government, Limit (1 if amount *explicitly* limited or if *lending not available*, 0 otherwise).

PGDP: Purchase of Government Debt, Primary market (1 if CB *not allowed* to participate, 0 otherwise).

CBBA: CB Board Appointment (1 if *majority* of Board members not appointed by Government, 0 otherwise).

CBBT: CB Board, Terms of office (1 if all Board appointed for more than 5 years, 0 otherwise).

CBBG: CB Board, Government representative (1 if presence not compulsory, 0 otherwise).

CLGA: CB Lending to Government, Availability (1 if not automatic, 0 otherwise).

DRSA: Discount Rate Setting Authority (1 if *only* CB can set discount rate, 0 otherwise).

CBSR: CB Supervisory Role (1 if CB *not involved* in banking supervision, 0 otherwise).

Table 1a Grilli, Masciandaro, and Tabellini (1991) index: original criteria

	AU	B	CH	CN	D	DK	FR	GR	IR	IT	J	NL	NZ	Ö	SP	UK	US
CBGA'	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
CBGT'	1	0	1	1	1	1	1	0	1	1	0	1	0	0	0	0	0
MPFA	0	0	1	0	1	1	0	0	1	0	0	1	0	1	0	0	1
PSSO	1	0	1	1	1	1	0	0	1	0	1	0	1	0	0	0	1
LDCR	1	0	1	1	1	0	0	1	0	0	0	1	0	1	0	0	1
CLGI'	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1
CLGM'	1	0	1	1	1	0	0	0	1	0	1	1	1	1	1	1	1
CLGL'	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1
PGDP'	1	1	1	0	1	1	1	0	0	0	1	1	0	1	0	0	1
CBBA'	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
CBBT	0	1	0	0	1	0	1	0	0	0	0	1	0	0	1	0	1
CBBG	0	0	1	0	1	0	0	0	0	1	0	1	0	0	1	1	1
CLGA	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	1	1
DRSA'	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1
CBSR	0	2	2	2	1	2	2	0	0	0	1	0	0	2	1	0	1
GMT	9	7	12	11	13	8	7	4	7	5	6	10	3	9	5	6	12

Note: GMT is a simple sum over the original values of the 15 criteria used in GMT (1991).

CBGA': CB Governor Appointment (1 if not appointed by Government, 0 otherwise).

CBGT': CB Governor, Terms of Office (1 if appointed for more than 5 years, 0 otherwise).

MPFA: Monetary Policy Formulation Authority (1 if no Government approval required, 0 otherwise).

PSSO: Price Stability as Statutory Objective (1 if at least among statutory objectives, 0 otherwise).

LDCR: Legal Directives for Conflict Resolution (1 if strengthen CB position, 0 otherwise).

CLGI': CB Lending to Government, Interest rate (1 if at market interest rate, 0 otherwise).

CLGM': CB Lending to Government, Maturity (1 if temporary, 0 otherwise).

CLGL': CB Lending to Government, Limit (1 if amount limited, 0 otherwise).

PGDP': Purchase of Government Debt, Primary market (1 if CB does not participate, 0 otherwise).

CBBA': CB Board Appointment (1 if Board members not appointed by Government, 0 otherwise).

CBBT: CB Board, Terms of office (1 if Board appointed for more than 5 years, 0 otherwise).

CBBG: CB Board, Government representative (1 if presence not compulsory, 0 otherwise).

CLGA: CB Lending to government, Availability (1 if not automatic, 0 otherwise).

DRSA': Discount Rate Setting Authority (1 if discount rate set by CB, 0 otherwise).

CBSR': CB Supervisory Role (2 if CB not involved in banking supervision, 1 if CB not alone to be involved in banking supervision, 0 otherwise).

Appendices of the original studies. The first nine characteristics appearing in each table are common to both indices, while all others are specific to the index under which they are listed. For each country, GMT9 and CUK9 denote the simple average over the values attributed to the first nine criteria by each study, while GMTN and CUKN describe the simple average over the values attributed to all criteria shown in each table.

To realise how useful these four supplementary measures can be, it is worth recalling in which precise respect synthetic indices of legal CBI can be considered as subjective. As Eijffinger and Schaling (1993) have correctly pointed out (p.50),

Table 2 Cukierman (1992) index, 1981–89: completed and normalised criteria

	AU	B	CH	CN	D	DK	FR	GR	IR	IT	J	NL	NZ	Ö	SP	UK	US
CBGA	0	0	0	1	1	0	0	1	0	1	0	0	0	0	0	0	0
CBGT	1	0	1	1	1	0	0	0	1	0	0	1	0	0	0	0	0
MPFA	0	0	1	0	1	1	1	0	1	0	1	0	0	1	0	0	1
PSSO	1	0	0	0	1	1	0	1	1	0	0	1	1	1	1	0	1
LDCR	0	0	1	0	1	1	1	1	0	0	0	0	0	1	0	0	0
CLGI	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
CLGM	1	1	1	1	1	1	0	1	1	0	0	0	0	1	0	1	1
CLGL	1	0	1	1	1	1	0	1	1	1	0	1	0	1	1	0	1
PGDP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CUK9	0.56	0.11	0.56	0.44	0.78	0.56	0.22	0.56	0.56	0.22	0.11	0.33	0.11	0.67	0.22	0.11	0.44
CBGD	1	0	1	1	1	0	1	1	1	1	1	0	1	1	0	1	0
CBGO	1	0	1	1	0	0	0	0	1	1	0	1	1	1	1	1	1
BPFA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PGDL	0	0	0	1	1	1	0	1	0	0	0	0	0	1	0	0	1
CLGC	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
GLGB	0	0	1	1	0	0	1	0	0	0	0	1	1	0	0	1	1
CLGT	0	NA	0	0	1	0	NA	1	0	0	NA	1	NA	0	0	NA	0
CUKN	0.44	0.07	0.56	0.56	0.69	0.44	0.27	0.50	0.44	0.25	0.13	0.38	0.27	0.56	0.19	0.27	0.44

Notes: Original criteria normalised to be expressed on binary scale, and when not available (13 instances), completed if possible using Appendix C in Grilli *et al.* (1991).

CUK9 is a simple average over the normalised values of the nine refined criteria common to both GMT and CUK, as measured by Cukierman (1992).

CUKN is a simple average over the normalised values of the nine refined and seven remaining criteria used in CUK (1992).

CBGD: CB Governor Dismissal (1 if cannot be decided by Government for policy reasons, 0 otherwise).

CBGO: CB Governor, Other governmental post tenure (1 if explicitly not allowed by law, 0 otherwise).

BPFA: Budgetary Policy Formulation Authority (1 if CB given active role, 0 otherwise).

PGDL: Purchase of Government Debt, Limit (1 if explicitly limited, 0 otherwise).

CLGC: CB Lending to Government, Conditions (1 if no Government influence on conditions, 0 otherwise).

CLGB: CB Lending to Government, Borrowers (1 if only political entities can borrow or if lending not available, 0 otherwise or if unclear).

CLGT: CB Lending to Government, Type of limit (1 if limit not expressed in terms of Government revenues or expenditures, 0 otherwise, NA if CLGL = 0).

See Table 1 for other abbreviations.

there are three types of choice involved when constructing any such index, in which some degree of personal judgement unavoidably intervenes: (i) which criteria should be included in the index? (ii) how should the legislation be interpreted with respect to each retained criterion (which leads to their individual valuation)? and (iii) what weight should be attributed to each criterion in the composite index?

Along the lines of this classification, it is therefore possible to describe the

Table 2a Cukierman (1992) index, 1981-89: original criteria

	AU	B	CH	CN	D	DK	FR	GR	IR	IT	J	NL	NZ	Ö	SP	UK	US
CBGA ^a	0.00	0.00	0.25	0.75	0.75	0.00	0.25	0.75	0.50	0.75	0.25	0.00	0.00	0.00	0.00	0.00	0.50
CBGT ^a	0.75	0.50	0.75	0.75	1.00	0.00	0.00	0.25	0.75	0.00	0.50	0.75	0.50	0.50	0.25	0.50	0.25
MPFA ^a	0.33	0.00	NA	0.33	0.67	NA	0.67	0.33	NA	NA	0.67	0.33	0.00	1.00	0.33	0.00	NA
PSSO ^a	0.40	0.00	0.00	0.20	1.00	0.60	0.00	0.80	0.80	0.20	0.00	0.80	0.40	0.60	0.60	0.20	0.40
LDCR ^a	0.20	0.20	1.00	0.20	1.00	1.00	0.60	0.60	NA	NA	0.00	0.20	0.00	0.60	0.00	0.00	0.20
CLGT ^a	1.00	0.50	0.25	0.75	0.25	0.25	NA	0.75	0.75	0.25	0.25	0.00	0.50	1.00	0.00	0.75	0.25
CLGM ^a	0.67	1.00	1.00	0.67	1.00	1.00	NA	1.00	0.67	0.00	0.00	0.00	0.00	1.00	0.00	1.00	1.00
CLGL ^a	0.33	0.00	1.00	0.33	0.67	1.00	0.00	0.67	NA	0.33	0.00	0.67	0.00	1.00	0.33	0.00	1.00
PGDP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CBGD ^a	0.83	0.00	NA	0.83	1.00	0.33	1.00	0.67	0.83	0.67	0.83	0.17	0.83	0.83	0.00	0.83	0.00
CBGO ^a	1.00	0.50	1.00	1.00	0.00	0.00	0.50	0.50	1.00	1.00	0.50	1.00	1.00	1.00	1.00	1.00	1.00
BPFA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PGDL ^a	0.00	0.00	NA	0.33	0.67	0.33	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.33
CLGC ^a	0.33	0.33	1.00	0.67	0.67	0.67	0.33	0.33	0.33	0.33	0.33	0.00	0.00	0.33	0.00	0.00	0.33
CLGB ^a	0.00	0.33	1.00	0.67	0.33	0.00	1.00	0.00	0.33	NA	NA	1.00	1.00	0.33	0.33	1.00	1.00
CLGT ^a	NA	NA	NA	0.33	1.00	NA	NA	1.00	NA	0.00	NA	1.00	NA	0.33	0.00	1.00	NA
CUK	0.31	0.19	0.68	0.46	0.66	0.47	0.28	0.51	0.39	0.22	0.16	0.42	0.27	0.58	0.21	0.31	0.51

Note: "CUK" is a weighted average over the original values of the 16 criteria used in Cukierman (1992).

<p>CBGA^a: CB Governor Appointment 1.00 if appointed by CB Board, 0.75 if appointed by legislative and executive branches of Government and by CB board, 0.50 if appointed by legislative branch, 0.25 if appointed by executive branch, 0.00 if appointed by 1 or 2 members of executive branch.</p> <p>CBGT^a: CB Governor, Terms of office 1.00 if longer than or equal to 8 years, 0.75 if between 6 years and less than 8 years, 0.50 if equal to 5 years, 0.25 if equal to 4 years, 0.00 if smaller than 4 years.</p> <p>MPFA^a: Monetary Policy Formulation Authority 1.00 if granted to CB alone, 0.67 if granted to both CB and Government, 0.33 if CB's capacity only advisory, 0.00 if granted to Government alone.</p> <p>PSSO^a: Price Stability as Statutory Objective 1.00 if only objective and CB has final authority, 0.80 if only objective, 0.60 if other non-conflicting objectives, 0.40 if other conflicting objectives, 0.20 if no objectives in CB charter, 0.00 if only other objectives in CB charter</p> <p>LDCR^a: Legal Directives for Conflict Resolution 1.00 if attributed to CB for CB's objectives, 0.80 if attributed to Government only for non-objectives, 0.60 if attributed to CB board, legislative and executive branches of Government, 0.40 if unconditionally attributed to legislative branch, 0.20 if conditionally attributed to executive branch, 0.00 if unconditionally attributed to executive branch.</p> <p>CLGT^a: CB Lending to Government, Interest-rate 1.00 if must be at market level, 0.75 if cannot be lower than a certain floor, 0.50 if cannot be higher than a certain ceiling, 0.25 if not restricted, 0.00 if no interest payment required.</p> <p>CLGM^a: CB Lending to Government, Maturity 1.00 if limited to 6 months, 0.67 if limited to 12 months, 0.33 if limited to more than 12 months, 0.00 if unlimited.</p>	<p>CLGL^a: CB Lending to Government, Limit 1.00 if lending prohibited, 0.67 if lending subjected to strict limits, 0.33 if lending subjected to accommodative limits, 0.00 if lending unlimited.</p> <p>PGDP: Purchase of Government Debt, Primary market 1.00 if CB not allowed to participate, 0.00 if CB allowed to participate.</p> <p>CBGD^a: CB Governor Dismissal 1.00 if not provided for, 0.83 if possible only for nonpolicy reasons, 0.67 if unconditionally possible by CB Board, 0.50 if conditionally possible by legislative branch, 0.33 if unconditionally possible by legislative branch, 0.17 if conditionally possible by executive branch, 0.00 if unconditionally possible by executive branch.</p> <p>CBGO^a: CB Governor, other governmental post tenure 1.00 if prohibited, 0.50 if subjected to approval by executive branch, 0.00 if not prohibited.</p> <p>BPFA: Budgetary Policy Formulation Authority 1.00 if active role for CB, 0.00 if no active role for CB.</p> <p>PGDL^a: Purchase of Government Debt, Limit 1.00 if purchase prohibited, 0.67 if purchase subjected to strict limits, 0.33 if purchase subjected to accommodative limits, 0.00 if purchase unlimited.</p> <p>CLGC^a: CB Lending to Government, Conditions 1.00 if controlled by CB, 0.67 if control attributed by law to CB, 0.33 if control left by law negotiations between CB and Government, 0.00 if controlled by executive branch.</p> <p>CLGB^a: CB Lending to Government, Borrowers 1.00 if only central Government, 0.67 if central and state Governments, 0.33 if all of the above plus public firms, 0.00 if all of the above plus private sector.</p> <p>CLGT^a: CB Lending to Government, Type of limit 1.00 if absolute cash amount, 0.67 if percentage of CB capital, 0.33 if percentage of Government revenues, 0.00 if percentage of Government expenditures.</p>
---	---

subjectivity spread from which every index of legal CBI suffers as consisting of three main parts:

- (i) a criteria spread, i.e. the extent to which the researchers' personal preferences affect the selection of the criteria to be included in the index;
- (ii) an interpretation spread, i.e. the extent to which the researchers' reading of the legislation pertaining to these criteria is different; and
- (iii) a weighting spread, i.e. the extent to which the final value of the index is affected by the choice of relative weights to be attributed to each individual criterion.

With the help of the supplementary indices computed above, each of these spreads can now be individually considered, and a tentative quantification of their value in the case of CUK and GMT can be offered.

2.2 Interpretation spread

Having identified nine criteria included in both CUK and GMT, and having converted to a common scale the values attributed by each study to those criteria (see Tables 1 and 2), it is now fairly straightforward to measure how strongly the authors of the two indices disagree in the criteria-valuation stage, i.e. their (average) interpretation spread.

Table 3 takes a closer look at these nine common criteria. It directly compares the value attributed by CUK and GMT to each criterion in each country: whenever both indices agree on that value, a 'Yes' entry is shown in the relevant position of the table; when they do not, 'No' is entered. For each country, the percentage of 'No' entries is then computed and shown in the last row of the table: it indicates the (average) Interpretation Spread by Country from which CUK and GMT seem to suffer. Similarly, for each criterion, the percentage of 'No' entries appears in the last column of the table, and is described as the (average) Interpretation Spread by Criterion exhibited by the indices in question. Finally, a measure of their (average) Overall Interpretation Spread appears in the bottom-right cell of the table.¹⁰

The results of this simple procedure reveal a significant degree of inconsistency between the two indices' valuation of their common criteria. Only in one country (out of 17) and in the case of one criterion (out of 9) have CUK and GMT translated the legislation in exactly the same way: their interpretation of the laws governing the Italian CB, and of the regulations concerning the CB Governors' terms of office in the countries sampled, exhibit no divergence. On the other hand, their average interpretation spread when examining Danish, French, Greek, and

¹⁰ Given the binary nature of each characteristic, it doesn't really matter which of the two indices is wrong and which is right when they disagree on the value x_j to be taken by a characteristic i in a country j : the Interpretation Spread (either by country or by criterion) still indicates the percentage of instances in which the relevant legislation was misinterpreted, either by one or by the other index. In that sense, this procedure neither allows nor attempts to decide which of the two indices is 'more wrong' in its reading of the law: it just indicates how wrong, on average, the two of them are.

Table 3 Criteria common to CUK and GMT: disagreement in the interpretation of CB legislation

	AU	B	CH	CN	D	DK	FR	GR	IR	IT	J	NL	NZ	Ö	SP	UK	US	IS (Cr)	
CBGA	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	12%
CBGT	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	0%
MPFA	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	24%
PSSO	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	29%
LDCR	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	24%
CLGI	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No	No	No	29%
CLGM	No	No	Yes	No	Yes	No	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes	41%
CLGL	Yes	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	Yes	35%
PGDP	No	No	No	Yes	No	No	No	Yes	Yes	Yes	No	No	Yes	No	Yes	Yes	No	No	59%
IS(Co)	22%	33%	33%	22%	22%	44%	44%	44%	11%	0%	44%	33%	33%	22%	11%	22%	33%	28%	

Notes: 'Yes' indicates agreement between CUKN and GMTN on the value of the relevant criterion in the relevant country; 'No' indicates disagreement.

IS (Cr) is the Interpretation Spread by Criterion; it indicates the proportion of countries in which the two studies do not agree on the value to be given to the relevant criterion.

IS (Co) is the Interpretation Spread by Country; it indicates the proportion of criteria over the value of which the two studies do not agree in the relevant country.

See Table 1 for abbreviations.

Japanese legislation is close to 50%, and they disagree in nearly 60% of countries when deciding whether the CB is legally allowed to purchase government debt in the primary market or not. Overall, it appears that in the 17 countries included in both CUK's and GMT's samples, virtually a third of the values attributed to their nine common criteria are subject to non-negligible interpretation problems.¹¹

While the scale of the problem had not previously been measured in such a systematic way, its existence had already been anticipated by a number of authors: the analysis in this section has allowed for a formal confirmation and a more precise quantification of those intuitive perceptions. Eijffinger and de Haan (1996) draw on their detailed knowledge of the Dutch CB's legislation to show (Section 3.2) that CUK attributes an incorrect value to five of the 16 characteristics by which he measures its legal independence (again, almost a third of the total). Cargill (1995), emphasising that 'the basic problem with indices is they ultimately rely on a researcher's interpretation of central bank laws' (p.163), admits having 'a difficulty to rationalise the ranking of the Bank of Japan and the Federal Reserve presented in the Bade and Parkin (1982) study' (p.161).¹² Eijffinger and Schaling

¹¹ This problem may even be worse than appears from these figures: they rely on the assumption that at least one of the two indices has interpreted the legislation correctly, but this is not necessarily the case.

¹² This does not prevent Cargill, however, from falling into the 'subjectivity trap' himself: his reasons for disagreeing with Bade and Parkin's (1982) attribution of the same rank to the BoJ and the Fed is that 'the Bank of Japan is generally regarded as one of the more formally dependent central banks' (p.164).

(1993) criticise Alesina's (1988) subjective re-interpretation of the Italian CB's statutes, stressing that it renders his index internally inconsistent (p.59).¹³

2.3 Criteria spread

Here again, it is impossible to discuss the appropriateness of the characteristics constituting individual indices without being exposed to some degree of subjectivity as well: on which objective grounds would a criterion be considered as acceptable for inclusion, and what would objectively justify the exclusion of another? It could always be argued that, intuitively, a few criteria should unquestionably qualify, while some others needn't even be considered; but where should the line be drawn between these two categories? Clearly, these questions cannot be answered satisfactorily: blaming authors for having incorporated one characteristic rather than another is displaying as much subjectivity as (if not more than) they do.

Instead, what the standardisation of the data attempted in Section 2.1 allows for is a relative approach to the problem. It is not used to criticise one or the other index for taking some characteristics into account, while neglecting others;¹⁴ it merely helps us to measure how strongly the authors of the indices disagree as to the necessity of including certain criteria. In that sense, the measure presented in this subsection can only capture the Criteria Disagreement Factor separating the two studied indices, rather than a genuine (but much more subjectivity-prone) Criteria Bias.

The procedure is simple enough: it merely involves computing the ratio, in both indices, of the number of index-specific criteria to their total number, and comparing these two ratios. Even on this very straightforward count, CUK and GMT reveal a significant degree of disagreement: as much as 40% of the characteristics included in GMT are not considered as relevant by CUK, while GMT disregards 45% of the criteria seen by CUK as essential to measure CBI.

Obviously, it would have been all the more interesting to be able to compare these proportions with those exhibited by other studies, but such a comparison is made particularly tricky by the fact that no other synthetic index of CBI is built upon such an extensive range of detailed criteria. Other published measures rely either on a series of loosely-defined policy types (Bade and Parkin, 1982; Eijffinger and Schaling, 1993; Eijffinger and van Keulen, 1995), or on indicators of political influence upon the CB (Cukierman and Webb, 1994), or even on answers to questionnaires sent out to CB officials (Cukierman, 1992, QVAU and QVAW).

¹³ Eijffinger and Schaling (1993) also attempt to measure what they call an 'interpretation effect', but choose to compare the Bade and Parkin (1982) index (rather than CUK) with GMT and rely on only two of these indices' common criteria to quantify such an effect.

¹⁴ In particular, nowhere in the analysis do I imply that the nine characteristics which CUK and GMT have in common are some sort of core features of CBI: that would clearly constitute an implicit criticism of other indices which might have chosen not to include some of these criteria.

Alternatively, a similar comparison could be made between the relative weight attributed by CUK and GMT to their respective index-specific criteria:¹⁵ it would indicate what importance the authors attribute in their index to the criteria on which they disagree. For obvious reasons, however, such a comparison would be directly affected by the third type of spread identified earlier: it would incorporate disagreement not only on the type of criteria to include, but also on the relative importance of each of those criteria.

As it turns out, the significance of the weighting spread¹⁶ inherent in CUK and GMT is difficult to estimate. It can be approximated by computing the partial correlations between the normalised versions of the indices and their original counterparts.¹⁷ However, the results suggest that the choice of a specific weighting scheme is probably unlikely to add much to an index's (already significant) degree of subjectivity.¹⁸ In any case, the weighting differential between GMT and GMTN is minimal, since it is only caused by a single criterion to which a value of up to two is attributed in GMT, while it is restricted to a maximum of 1 in GMTN.

In summary, significant signs of the presence of an impressive interpretation spread, of a major criteria spread but of a negligible weighting spread in existing CBI indices have been discovered through the above analysis. It is worth stressing, however, that all the above estimations find themselves limited by two factors: the rather small sample of countries examined in both indices, and the very fact that only two existing indices offer a wide enough scope for comparison. These limitations should be borne in mind when (and if) future references to the above findings are made.

2.4 Introducing country rankings

Table 4 summarises the values, for each of the 17 countries in the sample, of the 6 indices considered so far. It also introduces the country rankings obtained using each of these indices. Partial correlations between some of their respective country

¹⁵ Bénassy and Pisani-Ferry (1994), for instance, calculate in their Appendix 1 the relative weight attributed by GMT and CUK to all their respective characteristics. Before doing so, however, they group some of CUK's criteria into broader categories and compute those criteria's joint weight, which makes their results difficult to interpret.

¹⁶ In a sense, the criteria spread is an extension of the weighting spread: when the authors of an index decide not to include a certain characteristic, they implicitly attribute a weight of 0 to it. However, the information set upon which those decisions are made is not exactly the same: had additional criteria been included, the relative weights of all or some others might have been different.

¹⁷ Given that GMTN and CUKN include the same characteristics as GMT and CUK (respectively), but constitute simple rather than weighted averages of these characteristics' values, this procedure could reveal the significance of the distortion introduced by the particular weighting method chosen by each index's authors.

¹⁸ Both pairs of indices exhibit an identical partial correlation of 0.92, and the partial correlation between CUKN and GMTN (at 0.71) is virtually identical to the one between CUK and GMT (0.70), which confirms that even unifying the weighting schemes adopted by the two indices does not make them radically more consistent with each other.

Table 4 GMT, CUK, and derived CBI indices

	Values						Rankings (17-country sample)					
	GMT	GMTN	GMT9	CUK	CUKN	CUK9	GMT	GMTN	GMT9	CUK	CUKN	CUK9
AU	0.60	0.47	0.56	0.31	0.44	0.56	6	6	7	9	6	3
B	0.47	0.33	0.22	0.19	0.07	0.11	9	11	14	16	17	14
CH	0.80	0.73	0.89	0.68	0.56	0.56	2	1	1	1	2	3
CN	0.73	0.47	0.44	0.46	0.56	0.44	4	6	8	6	2	8
D	0.87	0.73	0.78	0.66	0.69	0.78	1	1	2	2	1	1
DK	0.53	0.40	0.33	0.47	0.44	0.56	8	9	9	5	6	3
F	0.47	0.33	0.22	0.28	0.27	0.22	9	11	14	12	11	11
GR	0.27	0.33	0.33	0.51	0.50	0.56	16	11	9	4	5	3
IR	0.47	0.47	0.67	0.39	0.44	0.56	9	6	4	8	6	3
IT	0.33	0.27	0.22	0.22	0.25	0.22	14	16	14	14	14	11
J	0.40	0.33	0.33	0.16	0.13	0.11	12	11	9	17	16	14
NL	0.67	0.60	0.67	0.42	0.38	0.33	5	4	4	7	10	10
NZ	0.20	0.20	0.22	0.27	0.27	0.11	17	17	14	13	11	14
Ö	0.60	0.60	0.67	0.58	0.56	0.67	6	4	4	3	2	2
SP	0.33	0.33	0.33	0.21	0.19	0.22	14	11	9	15	15	11
UK	0.40	0.40	0.33	0.31	0.27	0.11	12	9	9	9	11	14
US	0.80	0.73	0.78	0.51	0.44	0.44	2	1	2	9	6	8

	Partial correlations between values						Partial correlations between rankings					
	GMT	GMTN	GMT9	CUK	CUKN	CUK9	GMT	GMTN	GMT9	CUK	CUKN	CUK9
GMT	1.00	0.92	0.82	0.70			1.00	0.91	0.74	0.58		
GMTN		1.00	0.95	0.71				1.00	0.93		0.70	
GMT9			1.00		0.71				1.00			0.66
CUK				1.00	0.92	0.85				1.00	0.93	0.83
CUKN					1.00	0.92					1.00	0.85
CUK9						1.00						1.00

Note: Value of GMT (including one used to compute country ranking) normalised to be expressed in the [0, 1] interval.

rankings, as well as between some of their respective values (among which are those briefly discussed above), are displayed in the lower part of the table.

Reassuringly for the argument, the general picture offered by the comparison of values is consistent with the findings of the preceding subsections. The degree of correlation between GMT9 and CUK9, for instance, is not much higher than the one between GMT and CUK (0.71 and 0.70, respectively). This further confirms the seriousness of the measurement problem (interpretation spread) from which the two original indices suffer: if the nine criteria which they both include had been measured properly, GMT9 and CUK9 should have been identical (i.e. the correlation between them should have been very close to 1.00). Similarly, the fact that the correlation is stronger between GMTN and GMT9 than between GMT and GMT9 (0.95 vs 0.82), and that the same is observed on the CUK side (0.92 vs 0.85), may be seen as a confirmation of the presence of some weighting spread.

There are reasons to believe, however, that comparing the country rankings produced by the six indices, rather than their actual values, should yield even more valuable results. Most importantly, this procedure avoids at least some of the subjectivity involved in the measurement of CBI. It is not affected by the fact that, say, the most dependent CB according to one index carries a precise value of 0.01, while in another, in which it is also seen as the most dependent, it is attributed a different value of 0.1; all it takes into account is that the same CB is actually considered as the least independent by both of them. The partial correlations between rankings are therefore not influenced by extreme values, and should prove a more reliable indicator of how consistent the indices really are with each other.¹⁹

The impression left by these calculations is even more striking than when values were considered. For a start, the correlation between CUK's and GMT's classification of countries is a mere 0.58, which tends to reinforce the impression of general inconsistency gathered earlier; furthermore, the ranking implied by GMT9 is again poorly linked to its CUK9 counterpart (0.66), strengthening former observations about the indices' interpretation spread; finally, the country orderings offered by GMTN and CUKN are now markedly more consistent than those offered by GMT and CUK (0.70 vs 0.58), suggesting that the weighting spread may have deeper consequences than was previously thought.²⁰

There is an important conclusion to be drawn from the significant subjectivity spread and lack of consistency of the two most widely-used CBI indices as revealed in this section. Given that quite a few empirical studies on the effects of CBI on inflation (as well as on other variables) rely on one or the other of these measures, the findings of those studies may not be particularly reliable. This empirical doubt echoes the theoretical scepticism about the same relationship expressed by Posen (1994).

Obviously, further investigations are required to confirm (or dispel) these doubts. In particular, it would be interesting to test for the sensitivity of the CBI-inflation relationship (among others) to changes in measures of CBI, while possibly avoiding some of the subjectivity inherent in these measures.²¹ This is precisely the aim of the next section.

¹⁹ Some information is inevitably lost when switching from values to ranks. The latter, for instance, do not reflect the fact that, say, the institutional difference between the most and second-most independent CB may be much slimmer than between the least and next-to-least independent. But since the only objective of this exercise is to measure the consistency of the two retained indices, such information is irrelevant for our purposes, in particular in a sample as relatively homogeneous as 17 OECD countries.

²⁰ The only correlation which is stronger when rankings rather than values are considered is the one between CUK and CUKN, but even then the improvement is of the order of only one percentage point (0.93 against 0.92).

²¹ One of the major dangers of this type of exercise is the tendency to describe one of the indices as 'the most appropriate' simply because it exhibits the closest relationship with average inflation; as Walsh (1993) and Forder (1996) have already pointed out, many authors of existing CBI indices have tended to fall into this inflation-biased judgement.

3. A reconsideration of the empirical evidence on CBI

It is well known that the majority of empirical studies which have tried to examine the effect of statutory independence on various macroeconomic variables have produced fairly consistent results.²² Using various samples, periods and CBI indices (with GMT and CUK a recurrent choice), 17 out of the 19 studies listed in Table 8 of Eijffinger and de Haan (1996) have concluded that there is a statistically-significant inverse relationship between CBI and average inflation; seven out of the 12 studies in that list which also tested for a link between CBI and inflation variability have reached conclusive results; and eight out of the ten studies which, according to Table 11 of Eijffinger and de Haan (1996), examined the effect of CBI on the average and the variability of real-output growth, have not found any statistically significant link.

The previous section, however, underlined the weaknesses of the two most widely-used indices of legal CBI, and argued that because of their rather large subjectivity spread, they probably do not constitute a totally reliable measure of a CB's statutes. As a consequence, even the most compelling evidence on the effects of CBI may be weaker than previously thought.

Admittedly, some authors have relied on more than one measure of independence in their empirical research, or on a combination of several of them, and have generally obtained consistent results. But any combination of unreliable measures, no matter how elaborate, is still an unreliable measure itself: the fact that a link between CBI and various macroeconomic variables is insensitive to the type of index used says nothing about the sensitivity of that link to the quality of the indices.

This section presents an empirical test of the relationship between CBI and inflation (as well as growth) which avoids at least some of the subjectivity discussed earlier. Rather than using the values of a number of indices as explanatory variables, it uses the country rankings produced by these indices. As argued in Section 2.4, rankings are less affected by the precise characteristics of each index: they are likely to minimise the adverse effect of weighting or criteria problems. The regressions should thus reveal whether the results of former empirical investigations were influenced by the subjectivity of the CBI indices they relied on.

At this stage, it becomes possible and appropriate to include in the analysis as many other proxies of CBI as possible: all that matters is how these indices rank the countries they consider, regardless of which characteristics they take into account to achieve that ranking.²³ Accordingly, four supplementary measures are introduced: the Alesina (1988) extension of the Bade and Parkin (1982) CBI index

²² For a comprehensive list of these studies, see Eijffinger and de Haan (1996), table 7.

²³ For the sake of comparability, however, every attempt is made to keep the joint sample of countries as large as possible, which means that some measures are left out because the number of countries for which they are calculated is too small.

(AL); the Eijffinger and Schaling (1993) index of policy independence, extended by Eijffinger and van Keulen (1995) (ES); the turnover rate of CB governors, as calculated by Cukierman (1992) (TOR); and the Cukierman and Webb (1994) index of the CB's political vulnerability (CW). There are only 12 countries over which these as well as the other six indices discussed in Section 2 are all computed, so that the rankings are distributed according to this 12-country sample (and not according to each measure's original sample).

Figure 1 presents a preliminary view of the consistency of the ten resulting rankings. Switzerland and Germany are the countries over which the indices agree most strongly, with only TOR and CW placing neither the Schweizerische Nationalbank nor the Bundesbank in their top two positions. In most other cases, however, the disagreement as to the relative position of each CB appears more widespread. The charts reveal a rather large dispersion of opinion, with the ranks of the Bank of Japan and of the Bank of England being the most controversial.

Table 5 offers a full list of these classifications, as well as the respective coefficients of partial correlation between them.²⁴ It appears from the bottom half of this table that the rankings produced by the four indices introduced in this section are generally rather poorly related to each other; ES and CW in particular exhibit a low degree of correlation with all other measures. This should not, however, be interpreted as a sign that one index is better or worse than any other: all it indicates is that the various measures do not capture identical aspects of a CB's status.

Each of these ten series is then used as the explanatory variable, in the 12-country sample, for each of four dependent variables: the average annual rate of growth of the GDP deflator over the years 1980–89, the average annual rate of real GDP growth over the same period, and the variances of these two rates.²⁵ A summary of the results of these 40 regressions is given in Table 6.

With one exception (CW), the sign of the link between the discussed rankings and average inflation (proxied by the average growth of the GDP deflator) is consistent with existing evidence: the higher the rank of a country's CB, i.e. the less independent the CB is considered, the higher the average rate of inflation in that country during the 1980s. However, the statistical significance of the relevant coefficients is erratic: out of ten regressions, only two (those in which the rankings of GMT and GMTN are the explanatory variables) produce a *t*-statistic which is above its null-hypothesis rejection value at a 5% level of significance, and only one of these *t*-statistics (associated with the AL classification) is above that value at a

²⁴ The rankings derived from GMT, GMTN, GMT9, CUK, CUKN, and CUK9 do not display the same correlation among them as in Section 2: this is obviously due to the fact that the size of the sample has been reduced from 17 to 12 countries.

²⁵ While TOR and CW were calculated by their authors over four decades (1950–89) to give them greater significance, the rankings they imply are regressed here with averages and variances calculated only over the period 1980–89; this could arguably have an influence on the magnitude of some coefficients, but it should not affect the general nature of the results.

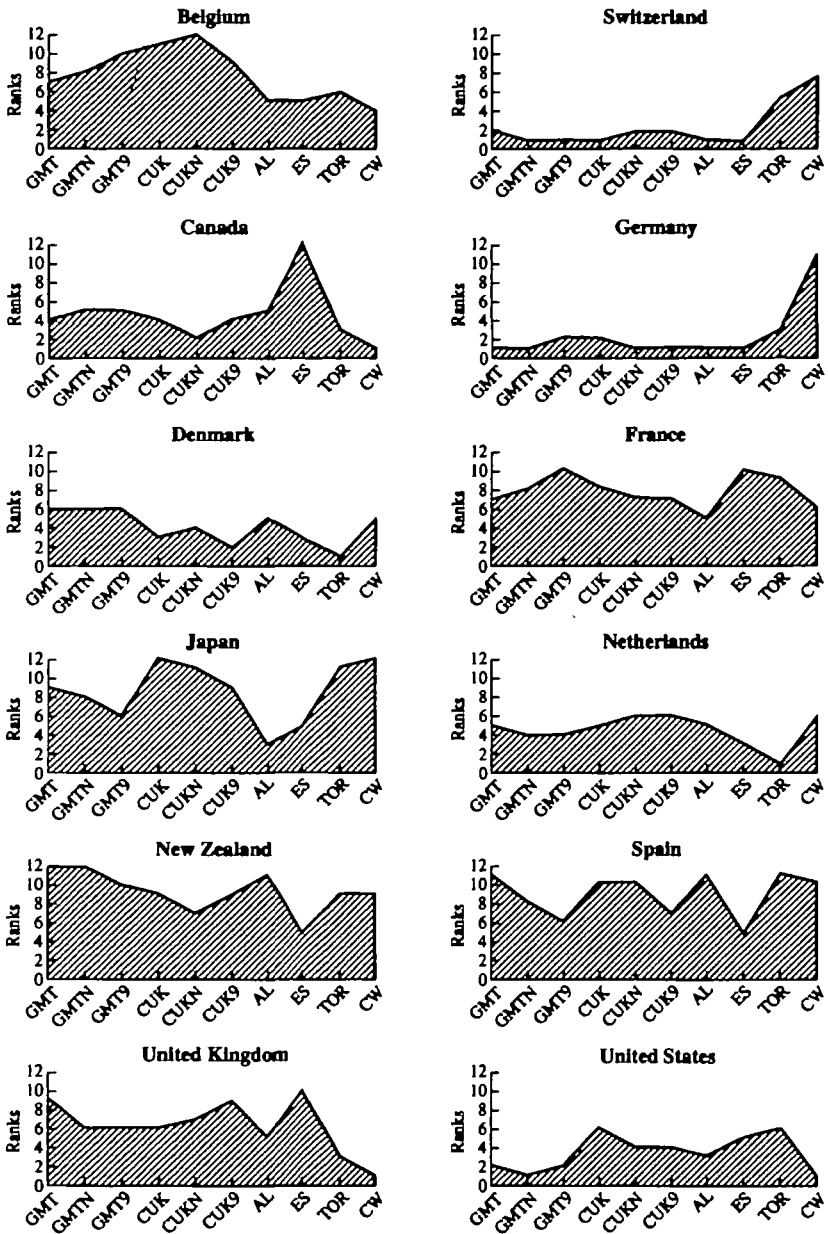


Fig. 1. CBI indices: consistency of rankings (12-country sample)

1% level of significance.²⁶ It thus seems that previously-established results on the strength of the CBI-inflation relationship cannot be confidently confirmed: the

²⁶ With 11 degrees of freedom, these rejection values are 2.201 and 3.106, respectively; see e.g. Pindyck and Rubinfeld (1991, p.563).

Table 5 CBI indices: comparison of rankings, 12-country sample

	GMT	GMTN	GMT9	CUK	CUKN	CUK9	AL	ES	TOR	CW
B	7	8	10	11	12	9	5	5	6	4
CH	2	1	1	1	2	2	1	1	6	8
CN	4	5	5	4	2	4	5	12	3	1
D	1	1	2	2	1	1	1	1	3	11
DK	6	6	6	3	4	2	5	3	1	5
FR	7	8	10	8	7	7	5	10	9	6
J	9	8	6	12	11	9	3	5	11	12
NL	5	4	4	5	6	6	5	3	1	6
NZ	12	12	10	9	7	9	11	5	9	9
SP	11	8	6	10	10	7	11	5	11	10
UK	9	6	6	6	7	9	5	10	3	1
US	2	1	2	6	4	4	3	5	6	1

Partial correlations between rankings, 12-country sample

	GMT	GMTN	GMT9	CUK	CUKN	CUK9	AL	ES	TOR	CW
GMT	1.00	0.92	0.75	0.75			0.82	0.31	0.55	0.23
GMTN		1.00	0.91		0.71		0.78	0.36	0.52	0.20
GMT9			1.00			0.73	0.62	0.45	0.38	0.00
CUK				1.00	0.94	0.87	0.52	0.27	0.73	0.21
CUKN					1.00	0.86	0.47	0.15	0.59	0.20
CUK9						1.00	0.54	0.43	0.52	0.01
AL							1.00	0.26	0.39	0.09
ES								1.00	0.09	-0.57
TOR									1.00	0.51
CW										1.00

Notes: AL is the Alesina (1988) 'extended and updated' version of the Bade and Parkin (1982) index of CBI;

ES is the Eijffinger and Schaling (1993) index of policy independence;

TOR is the average turnover rate of CB Governors (1950-1989), as computed by Cukierman (1992);

CW is the index of political vulnerability of the CB (1950-1989), as computed by Cukierman and Webb (1994).

above analysis suggests that there is a serious danger they were influenced by the significant subjectivity of the measures used to account for CBI.

Similar observations can be made about the results on the variability of inflation. In the ten relevant regressions, virtually all coefficients are of the expected (positive) sign, but only one *t*-value (with the ES ranking as the explanatory variable) allows rejection of the null hypothesis at a 5% level of significance: the evidence on the relationship between CBI and inflation variability is thus weaker than suggested by previous findings, again probably because of the subjective nature of the existing indices of CBI (and of related concepts).

The results on average growth are even more puzzling. Here the sign of nearly all coefficients suggests that, if anything, CBI is likely to have a negative effect on

Table 6 Regressions of respective rankings (12-country sample)

	(A) with average inflation, 1980-89 (measured by GDP deflator)			(B) with variability of inflation, 1980-89 (measured by variance of GDP deflator)			
	intercept	indep. var.	adj. R ²	intercept	indep. var.	adj. R ²	
GMT	2.39 (1.73)	*0.55 (2.82)	0.39	GMT	2.21 (0.66)	0.94 (2.02)	0.22
GMTN	2.80 (2.00)	*0.53 (2.49)	0.32	GMTN	3.54 (1.04)	0.80 (1.55)	0.11
GMT9	3.02 (1.86)	0.49 (1.94)	0.20	GMT9	3.29 (0.89)	0.85 (1.48)	0.10
CUK	*4.53 (2.47)	0.20 (0.79)	-0.03	CUK	6.56 (1.65)	0.24 (0.44)	-0.08
CUKN	*4.96 (2.79)	0.14 (0.55)	-0.07	CUKN	7.04 (1.84)	0.17 (0.32)	-0.09
CUK9	3.99 (2.13)	0.32 (1.10)	0.02	CUK9	2.47 (0.67)	0.98 (1.70)	0.15
AL	1.86 (2.07)	†0.79 (5.15)	0.70	AL	2.95 (0.95)	1.03 (1.93)	0.20
ES	*4.28 (2.69)	0.28 (1.14)	0.03	ES	1.50 (0.57)	*1.22 (2.95)	0.41
TOR	*4.28 (2.61)	0.27 (1.09)	0.02	TOR	7.60 (2.06)	0.09 (0.16)	-0.10
CW	†6.12 (3.57)	-0.05 (-0.21)	-0.10	CW	†12.18 (3.65)	-0.66 (-1.44)	0.09
	(C) with average real-GDP growth, 1980-89			(D) with variability of real-GDP growth, 1980-89			
	intercept	indep. var.	adj. R ²	intercept	indep. var.	adj. R ²	
GMT	†2.18 (4.90)	0.04 (0.59)	-0.06	GMT	4.62 (1.74)	0.02 (0.06)	-0.10
GMTN	†2.19 (5.13)	0.04 (0.60)	-0.06	GMTN	3.83 (1.52)	0.16 (0.42)	-0.08
GMT9	†2.29 (4.92)	0.02 (0.32)	-0.09	GMT9	2.25 (0.87)	0.44 (1.10)	0.02
CUK	†1.67 (4.43)	*0.12 (2.24)	0.27	CUK	2.67 (1.03)	0.32 (0.91)	-0.02
CUKN	†1.83 (4.74)	0.10 (1.73)	0.15	CUKN	2.26 (0.94)	0.41 (1.19)	0.04
CUK9	†1.90 (4.24)	0.09 (1.30)	0.06	CUK9	2.72 (1.00)	0.35 (0.84)	-0.03
AL	†2.51 (6.00)	-0.02 (-0.26)	-0.09	AL	4.22 (1.73)	0.11 (0.26)	-0.09
ES	†2.07 (5.26)	0.06 (1.04)	0.01	ES	4.02 (1.67)	0.14 (0.36)	-0.09
TOR	†1.79 (5.03)	0.11 (2.04)	0.22	TOR	*5.76 (2.35)	-0.17 (-0.48)	-0.08
CW	†2.34 (5.57)	0.01 (0.22)	-0.09	CW	†7.89 (3.65)	-0.51 (-1.70)	0.15

Note: *t*-stats are given in parentheses, under respective coefficient; * (†) indicates significance at the 5% (1%) level.

average growth, whereas in the few earlier studies in which any trace of such an effect was found, it carried a positive sign;²⁷ the lack of significance of most regressions, however (again, only one *t*-statistic implies significance at the 5% level), precludes unqualified confirmation of such a link.

Finally, even using rankings rather than values to describe the status of the CBs concerned, the variability of growth is not found to exhibit any significant relationship with CBI in the 12 countries of the sample.

The results of the above computations therefore tend to confirm the impression gathered in the previous section: as long as we are not sure that CBI is specified in a satisfactory way, any empirical test of its influence on other variables will lack statistical reliability (and thus credibility).²⁸ More fundamentally, one might question the very insistence on using such a controversial notion: since there seems to be no objective agreement on what independence really is, how exactly it should be measured and what precise effect it has. Why not concentrate instead on less ambiguous concepts? Such an alternative route is advocated in the next section.

4. Beyond independence: some suggestions for an alternative empirical approach

In the light of the statistics collected in Sections 2 and 3, it would be tempting to claim that previous evidence in favour of CBI was misleading, and that independence has no effect whatsoever on inflation or growth. My results, however, do not support such clear-cut implications. All they indicate is that, given the significant subjectivity spread affecting most (if not all) legislation-based indices of CBI, we cannot be sure that what was regressed against inflation or growth in earlier studies really was independence. Consequently, we do not know whether or not independence, if measured in an accurate way, would effectively exhibit a statistically-significant relationship with various macroeconomic variables.

One way out of such uncertainty would be to try and improve the way in which CBI is accounted for, i.e. to devise yet another measure which, ideally, would avoid the shortcomings of existing indices.²⁹ But this is not the task to which this section is devoted: the only support I can offer in this sense is the argument underlying Section 2.4, namely that a 'relative' approach, i.e. through rankings rather than values, probably offers less scope for subjectivity. The remainder of this section

²⁷ For a brief, yet clear exposition of the theories underlying these diverging results, see p.152 of Alesina and Summers (1993); for an elegant theoretical explanation of possibly diverging empirical results, see Alesina and Gatti (1995).

²⁸ Another issue which arises from these tests is linked to their causality implications: discovering a relationship between two variables is one thing, concluding that there is a one-way effect from the first to the second variable is quite another. A discussion of Granger causality, however, would lie beyond the scope of this paper.

²⁹ Forder (1996) is a strong proponent of the view that such an attempt is intrinsically doomed.

concentrates instead on a few suggestions in favour of an alternative route, which may yield promising results.

Independence is not a concept which academics should feel comfortable to work with. Not only do we find it quite hard to measure it objectively, bickering over its very components and their relative importance; we cannot even agree on a unique, undisputed way of defining it. Different authors mean different things when they discuss independence. In the models of Debelle and Fischer (1994) or Fratianni *et al.* (1993), it is equivalent, following Rogoff (1985), to the conservatism of the CB governor (i.e. his aversion to inflation as against real-output fluctuations); for Alesina (1988), Bade and Parkin (1982), Grilli *et al.* (1991), and many others, a distinction between political and economic independence needs to be drawn;³⁰ according to Debelle and Fischer (1994), the relevant distinction is rather between goal and instrument independence, while Cukierman (1992) prefers to speak of statutory and actual CBI.

Instead of trying to select from these alternative taxonomies, I suggest that we start with the theoretical insights of Walsh (1995) and Persson and Tabellini (1993), who first introduced notions of contract theory into the study of the institutional characteristics of monetary policy (on which more below). They argue that a CB's independence is, in reality, endogenous, a mere by-product of the incentive structure faced by the CB's governing body (Central Banker). On this view, empirical investigations should concern themselves with the determinants of incentives, rather than indulging in subjective estimations of the individual components of independence.

The idea is to examine not only the statutory factors which are likely to influence incentives (e.g., how do the CB's operational rules describe the formal tasks of the Central Banker?), but also (and crucially) on behavioural considerations, such as: what prevents the Central Banker from wanting to renege on his duties? Or, how well did he achieve what he claims he undertook? Or, what consequences does he face if his achievements do not live up to his promises?

In a sense, this shifts the focus from independence to the more tractable notion of the Central Banker's conservatism. The perspective, however, is different from that of Rogoff (1985), since the concept is introduced in an incentive-based rather than in a preference-based framework.³¹ In Rogoff (1985) or Lohmann (1992), a conservative Central Banker is defined as an individual with intrinsically asymmetric preferences, while in Walsh (1995) or Persson and Tabellini (1993), his conservatism is the endogenous result of a whole set of incentives he faces. These incentives may include statutory characteristics, such as the constitutional obligation to concentrate on price stability; but the ultimate, observed conservatism of the Central Banker is determined by the enforcement mechanism which constrains him to fulfil his obligations.

³⁰ Although I have argued elsewhere (Mangano, 1994) that it is far from clear which criteria these definitions should encompass, and what is their respective importance.

³¹ This point is made particularly clear in Walsh (1993).

The new theoretical literature on institutional aspects of monetary policy relies on principal-agent theory to formalise such a dual (legal/behavioural) framework. It shows that an optimal performance contract can be devised between the political authorities (the principal) and the Central Banker (the agent), which not only defines explicitly the latter's statutory objectives, but also provides him with the incentives to achieve these objectives by making him solely responsible for the outcome of monetary policy, and attaching well-publicised rewards or penalties to success or failure.³² The Central Banker is thus conservative, not necessarily because it is in his genes, not only because it is his duty to be so, but mainly because he knows that otherwise he will be punished.

I argue that this framework promises a better basis for empirical investigations than the usual independence-orientated framework. The following simple rules would take into account the theoretical insights just outlined, as well as the lessons learnt from the analysis of the previous sections. First, considering the major subjectivity problem uncovered in Section 2, priority should be given to straightforward, binary choices requiring no personal judgement for their application. Every selected variable should be allowed to take only two logical values, 'Yes' or 'No', and which of the two it takes in each instance should be unambiguous. Second, both institutional and behavioural aspects of the CB's situation should be included. Any element which influences the Central Banker's accountability, allows us to evaluate his performance or reveals the extent of his commitment to his official task should therefore be attributed as much importance as any other which is supposed to determine his formal duties.

Thus, at the sample-selection stage, one could choose to concentrate on institutional characteristics: the rule for inclusion might be that the attainment of some numerical objective is explicitly imposed upon the CB (i.e. a purely binary process: 'There exists an explicit objective', inclusion of the country; 'There exists no explicit objective', exclusion of the country). Once the sample is defined, one could then turn to behavioural indicators to assess each Central Banker's actual conservatism. For instance, is there a well-publicised reward/penalty system sanctioning the performance of the CB in terms of the pre-specified objective? And in the countries where such a system is in place, has the threat been systematically implemented when the (measureable) objective was not achieved (again two 'Yes/No' criteria)?

Even political influence on monetary policy-making, insofar as it also affects the observed conservatism of the CB, may be captured in this fashion. Instructive observations could include the following: after elections of the legislative or executive, was the CB more often kept in office, or replaced shortly after the election? Or,

³² The contract is optimal because, in the presence of stochastic supply shocks, it allows for the pre-determined objective to be achieved without increasing short-term output volatility. This result holds also in an asymmetric-information framework.

did reductions in short-term interest rates occur more often in the few months before such elections than at other times?³³

A possible objection to those proposals is that the theoretical literature on such performance contracts, let alone the practical application of its ideas, are rather new in the field of monetary policy-making.³⁴ As a result, one could be confronted with data-availability problems, both in the country-selection process and in the length of the observation period; the alternative empirical methodology I advocate may therefore have to wait until a larger and longer data-set is available to support it.

5. Conclusions

This study has been organised around three objectives. First, to estimate the subjectivity spread of two widely-used indices of CBI; second, to assess the robustness of established empirical results on CBI with respect to such subjectivity; and third, to recommend an alternative orientation for future empirical investigations.

Section 2 compared the indices of legal CBI computed by Grilli *et al.* (1991) and Cukierman (1992). With the help of four supplementary measures (derived from the two original ones), it revealed that: (i) on average, the authors diverged in their assessments of about 30% of the legislation consulted to construct their indices,³⁵ with the proportion rising to nearly 50% for some countries; (ii) they also disagreed markedly on which criteria should be included in an index of legal CBI, with only nine characteristics being common to both indices (out of a respective total of 15 and 16); and (iii) on the other hand, the fact that they used different weighting schemes to aggregate their preferred criteria into a synthetic index probably did not add much to their problems.

This identified subjectivity tends to undermine the findings of numerous empirical investigations, based on one or the other index, of a statistically-significant link between CBI and average inflation. The analysis of Section 3 reinforced such doubts. Instead of relying on the subjectivity-prone values of the indices, it focused on the rankings produced in a 12-country sample by these (and another four) measures; it then sought to use those rankings to explain, in turn, the average rate of inflation, the average rate of real-output growth, and the variance of these rates. While the sign of most coefficients in these regressions was consistent with previous evidence (the exception being those in which average real-output growth was the dependent variable), as many as 87.5% of the coefficients were not statistically significant. It thus appeared that former tests purporting to link an

³³ Cukierman and Webb (1994) have interesting suggestions in this respect, although they still rely on a precise valuation of each criterion examined, rather than on a 'Yes/No' classification.

³⁴ Although the theoretical and empirical literature is fast expanding; see in particular Svensson (1995), Leiderman and Svensson (1995).

³⁵ It was not possible to determine which of the two authors was wrong in each instance, but at least one of them was.

ill-defined measure of independence with various macroeconomic variables were not reliable.

On the basis of these findings and of new theoretical insights due to Walsh (1995) and Persson and Tabellini (1993), section 4 proposed a new approach to the empirical study of monetary institutions,³⁶ employing indicators of a CB's operational status³⁷ rather than direct (but biased) measures of CB independence.

By outlining major weaknesses in the prevailing empirical approach centred around CBI, and by motivating and suggesting a different perspective, this paper hopes to have made a constructive contribution to the ongoing debate on the optimal structure of monetary institutions.

Acknowledgements

This paper is based on a chapter of my PhD dissertation at the London School of Economics. I am indebted to my supervisor, Charles Goodhart, and two anonymous referees for many helpful suggestions on previous drafts of this paper, as well as to the E. Annaheim foundation (Villeneuve) and the University of Lausanne for financial support at the time of writing this paper. I also wish to thank Alvaro Almeida, Jakob Hørdér, and participants in various seminars at the London School of Economics and the University of Lausanne for constructive comments. The usual disclaimer applies.

References

- Alesina, A. (1988). 'Macroeconomics and Politics', *NBER Macroeconomics Annual*, 13–52.
- Alesina, A. and Gatti, R. (1995). 'Independent Central Banks: Low Inflation at No Cost?', *American Economic Review Papers & Proceedings*, 85, 196–200.
- Alesina, A. and Summers, L. (1993). 'Central Bank Independence and Macroeconomic Performance: Some Comparative Evidence', *Journal of Money, Credit and Banking*, 25, 151–62.
- Bade, R. and Parkin, M. (1982). 'Central Bank Laws and Inflation—A Comparative Analysis', mimeo, Department of Economics, University of Western Ontario.
- Barro, R. and Gordon, D. (1983). 'Rules, Discretion and Reputation in a Model of Monetary Policy', *Journal of Monetary Economics*, 12, 101–21.
- Bénassy, A. and Pisani-Ferry, J. (1994). 'Indépendance de la Banque Centrale et Politique Budgétaire', Working Paper No. 94–02, Centre d'Etudes Prospectives et d'Informations Internationales, Paris, June.
- Cargill, T. (1995). 'The statistical Association Between Central Bank Independence and Inflation', *Banca Nazionale del Lavoro Quarterly Review*, 193, 159–72.

³⁶ Interestingly, most CBs seem to have already adopted the principles of this alternative: they have abandoned attempts to measure, let alone control, some monetary aggregate (the definition of which is often at least as controversial as that of independence), and rely instead on various market-determined indicators to estimate the possible inflationary effect of their policies.

³⁷ A particularly attractive feature of these indicators is that far from being adversely affected by behavioural changes, they incorporate and thus endogenise any such changes.

- Cukierman, A. (1992). *Central Bank Strategy, Credibility and Independence*, MIT Press, Cambridge, MA.
- Cukierman, A. and Webb, S. (1994). 'Political Influence on the Central Bank—International Evidence', mimeo, American Political Science Association Meeting, New York, August.
- Debelle, G. and Fischer, S. (1994). 'How Independent Should a Central Bank Be?', mimeo, CEPR (Stanford)/FRBSF Conference on Monetary Policy in a Low Inflation Regime, San Francisco, CA, March.
- Eijffinger, S. and de Haan, J. (1996). 'The Political Economy of Central Bank Independence', mimeo, Center for Economic Research, Tilburg University.
- Eijffinger, S. and Schaling, E. (1993). 'Central Bank Independence in Twelve Industrial Countries', *Banca Nazionale del Lavoro Quarterly Review*, 184, 49–89.
- Eijffinger, S. and van Keulen, M. (1995). 'Central Bank Independence in Another Eleven Countries', *Banca Nazionale del Lavoro Quarterly Review*, 192, 39–83.
- European Commission (1991a). 'Protocol on the Statute of the European System of Central Banks and of the European Central bank', *Maastricht Treaty on European Union*, European Council, Brussels, 149–171.
- European Commission (1991b). 'Title VI: Economic and Monetary Policy', *Maastricht Treaty on European Union*, European Council, Brussels, 25–44.
- Forder, J. (1996). 'Can We Test for Benefits of Central Bank Independence?', mimeo, Institute of Economics and Statistics, Oxford University.
- Fratianni, M. von Hagen, J. and Waller, C. (1993). 'Central Banking as a Political Principal-Agent Problem', *Discussion Paper No 752*, CEPR, London, January.
- Grilli, V., Masciandaro, D. and Tabellini, G. (1991). 'Political and Monetary Institutions, and Public Finance Policies in the Industrial Countries', *Economic Policy*, 13, 341–92.
- Leiderman, L. and Svensson, L. (1995). *Inflation Targets*, Centre for Economic Policy Research, London.
- Lohmann, S. (1992). 'The Optimal Degree of Commitment: Credibility vs. Flexibility', *American Economic Review*, 82, 273–86.
- Mangano, G. (1994). 'Endogenising Central Bank Independence: An Appraisal of the Relevance of Relative Openness', mimeo, Graduate Institute of International Studies, Geneva, September.
- Persson, T. and Tabellini, G. (1993). 'Designing Institutions for Monetary Stability', *Carnegie-Rochester Conference Series on Public Policy*, 39, 53–84.
- Pindyck, R. and Rubinfeld, S. (1991). *Econometric Models and Economic Forecasts*, McGraw-Hill, New York, NY.
- Posen, A. (1994). 'Central Bank Independence and Disinflationary Credibility: a Missing Link?', *Brookings Discussion Papers in International Economics*, 109, August.
- Rogoff, K. (1985). 'The Optimal Degree of Commitment to an Intermediate Monetary Target', *Quarterly Journal of Economics*, 1169–89.
- Svensson, L. (1995). 'Optimal Inflation Targets, "Conservative" Central Bankers, and Linear Inflation Contracts', Discussion Paper No. 1249, CEPR, London, October.
- Walsh, C. (1993). 'Central Bank Strategy, Credibility and Independence: a Review Essay', *Journal of Monetary Economics*, 32, 287–302.
- Walsh, C. (1995). 'Optimal Contracts for Central Bankers', *American Economic Review*, 85, 150–67.