Monetary and Fiscal Policy Interaction With Various Degrees of Commitment[1]

Jan Libich[2]
La Trobe University and CAMA

Andrew Hughes Hallett[3]
George Mason University and University of St Andrews

Petr Stehlík[4]
University of West Bohemia

Abstract

Well before the global financial crisis, the stance of fiscal policy in a number of countries had raised concerns about risks for the outcomes of monetary policy. To provide some insights this paper examines the fiscal-monetary interactions in a novel game theory framework with asynchronous timing of moves. It generalizes the standard commitment concept of Stackelberg leadership by making it dynamic: it allows policies to be committed or rigid for different periods of time. We find that socially inferior medium-term monetary outcomes can occur due to fiscal spillovers. The bad news is that, unlike under static commitments, this may happen even if monetary policy acts as the leader for longer periods than fiscal policy. The good news is that appropriate institutional design of monetary policy may not only help the central bank resist fiscal pressure, but also discipline ambitious governments. Strong monetary commitment may therefore induce a reduction in the average size of the budget deficit and debt. The analysis implies that monetary policy in the United States, Switzerland, Japan, Eurozone, and other countries should be committed more explicitly to a numerical inflation target, and that this can improve the medium and long-term outcomes of both monetary and fiscal policy.

Keywords: commitment; monetary vs fiscal policy interaction; asynchronous games; Battle of the Sexes; explicit inflation targeting; **JEL classification**: E63, C73

---

[1]We are grateful to Don Brash, the Governor of the Reserve Bank of New Zealand during 1988-2002, for valuable views and for contributing the case study in Section 7. We would also like to thank Iris Claus, Damien Eldridge, Frank Hespeler, Jeff Sheen, Lawrence Uren, and the participants of several conferences for useful suggestions and comments. We gratefully acknowledge financial support from the Australian Research Council (DP0879638). The usual disclaimer applies.

[2]Corresponding author: La Trobe University, School of Business, Melbourne, Victoria, 3086, Australia. Phone: +61 3 94792754, Email: j.libich@latrobe.edu.au.

[3]School of Public Policy, MS 3C6, George Mason University Fairfax, VA 22030, USA, phone: 7039939123, fax: 7039932284, email: ahughesh@gmu.edu.

[4]University of West Bohemia, Univerzitní 22, Plzeň, 30614, Czech Republic. Phone: (+420) 777846682, Fax: (+420) 377632602, Email: pstehlik@kma.zcu.cz.

1. Introduction

Consider the following situation. A fiscally prudent political party makes the claim that interest rates and inflation would be higher under a 'less-prudent' rival party. The less-prudent party argues that the claim is misplaced since the country has an independent and responsible central bank solely in charge of interest rates and inflation. Which party is right? And under what circumstances?

This is far from a hypothetical situation, and it highlights the importance of understanding the interaction of fiscal (F) and monetary (M) policy on outcomes of *both* policies. The idea that M and F policies might interact goes back to Friedman (1948), Tinbergen (1954), Mundell (1962) and Cooper (1969). But until recently most models used for policy design treated each policy in isolation. The subsequent literature has mainly examined direct institutional interventions - the ability of the government (F policymaker) to affect M policy outcomes through the appointment of the central banker (Rogoff (1985)), optimal contract with the central banker (Walsh (1995)), or through overriding the central banker (Lohmann (1992)).

The focus of this paper is indirect interaction which is more subtle and less well understood.[5] It works through spillovers of economic outcomes – variables such as inflation, output, debt, exchange rate, asset prices, agents' expectations, or consumer confidence are all affected by both policies, and they in turn affect the optimal setting of both policies. Most obviously, excessive government spending commonly leads to a temporarily higher output and subsequent inflationary pressures that the central bank has to deal with.

Let us stress from the outset that our interest lies in *medium-run* outcomes of the interaction - averages over the business cycle. We will not examine the optimal mix of policy responses to a shock such as the global financial crisis. Our perspective follows Sargent and Wallace (1981), Alesina and Tabellini (1987), Nordhaus (1994) and the subsequent literature, as well as current debates about F sustainability, eg Leeper (2010).

Our analysis contributes to both macroeconomic policy and game theory. On the game theory front, we develop a novel framework with generalized timing featuring asynchronous moves. We show that the conventional conclusions made under the standard commitment concept of Stackelberg leadership, which is static, may not be robust. This highlights the importance of incorporating the time dimension into the timing of policy actions, ie using a dynamic commitment concept. On the policy front, our paper shows that concerns about fiscal excesses spilling over to monetary policy may be justified. To offer some recommendations we formally model, using the asynchronous structure of the game, how institutional remedies can prevent such spillovers under some circumstances, and ensure optimal outcomes not only for M policy, but also for F policy. In doing so we highlight the differences between a single country and a monetary union setting in which free riding of small members can occur.

[5] Examples of this type of indirect interaction appear in Leeper (2010), Cochrane (2009), Adam and Billi (2008), Eusepi and Preston (2008), Chadha and Nolan (2007), Persson et al. (2006), Benhabib and Eusepi (2005), Gali and Monacelli (2005), Hughes Hallett and Weymark (2005), Eggertsson and Woodford (2004), Dixit and Lambertini (2003, 2001), van Aarle, et al. (2002), Leeper (1991) or Sargent and Wallace (1981).

Game Theoretic Representation. The policy spillovers can be modeled effectively using game theoretic techniques since the F-M interactions are strategic in nature. Consider therefore two medium-run options for each policy: *discipline*, D, delivering the socially optimal levels on average; and *indiscipline*, I, delivering the discretionary levels that however are socially inferior. In terms of M policy, D and I can be interpreted as average low vs higher inflation.[6] In terms of F policy, D and I can be interpreted as running a balanced budget vs a structural deficit on average across the cycle. Importantly, the fiscal balance must incorporate intertemporal (eg demographic) considerations as well as the expected value of potential government liabilities future outlays (such as those arising from guarantees for financial institutions).

The payoff matrix below summarizes the policy interaction using a stylized 2x2 game theoretic representation. The variables $\{a, b, c, d, v, w, y, z\}$ denote the payoffs that are functions of the deep parameters of the underlying macroeconomic model.[7]

(1)

		F	
		D	I
M	D	$a = 1, v$	$b = 0, w = 0$
	I	$c = -1, y = -1$	$d, z = 1$

A number of scenarios are likely to occur in real world countries. We are primarily interested in the actions of a *responsible* central bank facing an *ambitious* government. In reduced form, a responsible central bank can be thought of as targeting low inflation and the natural rate of output, whereas an ambitious government does not mind over-stimulating the economy beyond its natural output rate.[8]

A responsible central bank is characterized in (1) by $a > \max\{b, c, d\}$. This, in combination with our payoff normalization, implies $d < 1$, and the fact that the central bank's preferred outcome is the socially optimal (D, D).[9] In contrast, an ambitious government can be defined by $z > \max\{v, w, y\}$, implying $v < 1$ and the fact that the government's preferred outcome is the socially inferior (I, I). Intuitively, the government prefers to spend excessively and/or to avoid unpopular reforms to secure votes, and likes the central bank to inflate some of the ensuing debt away similarly to Sargent and Wallace (1981).

[6]Our medium-run focus implies that D and I should be interpreted as setting *average* levels - having short term fluctuations around those levels is consistent with our setting (under zero mean shocks) and can be easily added.

[7]In order to keep the focus on the big picture of the policy interaction and our game theoretic insights - that are *not* model specific - we do not discuss a macroeconomic model in the main text. Appendix F lays out a simple macro model and offers an example of how analytically tractable models of policy interaction can be mapped into such a 2x2 game using the approach of Cho and Matsui (2005).

[8]As a driving force of the government's ambition one can think of an attempt to get re-elected (in the presence of naïve voters, lobby groups, unions etc). Alternatively, excessive spending may be 'unintentional' - caused by inherited F settings that automatically tip the budget into a deficit (eg unaffordable welfare/health/pension schemes, high debt, or liabilities implied by public guarantees for financial institutions).

[9]As usual, the first payoff refers to the row player (M), and the second to the column player (F). While the normalization in (1), $a = z = 1, b = w = 0, c = y = -1$, is done for parsimony and without loss of generality, we will also report results for the general payoffs without this normalization.

Due to such different preferences there exists a potential conflict between M and F policy. We can then observe the following three scenarios that differ in the probability that F excesses spill over onto M policy. First, it is the *Neglect scenario* in which $d \in (0,1), v < 0$, and thus (I,I) is the unique Nash equilibrium. Spillovers will surely occur. Second, we have the *Tug-of-war scenario* whereby $d < 0, v < 0$, and hence the (D,I) outcome is the unique Nash. Spillovers will not occur in the medium-run.[10] Third, there is a *Battle of the Sexes scenario* in which

$$d \in (0,1) \quad \text{and} \quad v \in (0,1)\,. \tag{2}$$

It features two pure Nash equilibria, (D,D) and (I,I), each preferred by a different player, and a mixed strategy Nash that is Pareto inferior to both pure Nash. In this case F spillovers onto M policy may or may not occur; in fact the outcome of F policy itself is uncertain.

Following the seminal work of Sargent and Wallace (1981) our attention will primarily be directed at the Battle of the Sexes scenario for three reasons. First, it is the most interesting scenario from the game theoretic point of view as there are equilibrium selection problems. Specifically, neither standard nor evolutionary game theoretic techniques can provide a clear choice between the pure Nash equilibria due to the symmetry of the game. Note that it is also the only scenario of the three in which the timing of the actions determines the equilibrium. Under the standard commitment concept the Stackelberg leader will ensure his preferred outcome, whereas in the other two scenarios leadership does not alter the set of possible equilibria.

Second, the game features both a *conflict* (to secure the preferred pure Nash equilibrium) and a *coordination problem* (to avoid the inferior mixed Nash). These two characteristics seem to occur in many real world cases as well as in a wide range of policy interaction models following Sargent and Wallace (1981): see Leeper (2010), Adam and Billi (2008), Branch, et al. (2008), Resende and Rebei (2008), Benhabib and Eusepi (2005), Dixit and Lambertini (2003) and (2001), Blake and Weale (1998), Nordhaus (1994), Sims (1994), Woodford (1994), Leeper (1991), Wyplosz (1991), Petit (1989) or Alesina and Tabellini (1987). The intuition of our findings will therefore apply to any of these diverse settings.

Third, the results derived in the Battle of the Sexes will imply analogous results for many alternative scenarios including those of a responsible government (such as a *Policy Symbiosis*), or those under an ambitious central bank (such as the *Prisoner's Dilemma* and the *Game of Chicken*). We discuss these scenarios in Section 6.[11]

[10] Nevertheless, the (D,I) outcome cannot obtain in the (very) long-run as the government's intertemporal budget constraint has to hold. In an important recent body of work, Davig and Leeper (2010) examine the combination of (I,I) and (I,D) that replaces the (D,I) outcome when the economy approaches/hits its fiscal limit.

[11] The M-F interactions have often been modelled as the Game of Chicken, eg Barnett (2001), Bhattacharya and Haslag (1999), Artis and Winkler (1998), and Alesina and Tabellini (1987). It is similar to the Battle of the Sexes in that it also features two pure Nash equilibria, (D,I) and (I,D), each preferred by a different player, and one inferior Nash in mixed strategies. Nevertheless, such an anti-coordination game cannot arise under a responsible central bank since it has no structural temptation to inflate if the government is disciplined over the medium-term.

Dynamic Commitment. We examine policy interactions in an *asynchronous game* that generalizes the alternating move games of Maskin and Tirole (1988) and Lagunoff and Matsui (1997).[12] In this framework, the players may not necessarily move every period in a simultaneous fashion, nor every other period in an alternating fashion. Instead, after a synchronized initial move in period $t = 1$, each player i moves with a certain *constant frequency*, namely every $r^i \in \mathbb{N}$ periods. Figure 1 in Section 2 offers an illustration of such timing.

This deterministic framework captures the observation of Tobin (1982) that '*Some decisions by economic agents are reconsidered daily or hourly, while others are reviewed at intervals of a year or longer...'*, and follows Tobin's call: *...'It would be desirable in principle to allow for differences among variables in frequencies of change...'*.

The variable r^i can be interpreted as the degree of commitment *or* rigidity of player i. These two concepts are formally identical in our framework, both referring to the players' inability to move. Nevertheless, in the real world such inability comes from different sources, which we will acknowledge by referring to r^M as *long-term M commitment* (as it is performed by a responsible central bank), and to r^F as *F rigidity* (as it is performed by an ambitious government). Our medium-term setup implies that r^M should *not* be interpreted as the frequency of the central bank's interest rate decisions, because it does not restricts the ability to make period by period policy decisions. Instead, r^M describes the ability to change the parameters that guide those decisions, most importantly the target for average inflation.[13]

The specification implies that our dynamic commitment concept is a natural generalization of the Stackelberg leadership concept. It also implies that the *stage game* of our asynchronous game is itself an extensive-form game lasting T periods, where $T \in \mathbb{N}$ is the 'least common multiple' of r^M and r^F. For instance, the dynamic stage game in Figure 1 with $r^M = 5$ and $r^F = 3$ is $T = 15$ periods long. We will throughout use the fact that if a *Pareto-efficient* outcome *uniquely* obtains on the equilibrium path of the dynamic stage game, one can ignore its further repetition without loss of generality.[14]

Findings. The standard static commitment provides us with the following policy prediction in the Battle scenario. If M is the committed player (Stackelberg leader) then the central bank will ensure its preferred and socially optimal (D, D) outcome - 'win' the Battle.[15] Note that there are no caveats, no contingencies, no strings attached. Our contribution lies in refining and in some aspects qualifying this conventional result

---

[12]The existing game theoretic work provides a strong motivation for our general approach. For example, Cho and Matsui (2005) argue that: *'[a]lthough the alternating move games capture the essence of asynchronous decision making, we need to investigate a more general form of such processes. . . '*.

[13]Arguably, this depends on how *explicitly* the target is grounded in the central banking statues or legislation. The same is true for r^F - it expresses the degree with which various fiscal schemes and settings that lead to excessive (welfare/health/pension) spending are grounded in the legislation. Let us note that if the roles were reversed the terminology would be too: under a responsible government and an ambitious central bank r^F and r^M would denote F commitment and M rigidity respectively.

[14]Repetition is commonly used to help alleviate inefficiency and enhance cooperation through reputational channels, see eg Mailath and Samuelson (2006), or in the monetary context Barro-Gordon (1983). The advantage of focusing on the dynamic stage game is that it provides the worst case scenario in which reputation cannot help in cooperation.

[15]Such situation has been termed dominant monetary policy in Sargent and Wallace (1981), active fiscal/passive monetary policy in Leeper (1991), or a Ricardian regime in Woodford (1995).

by allowing for dynamic commitment/rigidity of both players. We find that the insights obtained under standard commitment, ignoring the time dimension, are not robust and may in fact be misleading. The picture looks bleaker for M policy than what the static commitment concept would have us believe.

For example we demonstrate that, under dynamic commitment, there are situations where M policy is 'likely' to lose the Battle even if its commitment is stronger than F rigidity: $r^M > r^F$. Furthermore, we show that under some circumstances, namely if the central bank discounts the future heavily, $\delta_M < \overline{\delta_M}\left(r^F, d\right)$, then M policy's preferred outcome will not be guaranteed even if it is *infinitely strongly* committed, $r^M \to \infty$. This means that central bank independence and an explicit inflation target may not be sufficient on their own for long-term price stability and credibility. As Sargent and Wallace (1981) put it: *'...Friedman's list of the things that monetary policy cannot permanently control may have to be expanded to include inflation.'*

To safeguard its credibility and the preferred D level in the Battle scenario, M policy has to be both: (i) *sufficiently patient*: satisfy the necessary condition $\delta_M > \overline{\delta_M}$; and also (ii) *sufficiently strongly committed* relative to F rigidity: satisfy the necessary and sufficient condition $r^M > \overline{r^M}(r^F, \delta_M, \delta_F, d, v) > r^F$. This is implied by the asynchronous structure of the game where each player can reconsider its medium-run stance only once in so many periods. If the frequency is sufficiently different for each player, then the one with the longer fixed horizon can force the opponent to comply with his preferred stance.

To highlight the intuition, let us take the simplest case of $\frac{r^M}{r^F} = 2$, in which the dynamic stage game features one move of M and two moves of F (one revision). If M plays D and F starts with I, then the central bank suffers a 'conflict cost', payoff b. But this will only last for r^F periods, after which the government gets to revise its action and play its static best response D. This rewards the central bank with the a payoff for the rest of the stage game, ie for $\left(r^M - r^F\right)$ periods. Hence for the central bank to be willing to undergo a costly conflict with the government, the payoff from playing D, namely $br^F + a\left(r^M - r^F\right)$, has to be greater than the payoff from surrendering from the start and playing I, which is dr^M. In other words, M commitment must be sufficiently high relative to F rigidity and various factors influencing the bank's payoffs such as the structure of the economy.

Interestingly, if these conditions are satisfied, then M policy commitment can indirectly discipline F policy in the Battle scenario, and achieve the socially optimal D outcomes for *both* policies throughout the medium and long-run. Intuitively, if the inflation target is explicitly stated in the statutes or related legislation with the central banker being held accountable for achieving it, the government knows that the central banker is willing to engage in a costly tug-of-war, and would fully counter-act the excessive F actions by a strong M tightening. As this would eliminate any political gain, the government's incentive to engage in excessive F actions or avoid tough F reforms fades away - leading to an improvement in the budget and debt.[16]

[16] In relation to that, Section 7 presents a short case study written by Dr Don Brash, Governor of the Reserve Bank of New Zealand during 1988-2002. His contribution describes the developments in New Zealand shortly after the adoption of an explicit commitment to a low-inflation target, and highlights

There are two important caveats to this conclusion. First, while such disciplining effect of monetary commitment on fiscal policy can happen in the Battle, it can never occur in the Tug-of-war and Neglect scenarios. This is because the government is too ambitious in those cases, $v < 0$, and I is its strictly dominant strategy in the normal-form game. In such cases, dynamic commitment does not alter the outcomes of the game, and the socially optimal (D, D) outcome cannot obtain even if M is both fully patient, $\delta_M = 1$, and infinitely strongly committed, $r^M \to \infty$.

Moral Hazard in a Monetary Union. The second caveat is a membership in a currency union. We show formally in Section 5 how accession to a currency union may be associated with a free-riding problem. Intuitively, if a small member country engages in F indiscipline, its impact on the inflation outcomes of the union as a whole is small. Because of that, the M punishment by the common central bank will be of a much smaller magnitude. Further, it will be spread across all member countries, Masson and Patillo (2002). Therefore, if a small and fiscally ambitious country does not internalize the cost it imposes on others, it tends to spend excessively - more so than before joining the union when it had to bear the full weight of its own central bank's punishment. Recent events in Europe have provided a number of examples of that sort of problem. Note that this constitutes a different type of moral hazard to the one commonly discussed, which is relying on a bailout by the rest of the members, although that may have to follow.

In regards to both caveats, as M commitment cannot discipline the F policymakers, there is need for alternative arrangements that *directly* commit F policy in the long-term and anchor F expectations, as argued convincingly by Leeper (2010) and many others. In fact, our analysis implies that such a transparent and accountable medium-long-term F commitment is desirable in *all* scenarios: institutionalizing good policy provides an insurance against future excessively ambitious governments, and better guides expectations on the transition path.

2. Dynamic Commitment

This section postulates an asynchronous game framework to allow for various degrees of M commitment and F rigidity.[17] Our goal is to examine how the medium-run macroeconomic outcomes of the policy interaction depend on these variables.

2.1. Assumptions and Notation.

For maximum comparability our framework adopts all the assumptions of a standard repeated game. First, commitment and rigidity r^i are exogenous and constant throughout each game. Second, they are common knowledge. Third, all past periods' moves can be observed. Fourth, the game starts with a simultaneous move. Fifth, players are rational, have common knowledge of rationality, and for expositional clarity they have complete information about the structure of the game and the opponent's payoffs. These assumptions can easily be relaxed. They are introduced here so that the only difference from the standard repeated game is in allowing $r^i > 1$ values that differ across players.

---

the disciplining effect this M arrangement has had on F policymakers. Empirical evidence on this effect is presented in Franta, Libich, and Stehlík (2010) and discussed below.

[17] For more details on the game theoretic aspects of the framework see Libich and Stehlík (2010a).

Denoting n^i to be the i's player's n'th move (not period), and N^i the number of moves in the asynchronous stage game, it follows that $N^i = \frac{T(r^M, r^F)}{r^i}$. Also, M_n^l and F_n^l will denote a certain action $l \in \{D, I\}$ at a certain node n^i; eg F_2^I refers to F's indiscipline in her second move.

For the rest of this section we assume some $r^i > r^j$, where $i \in \{M, F\} \ni j$. We can then denote $\frac{r^i}{r^j} \geq 1$ to be the players' *relative* commitment/rigidity. Also, $\left\lfloor \frac{r^i}{r^j} \right\rfloor \in \mathbb{N}$ will be the integer value of relative commitment (the floor) and

$$R = \frac{r^i}{r^j} - \left\lfloor \frac{r^i}{r^j} \right\rfloor = [0, 1), \tag{3}$$

denotes the fractional value of relative commitment (the remainder).[18] It will be evident that R plays an important role as it determines the exact type of asynchrony in the game. Note that if $R > 0$ both players take the leadership role during the stage game.

Further, we denote $B(.)$ to be the best response. For example, $F_1^D \in B(M_1^D)$ expresses that F_1^D is F's best response to M's initial D move, and $\{F_1^D\} = B(M_1^D)$ expresses that it is the unique best response. Thus $F_1^* \in B(M_1)$ expresses that F's optimal play in move 1 is the best response to M's first move.

2.2. **Recursive Scheme.** The fact that we will be able to present proofs for general values of r^i's is due to the existence of a recursive scheme in the moves. Formally, let k_n be the number of periods between the n^i-th move of player i and the immediately following move of player j (for a graphical demonstration see Figure 1). Using this notation we can summarize the recursive scheme of the game as follows:

$$k_{n+1} = \begin{cases} k_n - Rr^j & \text{if} \quad k_n \geq Rr^j, \\ k_n + (1-R)r^j & \text{if} \quad k_n < Rr^j. \end{cases} \tag{4}$$

Generally, k_n is a non-monotonic sequence.

2.3. **History, Future, Strategies, and Equilibria.** By convention, history in period t, h_t, is the sequence of actions selected prior to period t. And the future in period t is the sequence of current and future actions. It follows from our perfect monitoring assumption that h_t is common knowledge at t. Let us refer, following Aumann and Sorin (1989), to moves in which a certain action $l \in \{D, I\}$ is selected for all possible histories as *history-independent*.

A strategy of player i is a vector that, $\forall h_t$, specifies the player's play $\forall n^i$. The asynchronous game will commonly have multiple Nash equilibria. To select among these we will use a standard equilibrium refinement, subgame perfection, that eliminates non-credible threats. Subgame perfect Nash equilibrium (SPNE) is a strategy vector (one strategy for each player) that forms a Nash equilibrium after any history h_t.

Given the large number of nodes in the game we focus on the equilibrium path of the stage game SPNE, ie the actions that actually get played. In doing so we will use the following terminology regarding two symmetric types of SPNE we are interested in.

[18] For example, in Figure 1 we have $\left\lfloor \frac{r^M}{r^F} \right\rfloor = \left\lfloor \frac{5}{3} \right\rfloor = 1$, and $R = \frac{2}{3}$.

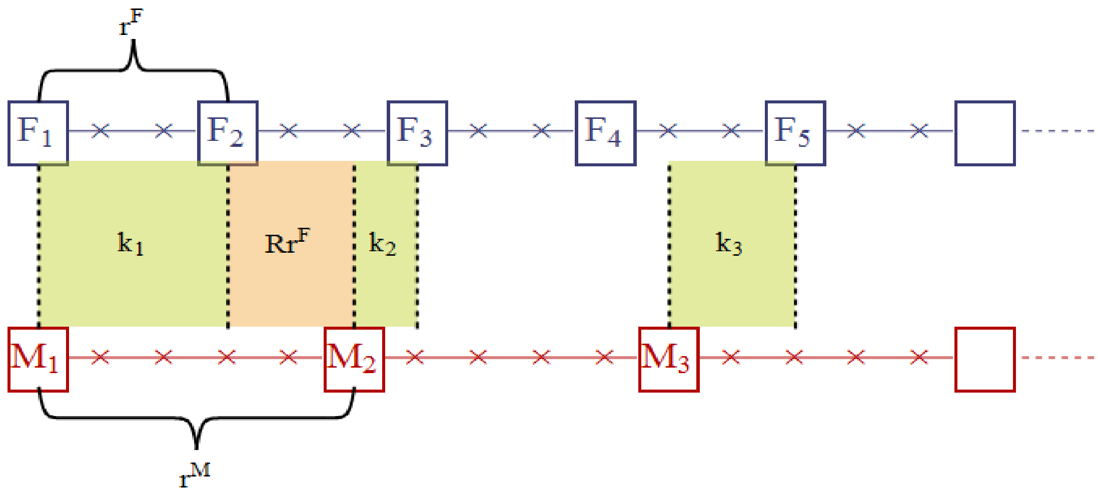

Figure 1. An asynchronous stage game with deterministic commitment/rigidity: illustration of the recursive scheme and the concepts of r^i, R, k and n^i under $r^F = 3, r^M = 5$.

Definition 1. *Any SPNE of the asynchronous stage game that has, on its equilibrium path, both policymakers playing D in all their moves* $(i_n^{D*}, \forall n, i)$ *will be called* ***Disciplined SPNE.*** *Those SPNE with I in all their equilibrium moves* $\left(i_n^{I*}, \forall n, i\right)$ *will be called* ***Indisciplined SPNE.***

Definition 2. *Player i will be said to (i)* ***surely-win*** *the Battle (and the opponent j to* ***surely-lose****) if all SPNE of the asynchronous stage game are of i's preferred type; and (ii)* ***likely-win*** *(and j* ***likely-lose****) if there exists at least one SPNE of i's preferred type, and there exists no SPNE of j's preferred type.*

2.4. **(Non)-Repetition.** As our focus is on conditions under which an efficient outcome uniquely obtains on the equilibrium path of the asynchronous stage game, its further repetition can be disregarded without loss of generality. Intuitively, if the effective minimax values of the players in the dynamic stage game (that are the infima of the players' subgame perfect equilibrium payoffs, see Wen (1994)) are unique and Pareto-efficient, then the effective minimax values of the repeated game (with any finite or infinite number of repetitions) will also be the same. Put differently, the set of Pareto-superior payoffs is empty as we are already on the Pareto-frontier. The uniqueness property also implies that we can only focus on pure strategies without loss of generality.

3. Results in the Battle Scenario Without Discounting

It is important to realize that the Battle of the Sexes type interaction featuring both a coordination problem and a policy conflict arises in *some* macroeconomic models under *some* parameter values - some alternative scenarios are discussed in Section 6. That section also discusses the mapping between the deep parameters of a macroeconomic model and our game theoretic representation.

The justification for why even a benevolent and responsible central bank may choose debt monetization and deviate from the socially optimal inflation level, ie $d > b$, may differ across the macroeconomic settings with each model potentially offering a different explanation. For example, Sargent and Wallace's (1981) unpleasant monetary arithmetic requires the central bank to generate seigniorage revenues to prevent the government's default. A parallel avenue is Leeper (1991) and the Fiscal Theory of the Price Level literature, where M policy is forced to be passive by an active F policy resulting in permanent changes in M responses to F shocks. Another explanation is that, if there are frictions in the economy and the policy instruments are substitutes in affecting output, they may be used according to comparative advantage to minimize the various distortions (Hughes Hallett (1986)). That might lead the central bank to deviate from D (see eg Adam and Billi (2008) or Resende and Rebei (2008)).

To develop the intuition of the game theoretic framework this section now reports results for (i) the normalized Battle payoffs in (1)-(2), and (ii) under fully patient players with discount factors $\delta_M = \delta_F = 1$. Both restrictions will be relaxed in Section 4.

Proposition 1. *Consider the Battle scenario without discounting in which (1)-(2) hold. (i)* ***(Disciplined SPNE)*** *M surely-wins the game and F surely-loses if and only if M commitment is* ***sufficiently strong*** *relative to F rigidity*

$$r^M > \overline{r^M}\left(\underset{+}{r^F}, \underset{+}{d}, \underset{+}{v}\right) > r^F. \tag{5}$$

(ii) ***(Indisciplined SPNE)*** *M surely-loses and F surely-wins iff M commitment is* ***sufficiently weak*** *relative to F rigidity*

$$r^M < \underline{r^M}\left(\underset{+}{r^F}, \underset{-}{d}, \underset{-}{v}\right) < r^F. \tag{6}$$

Proof. See Appendix A. □

Figure 2 summarizes these results graphically. The $\frac{r^M}{r^F}$ space is expanded compared to a one shot (or simultaneously repeated) game where $\frac{r^M}{r^F} = 1$. The space can be broken into three main regions, in which there are: (i) only the Disciplined type of SPNE - M's sure-win, (ii) only the Indisciplined type of SPNE - F's sure-win, (iii) multiple types of SPNE - neither player's sure-win as both D and I occur on the equilibrium path for one or both policies. Note that, in contrast to the static concept of commitment, our framework gives us additional valuable information. Specifically, it tells us the exact degree of M commitment that is required - as a function of several variables. In particular, $\overline{r^M}$ is increasing in r^F, d, and v, which can be traced back to the parameters of the particular underlying model - see Section 6.

Section 1 discussed the intuition of this result using $\frac{r^M}{r^F} = 2$ (which is a special case of $R = 0$). For the central bank to surely-win, it must be willing to engage in a costly tug-of-war; ie the following incentive compatibility condition for $B(F_1^I) = \{M_1^D\}$ has to

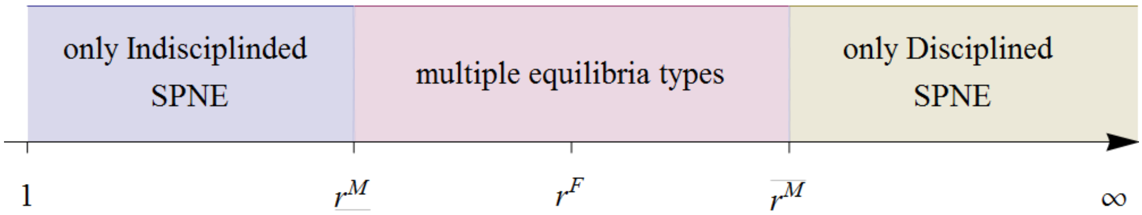

FIGURE 2. The r^M space featuring the thresholds and regions of SPNE.

hold

$$\underbrace{br^F}_{\substack{(D,I)\\ \text{conflict cost}}} + \underbrace{a(r^M - r^F)}_{\substack{(D,D)\\ \text{victory reward}}} > \underbrace{dr^M}_{\substack{(I,I)\\ \text{surrender payoff}}}. \tag{7}$$

Rearranging this and using (1)-(2) yields the following threshold

$$r^M(R=0) > \overline{r^M}(R=0) = \frac{r^F}{1-d}. \tag{8}$$

If this condition is satisfied then the victory reward more than offsets the initial conflict cost, and hence M is not willing to accommodate excessive F policy. Such determination to fight if necessary eliminates the incentive of F to run structural deficits and accumulate debt as they would not lead to any boost in output or other political gain. In other words since the M threat is credible there is in fact no fight in equilibrium.

This situation can be thought of as the case of dominant M policy regime of Sargent and Walace (1981); Leeper's (1991) active M and passive F policy; or Ricardian regime in Woodford (1995). It can therefore be concluded that, in the Battle scenario, such a sufficiently strong M commitment is not only capable of shielding the central bank from F pressure and spillovers, but also able to discipline F policy by improving the government's incentives and equilibrium play. Section 7 presents a short case study by Dr Don Brash documenting that this actually happened in New Zealand. He argues that adoption of a stronger M commitment gave him as Governor more ammunition to stand-up to excessive F policy, and that this in turn has had a disciplining effect on F policymakers.

In contrast, if $r^M \leq \overline{r^M}$ then the victory reward is insufficient to compensate M for the initial conflict cost, and hence it fails to guarantee M's sure-win. The threat of offsetting M policies is no longer fully credible. If also $r^M < \underline{r^M} < \overline{r^M}$, then the outcomes are reversed since it is now F who is willing to engage in a costly tug-of-war with M. Because of that, it is the government who surely-wins the Battle and M surely-loses. This is comparable to the dominant F regime in Sargent and Wallace (1981); the accommodating M policy in Sims (1988); active M/passive F policy in Leeper (1991); or a non-Ricardian regime in Woodford (1995).

The proof in the Appendix shows that the nature of this special case $R = 0$ carries over to the more asynchronous cases $R > 0$. This may appear surprising because in the latter case both players take the role of the leader during the dynamic stage game. For example in Figure 1 there are four changes in leadership (in F_2, M_2, F_3, and M_3), and because of that there are multiple periods of potential policy conflict with the decisions about them intertwined. The intuition for this result is twofold. First, for any R the player with lower r^i makes the last revision, which can be used by the opponent. Second, the most 'important' action happens in the initial simultaneous move since the conflict cost would last the longest relative to the victory reward. This move therefore yields to strongest (and hence sufficient) incentive compatibility condition. Formally proving this result - which is not obvious by any means - is one of the contributions of this paper.

The findings are in contrast to those under standard Stackelberg commitment, whereby the leader (committed player) wins the game independently of any structural or policy parameters. The results of Proposition 1 can be viewed as a refined version of the conventional result, and therefore a richer basis for policy recommendations. They also offer an explanation for the observed institutional differences across countries.

It is straightforward to show that in the intermediate region $r^M \in \left[\underline{r^M}, \overline{r^M}\right]$ there are either: (i) both Disciplined and Indisciplined SPNE, or (ii) only one of these two types, or (iii) neither of them, in which case in all SPNE both D and I occur on the equilibrium path - for one or both policies. This implies that in this region the *variability* of (trend) inflation and debt is commonly higher than under $r^M > \overline{r^M}$. It is here that we will observe (policy) cycles in the outcomes in some cases. We leave a more detailed investigation of this region for future research, and just report one finding that qualifies the intuition of the standard Stackelberg commitment in an important way.

Proposition 2. ***(Multiple Equilibria Region)*** *Consider the Battle scenario in which (1)-(2) hold, and some* $r^M \in \left(r^F, \overline{r^M}\right)$. *Despite* M *being the more strongly committed player there exist circumstances under which* M *likely-loses and* F *likely-wins.*

Proof. Appendix B shows that this happens if $d > \bar{d}\left(r^F, r^M\right)$ and $v < \bar{v}\left(r^F, r^M\right)$, ie the surrender payoffs of M (F) are sufficiently high (low). □

The fact that the player with a higher r^i is less 'likely' to achieve its preferred outcome than the opponent (see Definition 2) is in stark contrast with the standard commitment solution. Intuitively, this occurs if F is insensitive to the conflict cost, whereas M is highly sensitive to it. The mapping in Appendix F implies that this happens if the central bank places high weight on output stabilization relative to inflation stabilization, in which case the output cost discourages the bank from engaging in a tug-of-war.

The novel insight here is that insufficient M conservatism may reduce the effectiveness of an explicit M commitment. This implies partial substitutability of strict and explicit inflation targeting, also apparent in (5) where $\overline{r^M}$ is increasing in d. The more explicitly committed the M regime is, the less strict it needs to be.[19] Such result is at odds with concerns by inflation targeting sceptics such as Greenspan (2003) or Kohn (2005) who believed that an explicit inflation target reduces M policy flexibility. But it is in line with Woodford who called such concerns to be the '*traditional prejudice of central bankers*'.

[19]This is in line with Svensson (2008) who argued that: '*it is desirable to do flexible inflation targeting more explicitly*'. For some empirical evidence that explicit inflation targets have *not* lead to stricter M policy, see Creel and Hubert (2010).

4. Extension I: Discounting and General Payoffs

This section introduces discounting for both players, $\delta_M \leq 1, \delta_F \leq 1$, and solves the Battle of the Sexes scenario for general payoffs, namely

$$a > c, a > d > b \text{ and } z > y, z > v > w. \tag{9}$$

It will become apparent that while the nature of the above game theoretic analysis is robust to discounting, the players' impatience may change the outcomes in an important way. We will focus on deriving conditions for a sure-win of the M policymaker since the resulting Disciplined SPNE is the socially optimal outcome. But as Proposition 1 demonstrates, the results apply analogously for F's sure-win.

Proposition 3. *Consider the Battle scenario with discounting and general payoffs (9). M surely-wins the game iff he is sufficiently patient*

$$\delta_M > \overline{\delta_M} = \sqrt[r^F]{\frac{d-b}{a-b}}, \tag{10}$$

and sufficiently strongly committed

$$r^M > \overline{r^M}\left(r^F, \delta_M, \delta_F, a, b, d, v, w, y, z\right) \geq r^F. \tag{11}$$

If M is insufficiently patient, $\delta_M < \overline{\delta_M}$, then $\overline{r^M}$ does not exist, ie even $\frac{r^M}{r^F} \to \infty$ does not deliver M's sure-win.

Proof. See Appendix C.[20] □

The exact form of the necessary and sufficient threshold in (11) is derived in Appendix C, equation (31). It implies that $\overline{r^M}$ is a step function of F's payoffs, specifically increasing in v and y, and decreasing in w and z. In terms of the other variables while we cannot formally prove the relationships for all R, valuable insights can nevertheless be obtained from the special case of $R = 0$ which was shown above to be representative of the more asynchronous cases $R > 0$. In this case $\overline{r^M}(0)$ is increasing in r^F, δ_F and d, and decreasing in δ_M, a and b: see Figure 3 for a graphical demonstration.[21]

Intuitively, M's impatience strengthens the necessary and sufficient condition; ie it makes it more difficult for M to surely-win and for the policymakers to coordinate. The intuition is similar to a standard repeated game in which it is harder to deter an impatient player from defecting since the future reward for not defecting has a smaller present value. The policy implication is therefore the following: a less patient central banker needs to commit more explicitly to guarantee his preferred medium-term outcomes. This implies partial substitutability between an explicit inflation target and longer mandates for central banks.

Proposition 3 not only refines the standard result obtained under Stackelberg leadership, it also qualifies its intuition substantially. Under static commitment, the (more)

[20] The patience and commitment thresholds for F's sure-win are again mirror images of $\overline{\delta_M}$ and $\overline{r^M}$. In terms of the former we have

$$\overline{\delta_F} = \sqrt[r^M]{\frac{v-w}{z-w}}.$$

[21] Formal proofs of these relationships appear in the working paper version of this article.

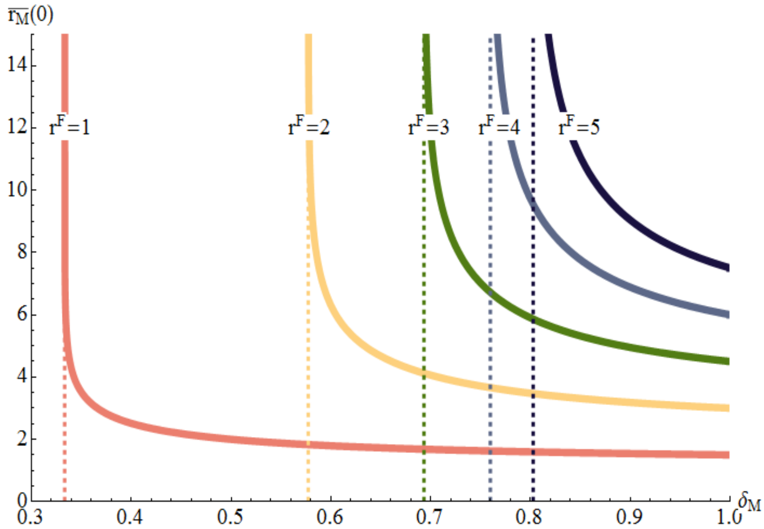

FIGURE 3. Dependence of $\overline{r^M}(0)$ on δ_M for various r^F (from (23) for the game in (1) with $d = \frac{1}{3}$). M's sure-win is delivered in the area to the right of the curves. The dotted asymptotes correspond to bounds $\overline{\delta_M}(0)$ for each particular r^F.

committed player *always* wins the game. In contrast, under dynamic commitment there are parameter regions in which he *may* not surely-win, and parameter regions in which he *never* surely-wins. If the more committed player is highly impatient then even an infinitely strong commitment is insufficient to deliver his sure-win. This is regardless of the opponent's discount factor. Hence the insights obtained under standard commitment are not robust.

Relating this result back to Figure 2, if $\delta_M < \overline{\delta_M}$ there are only two regions of SPNE; or only one region if both of $\delta_M < \overline{\delta_M}$ and $\delta_F < \overline{\delta_F}$ since neither $\underline{r^M}$ nor $\overline{r^M}$ exist.

5. Extension II: Fiscal Heterogeneity in a Monetary Union

Our dynamic commitment framework can easily incorporate any number of players. To demonstrate, let us examine the case in which F policy is heterogeneous focusing on two types of heterogeneity: in economic size and in F rigidity. This describes the situation in the European Monetary Union, and the United States to some extent, with a common currency and hence common M policy, but somewhat independent F policies.

Formally, the set of players is now $I = \{M, F^j\}$ where $j \in [1, J]$ denotes a certain member country, r^F_j its degree of F commitment, and s_j its relative economic size such that $\sum_{j=1}^{J} s_j = 1$. We assume that the overall payoff of M is a weighted average of the bank's payoffs gained from the interaction with each F^j - with weights s_n. The payoff of each independent government is, however, directly determined by its own actions and those of the common central bank as shown in (1).[22]

[22]Indirectly, the actions of other governments also have an impact since they determine the action of the central bank, and hence the equilibrium outcomes.

In extending our analysis to this case we will first assume the absence of free-riding by union members, whereby the nature of the above results remains unchanged.

Remark 1. *In a monetary union, the necessary and sufficient threshold $\overline{r^M}$ in the Battle scenario is increasing in the weighted average of the F rigidities of member countries with the weights being the country sizes s_j. Formally, r^F in (19),(21), and (30)-(31) is replaced by $\sum_{j=1}^{J} s_j r_j^F$.*

To offer a numerical example, assume $d = \frac{2}{3}$ and a union of two countries: one having double the size and double the F rigidity of the other (assumed to be $r^F = 2$). Then for M's sure-win in the $R_j = 0, \forall j$ case it follows that $r^M > \overline{r^M} = \frac{\sum_{j=1}^{J} s_j r_j^F}{1-d} = \frac{\frac{2}{3}4+\frac{1}{3}2}{\frac{1}{3}} = 10$ is required.

However, a moral hazard problem may occur on the part of individual governments. This is because the political benefits of F spending accrue primarily to the fiscally indisciplined country, whereas the economic costs in terms of tighter M policy are spread across all countries (Masson and Patillo (2002)).

In particular, the smaller a country is relative to the union, the less impact its F policy has on average inflation and output forecasts in the union - and hence on the interest rate response of the common central bank. Furthermore, the punishment in the form of an ensuing M contraction is also spread across the union as a whole. Hence even disciplined governments are penalized. The incentives for free-riding, whether deliberate or out of myopia or neglect, can therefore arise rapidly especially in smaller countries.

To formalize this denote by $m_j \in [0,1]$ the *degree of free-riding*, ie the extent to which a member country j does *not* internalize the negative impact of its F excesses onto the rest of the union members. The value $m_j = 0$ denotes no free-riding, whereas $m_j = 1$ denotes extreme free-riding in which country j totally ignores its impact on others. This can be incorporated through the free parameter v_j. Recall that it is the surrender payoff relative to the conflict cost and victory reward. It seems natural to assume the j government's conflict cost to be increasing in the country's weight s_j, and decreasing in its degree of free-riding m_j:

$$\frac{\partial v_j(m_j, s_j)}{\partial m_j} < 0, \qquad \text{and} \qquad \frac{\partial v_j(m_j, s_j)}{\partial s_j} > 0 \text{ for all } m_j > 0. \tag{12}$$

Let us further assume that v_j is a monotone function of m_j and that $v_j(m_j = 1) < w_j$. We can now contrast the outcomes in some country j before (B) and after (A) joining the M union. The case before joining the union is naturally $m_j^B = 0$ and $s_j^B = 1$. After joining, we have some $m_j^A \in [0,1]$ and $s_j^A \in (0,1)$.

Proposition 4. *(i) Consider country j described by the Battle scenario, (1)-(2), and $r^M > \overline{r^M}$, ie D uniquely obtains for both policies. After joining a monetary union, if the degree of j's free-riding is above a certain (country specific) threshold, $m_j > \overline{m_j}\left(s_j, v_j^B\right)$, then j's accession leads to deterioration of its medium-term F outcomes from D to I.*
(ii) The common central bank's commitment threshold $\overline{r^M}$ that ensures M discipline is increasing in the combined size of the countries with $m_j > \overline{m_j}$. If this size is sufficiently

high then F excesses may spill over to M policy even if the bank is both fully patient and infinitely strongly committed: $\delta_M = 1, r^M \to \infty$.

Proof. See Appendix D. For a solution using a specific example of (12) see Appendix E. □

Intuitively, the threat of punishment by the common central bank is no longer enough to discipline the government of a sufficiently small union member country with a sufficiently high degree of free-riding, and discourage it from F excesses.[23] This is because I becomes a strictly dominant strategy for F in the underlying game, and the scenario (from the free-riding government's perspective) changes from the Battle of the Sexes to Neglect. In terms of claim (ii), the central bank is worried about the cost of the policy conflict and therefore once the combined F indiscipline reaches a certain level the bank will no longer play D as the resulting conflict would be too widespread and costly. The bank therefore starts accommodating such F policy, which leads to sustained over-shooting of its optimal inflation level - even if it is highly explicit. In such case the central bank's instrument independence has been seriously compromised.

6. Alternative Scenarios and The Real World

6.1. **Alternative Scenarios.** It should by now be apparent that in the *Tug-of-war* and *Neglect scenarios,* dynamic commitment will not alter the outcomes of the game. This is because player F has a strictly dominant strategy in the underlying normal-form game.[24]

If the government is, like the central bank, responsible, $v > \max\{w, y, z\}$, then we will observe one of two scenarios depending on the value of d. The *Symbiosis scenario* (Dixit and Lambertini, 2003) features $d < 0$, and hence the socially optimal (D, D) outcome is the unique Nash equilibrium of the game. Alternatively, the *Pure coordination scenario* is characterized by $d \in (0, 1)$ meaning that there are two pure Nash equilibria, (D, D) and (I, I). But since the former Nash equilibrium is preferred by both players, such a game can be solved using the focal point argument. Therefore, there is no major problem in these scenarios, and dynamic commitment will again have no bearing on the outcomes of the policy interaction. Nevertheless, if there exists uncertainty about the type of government (as it may change over time with elections over the political cycle), implementing high M commitment acts as 'credibility insurance' for M policy against future ambitious governments.

In the unlikely case of an ambitious central bank, $d > \max\{a, b, c\}$, facing a responsible government, $v > \max\{w, y, z\}$, the conclusions as well as the terminology would be reversed. For the socially optimal outcome it is required that F commitment r^F is sufficiently strong relative to M rigidity r^M.

Finally, if both policymakers are ambitious we can, under some circumstances, observe two additional scenarios of interest. First, in the *Game of chicken scenario* there are two pure Nash equilibria, (D, I) and (I, D), each preferred by a different player. Second, in the *Prisoner's dilemma scenario* (I, I) is the unique Nash despite being Pareto dominated by the (D, D) outcome. In the latter scenario static commitment does not

[23]Greece comes to mind as an example of this type of behaviour.

[24]As Bernanke (2005) argued: '*No monetary-policy regime, including inflation targeting, will succeed in reducing inflation permanently in the face of unsustainable fiscal policies - large and growing deficits.*'

alter the outcomes, and thus dynamic commitment cannot help escape the inefficient equilibrium either. In the Game of Chicken the intuition is the same as in the Battle scenario: a sufficiently patient player that is sufficiently strongly committed (relative to the opponent) will ensure his preferred outcome. While Pareto-efficiency is ensured in such case, we never obtain the socially optimal outcome (D, D).

6.2. **Interpretation: from Macro Models to the Real World.** It is difficult to connect real world countries unambiguously with the above scenarios. This is not only because policy preferences and payoffs change over time. It is also because we observe the actual outcomes rather than the underlying preferences, and these may already be influenced by legislated commitment devices. Moreover these outcomes are not necessarily the equilibrium ones, they may reflect a transitory (off-equilibrium) phase.

To give an example, countries with observed medium-run (D, D) such as Australia, New Zealand, and most Nordic countries could be described by the Symbiosis or Pure coordination scenarios under any $\frac{r^M}{r^F}$, *or* by the Battle scenario under $r^M > \overline{r^M}\left(r^F\right)$. Similarly, countries in which we observe (D, I) - most industrial ones including the United States and many European Monetary Union members - could fall in the Tug-of-war scenario, *or* the initial 'conflict phase' of the Battle scenario.

Despite these caveats, it is important to note that such uncertainty does not alter the main prescription of our analysis: a long-term M commitment should be made as strong and explicit as possible. This will increase the range of circumstances under which the socially optimal outcomes obtain. Nevertheless, we have seen that this is not sufficient in all scenarios. Therefore, in order to 'cover all bases' and guarantee D for both policies regardless of the type of government, transparent and accountable commitments should apply to F policy as well, with oversight by an independent F policy council. This has been argued forcefully by Leeper (2010) and others before him, but only implemented in a minority of countries. Debt targeting would be one possible way to do this in practice.

The above analysis offers insights with respect to the partial substitutability between explicit inflation targeting (high r^M), strict inflation targeting (low d), and central bank goal independence (high δ_M). In order to obtain additional policy insights one needs to use a specific macro model and map it to our game theoretic representation. Since each scenario can be produced via fundamentally different macro models (see the references for the Battle scenario cited in the introduction), one obviously cannot write down a unique mapping between the deep parameters and the payoffs $\{a, b, c, d, v, w, y, z\}$. It can nevertheless be done separately for each macroeconomic model.

Appendix F offers an example of this procedure following the Cho and Matsui (2005) approach. Using a reduced-form model like that of Nordhaus (1994) with a standard quadratic utility function for both policymakers (but a higher output target of F), the analysis implies that we can interpret the payoffs of the Battle scenario as follows.

The conflict costs b and w are caused by greater variability in both nominal and real variables due to the offsetting tug-of-war behaviour of the policies. In micro-founded models these depend on variables such as the sensitivity of investment to output variability, the degree of price stickiness, wage rigidity and other labour market frictions, and the way agents form expectations and process information. The victory rewards a and z are generated by achieving the policymakers' preferred outcomes. For M this is due to stable inflation and output, for F it is due to securing votes through various spending programs or avoiding unpopular reform. Therefore, their magnitude depends on variables such as the proportion of naïve voters, the strength of the unions, the extent of inflation or unemployment aversion, and other cultural and historical specifics. Finally, the surrender payoffs d and v are determined by the cost of high inflation as well as the cost of debt repayments and other associated imbalances. These obviously depend on variables such as the size and development of the economy, the degree of indexation, and the completeness of financial markets.

Such mapping therefore enables one to propose more specific policy recommendations because the (qualitative as well as quantitative) effect of a particular deep parameter on the outcomes can be identified. It will also allow one to take the model to the data and attempt to identify the relevant scenario and the appropriate institutional remedy.

7. Summary and Conclusions

The stance of F policy in a number of countries has raised concerns about the degree of discipline, and about the risks for the credibility and outcomes of M policy. While the global financial crisis contributed to the problem, the underlying causes of these concerns had existed long before the crisis.[25] To contribute to this debate we use a novel asynchronous game theory framework that generalizes the standard (Stackelberg leadership) commitment concept from static to dynamic. We show that the conventional wisdom derived under standard static commitment is not robust, and that the risks of F spillovers to M policy may be greater than the conventional analysis has suggested.

Our investigation shows that the effect of M commitment on economic outcomes of the policy interaction crucially depends on its explicitness *relative* to the degree of F rigidity and ambition, as well as other structural and policy parameters. The problem is that under a range of circumstances inferior M policy outcomes (higher inflation and lower credibility) can occur due to spillovers from an excessive F policy - even if the central bank is independent, responsible, patient, and strongly committed. As Davig et al (2010) note: *'Without significant and meaningful fiscal policy adjustment, the task of meeting inflation targets will become increasingly difficult.'*

To offer some constructive conclusions, we have identified the scenarios and circumstances under which M policy outcomes will not be compromised by F excesses. They require the central bank to be sufficiently patient as well as sufficiently strongly committed (and we derive thresholds for both). Interestingly, under those conditions M policy may not only resist F pressure coming from an ambitious F setting, but its commitment may also discipline the government by reducing its payoff from excessive spending through a credible threat of a costly tug-of-war. We formally examine how the explicitness of long-run M commitment r^M can tip the balance between the two policies. Our proposed channel is different from Walsh (1995). It highlights the (desirable) logistic constraints associated with a legislated long-term objective, and may explain why many inflation targeting countries achieved sound outcomes even without a formal incentive contract/dismissal procedure.

[25]For example the IMF (2009) estimates the contribution of the crisis to the observed fiscal stress to only be 10.8% of that of aging population related spending in G20 countries.

This disciplining effect of M commitment on F policy can be observed in the real world. Dr Don Brash, Governor of the Reserve Bank of New Zealand during 1988-2002, in which period the Bank pioneered its explicit inflation targeting framework, wrote in private correspondence the following in response to our analysis (quoted with permission):

'*New Zealand provides an interesting case study illustrating the arguments in the article. We adopted a very strong commitment by the monetary authority, the Reserve Bank of New Zealand, when the Minister of Finance signed the first Policy Targets Agreement (PTA) with me as Governor under the new Reserve Bank of New Zealand Act 1989 early in 1990. The PTA required me to get inflation as measured by the CPI to between 0 and 2% per annum by the end of 1992, with the Act making it explicit that I could be dismissed for failing to achieve that goal unless I could show extenuating circumstances in the form, for example, of a sharp increase in international oil prices. At the time, inflation was running in excess of 5%.*

In the middle of 1990, the Government, faced with the prospect of losing an election later in the year, brought down an expansionary budget. I immediately made it clear that this expansionary fiscal policy required firmer monetary conditions if the agreed inflation target was to be achieved, and monetary conditions duly tightened.

Some days later, an editorial in the "New Zealand Herald", New Zealand's largest daily newspaper, noted that New Zealand political parties could no longer buy elections because, when they tried to do so, the newly instrument-independent central bank would be forced to send voters the bill in the form of higher mortgage rates.

I was later told by senior members of the Opposition National Party that the Bank's action in tightening conditions in response to the easier fiscal stance had had a profound effect on thinking about fiscal policy in both major parties in Parliament.

Some years later, in 1996, the Minister of Finance of the then National Party Government announced that he proposed to reduce personal income tax rates subject to this being consistent with the Government's debt to GDP target being achieved, to the fiscal position remaining in surplus, and to the fiscal easing not requiring a monetary policy tightening. The Minister formally wrote to me asking whether tax reductions of the kind proposed would under the economic circumstances then projected, require me to tighten monetary conditions. Given how the Bank saw the economy evolving at that time, I was able to tell the Minister that tax reductions of the nature he proposed would not require the Bank to tighten monetary conditions in order to stay within the inflation target.'

What are the policy implications of this finding? The lesson for M policymakers is that, to discourage and/or counter-act over-expansionary F policies, they should if possible commit to low average inflation more explicitly. This is desirable primarily in countries without a legislated numerical inflation target facing long-term F sustainability issues such as the United States, Switzerland, and Japan. The implication for F policymakers is that imposing such M commitment onto their central banks may provide a way to indirectly tie their hands, and gain political support for reforms towards F sustainability.

We identify an important caveat to this finding. We show that the disciplining channel is unlikely to be effective in a currency union where a moral hazard problem due to free-riding of small member countries occurs naturally. If indisciplined countries ignore the negative externality they impose on others, the M punishment they face from the common central bank is not strong enough. In such cases direct F commitment arrangements, ie legislated and enforceable F rules, are necessary to discipline F policy over the long term. In fact, these seem to be desirable - as an 'insurance' - in all countries given that political preferences and realities often change. Such reasoning provides a formal justification for the The European Financial Stability Facility and planned amendments to the Lisbon Treaty.

The paper also has several implications that can be taken to the data: see Appendix G. Specifically, our analysis implies that for some but not all parameter values, a more explicit long-term M commitment can have three effects. First, it can reduce the average level and the variability of inflation, and increase M policy credibility. This is consistent with results due to Fang and Miller (2010), Neyapti (2009), Corbo, Landerretche and Schmidt-Hebbel (2001) or Debelle (1997) among others.

Second, it can act as a partial substitute for central bank goal independence (patience δ_M and conservatism $\frac{1}{d}$) in achieving credibility. This is in line with the negative correlation between central bank (goal) independence and accountability reported by Briault, Haldane and King (1997), de Haan, Amtenbrink and Eijffinger (1999), and Sousa (2002).

Third, M commitment may be able to discipline F policy and induce reductions in the average level and the variability of budget deficits and debt (except for small free-riding members of a M union). In addition to Don Brash's account above and other narrative evidence, Franta, Libich and Stehlík (2010) show formally that F outcomes in most inflation targeting countries have improved shortly after adoption of the regime, and have largely remained in a good shape thereafter (see Appendix G for a sample of the results). In contrast the main non-targeters (the United States, Japan, and Switzerland), as well as most small EMU members have seen their F outcomes deteriorate over the same period.

There are two issues regarding robustness and extensions worth noting. First, our long-run M commitment is flexible in the sense that the central bank is still able to choose the desired long-run policy level (every r^M periods) and the underlying unmodelled short-run stabilization actions (every period) without any restrictions on *how* these choices need to be made. Put differently, since shocks have a zero-mean over the medium-term (business cycle), our long-term M commitment is compatible with a discretionary solution, an instrument rule such as Taylor (1993), as well as the timeless perspective type of commitment of Woodford (1999) or the quasi commitment of Schaumburg and Tambalotti (2007).

Second, commitment and rigidity can easily be endogenized in our framework. We could incorporate into the payoffs some cost of increasing M policy commitment (such as implementation cost of inflation targeting), $\frac{\triangle C^M}{\triangle r^M} > 0$, and some political cost of reducing F rigidity (such as loss of votes from an unpopular welfare or pension reform), $\frac{\triangle C^F}{\triangle r^F} < 0$. This would enable us to derive the equilibrium values of r^M and r^F that are optimally selected by the policymakers.[26]

[26] We do not do so as this would merely alter the key variable from r^i to C^i without obtaining additional theoretic insights: M's sure-win would obtain if $C^M < \underline{C^M}(C^F)$. In addition, r^i seems easier to identify and interpret than C^i.

8. References

ADAM, K., and R. BILLI (2008): "Monetary Conservatism and Fiscal Policy," Journal of Monetary Economics, 55(8), 1376-1388.

ALESINA, A., and G. TABELLINI (1987): "Rules and Discretion with Noncoordinated Monetary and Fiscal Policies," Economic Inquiry, 25(4), 619 - 630.

ARTIS, M. J., and B. WINKLER (1998): "The Stability Pact: Trading Off Flexibility for Credibility?," in Fiscal Aspects of European Monetary Integration, ed. by Hughes Hallett, Hutchison, and Jensen: Cambridge University Press, 157-188.

AUMANN, R., and S. SORIN (1989): "Cooperation and Bounded Recall," Games and Economic Behaviour, 1, 5-32.

BALL, L., and N. SHERIDAN (2003): "Does Inflation Targeting Matter?," International Monetary Fund IMF Working Papers: 03/129.

BARNETT, R. (2001): "Inflation, Taxes, and the Coordination of Monetary and Fiscal Policy by Use of a Game of Chicken," Canadian Journal of Economics, 34, 82-99.

BARRO, R. J., and D. B. GORDON (1983): "A Positive Theory of Monetary Policy in a Natural Rate Model," Journal of Political Economy, 91, 589-610.

BENHABIB, J., and S. EUSEPI (2005): "The Design of Monetary and Fiscal Policy: A Global Perspective," Journal of Economic Theory, 123(1), 40-73.

BERNANKE, B. (2005): "Inflation in Latin America - a New Era?," Remarks at the Stanford Institute for Economic Policy Research Economic Summit, Stanford, California, 11 February 2005.

BHATTACHARYA, J., and J. H. HASLAG (1999): "Monetary Policy Arithmetic: Some Recent Contributions," Economic and Financial Policy Review, Federal Reserve Bank of Dallas, Q III, 26-36.

BLAKE, A. P., and M. WEALE (1998): "Costs of Separating Budgetary Policy from Control of Inflation: A Neglected Aspect of Central Bank Independence," Oxford Economic Papers, 50(3), 449-67.

BRANCH, W. A., T. DAVIG, and B. MCGOUGH (2008): "Monetary–Fiscal Policy Interactions under Implementable Monetary Policy Rules," Journal of Money, Credit and Banking 40(5).

BRIAULT, C., A. HALDANE, and M. KING (1997): "Independence and Accountability," in Towards More Effective Monetary Policy, ed. by Kuroda. London: Macmillan Press in association with Bank of Japan, 299-326.

CALVO, G. A. (1983): "Staggered Prices in a Utility-Maximizing Framework," Journal of Monetary Economics, 12, 383-98.

CHADHA, J. S., and C. NOLAN (2007): "Optimal Simple Rules for the Conduct of Monetary and Fiscal Policy " Journal of Macroeconomics, 29(4), 665-689.

CHO, I., and A. MATSUI (2005): "Time Consistency in Alternating Move Policy Games," Japanese Economic Review, 56(3) 273–294.

CHORTAREAS, G., D. STASAVAGE, and G. STERNE (2002): "Monetary Policy Transparency, Inflation and the Sacrifice Ratio," International Journal of Finance and Economics, 7, 141-55.

COCHRANE, J. H. (2009): "Fiscal Theory, and Fiscal and Monetary Policy in the Financial Crisis," unpublished.

COOPER, R. N. (1969): "Macroeconomic Policy Adjustment in Interdependent Economies " The Quarterly Journal of Economics, 83, No. 1, 1-24.

CREEL, J., and P. HUBERT (2010): "Has Inflation Targeting Changed Monetary Policy Preferences?," unpublished.

DAVIG, T., E. M. LEEPER, and W. T. B. (2010): "Inflation and the Fiscal Limit," prepared for European Economic Review conference.

DEBELLE, G. (1997): "Inflation Targeting in Practice," International Monetary Fund IMF Working Papers: 97/35.

DIXIT, A., and L. LAMBERTINI (2001): "Monetary–Fiscal Policy Interactions and Commitment Versus Discretion in a Monetary Union," European Economic Review, 45(4), 977-987.

DIXIT, A., and L. LAMBERTINI (2003): "Interactions of Commitment and Discretion in Monetary and Fiscal Policies," American Economic Review, 93(5), 1522-1542.

EIJFFINGER, S. C. W., and P. M. GERAATS (2006): "How Transparent Are Central Banks?," European Journal of Political Economy, 22(1), 1-21.

EUSEPI, S., and B. PRESTON (2008): "Stabilizing Expectations under Monetary and Fiscal Policy Coordination," NBER Working Papers 14391, National Bureau of Economic Research, Inc.

FANG, W., and S. M. MILLER (2010): "The Lag in Effect of Inflation Targeting and Policy Evaluation," Working papers 2010-01, University of Connecticut, Department of Economics.

FRANTA, M., J. LIBICH, and P. STEHLÍK (2010): "The Big Picture of Monetary-Fiscal Interactions," in Monetary-Fiscal Interactions: How to Improve Policy Outcomes?, ed. by Libich, Savage, and Walsh. Sydney: Symposion at the Australian Conference of Economists.

FRIEDMAN, M. (1948): "A Monetary and Fiscal Framework for Economic Stability.," American Economic Review, 38(2), 245–264.

FRY, M., D. JULIUS, L. MAHADEVA, S. ROGER, and G. STERNE (2000): "Key Issues in the Choice of a Monetary Policy Framework," in Monetary Frameworks in a Global Context, ed. by Mahadeva, and Sterne. London: Routledge.

GALI, J., and T. MONACELLI (2005): "Optimal Fiscal Policy in a Monetary Union," Proceedings, Federal Reserve Bank of San Francisco.

GREENSPAN, A. (2003): "Monetary Policy and Uncertainty: Adapting to a Changing Economy, Symposium Sponsored by the Federal Reserve Bank of Kansas City," Jackson Hole, Wyoming, August 28 - 30, 2003.

HAAN, J. D., F. AMTENBRINK, and S. C. W. EIJFFINGER (1999): "Accountability of Central Banks: Aspects and Quantification," Banca Nazionale del Lavoro Quarterly Review, 52, 169-93.

HUGHES HALLETT, A. (1986): "Autonomy and the Choice of Policy in Asymmetrically Dependent Economies," Oxford Economic Papers, 38, 516-44.

HUGHES HALLETT, A., and D. N. WEYMARK (2005): "Independence before Conservatism: Transparency, Politics and Central Bank Design," German Economic Review, 6, 1-21.

IMF (2009): "Fiscal Implications of the Global Economic and Financial Crisis," IMF Staff Position Note, SPN/09/13.

KOHN, D. L. (2005): "Comments on Marvin Goodfriend's "Inflation Targeting in the United States?," in The Inflation-Targeting Debate, ed. by Bernanke, and Woodford. NBER Studies in Business Cycles, vol. 32. Chicago and London: University of Chicago Press.

LEEPER, E. M. (1991): "Equilibria under 'Active' and 'Passive' Monetary and Fiscal Policies," Journal of Monetary Economics 27 (1), 129-147.

LEEPER, E. M. (2010): "Monetary Science, Fiscal Alchemy," Jackson Hole Symposium, Wyo: The Federal Reserve Bank of Kansas City.

LIBICH, J. (2009): "A Note on the Anchoring Effect of Explicit Inflation Targets," Macroeconomic Dynamics, 13(5).

LIBICH, J., and P. STEHLÍK (2010): "Incorporating Rigidity in the Timing Structure of Macroeconomic Games," Economic Modelling, 27, 767–781.

LOHMANN, S. (1992): "Optimal Commitment in Monetary Policy: Credibility Versus Flexibility," American Economic Review, 82, 273-86.

MAILATH, G. J., and L. SAMUELSON (2006): Repeated Games and Reputations: Long-Run Relationships. Oxford University Press.

MASKIN, E., and J. TIROLE (1988): "A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs," Econometrica, 56, 549-69.

MASSON, P., and C. PATTILLO (2002): "Monetary Union in West Africa: An Agency of Restraint for Fiscal Policies?," Journal of African Economies, 11 387-412.

MUNDELL, R. (1962): "Appropriate Use of Monetary and Fiscal Policy for Internal and External Stability," IMF Staff Papers, 9, 70-79.

NEYAPTI, B. (2009): "Performance of Monetary Institutions: Comparative Evidence," Departmental Working Papers 0902, Bilkent University.

NORDHAUS, W. (1994): "Policy Games: Coordination and Independence in Monetary and Fiscal Policies," Brookings Papers on Economic Activity, 1994:2, 139-216.

PERSSON, M., T. PERSSON, and L. E. O. SVENSSON (2006): "Time Consistency of Fiscal and Monetary Policy: A Solution," Econometrica, 74(1), 193-212.

PETIT, M. L. (1989): "Fiscal and Monetary Policy Co-Ordination: A Differential Game Approac H," Journal of Applied Econometrics, 4(2), 161-79.

RESENDE, C., and N. REBEI (2008): "The Welfare Implications of Fiscal Dominance," Bank of Canada Working Paper 2008-28.

ROGOFF, K. (1985): "The Optimal Degree of Commitment to an Intermediate Monetary Target," Quarterly Journal of Economics, 100, 1169-89.

SARGENT, T. J., and N. WALLACE (1981): "Some Unpleasant Monetarist Arithmetic," Federal Reserve Bank of Minneapolis Quarterly Review, 5, 1-17.

SCHAUMBURG, E., and A. TAMBALOTTI (2007): "An Investigation of the Gains from Commitment in Monetary Policy," Journal of Monetary Economics, 54(2), 302-324.

SIMS, C. (1994): "A Simple Model for Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy," Economic Theory, 4(3), 381-99.

SOUSA, P. (2002): "Central Bank Independence and Democratic Accountability," mimeo, Portucalense University, available at http://www.univ-orleans.fr/deg/GDRecomofi/Activ/doclyon/desousa.pdf.

SVENSSON, L. E. O. (2008): "What Have Economists Learned About Monetary Policy over the Past 50 Years?," Sveriges Riksbank

TAYLOR, J. B. (1993): "Discretion Versus Policy Rules in Practice," Carnegie-Rochester Conference Series on Public Policy, 39, 195-214.

TINBERGEN, J. (1954): Centralization and Decentralization in Economic Policy. Amsterdam.

TOBIN, J. (1982): "Money and Finance in the Macroeconomic Process," Journal of Money, Credit and Banking, 14 (2), 171-204.

VAN AARLE, B., J. ENGWERDA, and J. PLASMANS (2002): "Monetary and Fiscal Policy Interaction in the Emu: A Dynamic Game Approach " Annals of Operations Research, 109(1-4), 229-264.

WALSH, C. E. (1995): "Optimal Contracts for Central Bankers," American Economic Review, 85, 150-67.

WEN, Q. (1994): "The Folk Theorem for Repeated Games with Complete Information," Econometrica, 62, 949-954.

WILLARD, L. (2006): "Does Inflation Targeting Matter: A Reassessment," CEPS WP 120, Princeton University.

WOODFORD, M. (1994): "Monetary Policy and Price Level Determinacy in a Cash-in-Advance Economy," Economic Theory, 4(3), 345-80.

WOODFORD, M. (1999): "Optimal Monetary Policy Inertia," Manchester School, 67, 1-35.

WYPLOSZ, C. (1991): "Monetary Union and Fiscal Policy Discipline," CEPR DP488.

Appendix A. Proof of Proposition 1

Proof. We solve the game backwards and prove the claims by mathematical induction, initially focusing on $r^M > r^F$. First, we derive conditions under which D will be played in M's last move on the equilibrium path, $n^M = N^M$ (the inductive basis). Specifically, part A) of the proof will examine the case $R = 0$, and part B) the case $R > 0$. Second, supposing that this holds for some $n^M \leq N^M$, we show in part C) the conditions under which the same is true for $n^M - 1$ as well.[27]

A) $\mathbf{n^M = N^M}$ **under** $\mathbf{R = 0}$. Here we have $T(r^M, r^F) = r^M$, and therefore $N^M = 1$ and $N^F = \frac{r^M}{r^F}$. Solving backwards, we know F would like to play the best response to M's initial action, $F_n^* \in B(M_1), \forall n^F$. From her second move till the end of the dynamic stage game F can observe M_1, and will hence rationally respond with D to M_1^D, and I to M_1^I.

Moving backwards, M uses this information and hence knows that if he opens with D he will from period r^F onwards be rewarded by payoff a. But M also knows that such inducement play may be costly, payoff b, if F plays F_1^I. Therefore, for M to surely-win his victory reward must more than offset his conflict cost, in which case M's optimal play in period 1 will be D even if he knows with certainty that F_1^I will be played. Formally, the incentive compatibility condition (7) reported in the main text needs to hold. Using (1)-(2) and rearranging yields equation (8). The $\overline{r^M}(0)$ threshold is therefore the necessary and sufficient degree of M commitment that delivers M's sure-win for the case $R = 0$.

B) $\mathbf{n^M = N^M}$ **under** $\mathbf{R > 0}$: We know that the number of M's moves is $N^M = \frac{T(r^M, r^F)}{r^M} > 1$. A condition analogous to (7) is $br^F R + a(r^M - r^F R) > dr^M$ which

[27] It will become evident that for most parameter values satisfying (5) there will be a *unique* Disciplined SPNE. Nevertheless, since our attention is on the equilibrium path we will not examine the exact number of SPNE (off-equilibrium behaviour).

implies, using (1)-(2) and rearranging,

$$r^M > \frac{Rr^F}{1-d}. \tag{13}$$

C) $\mathbf{n^M+1 \to n^M}$ (if applicable, ie if $1 \le n^M < N^M$): The proof proceeds by induction. We first assume that M's unique best play in the (n^M+1)-th step is D regardless of F's preceding play (ie that M_{n+1} is history-independent), and we attempt to prove that this implies the same assertion for the n^M-th step. Intuitively, this means that if M inflates he finds it optimal to immediately disinflate. Two scenarios are possible in terms of the underlying F behaviour that determines the costs of the disinflation. If F runs a deficit, F^I, the conflict costs b and w will occur for at least one period. In contrast, if F switches from deficits to a balanced budget pre-emptively, ie play F^D before the start of the disinflation (in its anticipation), the disinflation will only be accompanied by the payoffs a and v and hence costless. This implies that one of the following two conditions analogous to (7) will apply at any move n^M

$$bk_n + a(r^M - k_n) + a[r^F - (r^F - k_{n+1})] > dr^M + \underbrace{b[r^F - (r^F - k_{n+1})]}_{(D,I)\text{: costly disinflation}}, \tag{14}$$

$$bk_n + a(r^M - k_n) > d[r^M - (r^F - k_{n+1})] + \underbrace{a(r^F - k_{n+1})}_{(D,D)\text{: costless disinflation}}. \tag{15}$$

Which of these two conditions is relevant to a certain n^M depends on F's payoffs $\{v, w, y, z\}$, and importantly on k_{n+1}. In particular, if

$$\underbrace{z(r^F - k_{n+1})}_{(I,I)} + \underbrace{wk_{n+1}}_{(D,I)} \ge \underbrace{y(r^F - k_{n+1})}_{(I,D)} + \underbrace{vk_{n+1}}_{(D,D)}, \tag{16}$$

then (14) obtains, otherwise (15) is the relevant condition. Now, we will show that if the conditions (14) and (15) are satisfied at $n^M = 1$, then they hold in all other n^M as well. This interesting feature notably simplifies the solution of the game.

Lemma 1. *Consider the Battle scenario in which (1)-(2) hold and $\delta_F = \delta_M = 1$. For any given R, out of the necessary and sufficient conditions for M to surely-win, $\underline{\{M_n^D\}} = B(F^I)$, the one regarding the initial move $n^M = 1$ yields as least as high $\overline{r^M}(R)$ as any other n^M. Therefore, $\{M_1^D\} = B(F_1^I)$ is the sufficient condition.*

Proof. Equations (14) and (15) can be, respectively, rearranged into

$$r^M > \frac{(k_n - k_{n+1})}{1-d} \quad \text{and} \quad r^M > \frac{k_n}{1-d} + (r^F - k_{n+1}). \tag{17}$$

The strength of both conditions is increasing in k_n and decreasing in k_{n+1}. Thus the strongest condition is guaranteed by the maximum of $(k_n - k_{n+1})$. From (4) it follows that $k_n - k_{n+1} \le Rr^F$. The fact that $k_1 - k_2 = Rr^F$ then proves the claim for $R > 0$. Realizing that for $R = 0$ we have $N^M = 1$ finishes the proof. □

Continuing the proof of Proposition 1, Lemma 1 means that regardless of the exact dynamics/asynchrony R, it suffices to focus on the initial simultaneous move (similarly to a one-shot game) *assuming* that all further relevant conditions hold. If the strongest condition for $n^M = 1$ is satisfied we then know that a unique (type of) equilibrium outcome obtains throughout. Lemma 1 therefore implies, in combination with the recursive scheme, that throughout the proof we can use the following:

$$k_n = k_1 = r^F \quad \text{and} \quad k_{n+1} = k_2 = (1-R)r^F. \tag{18}$$

Substituting this into (14)-(15) or (17) we obtain, together with (8)

$$r^M > \overline{r^M}(R) = \begin{cases} \frac{r^F}{1-d} & \text{if} \quad R = 0, \\ \left(\frac{1}{1-d} + R\right) r^F & \text{if} \quad R \leq \bar{R} = \frac{v}{v+2}, \\ \frac{Rr^F}{1-d} & \text{if} \quad R > \bar{R} = \frac{v}{v+2}, \end{cases} \tag{19}$$

where the threshold $\bar{R} \in (0,1)$ is derived from (16). $\overline{r^M}(R)$ is the necessary and sufficient threshold for uniqueness of the Disciplined type of SPNE (note that all three are at least as strong as the condition for N^M in (13)). By inspection, $\overline{r^M}(R)$ is increasing in r^F and d for all R.[28] It is also increasing in v which follows from the fact that the condition for $R \leq \bar{R}$ is stronger than the one for $R > \bar{R}$, and hence a higher v increases $\bar{R}$ and leads to strengthening of (19). This completes the proof of claim (i).

In terms of claim (ii), by symmetry the necessary and sufficient condition for F to surely-win is

$$r^F > \overline{r^F}(R) = \begin{cases} \frac{r^M}{1-v} & \text{if} \quad R = 0, \\ \left(\frac{1}{1-v} + R\right) r^M & \text{if} \quad R \leq \bar{R} = \frac{d}{d+2}, \\ \frac{Rr^M}{1-v} & \text{if} \quad R > \bar{R} = \frac{d}{d+2}. \end{cases} \tag{20}$$

Let us realize that the threshold $\overline{r^F}(R)$ is just a 'mirror-image' of the threshold $\overline{r^M}(R)$. Furthermore, the former threshold can be expressed in terms of r^M rather than r^F to obtain the threshold $\underline{r^M}(R)$ in the main text. Specifically, (20) can be re-written as

$$r^M < \underline{r^M}(R) = \begin{cases} (1-v)\, r^F & \text{if} \quad R = 0, \\ \left(\frac{1-v}{1+R-Rv}\right) r^F & \text{if} \quad R \leq \bar{R} = \frac{d}{d+2}, \\ (1-v)\, Rr^F & \text{if} \quad R > \bar{R} = \frac{d}{d+2}. \end{cases} \tag{21}$$

By inspection, $\underline{r^M}(R)$ is increasing in r^F and decreasing in v. Given that the condition for $R \leq \bar{R}$ is now weaker than the one for $R > \bar{R}$, the threshold $\underline{r^M}(R)$ is also decreasing in v. This completes the proof of Proposition 1. □

Appendix B. Proof of Proposition 2

Proof. To prove this existence claim it suffices to provide a specific example. Let us consider the simplest case of $R > 0$, namely $r^M = 3, r^F = 2$ (implying $R = \frac{1}{2}$) and the payoffs of (1)-(2). To prove that there exists no Disciplined SPNE it is sufficient to show that F will play F^I in one of her moves regardless of the preceding move of M. To prove that there exists at least one Indisciplined SPNE it suffices to show that in neither of his moves M will play M^D regardless of F's preceding move.

[28]This mplies that the conditions for the $R > 0$ cases only differ quantitatively from the $R = 0$ case, not qualitatively.

Focus on the condition for M's last move to be uniquely D in equation (13), $\frac{r^M}{r^F} > \frac{R}{1-d}$. Notice that since $R = \frac{1}{2}$, under $d > \bar{d} > \frac{2}{3}$ the condition is *not* satisfied. Therefore, M_2 is not history-independent and it will be the best response to F's preceding move, F_2. Moving backwards, player F takes this into account in comparing the continuation payoffs from F_2^D and F_2^I. Under M_1^D the continuation payoff from playing F_2^D is $4v$, whereas from playing F_2^I it is 3. Therefore, if $v < \bar{v} = \frac{3}{4}$ then F_2 is history-independent - regardless of M's preceding move, M_1, F will uniquely play F_2^I in order to ensure the I levels for the remaining four periods of the stage game. This proves that in this case there exists no Disciplined SPNE as there will never be F_2^D on the equilibrium path.

In order to prove that there exists an Indisciplined SPNE it suffices to note that, similarly to M_2, in M_1 the level D is not a unique play regardless of the level played in F_1. Put differently, we have $M_1^I \in B(F^I)$ since M knows that F_2^I is always played and there would be no victory reward from M_1^D. This implies that an outcome with $\left(F_1^I, M_1^I, F_2^I, M_2^I, F_3^I\right)$ on the equilibrium path belongs to the set of SPNE. □

Appendix C. Proof of Proposition 3

Proof: The derivation of the generalized necessary and sufficient threshold is analogous in all its aspects to that of Proposition 1. In part A) the condition corresponding to (7) under M's impatience, $\delta_M < 1$, is

$$b\sum_{t=1}^{r^F}\delta_M^{t-1} + a\sum_{t=r^F+1}^{r^M}\delta_M^{t-1} > d\sum_{t=1}^{r^M}\delta_M^{t-1}. \tag{22}$$

This can, using the formula for a sum of a finite series and rearranging, be written as

$$\delta_M^{r^M} < \frac{(a-b)\,\delta_M^{r^F} + b - d}{a-d}.$$

Taking the logarithms yields

$$r^M > \overline{r^M}\,(0) = \log_{\delta_M}\frac{(a-b)\,\delta_M^{r^F} + b - d}{a-d}. \tag{23}$$

The condition of part B) is again weaker than that. To prove part C) let us extend the result of Lemma 1 under the general payoffs and players' impatience.

Lemma 2. *Lemma 1 holds $\forall \delta_M \leq 1, \forall \delta_F \leq 1$, and any general payoffs satisfying (9).*

Proof. Lemma 1 shows this claim to hold under $\delta_M = \delta_F = 1$. The proof of Proposition 1 showed that the payoffs of the less committed player, F in our case, only affect the necessary and sufficient condition through the threshold $\bar{R}$. The same will thus be true for the value of δ_F. Let us therefore consider the effect of M's impatience. Under $\delta_M < 1$, the inequality in (14) that applies to the case of $R > \bar{R}$ becomes

$$b\sum_{t=1}^{k_n}\delta_M^{t-1} + a\sum_{t=k_n+1}^{r^M}\delta_M^{t-1} + a\sum_{t=r^M+1}^{r^M+k_{n+1}}\delta_M^{t-1} > d\sum_{t=1}^{r^M}\delta_M^{t-1} + b\sum_{t=r^M+1}^{r^M+k_{n+1}}\delta_M^{t-1}. \tag{24}$$

This can be, after some manipulation, rearranged into

$$(a-b)\sum_{t=1}^{r^M}\delta_M^{t-1}-(a-d)\sum_{t=1}^{r^M}\delta_M^{t-1}<(a-b)\delta_M^{k_n}\frac{1-\delta_M^{r^M+k_{n+1}-k_n}}{1-\delta_M}. \tag{25}$$

Since $\delta_M < 1$ we see that, analogously to Lemma 1, the strength of the condition is increasing in k_n and decreasing in k_{n+1}. Hence the same argument applies. We can readily check, using (15) under $\delta_M < 1$, that the same is true for $R \leq \bar{R}$. □

We will now complete the proof of Proposition 3 using this result. Lemma 2 implies that we need to substitute (18) into (25) for the costly disinflation case. Using formulas for finite sums, rearranging, and taking the logarithms yields

$$r^M > \log_{\delta_M}\frac{(a-d)-(a-b)\left(1-\delta_M^{r^F}\right)}{(a-d)-(a-b)\left(1-\delta_M^{r^F(1-R)}\right)}. \tag{26}$$

For the costless disinflation case, the analog of (15) under $\delta_M < 1$ is, using Lemma 2

$$b\sum_{t=1}^{r^F}\delta_M^{t-1}+a\sum_{t=r^F+1}^{r^M}\delta_M^{t-1}>d\sum_{t=1}^{r^M-r^F R}\delta_M^{t-1}+a\sum_{t=r^M-r^F R+1}^{r^M}\delta_M^{t-1}, \tag{27}$$

and after rearranging

$$r^M > \log_{\delta_M}\frac{b\left(1-\delta_M^{r^F}\right)+a\delta_M^{r^F}-d}{a\left(1+\delta_M^{-r^F}-\delta_M^{-r^F(1-R)}\right)-d\delta_M^{r^F R}}. \tag{28}$$

The threshold $\bar{R}$ determining whether the costly disinflation case of (26) or the costless disinflation case of (28) applies is derived from the generalization of (16) under F's impatience. Specifically, under $\delta_F < 1$ if

$$z\sum_{t=1}^{r^F R}\delta_M^{t-1}+w\sum_{t=r^F R+1}^{r^F}>y\sum_{t=1}^{r^F R}\delta_M^{t-1}+v\sum_{t=r^F R+1}^{r^F}, \tag{29}$$

then (26) obtains, otherwise (28) is the relevant condition. After rearranging this implies the following threshold

$$\bar{R}=\frac{1}{r^F}\log_{\delta_F}\frac{z-y+(v-w)\,\delta_F^{r^F}}{z-y+v-w}. \tag{30}$$

Combining (23), (26), (28), and $\bar{R}$ from (30) yields the following generalized necessary and sufficient condition for M's sure-win

$$r^M > \overline{r^M}(R)=\begin{cases}\log_{\delta_M}\frac{(a-b)\delta_M^{r^F}+b-d}{a-d} & \text{if } R=0,\\ \log_{\delta_M}\frac{(a-d)-(a-b)\left(1-\delta_M^{r^F}\right)}{(a-d)-(a-b)\left(1-\delta_M^{r^F(1-R)}\right)} & \text{if } R\leq\bar{R},\\ \log_{\delta_M}\frac{b\left(1-\delta_M^{r^F}\right)+a\delta_M^{r^F}-d}{a\left(1+\delta_M^{r^F}-\delta_M^{r^F(R-1)}\right)-d\delta_M^{Rr^F}} & \text{if } R>\bar{R}.\end{cases} \tag{31}$$

We can now use this condition to prove the claims of Proposition 3. Examining (30) and (31) reveals that $\overline{r^M}(R)$ is a function of r^F, both players' discount factors δ_M and δ_F, and all the payoffs except c.

In terms of the patience threshold, consider the logarithm's numerator of (23), (26), and (28). For the threshold $\overline{r^M}(R)$ to exist for all R it must hold that $(a-b)\,\delta_M^{r^F}+b-d>0$. Rearranging this inequality yields the necessary patience threshold $\overline{\delta_M}(.)$ in (10).

Finally, note that if $\delta_F < \overline{\delta_F}$ then there are cases in which $\overline{r^M} = r^F$ as claimed in (11). In such case any $r^M > r^F$ uniquely ensures discipline of both policies. The easiest way to see this is to consider $\delta_F = 0$. Such an impatient F will never reduce spending before the start of disinflation as she fully ignores the future. Therefore, disinflation will always be costly for both players, ie (28) no longer applies and (26) becomes the relevant condition $\forall n^M, R \in (0,1)$, and for all $\{a,b,d,v,w,y,z\}$ satisfying (9). This completes the proof of Proposition 3. □

Appendix D. Proof of Proposition 4

Proof. After joining the union, there is a decrease in s_j and a possible increase in m_j. If in country j the degree of free-riding is above a certain threshold

$$(32)\qquad m > \overline{m_j}\left(s_j, v_j^B\right) \quad \text{where} \quad \frac{\partial \overline{m_j}\left(s_j, v_j^B\right)}{\partial s_j} > 0 \quad \text{and} \quad \frac{\partial \overline{m_j}\left(s_j, v_j^B\right)}{\partial v_j^B} > 0,$$

then the value of v_j^A will fall below $w_j = 0$. This follows, using a continuity argument, from the monotonicity of $v_j(m_j)$ and the assumed $v_j(m_j = 1) < w_j = 0$. In such case the underlying game after accession for country j is no longer the Battle of the Sexes but the Neglect scenario since I becomes a strictly dominant strategy for F. Therefore, it will feature I for any level of $\delta_{M,j}$ and r_j^F, even if the common central bank has $\delta_M = 1$ and $r^M \to \infty$.

In terms of claim (ii), denote the number of countries in which $m_j > \overline{m_j}\left(s_j, v_j^B\right)$ by $\gamma \in \mathbb{N}$, and order the member countries such that those $\{1,...,\gamma\}$ feature $m_j > \overline{m_j}$, and those $\{\gamma+1,...,J\}$ feature $m_j \leq \overline{m_j}$. Let us report the conditions only for the special case $R=0$ and $\delta_M = 1$ as it was shown to be representative of the other cases as well. The condition analogous to (7) becomes

$$\underbrace{b\sum_{j=1}^{J} s_j r_j^F}_{(D,I)} + \underbrace{b\sum_{j=1}^{\gamma} s_j(r^M - r_j^F)}_{(D,I):\ \forall j\leq\gamma} + \underbrace{a\sum_{j=\gamma+1}^{J} s_j(r^M - r_j^F)}_{(D,D):\ \forall j>\gamma} > \underbrace{dr^M}_{(D,I):\forall j}.$$

Note that the condition only differs from the no free-riding case in the second element on the left hand side, which is now the payoff b rather than a since the γ countries with $m_j > \overline{m_j}$ will not switch to D. Intuitively, the cost of conflict is higher and the victory reward is lower. Rearranging yields

$$(33)\qquad r^M > \overline{r^M} = \frac{(a-b)\sum_{j=\gamma+1}^{J} s_j r_j^F}{b\sum_{j=1}^{\gamma} s_j + a\sum_{j=\gamma+1}^{J} s_j - d}.$$

By inspection, $\overline{r^M}$ is increasing in the total size of the γ free-riders, $\sum_{j=1}^{\gamma} s_j$. Since the numerator is positive, if the denominator is negative then the threshold $\overline{r^M}$ does not exist. This means that even $\delta_M = 1$ and $r^M \to \infty$ do not guarantee the D outcome for M policy. By inspection this happens if $\sum_{j=1}^{\gamma} s_j$ is above a certain threshold that is a function of $\left(\underset{-}{a}, \underset{+}{b}, \underset{-}{d}\right)$. Alternatively, this condition can be expressed as $d > \bar{d}\left(\underset{+}{a}, \underset{-}{b}, \underset{-}{\gamma}, s_j\right)$, where $\bar{d}$ is decreasing in s_j for all $j \leq \gamma$, and increasing in s_j for all $j > \gamma$. □

Appendix E. A Specific Example of the Free-riding Function (Can Be Removed)

Let us demonstrate the intuition using the following functional form of (12)

$$v_j^A = \left(v_j^B + 1\right)(1 - m_j)^{\frac{1}{s_j}} - 1,$$

and assume a value of $v_j^B = \frac{3}{4}$. From $v_j^A < 0$ one can derive the free-riding threshold over which F outcomes in country j deteriorate after accession

$$m_j > \overline{m_j} = 1 - \left(\frac{4}{7}\right)^{s_j} > 0.$$

In this case for a country that forms 10% of the M union the free-riding threshold is $\overline{m_j} \simeq 0.055$. To give an example of when this will spill over to M policy consider the special case of a union in which all countries are the same size, and further assume $R = 0$ and the normalized payoffs in (1)-(2) prior to joining. Then if the number of member countries with $m_j > \overline{m_j}$, denoted γ, is sufficiently large, $\frac{\gamma}{J} \geq (1 - d)$, the central bank's initial D action is no longer history-independent. Hence it may play I on the equilibrium path even if $\delta_M = 1$ and $r^M \to \infty$.

Appendix F. An Example of Mapping from a Macro Model to the Game (Can Be Shortened or Removed)

The intuition of our policy interaction follows Nordhaus (1994). We will summarize it in schematic terms, but it should be noted that this only serves as a simple example to demonstrate *the techniques* in the main text that can be applied to analytically tractable macro models.

F.1. **Policy Preferences and Economy.** Each policymaker $i \in \{F, M\}$ minimizes the fluctuations of inflation π and output gap x from their respective targets π_i^T and x_i^T. Let us depict the standard quadratic loss which Woodford (2003) showed to be derivable from micro-foundations

$$u_i = -\beta_i \left(x - x_i^T\right)^2 - \left(\pi - \pi^T\right)^2, \tag{34}$$

where $\beta > 0$, and where we reduced the amount of heterogeneity by assuming a common inflation target for the policies (and for parsimony we will set it to zero, $\pi_M^T = \pi_F^T = \pi^T = 0$). The assumed responsible central bank and ambitious government can thus be described by

$$x_F^T > x_M^T = 0. \tag{35}$$

We will use the simplest reduced-form economy in which both policies can affect both targeted variables, either directly through the constraints of the economy or indirectly influencing the optimal choice of the other policymaker

$$x = \mu(\pi - \pi^e) + \rho\,(G - \pi)\,. \tag{36}$$

The π^e variable denotes inflation expectations for the coming period that are formed rationally by private agents. Since our focus is on the medium-run outcomes, neither the exact details of expectations formation nor inclusion of shocks would affect our conclusions.

The G variable is the instrument of F policy, which should be interpreted broadly as the medium-run stance of F policy.[29] In terms of M policy, we assume the central bank to directly use π as its instrument. The parameters $\mu > 0$ and $\rho > 0$ will hence be referred to as the *potency* of M and F policy respectively. Excessively expansionary, excessively contractionary, and balanced policies can therefore be described by $G > 0$, $G < 0$, and $G = 0$ respectively for F policy, and $\pi > 0$, $\pi < 0$, and $\pi = 0$ respectively for M policy. We will assume the $G = 0$ and $\pi = 0$ levels to be socially optimal.

Using (34)-(36) and rational expectations we first derive the policy reaction functions, and then solve them jointly to obtain the following equilibrium outcomes in the medium-run

$$\pi^* = \beta^M x_T^F(\rho - \mu),\;\; G^* = \frac{x_T^F}{\rho} + \beta^M x_T^F(\rho - \mu),\;\; \text{and}\;\; x^* = x_T^F. \tag{37}$$

F.2. **Game Theoretic Representation.** In truncating the players' action sets from continuous to only two levels for each policymaker (D and I) we will follow Cho and Matsui (2005). They choose the two natural candidates - the socially optimal π and G levels, and the time-consistent (but socially sub-optimal) levels from (37)[30]

$$M^D = F^D = 0 \;\; \text{and}\;\; M^I = \pi^*, F^I = G^*. \tag{38}$$

Substituting (37)-(38) into (36) together with rational expectations yields the output gap x for each strategy profile in the payoff matrix. Using this x together with π from (37)-(38) in (34)-(35) we obtain the values of the payoffs in (1)-(2), $\{a, b, c, d, v, w, y, z\}$, as functions of the macro model's parameters. Those of the central bank are as follows:

(39)

		F	
		D	I
M	D	$a = 0$	$b = -\beta^M \left[x_T^F + \rho\beta^M(\rho - \mu)x_T^F\right]^2$
	I	$c = -\left[\beta^M(\rho - \mu)x_T^F\right]^2 \left(1 + \rho\beta^M\right)$	$d = -\beta^M \left(x_T^F\right)^2 - \left[\beta^M(\rho - \mu)x_T^F\right]^2$

This for example shows that the central bank's conflict cost b is increasing in β^M (ie decreasing in the degree of conservatism/strictness) as argued in the main text.

[29]The specification in (36) postulates that the real economy is affected by F policy in *real* terms - due to the medium-run focus. Nevertheless, the general picture of the policy interaction would not change if we assumed only nominal effects as does Nordhaus (1994). The specification also implicitly assumes that the economy exhibits some non-Ricardian features (eg naïve voters or borrowing constraints).

[30]To ensure that M has two distinct actions assume $\rho \neq \mu$.

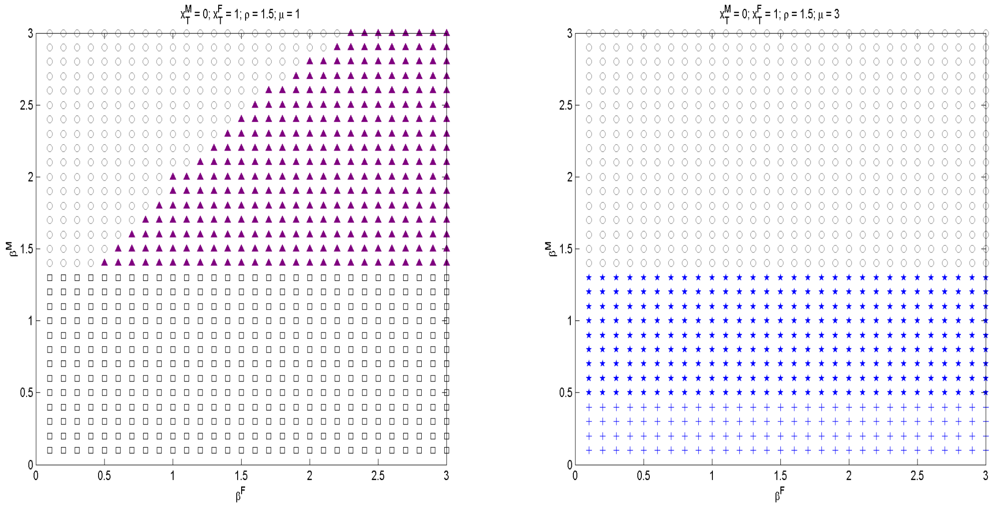

FIGURE 4. Outcomes under various β^M (vertical axis) and β^F (horizontal axis) for $x_T^M = 0, x_T^F = 1, \rho = 1.5$, and $\mu = 1$ (left) or $\mu = 3$ (right). The symbols denote the following scenarios: square is Neglect, pyramid is the Battle, circle is Pure coordination, star is Symbiosis, and cross is Tug-of-war.

F.3. **Scenarios.** We can now examine the scenarios that may arise as the underlying parameters change. Under a responsible M and ambitious F in (35) we can obtain the five scenarios discussed in the main text: Battle, Tug-of-war, Pure coordination, Neglect, Symbiosis, and Pure coordination.[31] Figure 4 offers a graphical illustration. It suggests that if M policy is more potent than F then the more favourable scenarios obtain, namely Symbiosis, Pure coordination, and Tug-of-war. If the reverse is true F spillovers are more likely to occur as we have the Neglect or Battle scenarios for most parameter values.

We can now use the mapping to derive the necessary and sufficient M commitment threshold as a function of the macro variables. For example assuming parameter values that yield the Battle scenario, and focusing on the special case of $R = 0$ with fully patient players and the payoffs in (1), the threshold (8) becomes

$$\overline{r^M}(0) = \frac{r^F}{1 - \beta^M \left(x_T^F\right)^2 \left[1 + \beta^M (\rho - \mu)^2\right]}. \tag{40}$$

This offers three insights related to the fact that $\overline{r^M}(0)$ is increasing in x_T^F, β^M, and $(\rho - \mu)^2$. First, a central banker facing a more ambitious government needs to be more strongly committed to withstand F spillovers. Second, explicit M commitment can partly substitute for M conservatism. Third, the greater the difference in the effectiveness of the policies, the greater the extent of the policy conflict, and hence the stronger

[31] Allowing for $x_F^T = 0$ and/or $x_M^T > 0$ obviously yields additional scenarios.

the M commitment required for policy coordination and socially optimal outcomes. Nevertheless, as stressed above the robustness of these findings would need to be examined further in a range of richer macroeconomic models.

Appendix G. Empirical Evidence (Can Be Shortened or Removed)

Our analysis implies that for *some* but not all parameter values, a more explicit long-term M commitment can have three effects.

G.1. **Effect of r^M on Inflation.** It was shown above that an increase in r^M may reduce the average level and the variability of inflation, and increase M policy credibility, but *only if* the initial level of explicitness had been insufficient to achieve low and credible inflation, $r^M < \overline{r^M}$. Otherwise r^M may have no medium-run effect. Our results are therefore not equivalent to the claim that adoption of inflation targeting will necessarily lower the level and variability of inflation. Unfortunately, this is what the empirical literature has commonly tested for, and therefore reached conflicting conclusions.

Our analysis implies that: (i) papers only including industrial countries are likely to find weak or insignificant effects of inflation targeting on inflation and its volatility (Ball and Sheridan (2003) and Willard (2006), whereas (ii) papers with larger country samples also including emerging and developing countries are likely to find strong and significant effects (eg Fang and Miller (2010), Neyapti (2009), Corbo, Landerretche and Schmidt-Hebbel (2001)).

Furthermore, in line with the prediction of our model, inflation has been found negatively correlated with two common proxies for M commitment: accountability (Briault, Haldane and King (1997)) and transparency (Chortareas, Stasavage and Sterne (2002), Fry et al. (2000)). See also Debelle (1997) who finds inflation targeting to increase the M policy's credibility. All these papers include either pre-1980 inflation data and/or developing countries. In contrast, papers that only focus on industrial countries and use post 1990 data often find no correlation between transparency and inflation, see eg Eijffinger and Geraats (2006).

G.2. **Substitutability of r^M.** The paper implies that long-term M commitment can act as a partial substitute for central bank (goal) independence (patience δ_M and conservatism $\frac{1}{d}$) in achieving credibly low inflation. The testable implication is that countries with initially low degree of central bank independence were more likely to adopt an explicit inflation target: low δ_M and high d lead to higher $\overline{r^M}$.

This is supported by a negative correlation between indices of central bank (goal) independence and accountability, which has been reported by eg Briault, Haldane and King (1997), de Haan, Amtenbrink and Eijffinger (1999) and Sousa (2002).

If we plot the Sousa (2002) final responsibility index against the length of term in office (which is one of the criteria in his independence index) the same conclusion is reached. Furthermore, in a comprehensive data set of Fry et al. (2000) the length of term in office is negatively correlated to accountability procedures in both industrial and transition countries. Finally, Hughes Hallett and Libich (2007a) present evidence that transparency, too, is negatively correlated to goal-CBI. For example, it is shown that

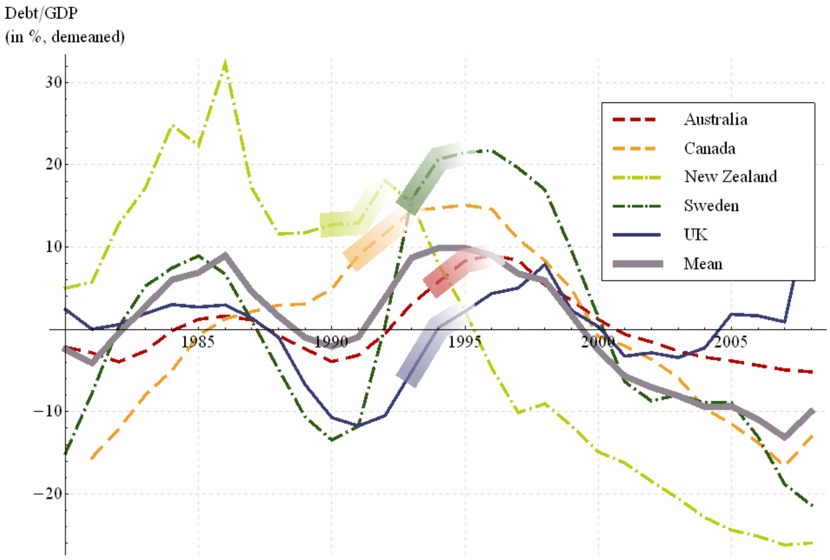

FIGURE 5. Government debt to GDP ratio (in %, demeaned), the start of the shaded area indicates adoption of an explicit inflation target.

the correlation between transparency in Eijffinger and Geraats (2006) and goal-CBI in Briault, Haldane and King (1997) is −0.86 (and the t-value equals −4.46).[32]

G.3. **Effect of r^M on Fiscal Outcomes.** As our third testable implication, the paper implies that M commitment may be able to discipline F policy and induce reductions in the average level and variability of the budget deficit and debt.

In our companion paper Franta, Libich and Stehlík (2010) we examine this hypothesis formally using various empirical strategies (time varying parameters VARS, SVARS, and an estimated DSGE model). We show that countries that have adopted an explicit inflation target (such as Australia, New Zealand, Canada, Nordic countries etc) have improved their F outcomes relative to comparable non-targeters (such as the United States, Japan, or Switzerland). To provide one piece of supporting evidence Figure 5 shows the behaviour of central government debt to GDP ratio in five early adopters of inflation targeting. In contrast to non-targeters, all five countries have seen a reduction in the ratio starting about 1-3 years post-adoption.[33]

As a demonstration of the mechanism that lies behind, and that we modeled above, Figure 6 shows the reaction of the interest rate instrument in Australia to an unexpected government spending shock from an SVAR model with five endogenous variables: government spending, output, private consumption, short term interest rate, and

[32]This paper also demonstrates that the Debelle and Fischer (1994) distinction between goal and instrument CBI is important. Since instrument CBI has come hand in hand with inflation targeting (as one of the prerequisites of the regime, see eg Masson, Savastano and Sharma (1997) or Blejer and et al. (2002)) its correlation with transparency and accountability is positive in most cases, see eg Chortareas, Stasavage and Sterne (2002).

[33]While these findings are consistent with the disciplining effect discussed above, it should be stressed that they do not constitute evidence of causality.

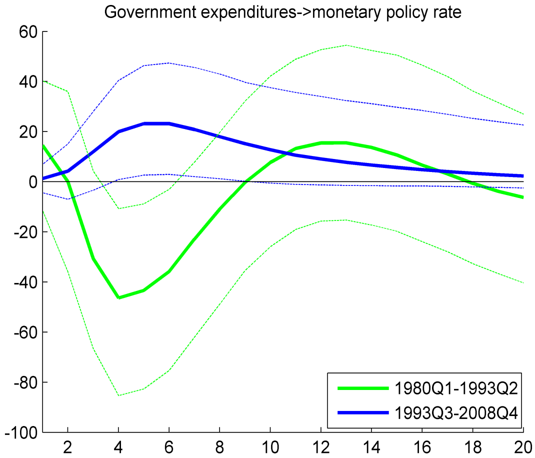

FIGURE 6. Impulse response of the Australian cash rate to a unit unexpected government spending shock from an SVAR model, for details see Franta, Libich and Stehlík (2010).

government debt. It shows that M policy reaction is qualitatively different in the pre and post-inflation targeting period. In the latter M policy accommodates unexpected government spending, whereas in the post period it counter-acts it. This is in stark contrast from the non-targeters where such offsetting is not present. The same difference is present when comparing government debt (the above targeters have usually seen a marked improvement starting about 1-3 years after the regime's adoption, unlike the non-targeters).