Tiago Tavares Flórido

Transitions in Central Bank Leadership: Empirics and a Simple Theory

DISSERTAÇÃO DE MESTRADO

DEPARTAMENTO DE ECONOMIA
Programa de Pós–graduação em Economia

Rio de Janeiro
March 2015

Pontifícia Universidade Católica do Rio de Janeiro

Tiago Tavares Flórido

Transitions in Central Bank Leadership: Empirics and a Simple Theory

Dissertação de Mestrado

Dissertation presented to the Programa de Pós-Graduação em Economia of the Departamento de Economia, PUC-Rio as partial fulfillment of the requirements for the degree of Mestre em Economia.

Advisor : Prof. Carlos Viana de Carvalho
Co–Advisor: Prof. Eduardo Zilberman

Rio de Janeiro
March 2015

Pontifícia Universidade Católica
do Rio de Janeiro

Tiago Tavares Flórido

Transitions in Central Bank Leadership: Empirics and a Simple Theory

Dissertation presented to the Programa de Pós-Graduação em Economia of the Departamento de Economia do Centro de Ciências Sociais da PUC– Rio as partial fulfillment of the requirements for the degree of Mestre em Economia. Approved by the following commission:

Prof. Carlos Viana de Carvalho
Advisor
Departamento de Economia — PUC–Rio

Prof. Eduardo Zilberman
Co–Advisor
Departamento de Economia — PUC–Rio

Prof. Leonardo Bandeira Rezende
Departamento de Economia – PUC–Rio

Prof. Bernardo de Vasconcellos Guimaraes
Fundação Getúlio Vargas -SP, EESP

Prof. Monica Herz
Coordenadora Setorial do Centro de Ciências Sociais – PUC–Rio

Rio de Janeiro, March 10th, 2015

Tiago Tavares Flórido

Graduated in Economics from PUC–Rio in 2013.

Bibliographic data

Flórido, Tiago

Transitions in Central Bank Leadership: Empirics and a Simple Theory / Tiago Tavares Flórido; advisor: Carlos Viana de Carvalho; coadvisor: Eduardo Zilberman. — Rio de Janeiro : PUC–Rio, Departamento de Economia, 2015.

v., 88 f: il. ; 29,7 cm

1. Dissertação (Mestrado em Economia) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento De Economia.

Inclui Bibliografia.

1. Economia – Dissertação. 2. Política Monetária. 3. Banco Central. 4. Transições. 5. Reputação. I. Carvalho, Carlo Viana. II. Zilberman, Eduardo. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. IV. Título.

CDD: 330

Acknowledgments

First, I would like to thank my advisors, Carlos and Dudu, for their continued guidance throughout the master's program. Their teachings on economic research were invaluable to this thesis and will carry on with me through my next academic steps. This thesis also greatly benefited from Profs. Leonardo and Bernardo's suggestions: from Leo analyzing my early scribbles at his whiteboard to the comments they both gave during the thesis defense. I am also thankful to my other professors, in particular Gustavo, who encouraged me to join the master's program, and Rogério, who was always available to chat with and give advice on whichever topic I required.

On a personal level, I thank all of my friends. Those who never understood what I could possibly be doing at PUC during weekends helped me (often by teasing me) not to forget that there was a life outside PUC. However, for better or worse, it was at PUC that I spent most of the past two years. Therefore the friends who shared these experiences with me are not only recipients but also objects of my gratitude. From the endless hypothetical dilemmas at Bandejão to the occasional chaos and coup d'état attempts at Favelinha, they made these two years very fun.

I am also thankful to Alice, whom I could have never done without. In regard to this thesis, she mercilessly corrected my writing inumerous times. Nonetheless, all these improvements are inconsequential when contrasted with her importance in my life: she is the one I owe my happiest times to. And for that and everything else I am grateful.

Finally, I thank my family and most importantly my parents, Luiz and Paula, both of whom have always had an irrational confidence in me. It is impossible to be fair to them, so an understatement will have to suffice: any achievement and every challenge overcome are consequences of the love and encouragement I received from them at each moment of my life. And, as in this past year they could take comfort in the fact that I had an income, I gratefully acknowledge financial support from CAPES, CNPq and Vinci Partners.

Resumo

Flórido, Tiago; Carvalho, Carlo Viana; Zilberman, Eduardo. **Transitions in Central Bank Leadership: Empirics and a Simple Theory**. Rio de Janeiro, 2015. 88p. Dissertação de Mestrado — Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

A importância dada à identidade de um banqueiro central demonstra que transições na liderança do banco central são um aspecto importante da política monetária. Analisamos essas questões usando um painel de países construído especificamente para esse propósito. Provemos evidência que períodos de transição são associados a uma politica monetária mais contracionista e avaliamos empiricamente algumas explicações. Os resultados não podem ser explicados por ciclos eleitorais, política fiscal ou endogeneidade das transições. Os resultados são mais fortes quando o banco central é menos transparente, menos independente ou quando a qualidade regulatória do país é menor. Além disso, os resultados são mais fracos quando o banqueiro central já era membro do comitê e são mais fortes quando o banqueiro central cujo mandato está acabando é mais forte. Propomos uma explicação baseada em transferência de reputação. Para esse fim, desenvolvemos um modelo simples de inconsistência temporal da politica monetária em que o banqueiro central cujo mandato está acabando pode, ao distorcer suas decisões finais, afetar as crenças que o público tem a respeito de seu sucessor: uma transferência reputacional ocorre.

Palavras–chave

Política Monetária; Banco Central; Transições; Reputação;

Abstract

Flórido, Tiago; Carvalho, Carlo Viana; Zilberman, Eduardo. **Transitions in Central Bank Leadership: Empirics and a Simple Theory**. Rio de Janeiro, 2015. 88p. Dissertação de Mestrado — Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

The importance assigned to the identity of a central banker suggests that transitions in central bank leadership are important times for monetary policymaking. We analyze such transitions empirically using a country panel that we assemble specifically for that purpose. We provide evidence that transition periods are associated with a more contractionary monetary policy stance, and empirically assess a few natural explanations. Our findings cannot be explained by electoral cycles, fiscal policy or endogenous transitions. Results are stronger when the central bank has less independence, is less transparent, and when the country's regulatory quality is lower. Results are weaker when the new governor was previously in the committee, and are stronger when the outgoing governor had more power. We offer an explanation based on reputation transfer. To that end we develop a simple model of monetary policy with time inconsistency in which a departing central banker can, by distorting his final decisions, affect the public's beliefs about his successor and engender a reputation transfer

Keywords

Monetary Policy; Central Bank; Transitions; Reputation;

Contents

List of figures

List of tables

1 Introduction

The importance of monetary policy to economic outcomes is uncontroversial. Consequently, a lot of attention has been given to central banks, who conduct monetary policy with varying degrees of independence, throughout the world. Moreover, this attention is not limited to central banks as institutions, but also encompasses the people in the leadership positions of these institutions.

Recently this fact came into particular prominence due to the end of Fed Chairman Ben Bernanke's term and the consequent speculations about possible successors. Before it was confirmed that vice-chairwoman Janet Yellen would be the next leader of the United States' central bank, there were several articles in the media discussing pros and cons of different 'candidates' [1]. This reflects the great importance that is assigned to the identity of a central bank's leader.

This view also finds resonance in the academic literature in different forms. Romer and Romer (2004) analyse historical transcripts and past speeches to offer support to the idea that a central banker's views about the economy are a key determinant of the success of the monetary policy. With a theoretical approach, Rogoff (1985) analysed the preferences of a central bank and concluded that, in order to address the temporal inconsistency problem, the government should appoint a central banker who is more conservative than society as a whole.

Nevertheless, little has been done in the academic literature to study the transition period itself. The main objective of this paper is thus to analyse how monetary policy changes during transition periods. In particular, we are interested in studying the reputation aspects of transition periods. After all, given the importance of keeping inflation expectations anchored, it is vital that the public trusts that the central banker will keep inflation under control - the central banker must have a solid reputation.

Naturally, the public has great uncertainty about new central bankers: "So, Mr. Carney, Hawk or Dove" at the WSJ and "ECB: Clearing the way for an Italian hawk?" at the BBC demonstrate such uncertainty. Indeed, central bankers are aware of this special uncertainty, as the quote bellow illustrates.

[1] For a particularly stark opinion piece, see: "Why Janet Yellen, Not Larry Summers, Should Lead the Fed" by Joshph Stiglitz in the New York Times at September 6, 2013.

> "I think any action we take–because we are certainly in the **spotlight today**–will be looked at very eagerly and there are **psychological reactions** coming from what we do." Winn, Cleveland Fed President, at Volcker's first meeting.

The public's uncertainty and central bankers' awareness motivate our two main questions. First, what are the incentives of the new central banker? Can he signal to the public that he is a Hawk? Second, what are the incentives of the departing central banker? Namely, how a central banker with an established reputation and at the end of his tenure can affect future economic outcomes and potentially help his successor signal commitment to sound monetary policy. We address these questions both empirically and theoretically.

Empirically, we construct a novel dataset containing 35 countries in order to establish how central bankers' behaviour differ in the first monetary policy meetings of a new central banker and in the last meetings of a departing one from the other "usual" meetings. We find that both the first and last meetings are associated with a more contractionary stance. We also investigate how the results change with different specifications and when we consider heterogeneity in leadership transitions. The overall empirical evidence is consistent with signalling incentives in the first meetings and with "reputation transfer" in the final meetings.

To the best of our knowledge, the only paper which has offered empirical evidence on signalling in monetary policy is Hansen and McMahon (2013). They use data from the Bank of England monetary policy committee to show that new members tend to be tougher on inflation initially to signal they are not dovish. As they use data from one country they cannot, however, answer the question of how a departing central bank leader can affect his successor. Our focus on governors (the leaders)'s transition allows us to do that. The incentives of both entering and departing central bankers alike are mapped. It is also important to stress that we abstract from committee voting considerations, an important part of Hansen and McMahon (2013), and treat monetary policy decisions as made by one entity embodied by the governor.

Thus we provide reduced form empirical evidence for signalling by new governors in the spirit of Hansen and McMahon (2013) for multiple countries. Also, we are the first to provide empirical evidence on the incentives of the departing governor, who has also a role to play during transition dynamics. Final meetings are associated with tighter monetary policy.

As we are the first to document how a governor distorts his decision at final meetings, we develop a model that shows how this behaviour might be rationalized. Extending the Barro and Gordon (1983) model in order to study transition dynamics, we show how type uncertainty about the central bank leads to a signalling game as in the seminal work of Spence (1973). Furthermore, the model dynamics allow a departing central bank to help the signalling done by the new central banker.

Besides offering a rationalization for the behaviour found empirically, our model provides a simple framework to study transition periods in monetary policy, which the theoretical literature overlooks. The main exception is Debortoli and Nunes (2011), who study the effects of exogenous regime switches between a hawkish and a dovish central banker in an infinite period model with loss functions akin to the standard Barro and Gordon (1983) model but with a New Keynesian Phillips Curve. One of the advantages of this approach is that they can write the problem recursively and thus they connect better with the bulk of modern macroeconomics literature.

Nevertheless, Debortoli and Nunes (2011) is unsuitable for the study of reputation effects and reputation transfers. There the shifts in central bank leadership are exogenous and there is no type uncertainty - agents always know which type of central banker they are dealing with. Also, the absence of endogenous state variables eliminates the possibility of the present central banker to strategically affect future policy makers. Alas, adding such features to the regime-switch model would make it intractable.

It is important to stress what we mean by 'reputation'. Reputation is how likely the public thinks that the governor is a Hawk. Thus we do not model a repeated game with full information where 'reputation' arises in the sense that the agents expect a low level of inflation while the central bank 'behaves' by delivering inflation equal to expectations and punish the central banker otherwise - a trigger strategy. Our choice is based on several reasons. First, the reputation dynamics we are interested derive from the uncertainty surrounding a new central banker as illustrated by the news articles above; second, our focus on transitions warrants few time periods instead of a long repeated game; third, repeated games can usually sustain almost any kind of equilibrium as the discount factor approaches one - Folk theorems, as in Fudenberg and Maskin (1986), kick in.

Thus, reputation transfer is modelled as the way a departing central banker can affect, through his choice of monetary policy, the process of belief formation by economic agents regarding the new central banker's type. Although our set up and approach to uncertainty relates to many vast literatures, our model contributes to the theoretical literature by addressing an important real monetary policy question not yet answered - reputation incentives in transitions. Moreover, the model sheds light on our empirical results, which are our main contribution: the first paper to document how a departing central banker's behaviour changes and the first to document new central banker's behaviour with data from many countries.

This paper is divided as follows: in section 2 we develop the empirical analysis, which is the core of the paper; in section 3, we present the model that illustrates how the empirical results can be rationalized; in section 4 we conclude.

2
Empirical Analysis

2.1
The Data

We construct a novel dataset: a 'panel' composed of 35 countries, where each observation ct consists of a country c and a monetary policy meeting t. One should note that t does not correspond to the same time period. After all, the t-th meeting we have for, say, the United States FOMC is not at the same date as the t-th meeting of the UK monetary policy committee. In fact, they do not even have to have the same periodicity: countries vary in the number of meetings held per year - spanning from monthly to quarterly meetings. In addition, the countries enter the sample at different years, starting in 1984 with the US while Georgia is the last country to enter the sample in 2008. Table 6.2 in Appendix A lists all countries and their number of meetings and governors.

The panel is unbalanced because we only use data from countries and periods where there are published decisions from every monetary policy meeting, and where there is a meeting calendar. For instance, until the late 1990's many countries did not announce whether a meeting took place - they simply announced when there was a change in the target interest rate and thus we cannot know when it was actively decided to keep policy constant. Furthermore, we use the data only when the policy target is an interest rate (in contrast with money supply growth rates). Finally, we drop the financial crisis period - 2008 and 2009 - since this would confound our results on transition effects. There were many things affecting monetary policy during that period so it is hard to establish how monetary policy changed due to the transitions per se in comparison to regular policy.

The average number of meetings used per country is 111 and the median is 102. The total number of transitions in the sample is around 70. Our main variables[1] are:

– Policy interest rate decisions(in %): i_{ct}

– Inflation (YoY): π_{ct}

– Activity level : y_{ct} - mostly unemployment when available, output growth otherwise.

[1]For the data which was not seasonally adjusted from the source, we use the ARIMA X12 deaseasoning proceeding.

- Dummy for First Meeting of a CB governor: FM_{ct}
- Dummy for Last Meeting of a CB governor: LM_{ct}

The data comes from four main sources: the OECD database, *Bloomberg* terminals, *Datastream - Thomson Reuters* terminals and individual central banks' websites. Typically, the policy rate decisions and governors' changes were taken from each central bank's website, since we wanted the specific meeting dates which are typically irregular. The macroeconomic series were mainly obtained from the OECD database and the data terminals. However, there were countries whose time series available on the terminals were too short or inexistent. For these countries, we searched central banks' websites and national data bureaus to obtain the macroeconomic series.

We match the macroeconomic series with each central bank meeting of a given country according to the following algorithm. First, we identify the calendar month of each meeting t. For instance, a meeting in the 17th of April counts as April. Then we match with the inflation and unemployment referring to that calendar month. However, some countries do not report unemployment monthly. In these cases we check the availability of quarterly data for unemployment and GDP growth. We use the quarterly value for the three months of the corresponding quarter, as if it was a monthly variable. For instance, if the rate of unemployment was 7 % for the second quarter of a given year, we input 7% in the cells referring to April, May and June. Then we proceed as before matching the quarterly rate to meeting in the corresponding month. Thus, when there is no monthly unemployment rate, we use the quarterly unemployment rate and, when even such periodicity is lacking, we use GDP growth as the activity level variable.

The data on first and last meetings of a governor is found in each central bank website. Normally, there is a webpage reporting the list of former governors with the dates of their tenures. In cases where this page was ambiguous about in which exact meeting the transition happened, we check the minutes of the relevant meetings in order to make sure when the transition took place. It is important to note that it will not always be the case that the final meeting of a governor is the one exactly before the first meeting of his successor - in other words, LM_{ct} will not always be the lag of FM_{ct}. This happens because sometimes a governor's tenure ends before the appointment of his successor so that there may appear a acting governor for a couple meetings. In the Appendix 6 we discuss in detail how we code these transitions, but in any case our results vary little within reasonable code changes.

Figure 2.1: Transitions per Biennium

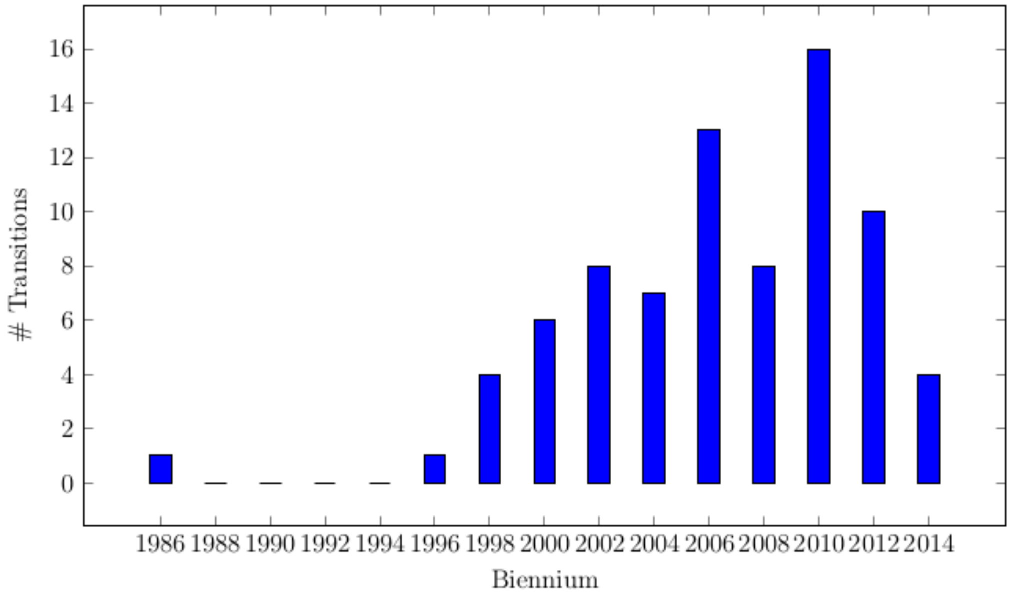
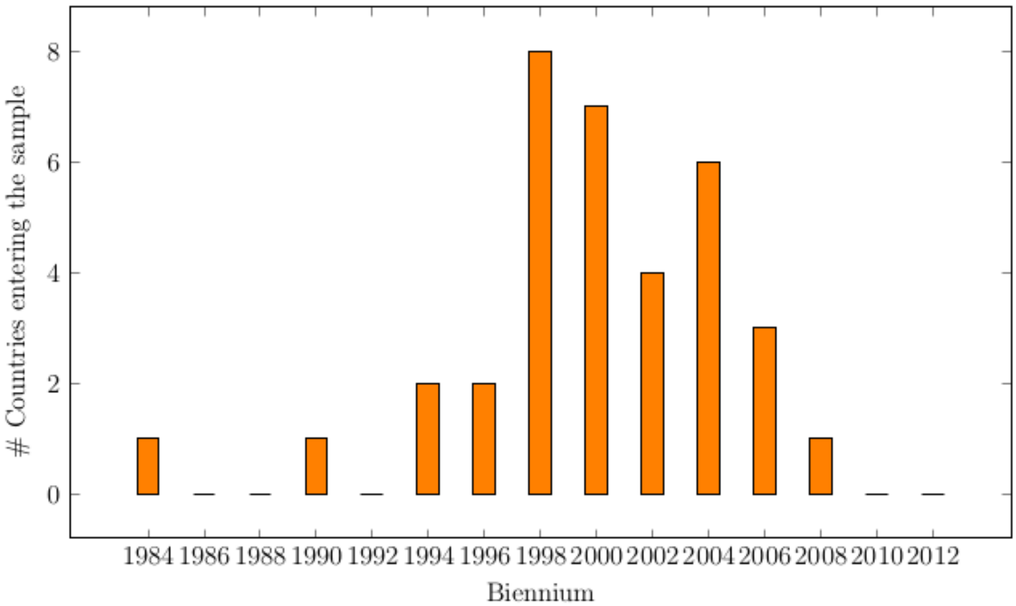


Figure 2.2: Countries Entering the Sample per Biennium



Finally, before we begin the empirical analysis, it is important to assuage a possible concern about our dataset. That is, the possibility that most transitions are clustered around a couple of years. Figure 2.1 shows that the transitions are scattered, with most of them after the late 1990's. In fact, most of our sample begins in the 1990's as shown in Figure 2.2.

2.2 Empirical Strategy

Recall that $FM_{c,t}$ ($LM_{c,t}$) is a dummy which takes value one when meeting t is a first (last) meeting of a given CB governor c. In order to estimate the effects of transitions in monetary policy, we add $FM_{c,t}$ and $LM_{c,t}$ to a simple Taylor rule in which inflation $\pi_{c,t}$, economic activity $y_{c,t}$ (either unemployment or GDP growth as explained above) and lagged interest rate $i_{c,t-1}$ are accounted for. In particular, we pool all observations and allow the coefficients on each of these variables to vary with countries. Moreover, we allow each country to have a different intercept δ_c and we include year dummies δ_y for $y = 1984, ..., 2014$. Hence, we estimate the following Taylor rule by OLS:

$$i_{c,t} = \rho_c i_{c,t-1} + \alpha_{\pi,c}\pi_{c,t} + \alpha_{y,c}y_{c,t} + \beta_F FM_{c,t} + \beta_L LM_{c,t} + \delta_c + e_{c,t} \qquad (2\text{-}1)$$

Our coefficients of interest are β_F and β_L. The idea is that, once changes in monetary policy warranted by macroeconomic factors are accounted for, β_F and β_L capture the effect of transitions in the interest rate $i_{c,t}$. In order to the exercise be meaningful, the bulk of variation in monetary policy due to macroeconomic factors must be accounted for. It is reassuring that, despite the simplicity of the functional form above, the R^2 of our baseline specification is 99.6. In fact, the smoothing term improves a lot the fit of the regression. In the Appendix 6, we report and discuss how the results would change were two lags included in the Taylor rule.

2.3 Results

It is important to underscore that the transition incentives faced by departing and incoming governors do not necessarily have to be limited to only the first and last meetings. For instance, a departing governor could influence his successor by changing policy at the penultimate meeting and not making any changes at the very last meeting. Hence, we report the results from regression (2-1) for different specifications. For example, the variable $FM_{c,t}$

($LM_{c,t}$) may include the first (last) n meetings. If $n = 2$, for instance, the specification consider the first (last) two meetings.

Before reporting the results, one word on inference: throughout the empirical section we use robust standard errors as usual. Even so, in the Appendix 6 we engage in a Monte Carlo exercises which seeks to shed light on when robust standard errors might not be adequate and we then report how our main results remain essentially the same when we use Driscoll and Kraay (1998) errors, which are resistant to many criticisms against robust errors.

2.3.1 Baseline

At Table 2.1, we report the results from regression (2-1) for specifications with n (# Meetings) varying from one to four.

Table 2.1: Baseline Regression: $i_{c,t}$ is the dependent variable

# Meetings	1	2	3	4
$FirstMeet(\beta_F)$	0.052*	0.075***	0.061**	0.036
	[0.072]	[0.002]	[0.048]	[0.159]
$LastMeet(\beta_L)$	0.088*	0.076**	0.088***	0.089***
	[0.098]	[0.025]	[0.001]	[0.001]
Country FE	Y	Y	Y	Y
Year Dummy	Y	Y	Y	Y
# Obs	3881	3881	3881	3881

P-value between [ ], calculated with robust standard errors.

Table 2.1 shows that both the first few and last[2] few meetings are associated with higher interest rates than those prescribed by the Taylor Rule in comparison to regular meetings. These results are highly statistically significant and seem robust across specifications. Moreover, they are economically relevant: take column 2, it is an increase of around 0.075 percentage points in the first meetings and 0.070 in final ones. To provide some idea of this magnitude, we note that over 50% of interest changes are of 0.25 percentage points.

As we will detail in our model at section 3, these results are consistent with a signaling process. A new governor who knows that the public is uncertain whether he is a Hawk or a Dove has incentives to tighten monetary

[2]We shall use 'final' interchangeably with 'last' when applied to 'meeting'.

policy in order to signal he is a Hawk and face lower inflation expectations in the future - a positive β_F. As to the departing governor, if he wants to help a Hawk successor to signal his type, he should tighten monetary policy to make it harder for a Dove to pretend he is a Hawk. After all, Doves should find it more costly to tighten monetary policy further after an interest rate increase. It is important to stress that for this interpretation about the departing central banker's actions to make sense, we have to assume that the departing central banker knows more about his successor than the economic agents. That is why departing banker's actions can be informative about the type of the new central banker. We believe this to be quite reasonable. First, in many cases it is possible than a departing governor's prestige might give him some say in the choice of his successor. Even if this is not the case, it is possible for both central bankers to have met in informal talks before and during the transition period. Furthermore, it will almost always be the case that the departing governor knows more about the public's opinion on the new governor than the public knows about the departing banker's opinion on his successor. Consequently, if the departing governor believes the public is being too harsh, he might want to affect their beliefs - igniting a reputation transfer.

Besides being an important aspect of empirical Taylor rules, the interest rate smoothing address a concern that the new central banker might not be tightening policy. Assume that the departing banker increases interest rates above the prescribed by the Taylor rule, generating a positive residual. Even if the new governor did not move policy, it is likely that macroeconomic conditions would have little changed from one meeting to the other so that the residual of the Taylor rule would remain positive. However, smoothing prevents this to be the case. As the smoothing coefficient is quite high (almost always above 0.9), most of the interest rate hike engendered by the departing banker is absorbed by the Taylor rule. Consequently, a positive coefficient of similar magnitude means that there was indeed a further tightening during the first few meetings - precisely as our reputation theory implies.

In the remainder of the empirical section, we use the heterogeneity of transitions to address concerns of endogeneity and to try to pin down the mechanism of the results: whether the heterogeneities are still consistent with signalling dynamics and inconsistent with a few competing interpretations. Then, in the remainder of the empirical section, we report the results considering $n = 2$ and $n = 3$ meetings. In this way we are not limited to a too short time interval nor we are allowing the transition to last too long. In addition, reporting both way shows how robust each result is when one shifts from using the first (last) two meetings to using the first (last) three meetings.[3]

2.3.2 Transition Timing

Even after establishing that transitions are associated with tighter monetary policy stance, one could argue that they are endogenous. Transitions could be more likely to occur at times when interest rates are above the prescribed by the Taylor rule. We make sure that this is not driving our results by exploring transitions' timing. We analyse the heterogeneity between transitions where the governor mandate is fixed (e.g. US) and those where there is no such regularity (e.g. Brazil).

The idea here is that, for fixed mandates, the transition should not have been caused by other factors that could also cause the increase in the interest rate - the end of the mandate was determined early on. Consequently in these cases we can rule out the channel of tighter monetary policy triggering leadership transitions. In the first column of Table 2.2, we report the coefficients of regression (2-2) (always including year and country fixed effects).

$$i_{c,t} = \rho_c i_{c,t-1} + \alpha_{\pi,c}\pi_{c,t} + \alpha_{y,c}y_{c,t} + \beta_F FM_{c,t} + \beta_{FN} FM_{c,t} \times NotFix_c + \beta_L LM_{c,t} + \beta_{LN} LM \times NotFix + e_{c,t} \tag{2-2}$$

On the equation above, β_F and β_L capture the effect for the transitions where there is a fixed governor's mandate and $\beta_F + \beta_{FN}$ and $\beta_L + \beta_{LN}$ are the effect of those transitions where the governor's mandate is not fixed. As before, β_F and β_L are the coefficients we are mostly interested in.

Table 2.2: Fixed Regimes and Unannounced Resignation

# Meetings	2	3		2	3
FM	0.046**	0.039	FM	0.043*	0.036
	[0.036]	[0.254]		[0.069]	[0.330]
$FM \times NotFix$	0.179**	0.150*	$FM \times Unan$	0.130**	0.105*
	[0.043]	[0.075]		[0.044]	[0.096]
LM	0.048*	0.064***	LM	0.013	0.034*
	[0.062]	[0.003]		[0.577]	[0.083]
$LM \times NotFix$	0.190	0.156	$LM \times Unan$	0.221**	0.204**
	[0.267]	[0.203]		[0.025]	[0.012]
# Trans	$61-10$	$61-10$		55-16	55-16

P-value between [ ], calculated with robust standard errors.
Regressions still include FE and Year dummies.

The first two columns of Table 2.2 show that β_F and β_L remain positive, statistically significant, and with an economically relevant magnitude.

[3] Results for $n = 1$ and $n = 4$ are available upon request.

Therefore the transition effects found in Table 2-1 do not appear to stem from endogenous transition timing.

In addition, to be even more cautious, we create the variable $Unan$[4], which includes all the transitions which followed an unannounced resignation and all the ones without a fixed governor's mandate. The reason is that even in countries with fixed mandates, a governor could suddenly resign in a time of crisis, which would trigger a transition with endogenous timing. Therefore we are able to increase our shield against endogenous transitions effects. The coefficients of FM and LM capture the effects of fixed mandate transitions where there was no unannounced resignation.

Now, there are less transitions to estimate FM and LM, which makes it harder to obtain precise estimates - the coefficients of FM and LM are not always statistically significant. Nevertheless, it is possible to see in the last two columns Table 2.2 that FM's coefficient is significant using 2 meetings and LM's coefficients is with 3 meetings. Moreover, the coefficients' magnitudes are broadly similar to those found using all fixed mandate transitions. We interpret the less robust statistical significance of the last 2 columns as a natural consequence of the reduced number of transitions we have to estimate the coefficients of interest.

Considering Table 2.2 as a whole, the empirical evidence suggests that causation direction goes from transitions to tighter monetary policy: there is something going on during transitions that leads to interest rates above the Taylor rule's predictions; it does not seem to be the case that the results stem from transitions being more likely to occur during periods with tight monetary policy.

2.3.3 Governor Fixed Effects

Even if the leadership transitions are not endogenous, one could argue that the choice of the new governor is. After all, it is natural to think that if inflation was too high, the appointment of a hawkish governor to lead the central bank would be more likely. In this case, the Hawk could increase interest rates above those prescribed by the Taylor rule due to his preferences, which could explain the positive sign of FM's coefficient without appealing to signalling interpretations.

In this section we argue that though this criticism may play a role, it cannot account for the full effect - hence the results remain consistent with signalling dynamics. In order to control for different preferences across

[4]Details in Appendix A.

governors, we allow the intercept to also vary with the governor. Hence, each governor may differ in the average interest rate chosen within his country in a given year. In other words, we add governor's fixed effects. A Hawk governor, for instance, should have a higher fixed effect than a Dove. In this case, β_F captures the difference of first meetings with the 'normal' meetings for the same governor. A positive β_F implies that, on average, the same governor is more hawkish during his first meetings than throughout the rest of his tenure - evidence in favour of signalling incentives.

Table 2.3: Governor Fixed Effects

# Meetings	2	3
FM	0.098***	0.084**
	[0.002]	[0.021]
LM	0.039	0.053*
	[0.289]	[0.071]
Gov FE	Y	Y
Year Dummy	Y	Y
# Obs	3881	3881

P-value between [ ], calculated with robust standard errors.

Table 2.3 shows that the results survive after we control for governors' fixed effects. Since this control involves the inclusion of over an hundred governor dummies, it is natural to lose some precision. Nevertheless, the coefficients less precisely estimated were the LM ones, which were less threatened by the criticism of endogenous governors' appointments - the departing governor was chosen well before the transition approached. As the possible criticism we raised early concerned the FM's coefficients, it is reassuring that they remain positive, statistically significant and economically relevant. Consequently, different governors' preferences cannot explain our results. Having addressed the two most immediate sources of endogeneity, the rest of the empirical section analyses whether this result is consistent with signalling and reputation dynamics.

2.3.4 Independence, Transparency and Regulatory Quality

In this section we evaluate how the effects found in leadership transitions vary across countries. To do so, we interact the variables of interest with three indices: two refer specifically to the central bank, *Transparency* and I*ndependence*, and one refers to the whole country, *Regulatory Quality*. The reason why we analyse this heterogeneity is to shed light on the mechanisms behind the leadership transition effect.

In this paper, we have offered an interpretation based on signalling - the new governor who must prove his hawkishness and the departing one who wants to help this process. As we cannot rule out all other alternative interpretations, if the heterogeneity goes in the same direction one would expect were signalling true, this provides further evidence of this view and may help discard other confounding stories.

The assumption underlying this section is that signalling incentives should be stronger where there is greater institutional uncertainty. After all, in places where the public trusts that no governor will try to exploit the short term benefit of inflating the economy, there would be little reason for a new governor to signal that he is a Hawk by distorting monetary policy. On the other hand, in places where every leadership change brings back fears of bad policy, a new governor has strong incentives to signal that he is committed to fighting inflation.

This raises the question of how to measure what we referred to as institutional uncertainty. We opted to use the three indices mentioned before: *Transparency* from Crowe and Meade (2007) , *Independence* from Dincer and Eichengreen (2014) and *Regulatory Quality* from the World Bank Governance Indicators. Although every index has flaws, they should roughly capture this institutional uncertainty. As all indices were constructed by different authors, there is less risk of one methodology being the sole driver of results. Also, since the indices were not constructed to study leadership transitions, we do not worry about hindsight biasing our results.

We interact the variables of interest with these three indices. Results are reported in Table 2.4.

Heterogeneity goes in the direction we expect: the more independent, transparent or the better the regulatory quality, the smaller the tightening of monetary policy during first and last meetings is. The interaction coefficients are always negative and most are statistically significant Even in the cases the coefficients are less precise (when using 3 meetings), the size of the point estimates are very similar to their significant counterparts. In addition, note that the indices' coefficients are economically relevant (indices are normalized) and that transitions are still associated with tight monetary policy for countries with average institutional qualities.

2.3.5

Table 2.4: Independence, Transparency and Regulatory Quality

# Meet	2		2		2
FM	0.073***	FM	0.068***	FM	0.059***
	[0.003]		[0.004]		[0.007]
$FM \times Ind$	−0.038*	$FM \times Trp$	−0.059***	$FM \times RQ$	−0.068***
	[0.070]		[0.004]		[0.002]
LM	0.081**	LM	0.078**	LM	0.071**
	[0.024]		[0.024]		[0.020]
$LM \times Ind$	−0.051**	$LM \times Trp$	−0.040	$LM \times RQ$	-0.043
	[0.016]		[0.155]		[0.180]
# Meet	3		3		3
FM	0.062*	FM	0.062**	FM	0.053*
	[0.056]		[0.060]		[0.065]
$FM \times Ind$	−0.020	$FM \times Trp$	−0.018	$FM \times RQ$	−0.057**
	[0.523]		[0.487]		[0.028]
LM	0.091***	LM	0.089***	LM	0.083***
	[0.001]		[0.001]		[0.001]
$LM \times Ind$	−0.034**	$LM \times Trp$	−0.042*	$LM \times RQ$	−0.045*
	[0.049]		[0.077]		[0.084]

P-value between [ ], calculated with robust standard errors. Indices are normalized
[a] Independence (Crowe and Meade 2007)
[b] Transparency (Dincer and Eichengreen 2013)
[c] Regulatory Quality (World Bank Governance Indicators).

Poor countries

In this section we seek to establish that our results do not rely on one particular group of countries - the poorest in our sample. As we use data from 35 countries, there are varying levels of income (e.g., Norway, Brazil, Kenya). It would decrease the interest of our results were they driven solely by the poor countries. After all, these countries have more idiosyncrasies and often Taylor Rules describe less adequately their monetary policy. Hence it is worthwhile to decompose our results between the 9 poorest countries (GDP per capita less than US$10000 PPP) and the rest.

Table 2.5 shows that the results survive: the coefficients of FM and LM remain positive, statistically significant and economically relevant. These coefficients capture the leadership transition effect for the countries in the sample that are not poor. In addition the coefficients of $FM \times Poor$ and $LM \times Poor$ show how large the effect is in Poor countries. As it can be seen in Table 2.5, the effects seem larger (which could be explained by greater need of signalling) but very noisy: we cannot affirm that the transition effect is statistically different.

Table 2.5: Poor countries

# Meet	2	3
FM	0.063***	0.061*
	[0.008]	[0.067]
$FM \times Poor$	0.092	0.015
	[0.277]	[0.877]
LM	0.061*	0.069***
	[0.079]	[0.009]
$LM \times Poor$	0.113	0.135
	[0.310]	[0.128]
# Trans	56 − 15	56 − 15

P-value between [ ], calculated with robust standard errors.
Regressions still include FE and Year dummies.

2.3.6 Monetary Policy Committee

Although we focus only on governors' transitions, nowadays most monetary policy decisions in the world are made by committees. Without repudiating the working assumption that leadership transitions are the most relevant, we can use the committee structure to exploit heterogeneities and, consequently, test if they are consistent with the signalling interpretation offered for our results. Namely, a new governor wants to signal that he is a Hawk and the departing wants to help this process by making it harder for a Dove to pretend to be a Hawk.

Governor's strength

Earlier we stressed that this paper focuses only on leadership transitions (the change of governor). Therefore it is natural to expect the transition effect to be larger the stronger the governor is. To assess this strength, we use a CB committee typology proposed by Blinder (2007) (in increasing order of governor's strength):

1. Individualistic Committee.
2. Genuinely Collegial Committee.
3. Autocratically Collegial Committee.
4. Individual Governor.

According to Blinder, 1 is characterized by all members being exhorted to vote their own mind, with the governor often on the losing side of the vote (e.g. UK); 2 is the case in which there is an atmosphere that strives for consensus and thus the governor plays a role in forging this consensus (e.g.

ECB or Bernanke); in 3 the governor plays the dominant role and can shift the board to his preferred policy (e.g. Volcker or Greenspan[5]); 4 is obviously the case with the strongest governor as he is the sole determiner of policy.

The caveat with this typology is that it is inevitably subjective. For instance, both the UK and the US have similar committee structures on the surface - one vote per member, which is released to the public - but Blinder argues that tradition gives the US governor a much greater sway over the board than the UK one. Despite this caveat, our approach was to use Blinder's opinion for the countries he did categorize; to search in CB staff papers of each country how they categorize their own central bank; and, as a last resort, we took our best guess based on the committee structure and minutes. Appendix A discusses in detail how we constructed the index.

Table 2.6: Governor Power - Blinder Index

# Meet	2	3
FM	0.074***	0.064**
	[0.002]	[0.043]
$FM \times Blin$	0.023	0.019
	[0.159]	[0.256]
LM	0.078**	0.088***
	[0.024]	[0.001]
$LM \times Blin$	0.059***	0.057***
	[0.002]	[0.000]

P-value between [ ], calculated with robust standard errors.
It includes country FE and year dummies.

Table 2.6 shows that the coefficient of $LM \times Blin$ goes in the direction we expected: the stronger is the governor (higher $Blin$), the stronger is the monetary contraction at the last meetings. On the other hand, we did not find, as predicted by signalling dynamics, that the first meeting's effect is greater with stronger governors. In this sense, Blinder's typology provides partial evidence in favor of the signaling interpretation.

The Governor was previously part of the committee

The assumption behind this exercise is that the public should have a better idea of the type of a new governor if he was already part of the monetary policy committee before he held office. Hence, we create a dummy variable $PrevBoard$ indicating whether the governor was part of the previous board. If

[5]Blinder tells an anecdote where Greenspan started on the losing vote, but managed to persuade the committee to vote unanimously in favour of his choice.

that were the case, there would be less need of signalling by the new governor. Hence we expect a smaller tightening at the first meeting (smaller incentives to prove he is a Hawk) and at the last meeting (smaller incentives for the departing to help the signalling process.)

Table 2.7: Governor belonged to Previous Board

# Meet	2	3
FM	0.098***	0.043
	[0.002]	[0.158]
$FM \times PrevBoard$	−0.065	0.053
	[0.128]	[0.474]
LM	0.063	0.065*
	[0.188]	[0.074]
$LM \times PrevBoard$	0.033	0.060
	[0.593]	[0.222]
# Trans	46 × 25	46 × 25

P-value between [ ], calculated with robust standard errors.
It includes country FE and year dummies.

Overall, Table 2.7 shows that this exercise is inconclusive. We cannot reject the hypothesis that the effect at the first and last meetings is different in both specifications. However, there is some very weak evidence tightening found at the first meetings is smaller when the new governor was part of the board. In fact, there is a close to significant effect, with p-value of 13%, regarding the two but not three meetings per transition dummy. Given the decrease of the number of transitions used to estimate the effect of *PrevBoard*, we believe that this finding provides a slight support for our interpretation that signalling dynamics might be driving our results.

Consequently, when we assessed transitions in which governors belonged to the board, we found weak evidence of the expected effect in the first meetings. With Blinder's typology, we found strong evidence regarding the last meetings but no significant effect regarding the first meetings. Analyzing the empirical evidence as a whole, the idea is that together Tables 2.6 and 2.7 show that the committee structure of monetary policy decisions suggests the existence of signaling dynamics both at first and last meetings.

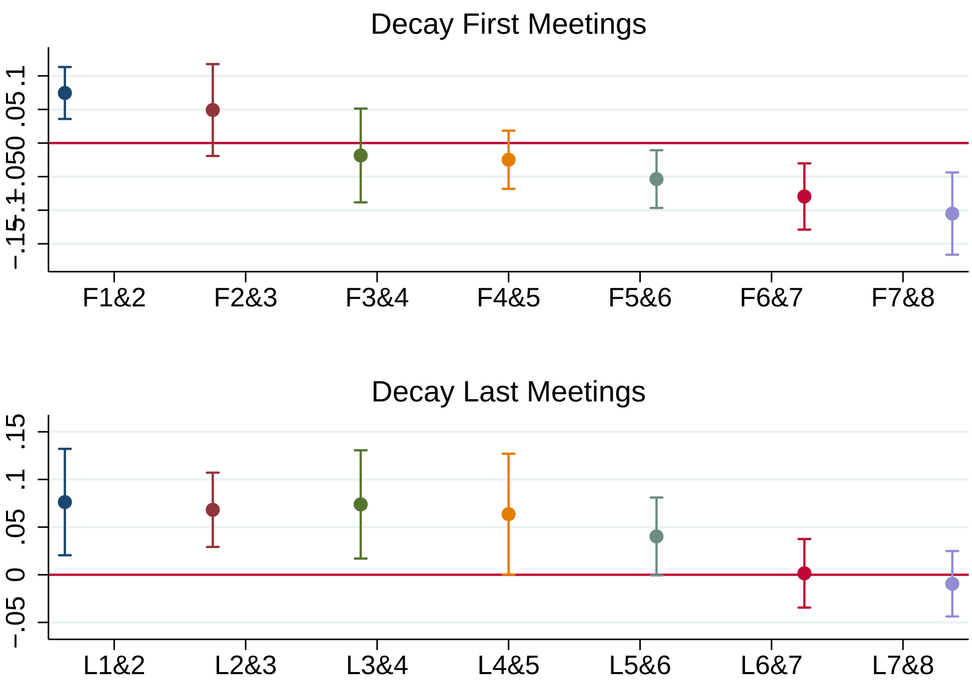
2.3.7 Decaying Effect

Earlier on, in section 2.3.1, we presented the main results for four specifications: only one meeting per transition dummy until four meetings per transition dummy. The idea was that a departing governor does not have to act precisely on the last meeting; he could tighten monetary policy a bit before and affect his successor in the same way.

This argument loses strength as meetings grow more distant from the actual governor's change - policy should return to normal. The objective of this section is to show that as the meetings distance themselves from actual change, the transition effects diminish. In particular, we drop $FM_{c,t}$ and $LM_{c,t}$ from the specification in (2-1), but add other two dummy variables across seven different specifications. The j-th specification includes one dummy variable that accounts for the j-th and $(j+1)$-th first meetings and another one that accounts for the last j-th and $(j+1)$-th meetings. For example, in the first specification, there are two dummies variables accounting for the first and last two meetings, respectively. Similarly, the second specification considers one dummy variable that accounts for the second and third meetings, as well as another dummy to account for the the penultimate and anti-penultimate meetings. The same logic applies for the subsequent specifications.

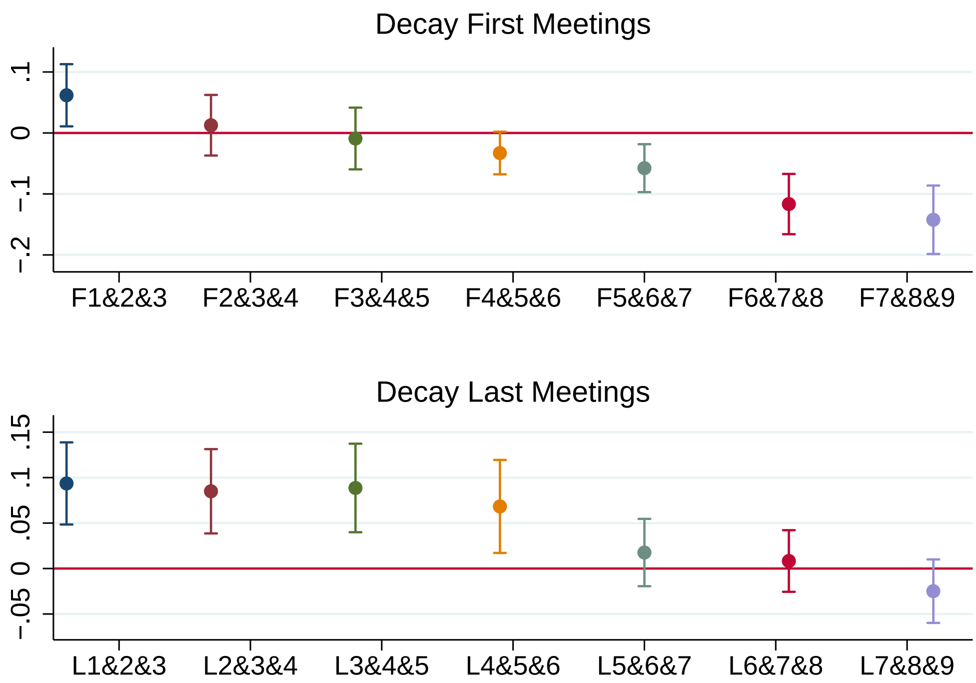
Figure 2.3 plots the coefficients for the dummy variables in these seven different specifications. The upper (bottom) part of the graph plots the value of the coefficients associated with the first (last) meetings.

Figure 2.3: Decay: Rolling Transition (2 meetings)



As it can be seen in Figure 2.3, the coefficients associated with last meetings fall as the meetings in question grow distant from the actual change until they become indistinct from zero, around the 6th meeting. Similarly, the coefficients attached to first meetings fall and become zero around the 4th meeting. In addition, they become statistically negative before returning to zero. We attribute this behaviour to the lagged term of the Taylor rule. After all, the increases above the Taylor rule from the real first meetings are incorporated to the Taylor rule due to i_{t-1}. Consequently, if the economically adequate interest rates were the originally prescribed by the Taylor rule, one would expect the residual to be negative for a while until the excess tightening absorbed by the rule dissipates. We find very similar results when we report the graph for a rolling window of 3 meetings, as shown by Figure 2.4.

Figure 2.4: Decay: Rolling Transition (3 meetings)



Hence there is indeed the decay we hoped to find around transitions, which enhances the evidence discussed so far on leadership transition effects. If there were other factors which do not spring from transitions driving our results, it would be reasonable to expect these factors to have influenced monetary policy for longer than 5 or 6 meetings before transitions. After all, these alleged confounders were strong enough to trigger transitions. In addition, in regard to first meetings effects, it is hard to come up with a confounder that initially is associated with tight monetary policy and then reverts to normal. For instance, a regime change that coincides with a transition should leave consequences for longer periods than what we report in Figures 2.3 and 2.4.

2.3.8 Political Transitions

In this section we discuss how transition in central bank leadership interact with political transitions. The concern is that if the changes at the central bank coincide with changes in government, our results could be driven by monetary policy responding to some government variable we do not account for. In a rough manner, the analysis on Transition Timing at section 2.3.2 helped somewhat in avoiding this confounder by showing that the results survive even when we focus on countries with fixed central banker's mandate. Indeed, usually the goal of such mandates is precisely to make monetary policy less susceptible to influence from the executive branch and thus the mandates's length is many times designed to not coincide with political cycles. Notwithstanding our previous efforts, we still feel that this concern warrants a thorough analysis. After all, it not difficult to imagine a scenario where even with fixed mandate, a central banker would schedule his resignation to coincide with a political cycle in order to smooth overall policy changes or even due to unobservable pressures from the government.

As a result, this section compares the effects of CB governor's changes when they coincide with political transitions with the effect of CB governor's changes during normal times. We will do so by interacting our main regressors of interest with dummies which mark election years and beginning of mandate years. Moreover, we distinguish between elections and reelections so that we create 4 dummies:

1. $ElecY_{ct}$: The meeting ct takes place in a election year when the winner is taking office for the first consecutive time.

2. $ReelecY_{ct}$: The meeting ct takes place in a election year when the incumbent wins the election.

3. $BegMandElecY_{ct}$: The meeting ct takes place in the year when a new head of government took office after a election.

4. $BegMandReelecY_{ct}$: The meeting ct takes place in the year when the incumbent head of government took office after a reelection.

Before we begin discussing the regressions, we clarify a few points regarding the data. First, the specific position of the head of government changes across countries. In presidential systems, it is naturally the president, who usually takes office in the year following the election (e.g. US, Brazil). In parliamentary systems, the head of government is the prime minister (even if the country does have a president) who is elected following a general election. In most cases, the prime minister takes office immediately after the election so that $ElecY_{ct}$ and $BegMandElecY_{ct}$ will coincide for most parliamentary systems. Second, it is also important to note that the dummies refer to calendar year and not the 12 month period before/after an election. For instance, if a presidential election in country c takes place in September 2014, all country c meetings in 2014 have $ElecY_{ct}$ or $ReelecY_{ct}$ equal 1. Finally, we note that, for convenience, reelection years refer to when a incumbent succeeded. If a president tries to get himself reelected and loses, the year in question will count as $ElecY_{ct}$ and not as $ReelecY_{ct}$.

Let $W_{ct} \in \{ElecY_{ct}, ReelecY_{ct}, BegMandElecY_{ct}, BegMandReelecY_{ct}\}$ be one of the four dummies described above. Results are reported in Table 2.8, where each column refers to a specification that considers a single political transition variable.

Notice that our empirical results are not driven by correlation to political cycles. Although we found that first election years are associated with higher interest rates, the coefficients of FM and LM in the first columns have similar magnitudes to the results of Table 2-1. One shortcoming is that the coefficient of LM in Table 2.8 is not statistically significant, though by a small margin. We note, however, that the LM is significant using 3 meetings as the transition period. In fact, this coheres with a pattern appearing throughout our empirical analysis: FM is more precisely estimated using 2 meetings as transition whereas LM is more precise with 3 meetings (see Tables 2.1, 2.3, 2.5). Therefore Tables 2.8 presents evidence consistent with our main results: first and final meetings have tighter monetary policy; they are economically significant (coefficients' sizes are similar to Table 2.1); and they cannot be explained away by appealing to political cycles involving elections or beginning of mandates.

Table 2.8: Political Cycles

Dummy W	*Elec*	*Reelec*	TookOf	Retook
# Meetings	2	2	2	2
FM	0.069***	0.068***	0.043*	0.061***
	[0.008]	[0.003]	[0.094]	[0.008]
$FM \times W$	0.033	0.072	0.119**	0.200
	[0.608]	[0.556]	[0.045]	[0.179]
LM	0.058	0.072*	0.040	0.071*
	[0.135]	[0.051]	[0.153]	[0.054]
$LM \times W$	0.087	0.037	0.136	0.075
	[0.290]	[0.700]	[0.233]	[0.375]
W	0.073***	0.010	-0.004	-0.010
	[0.005]	[0.682]	[0.828]	[0.667]
# Meetings	3	3	3	3
FM	0.057	0.050	0.044	0.051*
	[0.117]	[0.116]	[0.271]	[0.095]
$FM \times W$	0.026	0.141	0.067	0.142
	[0.708]	[0.170]	[0.272]	[0.337]
LM	0.063**	0.083***	0.044*	0.081***
	[0.029]	[0.003]	[0.056]	[0.004]
$LM \times W$	0.082	0.048	0.154**	0.081
	[0.183]	[0.562]	[0.040]	[0.289]
W	0.071***	0.006	-0.006	-0.011
	[0.001]	[0.820]	[0.769]	[0.636]
# Trans	59×12	65×6	56×15	66×5

P-value between [ ], calculated with robust standard errors.
Regressions still include FE and Year dummies.
2 First and 2 Last meetings per transition

2.3.9 Fiscal Policy

In this section we address the concern that fiscal policy could be driving our results. This could be the case if transitions were more likely to occur in times of fiscal build ups. Stories that justify this relation usually hinge on some change in government or policy, so we were already protected to some extent against this criticism by our results controlling for political transitions. However, we believe this to be a concern serious enough to warrant particular attention to fiscal developments and how they affect central bank transitions.

The variable we use to account for fiscal development is the ratio of government expenditures to GDP (henceforth GY). This is a quarterly measure which is available for most countries in our sample[6]. The matching of these quarterly data to each central meeting follows the algorithm detailed in section for quarterly unemployment rates. Then, we add $GY_{c,t}$ as an economic factor to the main equation (2-1). Importantly, we allow the coefficient on this variable to vary with countries. As we let GY to affect the interest rates, the fact that fiscal build ups make the central bank tighten monetary policy is controlled for.

Table 2.9 reports the results, which remain essentially the same as those of Table 2.1.

Table 2.9: Fiscal Policy Robustness

# Meetings	2	3
FM	0.087***	0.071**
	[0.000]	[0.022]
LM	0.092***	0.103***
	[0.009]	[0.001]
Gov FE	Y	Y
Year Dummy	Y	Y
# Obs	3781	3781

P-value between [ ], calculated with robust standard errors.

2.4 Discussion

In summary, the empirical analysis documented that transition periods in central bank leadership are robustly associated with tighter monetary policy

[6]See Appendix A for the list of countries for which we have only yearly GY.

and argued that this result is unlikely to stem from endogenous transitions, an argument built upon the timing of the transitions. Furthermore, we offered an explanation for this result: signaling dynamics and reputation transfer. The new governor tightens policy to signal he is a Hawk and the departing governor tightens policy to make it harder for a Dove to pretend he is a Hawk, thereby influencing the public's beliefs - what we called a reputation transfer. In order to test our explanation on the data, we exploited several heterogeneities between transitions to assess whether the results still go in the direction one would expect were the proposed mechanism true. Indeed, we showed that results are stronger when the central bank has less independence, is less transparent, and when the country's regulatory quality is lower. In addition, they are weaker when the new governor was previously in the committee, and are stronger when the outgoing governor had more power. Although we cannot affirm that our explanation in the only acceptable one, the fact that all of the heteregeneities go in the expected direction makes us confident that signaling dynamics and reputation transfer are an important part of the story. In the next section we propose a simple theory which rationalizes our explanation and sheds light on the mechanisms behind transfer of reputation.

3
The Model

We consider a model built on Kydland and Prescott (1977) and Barro and Gordon (1983) important contributions. Both papers study the time inconsistency of monetary policy. Before adapting their models to study monetary policy during transitions, we briefly summarize their main contribution through a basic setup.

3.1
Basic setup

Time is discrete and the horizon is finite, i.e. $t = 1, ..., T$. The relation between output y_t and inflation π_t is given by the following Phillips curve:

$$y_t = y_t^n + a(\pi_t - \pi_t^e), \tag{3-1}$$

where y_t^n is the natural level of output, π_t^e is the expected inflation, and $a > 0$ measures the output response to inflation surprises.

For each period t, taking π_t^e as given, the central bank (CB) chooses π_t in order to minimize the current loss function,

$$\frac{\pi_t^2}{2} - \lambda(y_t - y_t^n), \tag{3-2}$$

subject to the Phillips curve (3-1).

In equilibrium, rational expectations require that $\pi_t = \pi_t^e$. The classic result of inconsistency arises. In particular, the desire to stimulate output leads to positive inflation, $\pi_t = \pi_t^e = \lambda a > 0$, without output gains, i.e. $y_t = y_t^n$. In contrast, if in $t = 1$, the CB could credibly commit to $\pi_t = 0$ for $t = 1, ..., T$, then society would be better as $\pi_t = \pi_t^e = 0$ and $y_t = y_t^n$ arise in equilibrium.

3.2
Novel elements

In order to study monetary policy decisions during transitions, we add two ingredients to the basic setup.

First, inflation π_t comprises the sum of two components,

$$\pi_t = \gamma\pi_{t-1} + (1 - \gamma)\pi_t^c, \tag{3-3}$$

where $\gamma \in (0, 1)$ measures the degree of inertia in the economy and π_t^c is the inflation under control of the CB. Hence, π_{t-1} is the state variable and π_t^c is the control variable. This means that π_1 depends on an exogenous initial condition

π_0 in equilibrium. For simplicity, we assume $\pi_0 = 0$ and, thus, $\pi_1 = (1-\gamma)\pi_1^c$. This assumption does not alter the implications of the model, but it simplifies a bit the formulas.

This extension is necessary to connect the decisions of different central banks through time. Indeed, π_{t-1}^c chosen by the previous CB would affect current inflation π_t and, thus, the current CB's choice of π_t^c.

Second, the CB not only cares about current inflation π_t but also about inflation under control π_t^c. In particular, the current loss function reads

$$\frac{\theta(\pi_t^c)^2}{2} + \frac{\pi_t^2}{2} - \lambda(y_t - y_t^n), \tag{3-4}$$

where $\theta > 0$ measures the weight attributed to the controllable part of inflation. If θ is low (high), we say that the CB is dove (hawk).

This extension is necessary to generate non-trivial dynamics. Otherwise, if there is no cost to change inflation π_t^c (i.e. $\theta = 0$), then the CB could simply adjust π_t^c to set total inflation π_t at its preferred level. As a result, previous inflation π_{t-1} becomes irrelevant.

These two ingredients, inertial inflation and losses from changing π_t^c, allow us to transform the basic setup, inspired by Kydland and Prescott (1977) and Barro and Gordon (1983), into a dynamic model. Inertial inflation is an intuitive assumption, easily motivated by some degree of price stickiness and indexation. In contrast, the assumption that, apart from total inflation π_t, inflation under control also enters the loss function merits some digression.

We offer two interpretations for $\theta > 0$. The first is that it is costly to change inflation. In practice, the CB does not control inflation directly. Instead, it controls policy instruments, such as the interest rate, that affect inflation. One finds many reasons in the literature to avoid abrupt changes in the interest rate: to avoid financial stress (Cukierman, 1989); better control over long-term interest rates (Woodford, 2003); politico-economic costs associated with committee decision making (Riboni and Ruge-Murcia, 2010). If the central bank cares about any of these reasons, it will find costly to change the part of inflation under control today from its optimum level.

The second interpretation is that θ can capture career concerns. The public may consider the inherited state of the economy when judging the competence of a CB. Hence, central banks that deliver the same inflation rate, but inherit different ones, should be perceived differently. If the CB cares about how competent it is perceived to bring inflation close to zero, there is an extra cost associated with generating inflation under its control, π_t^c.

In the next sections, we model a transition between central banks by allowing θ to vary over time. In particular, we index θ by i and t, meaning that the CB i cares about the inflation under control of the CB in charge at period t. We also assume that the CB discounts the future according to $\beta \in (0,1)$. For example, suppose that the CB i took office at $t = j$. Hence, its loss function reads:

$$L^i = \sum_{t=j}^{T} \beta^{t-j} \left[\frac{\theta_{it}(\pi_t^c)^2}{2} + \frac{\pi_t^2}{2} - \lambda(y_t - y_t^n) \right]. \tag{3-5}$$

3.3
Full information benchmark

In this section, we assume that θ is publicly know. In order convey the main messages, we use a two period version of the model, i.e. $T = 2$. In each period, a different CB is in charge of monetary policy. At $t = 1$, the first central bank, say CB1, faces the following loss function (already substituting (3-1) in (3-5), and defining $\kappa \equiv a\lambda$):

$$L^1 = \frac{\theta_{11}(\pi_1^c)^2}{2} + \frac{\pi_1^2}{2} - \kappa(\pi_1 - \pi_1^e) + \beta \left(\frac{\theta_{12}(\pi_2^c)^2}{2} + \frac{\pi_2^2}{2} - \kappa(\pi_2 - \pi_2^e) \right).$$

At $t = 2$, the second central bank, say CB2, has the following loss function (again, after substituting (3-1) in (3-5)):

$$L^2 = \frac{\theta_{22}(\pi_2^c)^2}{2} + \frac{\pi_2^2}{2} - \kappa(\pi_2 - \pi_2^e).$$

We use backward induction to solve the model. First we solve CB2's problem, where π_1 is a state variable as made explicit by (3-3). Then, bearing in mind that optimal π_2^c is a function of π_1, we find π_1 in equilibrium by solving CB1's problem. Notice that CB1 has incentives to strategically influence CB2's decisions at $t = 2$ through the state variable π_1. This kind of mechanism is found elsewhere in the literature as in Alesina and Tabellini (1990) and in Debortoli and Nunes (2008).

3.3.1
CB2's problem

At $t = 2$, taking the expected inflation under control $E(\pi_2^c)$ and past inflation π_1 as given, CB2 solves:

$$\min_{\pi_2^c} \frac{\theta_{22}(\pi_2^c)^2}{2} + \frac{\pi_2^2}{2} - \kappa(\pi_2 - \pi_2^e)$$

$$\text{s.t.} \quad \pi_2 = \gamma\pi_1 + (1-\gamma)\pi_2^c \quad \text{and} \quad \pi_2^e = \gamma\pi_1 + (1-\gamma)E(\pi_2^c).$$

After inserting the restrictions into the objective function, the first order condition (FOC) with respect to π_c^2 yields:

$$\pi_2^c = \frac{\kappa(1-\gamma) - (1-\gamma)\gamma\pi_1}{\theta_{22} + (1-\gamma)^2}. \tag{3-6}$$

By plugging (3-6) at (3-3) with $t = 2$, one obtains:

$$\pi_2 = \gamma\pi_1 + (1-\gamma)\pi_2^c = \frac{\kappa(1-\gamma)^2 + \theta_{22}\gamma\pi_1}{\theta_{22} + (1-\gamma)^2}. \tag{3-7}$$

Notice that $\frac{\partial\pi_2^c}{\partial\pi_1} < 0$ and $\frac{\partial\pi 2}{\partial\pi_1} > 0$. The intuition is straightforward. An increase in π_1 raises the marginal cost of inflating the economy and, thus, entailing a lower π_2^c. However, this decrease in π_2^c is not large enough to compensate the direct increase in π_2 due to the inertial effect. Algebraically,

$$\frac{\partial\pi_2}{\partial\pi_1} = \underbrace{\gamma}_{\text{inertial effect} > 0} + \underbrace{(1-\gamma)\frac{\partial\pi_2^c}{\partial\pi_1}}_{\text{marginal cost effect} < \gamma} > 0.$$

3.3.2 CB1's problem

Before stating the CB1's problem, we emphasize some features of the timing of the model. Within each t, expected inflation π_t^e is set before the current CB chooses its control variable, i.e. π_t^c. In other words, the CB takes current inflation expectations as given. Nonetheless, π_{t+1}^e is set after π_t and, thus, CB1 knows it cannot stimulate output in the second period. In other words, in equilibrium, the CB in charge at t knows that $\pi_{t+1}^e = \pi_{t+1}$.

At $t = 1$, given π_1^e, BC1 takes into account the FOC of BC2 and solves its problem. That is,

$$\min_{\pi_1^c} \frac{((1-\gamma)^2 + \theta_{11})(\pi_1^c)^2}{2} - \kappa(\pi_1 - \pi_1^e) + \beta\left(\frac{\theta_{12}(\pi_2^c)^2}{2} + \frac{\pi_2^2}{2}\right).$$

$$\text{s.t.} \quad \pi_2 = \gamma(1-\gamma)\pi_1^c + (1-\gamma)\pi_2^c \quad \text{and} \quad \pi_2^c = \frac{\kappa(1-\gamma) - (1-\gamma)^3\gamma(\pi_1^c)^2}{\theta_{22} + (1-\gamma)^2}$$

Notice that we impose $\pi_2^e = \pi_2$ and $\pi_1 = (1-\gamma)\pi_1^c$, which follow from rational expectations in $t = 2$ and from (3-3) with $\pi_0 = 0$, respectively.

After some algebra, the FOC with respect to π_c^1 yields:

$$\pi_1 = (1-\gamma)\pi_1^c \text{ and } \pi_1^c = \frac{\kappa(1-\gamma) + \beta(1-\gamma)\left((\theta_{12} - \theta_{22})\frac{\kappa\gamma(1-\gamma)^2}{(\theta_2 + (1-\gamma)^2)^2}\right)}{\left((1-\gamma)^2 + \theta_{11} + \beta(1-\gamma)^2\left(\frac{\gamma^2\theta_{22}^2 + \gamma^2(1-\gamma)^2\theta_{12}}{(\theta_{22}+(1-\gamma)^2)^2}\right)\right)} \tag{3-8}$$

In other to develop some intuition, we rewrite (3-8) as follows:

$$((1-\gamma)^2+\theta_{11})\pi_1^c = \underbrace{\kappa(1-\gamma)}_{\text{inflationary bias}} + \beta\left(-\frac{\partial\pi 2}{\partial\pi_1^c}\underbrace{\pi_2}_{\text{marginal cost of }\pi_2} - \underbrace{\theta_{12}\pi_2^c}_{\text{marginal cost of }\pi_2^c}\frac{\partial\pi_2^c}{\partial\pi_1^c}\right) \tag{3-9}$$

Aside the inconsistency problem that generates the inflationary bias $\frac{\kappa(1-\gamma)}{(1-\gamma)^2+\theta_{11}}$, this decomposition shows the two main forces at work in the model. By choosing π_1^c, CB1 internalizes its effect on CB2, taking into account that: (1) a higher π_1^c (and consequently higher π_1) increases the marginal cost associated with π_2^c for CB2; (2) a higher π_1 increases π_2 due to inflation inertia. CB1 dislikes both π_2 and π_2^c because he knows output in $t=2$ will not be above its natural level. Consequently, there is a tradeoff: a higher π_1 decreases π_2^c but increases π_2.

3.3.3 Equilibrium analysis

Notice that equations (3-8) and (3-7) characterize the equilibrium levels of inflation in $t=1$ and $t=2$, respectively. Without inertia, that is $\gamma=0$, inflation would be $\pi_t=\pi_t^c=\frac{\kappa}{1+\theta_{tt}}$, which is the relative weight given to inflation surprises in the loss function. This is the inflationary bias that arise in the classical result discussed in Section 3.1. In other words, without inertia, CB1 cannot affect the future actions of CB2. Similarly, if CB1 does not care about the future, that is $\beta=0$, one also obtains $\pi_t=(1-\gamma)\pi_t^c$ and $\pi_1^c=\frac{\kappa(1-\gamma)}{(1-\gamma)^2+\theta_{11}}$, which is simply the inflationary bias result when $\pi_0=0$.

In the rest of the paper, we set $\theta_{11}=\theta_{12}$. That is, CB1 gives the same weight to π_1^c and, discounting aside, π_2^c. We rationalize it by assuming that θ represents the perceived costs of changing inflation under control. If this is the case, it seems reasonable that such perceived costs do not change over time. In contrast, if θ represents career concerns, then it is reasonable to set $\theta_{11}>\theta_{12}$. Indeed, this parametrization says that CB1 cares more about its career than CB2's. In the Appendix B.1.1, we discuss this case.[1] In order to simplify notation, let $\theta_1=\theta_{11}$ and $\theta_2=\theta_{22}$.

Recall that CB1 faces a trade-off when choosing π_1 as a higher π_1 decreases π_2^c but increases π_2. The following proposition tells us which force dominates. In particular, it states a necessary and sufficient condition that brings equilibrium inflation below the inflationary bias $\frac{\kappa(1-\gamma)}{(1-\gamma)^2+\theta_1}$ that arises

[1]In the Appendix B.1.1, we also consider the case in which κ vary across periods. In principle, different central banks can attach different weights to the output gap in the loss function.

in the classical result in a one period model. All proofs are relegated to the Appendix 6.

Proposition 1 *A sufficient and necessary condition for* $\pi_1 \leq \frac{\kappa(1-\gamma)}{(1-\gamma)^2+\theta_1}$ *is* $\theta_1 \leq \theta_2 C$*, where* C *is a constant greater than 1.*

Intuitively, if θ_2 is relatively small in comparison to θ_1, then CB2 does not care much about setting π_2^c too high. Hence, it is optimal for CB1 to increase π_1 above the inflationary bias to force CB2 to reduce π_2^c. Alternatively, if CB2 does care about π_2^c, then CB1 prefers to decrease π_1 in order to reduce π_2 through inflation inertia.

3.4 Incomplete information

In this section, we introduce uncertainty regarding θ_2. At the end of CB1's term (when the model starts), agents had already learnt CB1's type θ_1. However, in the beginning of CB2's term, they are uncertain about θ_2, which can take values in the set $\{\theta^H, \theta^D\}$, with $\theta^H > \theta^D$, where H and D stand for Hawk and Dove, respectively. Thus the Hawk CB finds inflation under control π_t^c more costly.

We also assume that CB1 knows CB2's type. It seems natural that, during transitions, the incumbent CB acquires information about the entrant that the general public still does not know. Given that agents know that CB1 knows CB2's type, they can extract information about θ_2. Hence, CB1 may also choose inflation to affect agents's beliefs about CB2's type, θ_2. In a similar vein, the choice of π_1 affects the trade-off faced by CB2 between revealing or not its type. To sum up, CB1's actions may help agents to uncover CB2's type.

In the context of this model, we propose a formal definition for reputation transfer that is in line with the interpretation we offered for our empirical results.

Definition 1 *Reputation Transfer: how, through the choice of* π_1^c*. the CB1 affects beliefs about CB2's type.*

In order to CB2's choice not be trivial, we increase the time length of the model to three periods. Otherwise, if $T = 2$, then in the last period, CB2 would simply choose its preferred level of inflation given π_1 and π_2^e. If $T = 3$ instead, then in the second period, a dovish CB2 faces a tradeoff. On one hand, CB2 can choose its preferred level of inflation and, thus, face higher inflation expectation in the last period. On the other hand, CB2 may pretend to be a Hawk in order to face lower expectations in the last period.

Notice that although agents know CB1's type, they are uncertain about its choice π_1^c. Indeed, different types of successors imply different levels of π_1^c. In particular, the public tries to infer CB2's type from CB1's choice of inflation under control. As a result, at $t = 1$, there may be either a separating equilibrium, in which agents discover CB2's type and the game becomes one of full information from now on, or a pooling equilibrium, in which beliefs about CB2's type are not updated and the game remains one of incomplete information.

Figure 3.1 illustrates all possible paths in this signaling game.

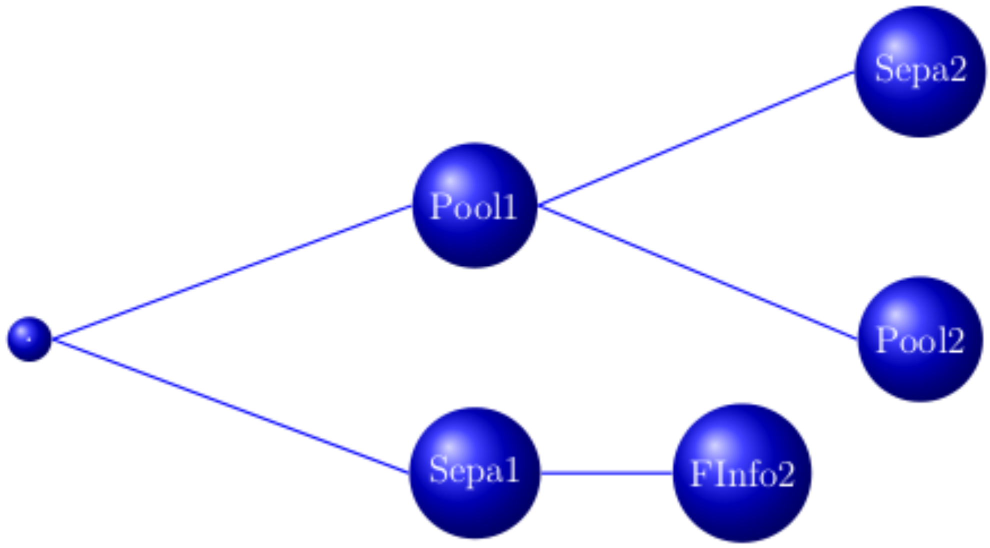


Figure 3.1: Equilibria Tree

In signalling models, different sets of beliefs can sustain multiple equilibria. This poses a challenge as CB1's problem cannot be defined if there are multiple equilibria at $t = 2$. In order to circumvent this problem, following Cukierman and Liviatan (1991) and Walsh (2000), we consider a specific set of beliefs. In particular, agents always expect a Hawk CB2 to choose its preferred action as if CB2 does not fear being mistaken for a Dove CB2. Thus the kind of equilibrium, pooling or separating, is determined by the Dove CB2 choice. If it prefers to mimic the Hawk CB2's choice of inflation, then the pooling equilibrium emerges. If, instead, it prefers to reveal its type by choosing its preferred level of inflation π_2^c, then the separating equilibrium arises.

Let π_{2S}^{cH} and π_{2S}^{cD}, respectively, be the Hawk and Dove CB2s' preferred choice of π_2^c when the equilibrium is separating. Similarly, let π_{2P}^c be chosen by the Hawk CB2 when it is expected that the Dove CB2 pool his actions. Finally, let $\mu \in (0, 1)$ be the prior probability at $t = 2$ that CB2's type is θ^H. The public has the following expectations.

- In a separating equilibrium: $E(\pi^c_{2S}) = \mu\pi^{cH}_{2S} + (1-\mu)\pi^{cD}_{2S}$
- In a Pooling Eq: $E(\pi^c_{2P}) = \pi^c_{2P}$

The equations above show that the expectations are formed from the optimization problems of each central banker, which is close to the spirit of Kydland and Prescott (1977). In the Appendix 6 we discuss how different refinement criteria may alter our results.

In the Appendix B.2.1 and B.2.2 we characterize the equilibrium levels of inflation in $t = 3$ and $t = 2$ for each of the two types of equilibrium: separating and pooling. Indeed, for either separating or pooling equilibrium, there is only one level of inflation compatible for each type of CB2 given our set of beliefs. This is because rational expectations require that not only CB2 act optimally given π^e but also that expectations are confirmed in equilibrium.

3.4.1 Equilibria Existence

Now we shall discuss the conditions which determine the existence of a particular equilibrium. Simply put, we know what an equilibrium would look like but it remains to be seen if it will in fact exist. We clarify once more that the two kinds of equilibria studied take their name from what happens at period 2. After all, as period 3 is the last, there will always be a separating equilibrium.

Our definition of equilibrium in the incomplete information framework has three requirements. First, each type of central bank minimizes its loss function taking current expectations and current beliefs as given but taking into account how its choice shall affect future beliefs and future expectations. Second, we require expectations to be rational; that is, expectations must equal the probability weighted (by μ) average of the hawkish bank's equilibrium strategy and the dovish bank's equilibrium strategy. Third, the beliefs update rule must follow Bayes' rule on the equilibrium path.

A- Separating equilibrium When considering possible deviations from equilibrium, we must make explicit how beliefs are updated off the equilibrium path. Let μ_{s3} be the belief at $t = 3$ that CB_2 is a Hawk.

$$\mu_{s3} = \begin{cases} 1, & \text{if } \pi_2 = \pi^H_{2S} \\ 0, & \text{if } \pi_2 = \pi^D_{2S} \\ 0, & \text{if } \pi_2 \neq \pi^H_{2S} \text{ or } \pi_2 \neq \pi^D_{2S} \end{cases}$$

The update rule above means that if agents see different inflation rates to equilibrium rates, they will believe that they are dealing with a dovish central bank and will update their beliefs to zero. Also, quite naturally, they will update to one after observing a hawkish equilibrium rate and to zero after observing a dovish one. This belief rule follows from the refinement criteria we adopted: agents expect $CB2^H$ to choose π_{2S}^H, which is its choice when it does not fear being mistaken for $CB2^D$. The precise formula of π_{2S}^i can be found in equation (6-6) of Appendix 6.

In order to check the threats against the existence of the separating equilibrium, one must check whether any of the CB types has incentives to deviate from its equilibrium strategy (here being its choice for inflation) given expectations and the beliefs' update rule. Consider first a hawkish central bank. If it chooses anything different from π_{2S}^{cH} it will face agents in period 3 who think he is a Dove and will thus expect higher inflation at $t = 3$. Besides worsening its welfare in period 3, it also worsens its welfare in period 2 since π_{2S}^{cH} was found by minimizing its loss function, taking expectations as given. Hence, if the Hawk deviates he harms himself in every period.

Alternately, a Dove could potentially improve its welfare by pretending to be hawkish (i.e. choose $\pi_2^c = \pi_{2S}^{cH}$) in order to face lower expectations and improve welfare in period 3. Hence the dovish bank faces a tradeoff: it can choose its preferred level of inflation at $t = 2$ given $E(\pi_{2S}^c)$, which would reveal its type, or it can pretend to be hawkish at $t = 2$ and improve its welfare at $t = 3$.

Let L_S^d be the loss of $CB2^D$ associated with being in a separating equilibrium and define L_{SD}^d as the loss associated with deviating from the prescribed equilibrium strategy and trying to disguise oneself as hawkish. For the separating equilibrium to exist, it cannot be more profitable for a dovish central bank to pretend to be hawkish and choose the hawkish equilibrium strategy given beliefs and expectations. Hence, it is required that:

$$L_S^d \leq L_{SD}^d$$

Proposition 2 *Holding π_1 and the parameters constant, for γ small enough, there exists β^S such that $L_S^d \leq L_{SD}^d$ $\forall$ $0 \leq \beta \leq \beta_S$ and that $L_S^d > L_{SD}^d$ $\forall$ $\beta > \beta^S$.*

The intuition of this proposition is straightforward. For β small, $CB2^D$ does not care much about period 3 and will choose its preferred inflation level, engendering a separating equilibrium. Alternatively, for β large, $CB2^D$ cares a lot about period 3 and therefore the benefits of facing lower inflation expectations in $t = 3$ are worth the costs of pretending to be dovish in $t = 2$ - it deviates from the separating equilibrium strategy.

B- Pooling equilibrium First, we explicit again the belief updating rule for pooling equilibrium. Let μ_{p3} be the belief at $t = 3$ if there was a pooling equilibrium at $t = 2$. Recall that π_2^c chosen by $CB2^H$ in a pooling equilibrium is π_{2P}^c.[2]

$$\mu_{p3} = \begin{cases} \mu, & \text{if } \pi_2 = \pi_{2P} \\ 0, & \text{if } \pi_2 \neq \pi_{2P} \end{cases}$$

If agents observe anything other than the optimum inflation for a Hawk which expects to be imitated, they will revise their beliefs so to be sure they face a Dove. With this update rule, it is straightforward that $CB2^H$ has no incentives to deviate. It would incur a cost at $t = 2$ and would face higher expectations at $t = 3$, which would make it worse off in both periods. On the other hand, a Dove may want to deviate from the pooling equilibrium because mimicking $CB2^H$'s actions might be too costly, as π_{2P} is not Dove's preferred inflation at $t = 2$. There is hence a tradeoff between revealing its type or not.

Let L_P^d be the loss from following the prescribed equilibrium and define L_{PD}^d as the loss from deviating from pooling equilibrium. To guarantee that the pooling equilibrium exists, it cannot be more profitable for $CB2^D$ to reveal his type than to mimic $CB2^H$. Hence, it is required that:

$$L_P^d \leq L_{PD}^d$$

Proposition 3 *Holding π_1 and the parameters constant, for γ small enough, there exists β_P such that $L_P^d \leq L_{PD}^d \ \forall \ \beta > \beta^P$ and that $L_P^d > L_{PD}^d \ \forall \ 0 \leq \beta \leq \beta^P$.*

The intuition behind this proposition is quite straightforward. A Dove mimics a Hawk's strategy at $t = 2$ in order to face lower expectations at $t = 3$. Consequently, if the weight assigned to $t = 3$ is too low, it will never play the pooling strategy and the equilibrium collapses. Alternatively, if the weight assigned to $t = 3$ is large enough, any extra loss borne at $t = 2$ will be acceptable because of the welfare gain at $t = 3$.

In summary, whether a separating or pooling equilibrium prevails depends mainly on the discounting factor β: for β low, it is not worth to mask oneself as Hawk in order to improve future expectations - separating equilibrium prevails; for β large, it is worth to sacrifice one's favorite choice in

[2] The formula of π_{2P}^c is given by equation (6-8) in Appendix 6.

$t = 2$ for more favorable expectations in $t = 3$ - pooling equilibrium prevails. For β where neither a separating nor a pooling equilibrium can be sustained, there will be a mixed strategies equilibrium where CB2^D randomizes between pooling and separating. Provided that the conditions of Propositions 2 and 3 are satisfied, the dependence on β can be depicted on the line below.

C - Boundary Cases To develop intuition, we analyze two extreme cases: $\mu = 1$ and $\mu = 0$.

Proposition 4 *If $\mu = 1$ then $L^d_S \leq L^d_{DS} \iff L^d_P \geq L^d_{PD}$.*

If agents are fully certain that they face a Hawk, the inflation chosen by this type of central bank in a pooling and in a separating equilibrium is, unsurprisingly, the same. Consequently, the Dove has the same payoff from pooling (separating) as it has from deviating from the separating (pooling) equilibrium. It is worth noticing that we do allow beliefs to be revised even with a prior equal 1, one could think of this as $1 - \epsilon$, as detailed in the belief updating rule of the previous section. After all, if agents observed inflation levels deemed impossible by their currents beliefs then it would make little sense not to update them. Hence, even when a Dove is enjoying the lowest inflation expectations possible, it still faces the trade-off of whether or not to reveal its type in each kind of equilibrium.

Proposition 5 *If $\mu = 0$ then $L^d_P > L^d_{PD}$, i.e., there is no pooling equilibrium.*

The intuition is simple: even if the Dove pools the Hawk's choice, the agents will still assign zero probability of CB2 being a Hawk. Regarding the separating equilibrium, it may exist or not depending on the other parameters even with $\mu = 0$ as long as we allow μ to be updated - one could consider $\mu = \epsilon$ arbitrarily small. In particular, Proposition 2 is still valid. As before, the Dove can try to pass itself off as a Hawk in order to gain at $t = 3$ at the expense of its welfare at $t = 2$. Thus there is a β_S that determines whether a separating equilibrium will exist. If $\beta > \beta_S$, there will not be a separating equilibrium. In this case, since a pooling equilibrium cannot exist with $\mu = 0$, we would be in a parameter region where there is no equilibrium in pure strategies.

3.4.2 Reputation Transfer

In the previous section, we showed how the kind of prevailing equilibrium depends on the discount factor β. However, it will also hinge on the inherited inflation π_1, which CB2 treats as exogenous. In this section we show how a monetary policy contraction ($\pi_1 \downarrow$) makes it easier to sustain a separating equilibrium and also makes it harder to sustain a pooling separating. In other words, that, by tightening policy, CB1 helps the public to discover that its successor is a Hawk: a reputation transfer takes place. This is captured by the Proposition 6, which is the main theoretical result of the paper. Define $\Delta^S(\pi_1) \equiv L^D_{DS} - L^D_S$ and $\Delta^P(\pi_1) \equiv L^D_{DP} - L^D_P$ and let $\bar{\pi}_1$ be an upper bound which is a function of the parameters of the model.

Proposition 6 *For $\pi_1 < \bar{\pi}_1$, $\Delta^S(\pi_1)$ decreases in π_1 and $\Delta^P(\pi_1)$ increases in π_1.*

Proposition 6 tells us that there are two effects when π_1 falls. First, there is an increase in the loss difference from deviating from the separating equilibrium, thereby making this deviation less attractive for CB2^D. Second, there is a reduction in the relative loss from deviating from a pooling equilibrium, thereby making this deviation more attractive. In other words, $\pi_1 \downarrow$ makes separating more attractive and pooling less attractive.

The intuition for Proposition 6 can be found at Figures 3.2 and 3.3. The preferred point for a Hawk (blue point) and a Dove (red point) are when the marginal cost of π_2^c equals its marginal benefit. As the Hawk's marginal cost of inflation is greater, we have that $\pi_{2H}^c < \pi_{2D}^c$. The area of the shaded grey triangle represents the cost of a Dove central banker for trying to pass himself as a Hawk - the area where his marginal benefit is greater than his marginal cost. When π_1 falls, there is an increase in the marginal benefit for both types (or, equivalently, a reduction in the marginal cost of π_2^c). As it can be seen in Figure 3.3, this increase the cost of a Dove passing himself as Hawk, which is represented by the larger dark grey triangle. This is the main mechanism that makes it easier to sustain a separating equilibrium and harder to sustain a pooling one.

The two graphs above provide intuition for the mechanisms in the model driving Proposition 6, but it is also worthwhile to map this intuition into central bank leadership transitions in real life. The context is of one departing central banker who knows the type of his successor. Both types of successors have incentives to act like a Hawk and thus if the departing central banker wants to help the true Hawk to separate himself he should raise interest rates.

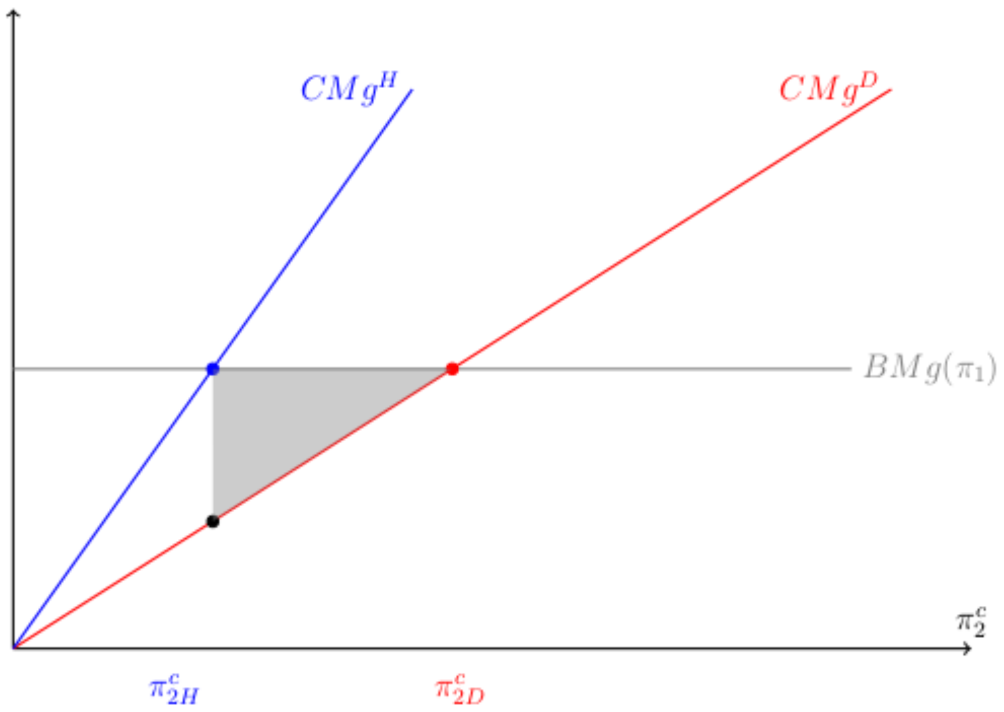


Figure 3.2: Intuition for Proposition 6 - Part I

After all, only a true Hawk would raise interest rates again on top of the recent increase. A Dove would find too costly to pretend to be Hawk if the monetary policy is already much too tight for his preferences.

This is the most important theoretical result because it rationalizes the results found during the empirical analysis. There we found that there were monetary contractions both at the first meetings of a new governor and at the final meetings of a departing governor. These empirical results are perfectly compatible with signalling dynamics in a simple model which added inter-period strategic interactions to a classical model of the theoretical literature on central banks' choices. Having discussed the intuition behind Proposition 6, we now detail what exactly we meant by saying it is harder or easier to sustain certain equilibria. We do so by focusing first on pure strategies equilibria and then discussing mixed strategies.

Pure Strategies Equilibria

The point here is that changes in π_1 alters the parameter space which sustains a type of pure strategies equilibrium. The tighter is the monetary policy, the more patient the central bankers can be without destroying the separating equilibrium. Conversely, they must have an even higher discount factor if a pooling equilibrium is to be sustained (See Figure 3.4). Naturally,

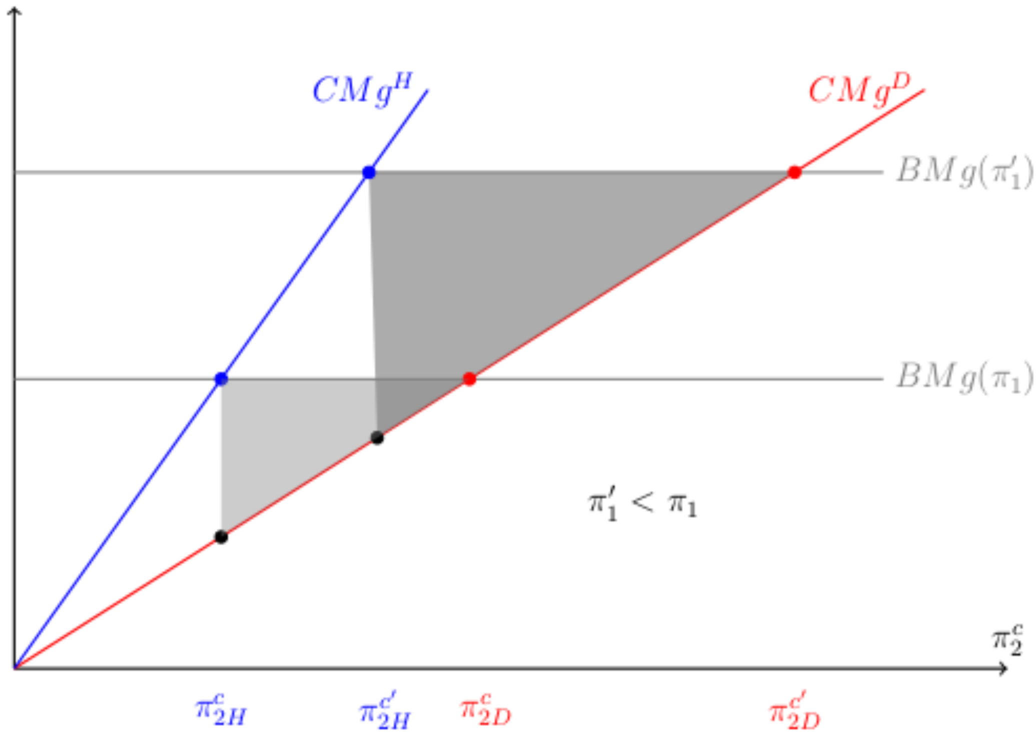


Figure 3.3: Intuition for Proposition 6 - Part II

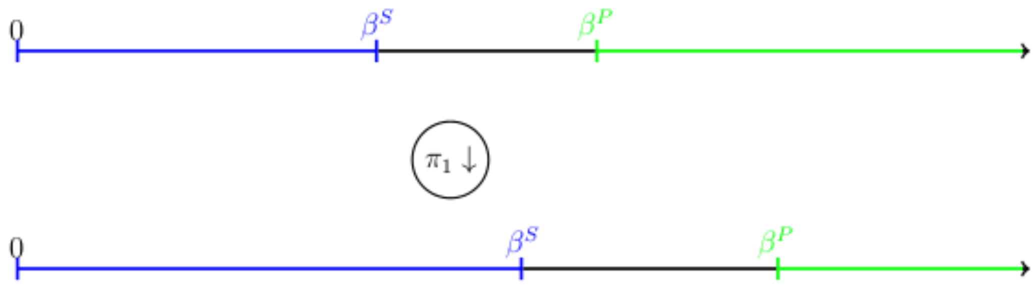


Figure 3.4: Parameter Space Shift (i)

π_1 shall also alter the space that sustains each of the two equilibria for other parameters of the model (e.g. μ, γ etc). We focus, however, on β because it has the most intuitive effect on the type equilibria: patience fosters the pooling equilibrium and undermines separating equilibrium.

Figure 3.5 plots the βs required for each type of pure equilibrium as a function of π_1. The blue (lower) line refers to the separating equilibrium whereas the green (upper) line refer to the pooling equilibrium. This makes easy to visualize what we mean by reputation transfer in pure strategies. The reduction of π_1 makes harder for a Dove to pass himself as Hawk and therefore, increases the parameter space where there is a separating equilibrium at the same time that reduces the space of pooling equilibrium. Consequently,

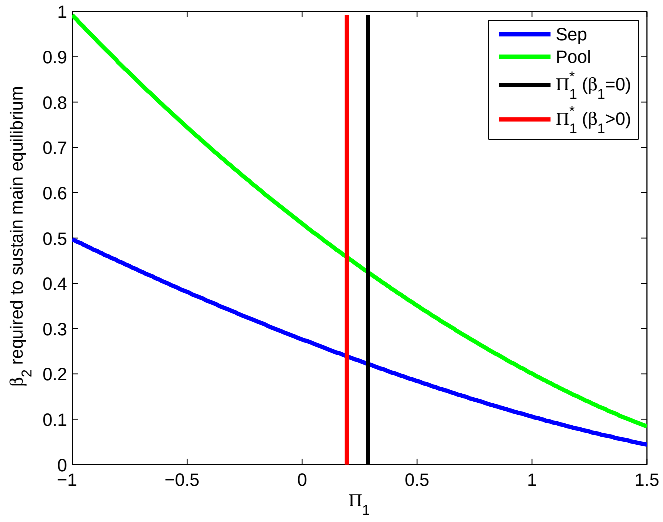


Figure 3.5: Parameter Space Shift (ii)

a reputation transfer can take place in the sense that CB1 can affect the type of equilibrium that CB2 will play. Naturally, the parameters are given when CB1 makes its decision, so that for extrema βs, it would take a too extreme movement of π_1 to change the type of equilibrium. Nevertheless, for values of β close to β^S or β^P, CB1 might, with very small movements of π_1, be able to ensure that a equilibrium is sustained or, alternatively, make it impossible for said equilibrium to be sustained. When we consider mixed strategies next, it will be easier to see how changes in π_1 affects the beliefs the public holds about CB2. Finally, we note that we relegate to the Appendix B.2.3 the characterization of CB1's problem and focus on CB2's choice to show that movements in π_1 can affect public's beliefs and help to determine the type of equilibrium to be played.

Mixed Strategies Equilibrium

As mentioned briefly before, for certain intermediate values of β, it is possible that neither a separating nor a pooling equilibrium are sustained. This means that $CB2^D$ is patient enough to pretend to be a Hawk if the equilibrium is separating, i.e. the agents expect different inflation for each type, but $CB2^D$ is not patient enough to pool the $CB2^H$'s choice in a pooling equilibrium. This is possible because there is greater payoff in deviating from a separating equilibrium than in subscribing to a pooling equilibrium. After all, in the first case the agents will think that $CB2^D$ is a Hawk with probability one and consequently will expect lower inflation at period three than in the second case where agents will keep their beliefs that the CB2 is a Hawk with probability $\mu < 1$.

In a mixed strategies equilibrium, the Dove randomizes between playing his separating strategy with probability α and pooling on the Hawk's choice with probability $1 - \alpha$. The value α will be what makes CB2^D indifferent between pooling on the CB2^H's choice, which depends on α, and playing his separating choice π_{2S}^{cD}. It is the indifference that makes optimal for CB2^D to randomize. If one course of action was strictly better, α would equal one or zero and we would be back to the pure strategies equilibria cases. As before, agents expect CB2^H to choose its preferred inflation level as if he did not fear being mistaken for a Dove. Given such expectations, CB2^H will indeed choose its preferred inflation π_{2M}^c. The formula for π_{2M}^c is given in equation (6-10) in Appendix 6. Below we explicit how mixed strategies equilibrium can be seen as a generalization of the pure strategies case since the former contains the latter:

$$\alpha = 1 \Rightarrow \pi_{2M}^c = \pi_{2S}^{cH} \qquad \text{(Pure Separating Eq.)}$$

$$\alpha = 0 \Rightarrow \pi_{2M}^c = \pi_{2P}^c \qquad \text{(Pure Pooling Eq.)}$$

Figure 3.6 plots α as a function of π_1 and shows that α decreases with π_1. The intuition is the same of Proposition 6: a decrease in π_1 makes it more costly for π_{2S}^{cD} to pretend it is a Hawk so that the probability of revealing its type, which is given by α, increases. In addition, for π_{2S}^{cD} to be indifferent between both actions, it is required that the pooling action becomes more attractive at $t = 3$ to compensate the increase in costs at $t = 2$. An higher α ensures this by increasing μ_{Post} since higher μ_{Post} implies lower inflation expectations at $t = 3$.

μ_{Post} is the posterior, after the belief updating between $t = 2$ and $t = 3$, that occurs in a mixed strategy equilibrium. Beliefs are updated according to:

$$\mu_{Post} = \frac{\mu}{\mu + (1-\mu)(1-\alpha)}$$

As μ_{Post} increases with α, a monetary policy contraction also raises μ_{Post} - CB1 affects the belief updating process. The idea is that as the probability of π_{2S}^{cD} choosing the separating action increased, if the agents see the pooling action π_{2M}^c they think it is more likely that they are facing a π_{2S}^{cH}. This is precisely what we had defined as reputation transfer and what we had suggested as an explanation for the empirical results. By tightening policy, the departing central banker makes the agent think it is more likely that the new governor is a Hawk.

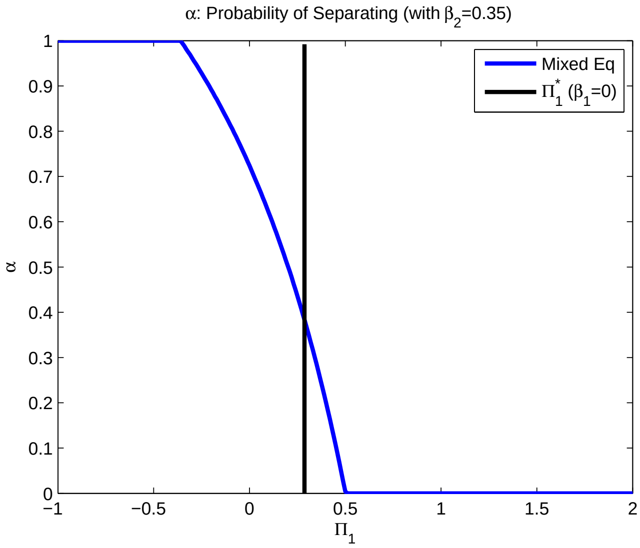


Figure 3.6: Reputation transfer: α as a function of π_1

3.5 Discussion

We have seen that there are two channels through which CB1 can shape agents' beliefs about CB2. First, it can directly affect the probability (μ) assigned at the beginning of $t = 2$ that CB2 is a Hawk. This happens when there is a separating equilibrium in $t = 1$. Second, CB1 can affect the existence conditions of separating and pooling equilibria and, as a result, CB1 is able to influence how the beliefs update process will unfold. Indeed, we showed how, by tightening policy, CB1 helps a hawkish successor to be seen as such.

We believe that these two channels capture the idea of reputation transfer that we are interested in. The first central bank can affect the public's inferences about its successor's type through the policy decisions taking place at the final period of its tenure. Moreover, the model rationalizes the interpretation we offered for the behaviour found in the data: there is a monetary contraction in the final meetings of a departing governor and also in the first meetings of a new central banker.

A shortcoming of our approach is that the in the cases in which CB2^D separates and reveal its type, it will set a higher level of inflation, which implies lower interest rates. However, in the date we observe that the average effect is an increase of interest rates. This does not mean that we need most central banks to be Hawks nor most equilibria to be pooling. The reason behind this is that in our model, if the Dove does not choose the exact same level of inflation chosen by the Hawk, the public will immediately discover its type. In reality the process takes longer as the economic agents update their beliefs over time until discovering one CB's type. It is be possible to modify our approach slightly in order to allow for this more gradual belief updating. One way to do so is to create noise in π_1 so that agents can never be sure which type they are facing even in a separating equilibrium. For instance:

$$\pi_1 = \pi_1^c(1-\gamma) + u, \text{ where } u \text{ i.i.d } N(0, \sigma^2)$$

Thus agents still update using Bayes' rule, but it is a gradual process where there is not absolute certainty. This means that even if a Dove CB does not profit from copying a Hawk CB, it will still have incentives to choose a lower inflation than would otherwise - now there is an incentive to approach slightly CB^H's choice so that the public has more favorable posterior beliefs. Before there was no benefit from a small change: or CB^2 went all the way until CB^H's choice or it would still be considered a Dove. Thus noisy inflation is able to reconcile the fact that many Doves will prefer a separating equilibrium with the fact that the data shows monetary contractions - separating does not fully reveal one's preferences and thus there are incentives for a small policy tightening.

4
Conclusion

In this paper we argued that although the economic literature gives great attention to the central bank's inconsistency problem and reputation building, there is too little work on transition periods between central bankers. The goal of this paper was to study this problem empirically within a country panel and then to propose a theoretical model which provides an interpretation for the empirical results.

The main result of the paper is that there is a monetary contraction on the final meetings of a departing governor and on the first meetings of a new governor that is not warranted by economic factors. This empirical result is robust to several specifications and we also provided evidence supporting the mechanism we proposed to explain the results.

We argued that transition periods are unlike others due to the inherent uncertainty associated with leadership changes. The uncertainty about a new central banker's preferences creates signalling incentives for the incoming governor. Moreover, a departing central banker who cares about economic outcomes after his tenure has incentives to try to mitigate this uncertainty - he could tighten policy and thereby help his successor to signal his type, what we called a transfer of reputation.

After presenting a mechanism consistent with the empircal results, we created a model which rationalized such mechanism. The model added a dynamic effect and uncertainty about a new central bank's preferences to a classical framework used to study inconsistency problems (Barro and Gordon (1983) and Kydland and Prescott (1977)). Within such incomplete information structure, we defined reputation transfer as the ability of the departing central bank to shape agents' beliefs about his successor.

We showed that there are two channels through which beliefs are shaped. The first is to directly affect the probability the agents assign to the second central bank being hawkish - the prior. The second channel is to affect how the agents interpret the actions of the central bank - the process of belief revision by helping to determine which kind of equilibrium takes place in the second period. Importantly, we showed that a departing central banker will help this belief formation process by contracting monetary policy at the end of his tenure, which is precisely the empirical result found before.

To conclude, our contribution lies in being the first to document the behaviour of a departing governor and the first to document signalling by a new governor using data from multiple countries. Furthermore, we provided a model which rationalizes the results found in the data and which can also be used to study further questions on reputation dynamics between transitioning central bankers.

5
Bibliography

ALESINA, A.; TABELLINI, G. A Positive Theory of Fiscal Deficits and Government Debt. *Review of Economic Studies*, v. 57, n. 3, p. 403–14, July 1990.

BARRO, R. J.; GORDON, D. B. A positive theory of monetary policy in a natural rate model. *Journal of Political Economy*, v. 91, n. 4, p. 589–610, August 1983.

BERTRAND, M.; DUFLO, E.; MULLAINATHAN, S. How Much Should We Trust Differences-in-Differences Estimates? *The Quarterly Journal of Economics*, v. 119, n. 1, p. 249–275, February 2004.

BLINDER, A. S. Monetary policy by committee: Why and how? *European Journal of Political Economy*, v. 23, n. 1, p. 106–123, March 2007.

COIBION, O.; GORODNICHENKO, Y. Why Are Target Interest Rate Changes So Persistent? *American Economic Journal: Macroeconomics*, v. 4, n. 4, p. 126–62, October 2012.

CROWE, C.; MEADE, E. E. The Evolution of Central Bank Governance around the World. *Journal of Economic Perspectives*, v. 21, n. 4, p. 69–90, Fall 2007.

CUKIERMAN, A. Why does the Fed smooth interest rates? *Proceedings*, p. 111–157, 1989.

CUKIERMAN, A.; LIVIATAN, N. Optimal accommodation by strong policymakers under incomplete information. *Journal of Monetary Economics*, v. 27, n. 1, p. 99–127, February 1991.

DEBORTOLI, D.; NUNES, R. Political disagreement, lack of commitment and the level of debt. International Finance Discussion Papers 938, Board of Governors of the Federal Reserve System (U.S.), june 2008.

DEBORTOLI, D.; NUNES, R. Monetary regime switches and unstable objectives. International Finance Discussion Papers 1036, Board of Governors of the Federal Reserve System (U.S.), june 2011.

DINCER, N. N.; EICHENGREEN, B. Central Bank Transparency and Independence: Updates and New Measures. *International Journal of Central Banking*, v. 10, n. 1, p. 189–259, March 2014.

DRISCOLL, J. C.; KRAAY, A. C. Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data. *The Review of Economics and Statistics*, v. 80, n. 4, p. 549–560, November 1998.

FUDENBERG, D.; MASKIN, E. The Folk Theorem in Repeated Games with Discounting or with Incomplete Information. *Econometrica*, v. 54, n. 3, p. 533–54, May 1986.

HANSEN, S.; MCMAHON, M. First Impressions Matter: Signalling as a Source of Policy Dynamics. CEPR Discussion Papers 9607, C.E.P.R. Discussion Papers, Aug. 2013.

KYDLAND, F. E.; PRESCOTT, E. C. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, v. 85, n. 3, p. 473–91, June 1977.

LJUNGQVIST, A.; SMOLYANSKY, M. To Cut or Not to Cut? On the Impact of Corporate Taxes on Employment and Income. NBER Working Papers 20753, National Bureau of Economic Research, Inc, Dec. 2014.

RIBONI, A.; RUGE-MURCIA, F. J. Monetary Policy by Committee: Consensus, Chairman Dominance, or Simple Majority? *The Quarterly Journal of Economics*, v. 125, n. 1, p. 363–416, February 2010.

ROGOFF, K. The Optimal Degree of Commitment to an Intermediate Monetary Target. *The Quarterly Journal of Economics*, v. 100, n. 4, p. 1169–89, November 1985.

ROMER, C. D.; ROMER, D. H. Choosing the Federal Reserve Chair: Lessons from History. *Journal of Economic Perspectives*, v. 18, n. 1, p. 129–162, Winter 2004.

SPENCE, M. Job Market Signaling. *The Quarterly Journal of Economics*, v. 87, n. 3, p. 355–74, August 1973.

WALSH, C. E. Market Discipline and Monetary Policy. *Oxford Economic Papers*, v. 52, n. 2, p. 249–71, April 2000.

WOODFORD, M. Optimal Interest-Rate Smoothing. *Review of Economic Studies*, v. 70, n. 4, p. 861–886, October 2003.

6
Appendix

Appendix A: Data and Empirical Robustness

A.1: Data

Transition Coding

In this section we detail how we coded different types of transitions and which were types used in the paper. Due to exposition concerns, we report in Table 6.1 the main types of transitions. Before we discuss our choices, consider one example which illustrates some of the challenges in our coding decisions. Assume that one governor's term expires but the body of government responsible for nominating the governor has not yet announced its decision. In these cases there will be an acting governor, who may be later appointed to be remain as the official governor. When should we label the first meeting? As soon as he becomes acting governor (D or E) or only after he is officially appointed (C)? Similarly, when would the last meeting be? These are the choices we had to make. In this section we explain what we opted to do and why.

Table 6.1: Coding of different transitions

Transition	AF	WBO	WAB	AP
A	Y	-	N	N
B	Y	-	N	Y
C	Y	-	Y	-
D	N	Y	-	Y
E	N	Y	-	N
F	N	N	-	-

AF: Assumed Office | WBO: Will Become Official
WAB: Was Acting Before | AP: Acting Predecessor

The types of transitions we consider for the variable FM (First Meeting) are A, B and C. The types of transitions we consider for the variable LM (Last Meeting) are A, B, E and F. Therefore we opted to consider the first meeting as the first one after the new governor took office. This option in the cases where the new central banker was acting governor stems from our view that an acting governor might not have political capital to change policy much so to print his own mark - there is a stand-by until the leadership appointment is settled. This is consistent with our choices of last meeting - it makes little sense to consider the meeting before the governor was appointed as the final meeting if the same person was the acting governor before. After all, we need different people for a reputation transfer to take place.

The way we construct the transition variables implies that it is possible that part of our effect comes from a departing governor who has not yet officially discovered the identity of his successor. This seems counterintuitive since our model assumes that the departing governor (CB1 in the model) knows the type of his successor (CB2). There are two arguments which justify our choice. First, in the cases where the acting governor will be appointed, it is possible that the departing banker already knows this by the time of his departure - he has more information about the government's choice than the public. Second, similar theoretical predictions arise if we assume that a departing governor wants to help the public to find the new banker's type. After all, the result that a policy tightening makes it harder for a Dove to pool appears before we consider the optimal choice of CB1.

Countries

Table 6.2 reports which countries we have in the sample; the total number of meetings per country; the first and last year each countries appear in the sample; and the number of different governors (excluding changes during the financial crisis) we have per country.

Fixed Regimes and Unannounced Resignation

We determine whether a country has a fixed regime for central bankers by checking whether an appointment also specifies how long a governor will remain in charge. The countries which do not have fixed mandates for central bankers (or did not at some point of the sample) are Brazil, Colombia, Peru and Thailand. There are of course some caveats. For instance, Tunisia has in name a fixed mandate for central bankers but in practise none of them stay until the end of their term and the government seems to have authority to change the central bank's leader at will. These caveats do not undermine our results because we also analyse the heterogeneity of unannounced resignations. Places where a fixed mandate does not happen in practice will have many unannounced resignations. In contrast, resignations in countries such as Norway will be announced with antecedence, many times to match the calendar year. This antecedence address the problem of endogenous timing - if the mandate was fulfilled or the resignation was announced in advanced, it is unlikely that the transitions are happening due to tighter monetary policy. We consider the resignation to be announced when there are more than 2 months of antecedence. Any choice of months will inevitably

Table 6.2: List of Countries

	Country	# Meetings	First Year	Last Year	#Governors**
1	Albania	103	2001	2014	2
2	Australia	270	1990	2014	3
3	Brazil	156	1999	2014	3
4	Chile	169	2000	2013	4
5	Colombia	242	1995	2014	2
6	Czech Rep	172	1998	2014	4
7	ECB	214	1999	2014	3
8	Georgia	61	2008	2014	1
9	Ghana	60	2002	2014	2
10	Guatemala	82	2005	2014	3
11	Hungary	141	2002	2014	3
12	India	60	2005	2014	2
13	Indonesia	110	2005	2014	3
14	Israel	231	1995	2014	4
15	Japan*	162	1998	2013	2
16	Kenya	48	2006	2014	2
17	Mexico	94	2005	2014	2
18	New Zealand	121	1999	2014	3
19	Nigeria	60	2003	2014	3
20	Norway	128	1999	2013	2
21	Pakistan	40	2005	2014	5
22	Peru	162	2001	2014	5
23	Philippines	124	2002	2014	2
24	Poland	182	1999	2014	4
25	Serbia	124	2007	2014	3
26	South Africa	78	2001	2014	1
27	South Korea	183	1999	2014	5
28	Sweden	174	1994	2014	3
29	Switzerland	62	2000	2014	4
30	Thailand	112	2001	2014	4
31	Tunisia	175	2000	2014	5
32	Turkey	115	2005	2014	3
33	United Kingdom	202	1997	2014	3
34	United States	300	1984	2013	3
35	Uruguay	25	2007	2013	1

* Between March 2001 and February 2006, Japan's monetary target was money growth. We drop these meetings from the sample
** This is the number of governors ignoring changes during 2007-2008.

be arbitrary, but the results change little when we consider 1 or 3 months as the cutoff for announced resignations.

Governor's Strength Index

The goal of this section is to make as clear as possible how we extended the typology proposed by Blinder (2007) to the set of countries present in our sample. Before we begin, we are the first to point out that this is a tentative extension which used more heuristics than ideal.

The procedure used in the paper was: first we checked whether Blinder himself had classified some of the countries; second we searched for papers (usually from central bank staff) where the authors applied Blinder's typology to their own country; third, lacking the previous options, we assessed the committee structure and its minutes and made our best guess regarding which of the 4 types is the best fit for the country in question. We assign number from 1 to 4 according to:

1. Individualistic Committee.

2. Genuinely Collegial Committee.

3. Autocratically Collegial Committee.

4. Individual Governor.

As there is no clear cut classification in some countries, we allow the index to vary in 0.5 increments to reflect such uncertainty. In addition, we allow different governors within a country to be classified differently, though we only do that for a couple of countries where there are strong reasons to do so: United States (following Blinder), Israel and South Korea.

Blinder (2007) classified 9 countries of our sample. In order of governor's strength: New Zealand, Canada, Australia, US, Japan, Switzerland, Euro Area, Sweden and UK. He also admits that his classification is a subjective one.

> "I have ranked the same nine banks on their degree of "democracy" in making monetary policy decisions - ranging from the individual governor in New Zealand to the Bank of England's highly-democratic Monetary Policy Committee. This ranking is admittedly subjective, but I checked it with several colleagues and made some modifications of my original views - an ersatz Delphi method."

In Table 6.3 we report the classification for each country following the proceeding outlined above. In the cases the classification derived from a paper/staff report, we also document the webpage of the paper in question. In the cases Blinder and the reports were silent, our best guess was based on the committee structure discussed in a central bank webpage (e.g. decomposition of nominal votes seems to indicate less governor's strength).

Fiscal Policy

For most countries we were able to assemble quarterly data on the ratio of government expenditures to GDP. The countries for which we found only yearly data on GY are: Bangladesh, Ghana, Kenya, Nigeria, Pakistan and Tunisia. In addition, at times there was not data available for the whole time series. This resulted in the loss of 100 observations, from 3881 in Table 2.1 to 3781 in Table 2.9.

Table 6.3: Governor's Strength Index per Country

	Country	Blinder Index	Webpage
1	Albania	2.5	
2	Australia	3	
3	Brazil	2.5	
4	Chile	2	
5	Colombia	2	
6	Czech Rep	1.5	
7	ECB	2	
8	Georgia	3.5	https://www.nbg.gov.ge/index.php?m=553
9	Ghana	3	
10	Guatemala	2.5	
11	Hungary	1	
12	India	3	http://rbi.org.in/scripts/BS_SpeechesView.aspx?Id=395
13	Indonesia	2	http://www.bis.org/publ/work262.pdf
14	Israel*	3.5 /2	
15	Japan	2.5	
16	Kenya	2	
17	Mexico	2.5	
18	New Zealand	4	
19	Nigeria	2.5	http://www.bis.org/events/fmda07.pdf
20	Norway	3	http://www.bis.org/publ/work274.pdf
21	Pakistan	2.5	
22	Peru	2.5	
23	Philippines	2	http://www.bsp.gov.ph/downloads/EcoNews/EN12-05.pdf
24	Poland	1.5	http://www.suerf.org/download/collmay11/ppt_/1sirchenko.pdf
25	Serbia	2.5	
26	South Africa	3	http://www.scielo.br/pdf/rep/v31n4/06.pdf
27	South Korea**	3/1	http://www.kmfa.or.kr/paper/econo/2008/12.pdf
28	Sweden	1	
29	Switzerland	2.2	
30	Thailand	2	http://www.bis.org/publ/work262.pdf
31	Tunisia	3	
32	Turkey	2	
33	United Kingdom	1	
34	United States***	3/2	
35	Uruguay	2	

* Israel changed from 3.5 to 2 in 2013 following a big change in how the committee was organized.
**South Korea's classification is 3 until 2002 and 1 starting in 2013 as explained in the paper cited in the webpage column.
*** US's classification is 3 for the Volcker and Greenspan period and 2 for the Bernanke period.

A.2: Standard Errors

Driscoll-Kraay Errors

For our empirical analysis, we reported the usual robust standard errors in every table. However, the features of the data are such that there could be reason to worry about serial correlation in the error term or spatial dependence, which is a problem that can arises in country panel as the cross-sectional unit is nonrandom and are likely to be subject to common disturbances. There are different ways of addressing these issues within a panel: the most common approach in the microeconometric literature is to control for clustering within the cross section variable (countries in our case) whereas a more popular approach when dealing with countries specifically is to use Driscoll and Kraay (1998) errors. In this section we will discuss the reasons we favoured Driscoll-Kraay error in this robustness exercise. Then we report in Table 6.4 the results analagous to Table 2.1 but with Driscoll-Kraay instead.

Table 6.4: Driscoll-Kraay errors: $i_{c,t}$ is the dependent variable

# Meetings	2	3
$FirstMeet$ (β_F)	0.075**	0.061**
	[0.014]	[0.041]
$LastMeet$ (β_L)	0.076*	0.088**
	[0.083]	[0.014]
Country FE	Y	Y
Year Dummy	Y	Y
# Obs	3881	3881

P-value between [ ], calculated with Driscoll-Kraay errors.

The advantage of Driscoll-Kraay errors is that they are robust to “very general forms of spatial and temporal dependence as the time dimension becomes large”. In other words, its asymptotic properties rely on large T holding N fixed, which is precisely our case. Our data comprises a small number of countries but long time periods for a given country. This also explains why clustering does not match our needs. Clustered erros are consistent as the number of clusters goes to infinite, which is hardly our case. Moreover, one needs a even greater number of cluster when they are from different cases - as in our case where some countries span decades while others span only a few years.

In Table 6.4 we report the main results with Driscoll-Kraay errors. As one can see, all the results are still significant. Moreover, while there were some increase in p-values, they were fairly small ones despite the large set of errors dependence this type of standard error is consistent for. Consequently, we are confident that the significance of our results does not stem from inconsistent standar errors and that our benchmark is justified.

Pseudo Transitions - Monte Carlo

Another way to validate the use of robust standard errors as benchmark is to do a Placebo transition exercise to check whether the size of our test is as it should be. This approaches derives from Ljungqvist and Smolyansky (2014). More precisely: we do a Monte Carlo exercise with 10000 series of

Placebo transitions and check how often we reject the null hypothesis. Apart from transitions, we use our real data to check whether some aspect of the time series (heteroskedascity, time dependence) are serious enough so that our robust standard errors are unsuitable for our purposes. If our inference is valid, we should expect 5% rejection frequency to be close to the threshold we choose (e.g. 5%).

Table 6.5: Pseudo Transitions: Rejection frequency

# Meetings	α	2	2	3	3
		HW	DK	HW	DK
	1%	1.2%	0.6%	3.7%	0.9%
Plac FM(β_F)	5%	8.5%	5.2%	13.2%	6.2%
	10 %	15.9 %	11.3%	21%	12.8%
	1%	1.8%	1%	5%	1%
Plac LM (β_L)	5%	10.8%	6.4%	15.5%	7.1 %
	10%	17.9 %	13.3%	25.3%	15.7%
Country FE	Y				
Year Dummy	Y				
# Obs	3881				

P-value between [ ], calculated with robust standard errors.

Table 6.5 shows that the benchmark robust standard errors (Huber-White or HW) worsen as we increase the number of meetings. The reason behind this is that with more meetings counting as transition, our regressors becomes more persistent over time - there are several “1s” in a row. This problem is similar in spirit with the one documented by Bertrand et al. (2004) on Differences-in-Differences estimation, though in a much milder form. It is also worth noting that the MC exercise will overestimate the problem since we programmed that the last meeting is always immediately before the first while in the real data it is possible for a brief interregnum to occur.

Having documented a possible issue in our inference, we also check whether DK errors mitigate this concern. Although the size is not still perfect, Table 6.5 shows that DK errors mitigate the problem considerably. As there seems to be little worry about inference once DK errors are used, section 6 reassure us that the significance of our main results do not stem from inadequate inference.

Alternative Taylor Rule

In the body of the paper, the Taylor rule of each country had the lag of the interest rate as it is standard in the literature. This lag capture the fact that interest rate decisions are very persistent due to many[1] factors which leads the central bank to smooth its policy changes as discussed in section 3.2. However, Coibion and Gorodnichenko (2012) have pointed out the empirical need to include 2 lags in the Taylor rule. Although this not the standard approach, we report here our main results when the Taylor rule used to generate the dependent variable had 2 lags of interest rates. In other words, this is analagous to Table 2.1 with a different Taylor rule.

Table 6.6: Taylor rule containing two lags of the interest rate

# Meetings	2	3
$FirstMeet$ (β_F)	0.060**	0.048
	[0.019]	[0.133]
$LastMeet$ (β_L)	0.056*	0.060**
	[0.063]	[0.014]
Country FE	Y	Y
Year Dummy	Y	Y
# Obs	3881	3881

P-value between [ ], calculated with robust standard errors.

Table 6.6 shows that the results are weakened but they survive. It is important to notice that this weakening is expected. Although there are intrinsic reasons which make monetary policy be smooth over time, much of the persistence in interest rates comes from the persistence in the macroeconomic variables which the central bank responds to. Thus the greater is the number of lags in a Taylor rule, the less informative will be the residuals as one will be removing part of the variation one should be interested in. Taken to limit, if ones adds too many lags, the Taylor rule will come closer to an AR(P) and the residuals will be white noise, which obviously cannot be used as a dependent variable in any kind of analysis whatsoever.

[1]See Cukierman (1989), Woodford (2003) and Riboni and Ruge-Murcia (2010).

Appendix B: Model

B.1 Full Information

B.1.1 θ as a career concern cost

In this case the values taken by θ_{12} can vary according to the relationship between central bankers as follows:

(i) **When θ_{12} equals zero.** In this case, where CB1 does not care about the reputation cost of CB2, the desire to influence the future unequivocally reduces the inflation bias of CB_1. How much the bias is reduced increases with: the discount rate; the partial derivative (it tells us how much CB1 can influence CB2); the size of π_2.

The reason this effect works in only one direction is that CB1 knows it cannot affect the inflation surprise at $t = 2$ since agents will incorporate π_1 in their expectations. Furthermore, it loses utility from $\pi 2$ when it is higher than zero and hence it reduces current inflation to distort the choice of CB2.

The intuition is that, without uncertainty, if CB1 does not care about CB2's perceived success as a central banker, CB1 has incentives only to reduce current inflation. Consequently, in this stylized case, one would expect a contraction in monetary policy in the periods preceding a transition.

(ii) **When θ_{12} is greater than zero.** In this case, CB1 cares about the reputation cost of CB2. This creates a new force in the model that pushes inflation up. In fact, under quite weak conditions - detailed below - the higher θ_{12} is, the higher CB1's equilibrium inflation will be. In fact, if θ_{12} is high enough, equilibrium inflation will be higher than in the classical result without multiple periods ($\pi_1^c > \frac{\kappa_1(1-\gamma)}{(1-\gamma)^2+\theta_1}$).

(iii) **When θ_{12} is smaller than zero.** Analytically, this will be a more extreme [2] version of (i) since now there is an extra force pushing inflation down. An intuition behind this unlikely set of values could be a case when the first central bank would like to cause the second central bank's failure

[2] θ_{12} must have a lower bound to keep the problem defined. We detail the lower bound at the end of the Proof of Preposition 7 in the Appendix.

- they could potentially belong to opposition parties. Although the model is flexible enough to allow this, one could argue that this 'schadenfreude' would hardly have a meaningful effect in reality.

Proposition 7 *A sufficient condition for* $\frac{\partial \pi_1}{\partial \theta_{12}} > 0$ *is* $\kappa_2 > \gamma\kappa_1$.

Given that this model aims to study reputation transfers during a transition in leadership, our main interest is in the case where the first central bank is hawkish and the second can be either hawkish or dovish (which is indexed by $\kappa^H < \kappa^D$). In this two types framework, a hawksih CB_1 implies $\kappa_1 \leq \kappa_2$, satisfying the condition in the hypothesis.

B.1.2 $\kappa_1 \neq \kappa_2$

Proposition 8 *Let* $z = \frac{\kappa_1}{\kappa_2}$. *A sufficient and necessary condition for* $\pi_1^c \geq \frac{\kappa_1(1-\gamma)}{(1-\gamma)^2+\theta_1}$ *is:*

$$\theta_1 \geq \frac{1}{2}\left(\theta_2 + (\gamma z - 1)(1-\gamma)^2 + \sqrt{(1-\gamma)^4(\gamma z - 1)^2 + 2\theta_2(1-\gamma)^2(1+\gamma z) + \theta_2^2(1+4\gamma z)}\right)$$

Corollary 8.1: *If* $\kappa_1 = \kappa_2$ *then the sufficient and necessary condition of Proposition 2 becomes* $\theta_1 \geq \theta_2 C$ *where* C *is a constant greater than 1.*

Corollary 8.2: *If* $z \to 0$ *then the condition becomes* $\theta_1 \geq \theta_2$.

Proposition 8 and its corollaries tell us when equilibrium inflation at $t = 1$ is higher than the no-dynamics-case. Since our main case of interest is a hawkish CB1 ($\theta_1 = \theta^H > \theta^D$), a hawkish CB2 implies $\kappa_1 = \kappa_2 = \kappa$ and $\theta_1 = \theta_2 = \theta^H$ resulting in $\pi_1^c < \frac{\kappa_1(1-\gamma)}{(1-\gamma)^2+\theta_1}$ by Corollary 2.1. On the other hand, with a dovish CB2, it may be the case that concern about the second period overwhelms the weight given to the first period. Indeed, by Corollary 8.2, if κ_2 becomes arbitrarily large or κ_1 becomes arbitrarily small, the condition for $\pi_1^c > \frac{\kappa_1(1-\gamma)}{(1-\gamma)^2+\theta_1}$ becomes $\theta^H > \theta^D$, which is satisfied by definition.

B.2 Incomplete Information

To solve the model under uncertainty, we proceed in a backwards induction manner. First, in section B.2.1, we characterize equilibrium inflation at $t = 3$ for both types of CB2 under two cases, depending on whether $t = 2$ was a pooling equilibrium or a separating equilibrium. Second, in section 6, we characterize the inflation levels at $t = 2$ for each type of bank in a separating and in a pooling equilibrium. Finally, we characterize CB1's inflation decision at $t = 1$ by taking into account how it will affect both the tradeoff faced by CB2 and agents' beliefs about CB2.

Given the importance of expectations in the model, we detail again how they are formed. Agents start $t = 2$ with a prior μ that CB2 is Hawkish before a CB chooses the level of inflation. After observing the economic outcomes, they update the prior belief using Bayes' rule. We note that as agents observe all economic variables, there are just two kinds of update for pure strategies equilibria. After a pooling equilibrium, the probability of CB2 being hawkish remains the same and after a separating equilibrium the beliefs degenerate: it equals 1 after observing the action that a hawkish CB would undertake in a separating equilibrium and it equals 0 after observing a dovish separating equilibrium action.

B.2.1 CB2's problem at $t = 3$

A - There was a separating equilibrium at $t = 2$. This is the simplest case. As the agents already know the type of the central bank they face, we have a problem analogous to the full information benchmark in period 2. Central Banker $i \in \{H, D\}$ solves the following problem taking π_2^i and $E(\pi_3^{ci})$ as given:

$$\min_{\pi_3^{ci}} L^i = \frac{(\pi_3^i)^2}{2} + \frac{\theta_i(\pi_3^{ci})^2}{2} - \kappa(\pi_3^i - \pi_3^{ei})$$

$$s.t. \quad \pi_3^i = \gamma\pi_2^i + (1-\gamma)\pi_3^{ci} \quad and \quad \pi_3^{ei} = \gamma\pi_2^i + (1-\gamma)E(\pi_3^{ci})$$

By doing the same kind of computations as in section 3.3.1 and imposing $E(\pi_3^{ci}) = \pi_3^{ci}$ after taking the F.O.C , we have:

$$\pi_{3S}^{ci} = \frac{\kappa(1-\gamma) - (1-\gamma)\gamma\pi_{2S}^i}{\theta^i + (1-\gamma)^2} \tag{6-1}$$

$$\pi_{3S}^i = \gamma\pi_{2S}^i + (1-\gamma)\pi_{3S}^{ci} = \frac{\kappa(1-\gamma)^2 + \theta^i\gamma\pi_{2S}^i}{\theta^i + (1-\gamma)^2} \tag{6-2}$$

B - There was a pooling equilibrium at $t = 2$. In this case, both types chose the same level of inflation and therefore agents' beliefs were not updated. Thus $E(\pi^c_{3P}) = \mu\pi^{ch}_{3P} + (1-\mu)\pi^{cd}_{3P}$. Note, however, that the F.O.C. do not depend on the expectations. Hence π^{ci}_{3P} will have the same functional form as π^{ci}_{3S} but with a different inherited π_2:

$$\pi^{ci}_{3P} = \frac{\kappa(1-\gamma) - (1-\gamma)\gamma\pi_{2P}}{\theta^i + (1-\gamma)^2} \tag{6-3}$$

Since $\theta^H > \theta^D$, it is easy to see the not surprising result that $\pi^{cd}_{3P} > \pi^{ch}_{3P}$. Facing the same expected and previous inflation rates, a dovish central bank (CB2^D) chooses a higher inflation. As a consequence of $E(\pi^c_{3P})$ being a weighted average of dovish and hawkish inflation, we know that CB2^D manages to stimulate output above its natural level while CB2^H ends up causing output below the natural level due to its higher inflation aversion.

B.2.2 CB2's problem at $t = 2$

A- Separating equilibrium In this case, CB2^i solves the following problem:

$$\min_{\pi^{ci}_{2S}} L^i_S = \sum_{t=2}^{3} \beta^{(t-2)} \left(\frac{(\pi^i_{tS})^2}{2} + \frac{\theta_i(\pi^{ci}_{tS})^2}{2} - \kappa\left(\pi^i_{tS} - E(\pi_{tS})\right) \right)$$

$$s.t.\quad \pi^i_{3S} = \text{(6-2)} \quad and \quad \pi^{ci}_{3S} = \text{(6-1)} \quad and \quad \pi^i_{2S} = \gamma\pi_1 + (1-\gamma)\pi^{ci}_{2S}$$

As usual, both types of CB take expectation as given, due to the timing of the model. Since equilibrium requires $E(\pi^c_{2S}) = \mu\pi^{ch}_{2S} + (1-\mu)\pi^{cd}_{2S}$, the F.O.C.s are:

$$(1-\gamma)\gamma\pi_1 + \left((1-\gamma)^2 + \theta^i\right)\pi^{ci}_{2S} + \beta(1-\gamma)\left(\pi^i_{3S}\frac{\partial\pi^i_{3S}}{\partial\pi^i_{2S}} + \theta^i\pi^{ci}_{3S}\frac{\partial\pi^{ci}_{3S}}{\partial\pi^i_{2S}}\right) = $$
$$= (1-\gamma)\kappa \text{ for } i \in \{h, d\} \tag{6-4}$$

One can note that there is no term referring to the usual desire to stimulate output at $t = 3$. This is because the central bank understands that due to B.2.1.A, a separating equilibrium at $t = 2$ means that nothing it can do at $t = 2$ can affect output in 3. Timing is important for this argument. The usual inflationary bias arises because expectations are set before the choice of inflation even though the output is not affected in equilibrium. However, in this case, the expectations about $t = 3$ are set after CB_2's decision at $t = 2$.

Therefore any action undertaken at $t = 2$ will be incorporated when agents form expectations at $t = 3$.

Define:

$$\left(\pi^i_{3S}\frac{\partial\pi^i_{3S}}{\partial\pi^{ci}_{2S}} + \theta^i\pi^{ci}_{3S}\frac{\partial\pi^{ci}_{3S}}{\partial\pi^{ci}_{2S}}\right) = \overbrace{\frac{\theta^i\gamma^2}{\theta^i + (1-\gamma)^2}}^{\equiv\varphi_i}\left(\pi^{ci}_{2S}(1-\gamma) + \gamma\pi_1\right)$$

Then:

$$\pi^{ci}_{2S} = \frac{(1-\gamma)\{\kappa - \gamma\pi_1(1+\beta\varphi_i)\}}{(1-\gamma)^2 + \theta^i + \beta(1-\gamma)^2\varphi_i} \tag{6-5}$$

$$\pi^i_{2S} = \gamma\pi_1 + (1-\gamma)\pi^{ci}_{2S} = \frac{(1-\gamma)^2\kappa + \gamma\pi_1\theta^i}{(1-\gamma)^2 + \theta^i + \beta(1-\gamma)^2\varphi_i} \tag{6-6}$$

B - Pooling equilibrium: By definition, a pooling equilibrium is one where both types choose the same action - π^c_{2P} - and thus that is what the agents will expect. Given $E(\pi^c_{2P})$, both kinds of central bank have preferred levels of inflation that differ from one another. However if they did act on these preference their types would be revealed, undermining the pooling equilibrium. Consequently, at least one of the central banks must choose a level of inflation different from its preferred level. Naturally, the only one that has incentives to do so is the dovish central bank so it can face a better tradeoff at $t = 3$.

Thus in order to find the pooling equilibrium, we focus on the problem faced by the hawkish CB, whose behavior will be mimicked by the dovish CB. When solving the hawkish bank's problem, we take into account that π^{ch}_{2P} will be a state variable at $t = 3$ for the subcase detailed in B.2.1.B. Analytically:

$$\min_{\pi^{ch}_{2P}} L^h_P = \sum_{t=2}^{3}\beta^{(t-2)}\left(\frac{(\pi^h_{tP})^2}{2} + \frac{\theta_h(\pi^{ch}_{tP})^2}{2} - \kappa(\pi^h_{tP} - E(\pi_{tP}))\right)$$

$$s.t.\quad \pi^h_{3P} = \gamma\pi^h_{2P} + (1-\gamma)\pi^{ch}_{3P} \quad and \quad \pi^{ch}_{3P} = \text{(6-3)} \quad and \quad \pi^h_{2P} = \gamma\pi_1 + (1-\gamma)\pi^{ch}_{2P}$$

This entails the following first order condition:

$$(1-\gamma)\gamma\pi_1 + \left((1-\gamma)^2 + \theta^H\right)\pi^c_{2P} + \beta(1-\gamma)\overbrace{\left(\pi^i_{3P}\frac{\partial\pi^h_{3P}}{\partial\pi^h_{2P}} + \theta^H\pi^{ch}_{3P}\frac{\partial\pi^{ch}_{3P}}{\partial\pi^h_{2P}}\right)}^{\varphi_H\left(\pi^c_{2P}(1-\gamma)+\gamma\pi_1\right)} =$$

$$= \kappa(1-\gamma)\left(1 + \beta(1-\mu)\frac{\partial\left(\pi^{cH}_{3P} - \pi^{cD}_{3P}\right)}{\partial\pi^c_{2P}}\right) \tag{6-7}$$

The condition above is quite similar to (6-4) but now CB2^H knows that there will be an inflation surprise in $t = 3$, which CB2^H can affect through his choice π^c_{2P}. The inflation surprise stems from the fact that the belief μ will not be updated in a Pooling Equilibrium. As a result, note that if $\mu = 1$, there would be no surprise and π^c_{2P} would equal π^{cH}_{2S}.

$$\pi^c_{2P} = \frac{(1-\gamma)\kappa\left(1 + \frac{\beta(1-\mu)\gamma(1-\gamma)^2(\theta^H-\theta^D)}{(\theta^H+(1-\gamma)^2)(\theta^D+(1-\gamma)^2)}\right) - \gamma(1-\gamma)\pi_1\left(1+\beta\varphi_H\right)}{\theta^H + (1-\gamma)^2\left(1+\beta\varphi_H\right)} \tag{6-8}$$

In a pooling equilibrium, therefore, both central banks shall choose inflation equal to π^c_{2P} and consequently the agents will not update their beliefs about the probability of the second central bank being hawkish. Also, as their expectations are proven correct, output remains at its natural level.

B.2.3 First Central Bank Decision

Now we characterize the monetary policy decision of the first central banker, that is hawkish, in the incomplete information framework. We assume agents already know the type of the first central bank because it has already established a hawkish reputation throughout its tenure. We also assume that, unlike the economic agents, the first central bank knows the type of its successor.

We analyse the incentives of a central banker that will be succeed by someone like itself and those of one who knows his successor is dovish and then we proceed to focus on the first case since it better captures the idea of reputation transfer. In other words: how a central bank with an established reputation can affect agents' beliefs about its successor in order to help this second central bank.

As discussed, there are two potential ways the first central bank can influence agents beliefs': the first is affecting the μ, the probability of being hawkish assigned at $t = 2$ to the second central bank. The second way is, under certain parameter conditions, to influence which equilibrium (separating vs pooling) the second central bank will be in.

We recall the point made when discussing the Equilibria tree (Figure 3.1): even though agents are sure about CB_1's type, the uncertainty about CB2's type is carried to $t = 1$ in the sense that central banks with different successors will act in different ways. Therefore $E(\pi^c_1) = \mu_0\pi^{cH}_1 + (1-\mu_0)\pi^{cD}_1$, where π^{ci}_1 denotes the inflation at $t = 1$ chosen by a hawkish central bank that will be succeeded by a bank with type $i \in \{H, D\}$ and μ_0 is the prior at $t = 1$ that CB2 is hawkish (μ_0 can potentially differ from the prior μ at $t = 2$). With uncertainty regarding CB1's actions, one can have a separating and pooling equilibrium at $t = 1$.

Separating Equilibrium

In the separating equilibrium case CB1 affects agents' beliefs via μ. By choosing different π_1^cs[3] CB1s with different successors reveal CB2's type. As a result, the model collapses into the full information benchmark from period 2 onwards. The problem to be solved is:

$$\min_{\pi_1^i} L_S^i = \sum_{t=1}^{3} \beta^{(t-2)} \left(\frac{(\pi_{tS}^i)^2}{2} + \frac{\theta_h(\pi_{tS}^{ci})^2}{2} - \kappa(\pi_{tS}^i - E(\pi_{tS})) \right)$$

$$s.t. \quad ;\ \pi_{3S}^i = \text{(6-2)};\ \pi_{3S}^{ci} = \text{(6-1)};\ \pi_{2S}^{ci} = \text{(6-5)}$$

As usual, both types of CB take expectation as given. The separating equilibrium is characterized by :

$$\left((1-\gamma)^2 + \theta^H + \beta \left(\frac{\gamma^2 \theta^i (1+\beta\varphi_i)}{\theta^i + (1-\gamma)^2(1+\beta\varphi_i)} \right) \right) (1-\gamma)\pi_{1S}^{ci} + \\ + \beta(\theta^H - \theta^i) \left[\pi_2^{ci} \frac{\partial \pi_{2S}^{ci}}{\partial \pi_{1S}^i} + \beta \pi_3^{ci} \frac{\partial \pi_3^{ci}}{\partial \pi_{2S}^i} \frac{\partial \pi_{2S}^i}{\partial \pi_{1S}^i} \right] = \kappa(1-\gamma) \quad \text{(6-9)}$$

Proposition 9 *A central bank whose successor is hawkish chooses lower inflation than one with a dovish successor in a separating equilibrium:* $\pi_{1S}^h \leq \pi_{1S}^d$.

The intuition is that a dovish successor cares less about π_t^c than $CB2^D$ and as $\frac{\partial \pi_t^c}{\partial \pi_1} < 0$ for $t > 1$, CB1 increases π_1 to reduce π_t^c in comparison with the central bank whose successor is also hawkish.

Pooling Equilibrium

In order to characterize a pooling equilibrium, we focus on the behaviour of the central banker whose successor is hawkish, given that the other type will just mimic it in a pooling equilibrium. Also, there will be no belief updates between $t = 1$ and $t = 2$ so that $\mu_0 = \mu$. However, even in the absence of updates, recall that the first central banker may be able to affect future beliefs

[3] As $\pi_1^i = (1-\gamma)\pi_1^{ci}$, choosing π_1^{ci} determines π_1^{ci} in a straightforward fashion.

about CB2 by helping to determine the type of equilibria that will occur at $t = 2$ through his choice of π_1. After all, CB1 can reduce or increase the cost of CB2^D trying to disguise itself as hawkish - this was the idea of Proposition 6.

The reason why this happens is that the agents do not know CB2's type. They do, though, understand CB2^D's incentives. As a result, they would not expect, say, a separating equilibrium if the dovish was going to pretend it was hawkish. Therefore, by affecting $L^d_S - L^d_{DS}$ or $L^d_P - L^d_{DP}$, CB1^H is affecting the equilibrium CB2^H will find itself into.

Naturally, whether CB1^H will in fact affect which equilibrium will be played at $t = 2$ depends on the other parameters in the model. For instance, if β is too close to zero, trying to support a pooling equilibrium would require CB1^H to detach so much of its preferred level of inflation that it will never be worthy it. Figure 3.6 helps to illustrate this point. For β very small, $\alpha = 1$ and small variations of π_1 will not change the fact that there will be a separating equilibrium in $t = 2$. Likewise, for β very large, $\alpha = 0$ and small reductions of π_1 will not alter the fact that a pooling equilibrium prevails. The interesting case is the middle: where π_1 can affect the probability of Dove choosing a inflation that reveals his type.

For a general $\alpha \in [0, 1]$, CB_2^H will choose:

$$\pi^c_{2M} = \frac{(1-\gamma)\kappa\left(1 + \frac{\beta(1-\alpha)(1-\mu_{Post})\gamma(1-\gamma)^2(\theta^H-\theta^D)}{(\theta^H+(1-\gamma)^2)(\theta^D+(1-\gamma)^2)}\right) - \gamma(1-\gamma)\pi_1\left(1+\beta\varphi_H\right)}{\theta^H + (1-\gamma)^2\left(1+\beta\varphi_H\right)} \tag{6-10}$$

where

$$\mu_{Post} = \frac{\mu}{\mu + (1-\mu)(1-\alpha)}$$

It is easy to see that:

$$\alpha = 1 \Rightarrow \pi^c_{2M} = \pi^{cH}_{2S} \qquad \text{(Pure Separating Eq.)}$$

$$\alpha = 0 \Rightarrow \pi^c_{2M} = \pi^c_{2P} \qquad \text{(Pure Pooling Eq.)}$$

Note also that:

$$E(\pi^c_{2M}) = \mu\pi^c_{2M} + (1-\mu)\left(\alpha\pi^{cD}_{2S} + (1-\alpha)\pi^c_{2M}\right)$$

$$\pi^{ci}_{3M} = \frac{\kappa(1-\gamma) - (1-\gamma)\gamma\pi_{2M}}{\theta^i + (1-\gamma)^2} \tag{6-11}$$

$$E(\pi^c_{3M}) = \mu_{Post}\pi^{cH}_{3M} + (1-\mu_{Post})\pi^{cD}_{3M}$$

CB_1^H understands that α and μ_{Post} are functions of π_1. Thus π_{1P} is given by the following F.O.C.:

$$\begin{aligned}&\left((1-\gamma)^2+\theta^H\right)\pi^c_{1P}+\\&+(1-\gamma)\beta_1\left(\pi_{2M}\frac{\partial\pi_{2M}}{\partial\pi_{1P}}+\theta^H\pi^c_{2M}\frac{\partial\pi^c_{2M}}{\partial\pi_{1P}}+\beta_1\left(\theta^H\pi^c_{3M}\frac{\partial\pi^c_{3M}}{\partial\pi_{1P}}+\pi_{3M}\frac{\partial\pi_{3M}}{\partial\pi_{1P}}\right)\right)=\\&=\kappa(1-\gamma)\left(1+\beta_1(1-\gamma)\left((1-\mu)\frac{\partial\left[\left(\pi^c_{2M}-\pi^{cD}_{2S}\right)\alpha\right]}{\partial\pi_{1P}}+\beta_1\frac{\partial\left[(1-\mu_{Post})\left(\pi^{cH}_{3M}-\pi^{cD}_{3M}\right)\right]}{\partial\pi_{1P}}\right)\right)\end{aligned} \tag{6-12}$$

As a result, when choosing π_1, CB_1^H considers his effect on future inflations π^c_{2M} and π^c_{3M} and on the inflation surprises. We stress that this is more than the direct effect - it includes the channel through α and μ_{Post}. Indeed, if other parameters of the model such as β and μ are not too extreme, CB1^H will affect the belief update process. If CB1's own discount β_1 is great enough, CB1^H will have incentives to contract monetary policy to increase the probability of the agents discovering his successor type in $t=2$ - the reputation transfer whose mechanism was discussed after Proposition 6.

B.3 Multiple Equilibria

In this section we discuss the space of sustainable equilibria in the model. By doing so, we shall recall the reasons behind our refinement criteria and analyse how different criteria would alter the model results. While some implications will obviously change, we wish to assess the robustness of what lies at the heart of reputation transfers: a reduction in π_1 makes it harder to sustain pooling equilibria and easier to sustain separating ones.

First, let us map the separating equilibria. In every separating equilibrium, the dovish central banker will choose his favorite choice π_{2S}^{cD}. After all, there are no advantages of choosing otherwise if his type will be revealed. Consequently, the different separating equilibria can be mapped into the Hawk's choice π_{2S}^{cH}. Different values for π_{2S}^{cH} can be sustained if agents' beliefs are such that if the Hawk does not play a given level π_2^* which agents expect, the Hawk will be seen as a Dove in $t = 3$. As long as π_2^* entails a smaller loss than playing π_{2S}^{cH} and be treated as a in $t = 3$, π_2^* can be sustained as a separating equilibrium.

How different equilibria can be sustained helps to shed light on the refinement criterion used in the paper. If one must choose what the agents will expect a Hawk to do, it is natural and intuitive to assume that a Hawk will choose his favorite option, without worrying about being mistake for a Dove. This is the logic of the criterion suggested by Cukierman and Liviatan (1991) and Walsh (2000) which we espouse in this paper.

In Figure 6.1 we map the separating equilibria space, between the blue lines, for a given parametrization[4] and different values of π_1. The upper blue line plots, as a function of π_1, the maximum level of inflation a Hawk can choose without making the Dove deviate and pretend he is a Hawk. The bottom blue lines plots the minimum level of inflation a Hawk is willing to choose in order not to be seen as a Dove - anything bellow will lead him to prefer playing his favorite level of inflation and be seen as a Dove in $t = 3$. As we can see from Figure 6.1, the blue lines decrease in π_1. As π_1 falls, the maximum level of π_2 that can be sustained without the Dove deviating increases. Hence, loosely speaking, monetary policy contraction makes it easier to sustain a separating equilibrium, which is at the heart of our reputation transfer results. This idea is also reflected on the fact that reduction in π_1 increases the separating equilibria space, i.e., the space between the blue lines. Naturally this kind of analysis only makes sense when one sees π_1 as a parameter. From the point of view of

[4]The specific parametrization is unimportant for the ideas conveyed here. Matlab codes are available under request.

$t = 1$, the model must specify a unique equilibrium (i.e. take into account the refinement) so that CB1 can choose optimally. Nevertheless, Figure 6.1 fulfills the goal of showing that the spirit of the reputation transfer result is present throughout.

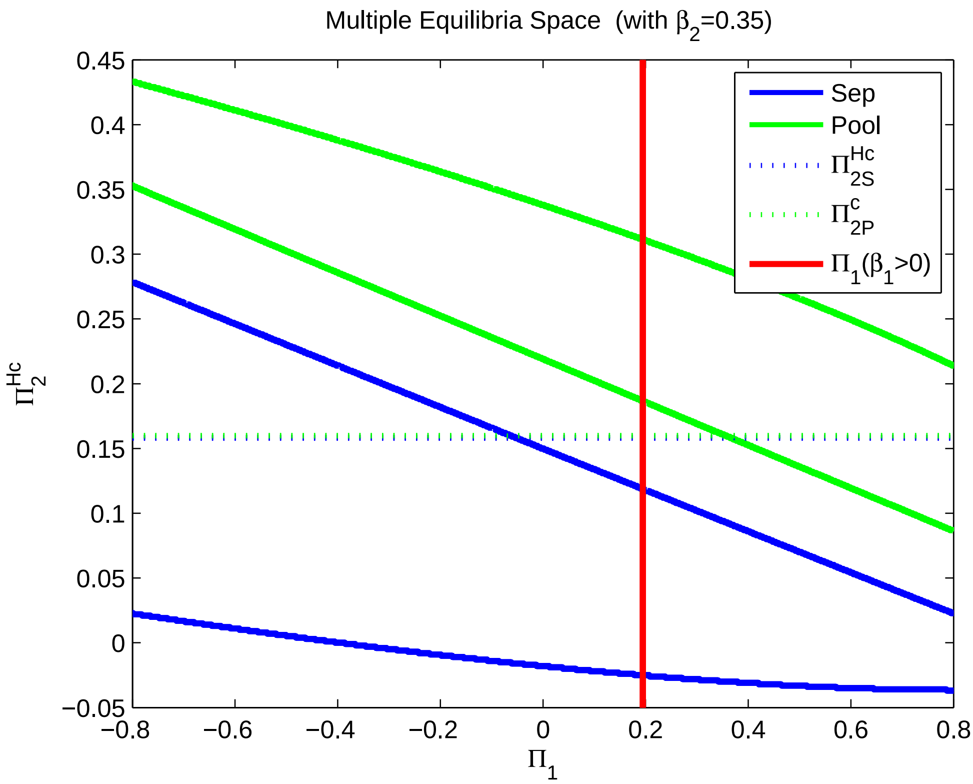


Figure 6.1: Multiple Equilibria

Now we turn to the pooling equilibria space, which is given by the space between the green lines in Figure 6.1. The lower green line plots the minimum level of inflation for which a Dove will pool the Hawk's decision; anything below the Dove will prefer to play π^{cD}_{2S} and have its type revealed. This logic is similar to the logic behind the upper blue line, which is no coincidence. The upper blue line was the limit of what a Dove can put up to to pass himself as Hawk in a separating equilibrium - a paralell downward shift from the lower green line. After all, deviating from a separating equilibrium is more attractive than a pooling equilibrium (of course the central bank does not have the option of choosing between the two) since the in the first agents believe he is a Hawk with probability one whereas in the second the agents will keep expecting a Hawk with probability μ. Despite the interest of this fact, the important is to notice how a reduction of π_1 reduces the size of the interval which could potentially sustain a pooling equilibrium. Therefore, the idea behind reputation transfer survives: a contraction in monetary policy makes it harder to sustain a pooling equilibrium.

Finally, it is worthwhile to point out that this discussion focused only on pure strategy equilibria. However, given one belief refinement criteria, a combination of β and π_1 may entail in the absence of a pure strategy equilibrium. Indeed, for the parametrization of Figure 6.1, the refinement of Cukierman and Liviatan (1991) and Walsh (2000) used in the paper has only a mixed equilibrium. In the graph this is shown by the fact that the horizontal lines (the levels of inflation for separating or pooling given our refinement) are outside the space between blue or green lines given the π_1 value marked by the vertical red line. In this case, we showed in the body of the paper that reputation transfer has a straightforward intuition: it increases the probability α of the Dove choosing the separating action and, consequently, increases the posterior probability of the central banker being seen as a Hawk.

Appendix C: Proofs

Proposition 1 *A sufficient and necessary condition for* $\pi_1 \leq \frac{\kappa(1-\gamma)}{(1-\gamma)^2+\theta_1}$ *is* $\theta_1 \leq \theta_2 C$ *where* C *is a constant greater than 1.*

*Prova.*See Corollary 8.1

Proposition 2 *Holding* π_1 *and the parameters constant, for* γ *small enough, there exists* β_S *such that* $L_S^d \leq L_{SD}^d$ $\forall\ 0 \leq \beta \leq \beta_S$ *and that* $L_S^d > L_{SD}^d$ $\forall\ \beta > \beta_S$.

Prova.

First note that

$$\lim_{\beta\to 0}\left[L_S^D - L_{SD}^D\right] = -\frac{(1-\gamma)^2(\theta^D-\theta^H)^2(\kappa-\gamma\pi_1)^2}{2\left((1-\gamma)^2+\theta^D\right)\left((1-\gamma)^2+\theta^H\right)^2} < 0$$

$$\Rightarrow L_S^D < L_{SD}^D$$

Hence, for β small, there is a separating equilibrium.

Also note that

$$\lim_{\beta\to+\infty}\left[L_S^D - L_{SD}^D\right] = +\infty$$

For β large, the separating equilibrium does not exist.

Due to the Intermediate Value theorem, there is β_S such that $L_S^D - L_{SD}^D = 0$. To conclude the proof, we need to show that for γ small, $\left[L_S^D - L_{SD}^D\right]$ is monotonous in β.

The derivatives are:

$$\frac{\partial L_S^D}{\partial \beta} = \frac{1}{2}\Bigg[\frac{\left((1-\gamma)^2\kappa\left((1-\gamma)^2(\beta\varphi_D+1)+(\gamma+1)\theta^D\right)+\gamma^2\pi_1\theta^{D2}\right)^2}{\left((1-\gamma)^2+\theta^D\right)^2\left((1-\gamma)^2(\beta\varphi_D+1)+\theta^D\right)^2} -$$
$$\frac{2\beta\gamma^2(1-\gamma)^2\theta^D\varphi_D\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^D\right)^2}{\left((1-\gamma)^2+\theta^D\right)\left((1-\gamma)^2(\beta\varphi_D+1)+\theta^D\right)^3} - \frac{2(\gamma-1)^4\theta^D\varphi_D(\kappa-\gamma\pi_1(\beta\varphi_D+1))^2}{\left((1-\gamma)^2(\beta\varphi_D+1)+\theta^D\right)^3} +$$
$$\frac{(1-\gamma)^2\theta^D\left(\kappa-\frac{\gamma\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^D\right)}{(1-\gamma)^2(\beta\varphi_D+1)+\theta^D}\right)^2}{\left((1-\gamma)^2+\theta^D\right)^2} + \frac{2\gamma(1-\gamma)^2\pi_1\theta^D\varphi_D(\gamma\pi_1(\beta\varphi_D+1)-\kappa)}{\left((1-\gamma)^2(\beta\varphi_D+1)+\theta^D\right)^2} -$$
$$\frac{2(1-\gamma)^2\varphi_D\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^D\right)^2}{\left((1-\gamma)^2(\beta\varphi_D+1)+\theta^D\right)^3}$$
$$-2(1-\gamma)^2\kappa\mu\left(\frac{\varphi\text{h}\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^H\right)}{\left((1-\gamma)^2(\beta\varphi\text{h}+1)+\theta^H\right)^2} - \frac{\varphi_D\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^D\right)}{\left((1-\gamma)^2(\beta\varphi_D+1)+\theta^D\right)^2}\right)\Bigg]$$

$$\frac{\partial L_{DS}^D}{\partial \beta} = \frac{1}{2}\Bigg[\frac{2(1-\gamma)^2\kappa(\theta^D-\theta^H)\left(\kappa\left(\theta^H-(1-\gamma)^2(-\beta\varphi_H+\gamma-1)\right)-\gamma^2\pi_1\theta^H\right)}{\left((1-\gamma)^2+\theta^D\right)\left((1-\gamma)^2+\theta^H\right)\left((1-\gamma)^2(\beta\varphi_H+1)+\theta^H\right)}$$
$$-\frac{2\beta\gamma(1-\gamma)^2\varphi_H\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^H\right)}{\left((1-\gamma)^2+\theta^D\right)\left((1-\gamma)^2+\theta^H\right)\left((1-\gamma)^2(\beta\varphi_H+1)+\theta^H\right)^3}\Bigg[$$
$$(1-\gamma)^2\kappa\left((\gamma-1)\theta^D((\gamma-1)(-\beta\varphi_H+\gamma-1)+\theta^H)+\theta^H\left((1-\gamma)^2(\beta\varphi_H+1)+\theta^H\right)\right)+$$
$$+\gamma^2\pi_1\theta^D\theta^H\left((1-\gamma)^2+\theta^H\right)\Bigg]$$
$$+\frac{\left((1-\gamma)^2\kappa\left(\gamma^2(\beta\varphi_H+1)+\beta\varphi_H+\gamma(-2\beta\varphi_H+\theta^D-2)+\theta^H+1\right)+\gamma^2\pi_1\theta^D\theta^H\right)^2}{\left((1-\gamma)^2+\theta^D\right)^2\left((1-\gamma)^2(\beta\varphi_H+1)+\theta^H\right)^2}$$
$$-2(1-\gamma)^2\kappa(1-\mu)\left(\frac{\varphi_D\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^D\right)}{\left((1-\gamma)^2(\beta\varphi_D+1)+\theta^D\right)^2} - \frac{\varphi_H\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^H\right)}{\left((1-\gamma)^2(\beta\varphi_H+1)+\theta^H\right)^2}\right)$$
$$-\frac{2(\gamma-1)^4\theta^D\varphi_H(\kappa-\gamma\pi_1(\beta\varphi_H+1))^2}{\left((1-\gamma)^2(\beta\varphi_H+1)+\theta^H\right)^3} + \frac{(1-\gamma)^2\theta^D\left(\kappa-\frac{\gamma\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^H\right)}{(1-\gamma)^2(\beta\varphi_H+1)+\theta^H}\right)^2}{\left((1-\gamma)^2+\theta^D\right)^2}$$
$$+\frac{2\gamma(1-\gamma)^2\pi_1\theta^D\varphi_H(\gamma\pi_1(\beta\varphi_H+1)-\kappa)}{\left((1-\gamma)^2(\beta\varphi_H+1)+\theta^H\right)^2} - \frac{2(1-\gamma)^2\varphi_H\left((1-\gamma)^2\kappa+\gamma\pi_1\theta^H\right)^2}{\left((1-\gamma)^2(\beta\varphi_H+1)+\theta^H\right)^3}\Bigg]$$

As we can see from above, $\frac{\partial L_S^D}{\partial \beta}$ and $\frac{\partial L_{DS}^D}{\partial \beta}$ are continuous in γ. Now note that:

$$\lim_{\gamma \to 0} \frac{\partial (L_S^D - L_{DS}^D)}{\partial \beta} = \frac{\kappa^2(\theta^H - \theta^D)}{(1+\theta^D)\,(1+\theta^H)} > 0$$

Therefore, for γ small enough, $L_S^D - L_{DS}^D$ will be strictly increasing in β, which concludes the proof.

It is interesting to note that:

$$\lim_{\beta \to +\infty} \frac{\partial (L_S^D - L_{DS}^D)}{\partial \beta} = \frac{\kappa^2(\theta^H - \theta^D)}{(\theta^D + (1-\gamma)^2)\,(\theta^H + (1-\gamma)^2)} > 0$$

So the restriction on γ is need for intermediate values of β. With γ high, it is possible some non monotonicity to be present. □

Proposition 3 *Holding π_1 and the parameters constant, for γ small enough, there exists β_P such that $L_P^d \leq L_{PD}^d \ \forall \ \beta > \beta_P$ and that $L_P^d > L_{PD}^d \ \forall \ 0 \leq \beta \leq \beta_P$.*

Prova.

This proof borrows a lot from the proof of Proposition 2.

First note that

$$\lim_{\beta \to 0} \left[L_P^D - L_{PD}^D\right] = \frac{(1-\gamma)^2(\theta^D - \theta^H)^2(\kappa\left((1-\gamma)^2(\gamma(\mu-1)+1)+\theta^H\right) - \gamma\left((1-\gamma)^2+\theta^H\right)\pi_1)^2}{2((1-\gamma)^2+\theta^D)((1-\gamma)^2+\theta^H)^4} > 0$$

$$\Rightarrow L_P^D > L_{PD}^D$$

Hence, for β small, there is no pooling equilibrium.

Also note that

$$\lim_{\beta \to +\infty} \left[L_P^D - L_{PD}^D\right] = -\infty$$

For β large, the pooling equilibrium shall exist.

As $\left[L_P^D - L_{PD}^D\right]$ is a continuous function of β, the Intermediate Value theorem implies that there is a β_P such that $L_P^D - L_{PD}^D = 0$. To conclude the proof, all we need to show is that, for γ small, $\left[L_P^D - L_{PD}^D\right]$ is monotonous in β. This means we can divide the parameter space for β in areas where the pooling equilibrium is sustained and areas where the Dove will deviate.

The derivatives $\frac{\partial (L_P^D)}{\partial \beta}$ and $\frac{\partial (L_{DP}^D)}{\partial \beta}$ are omitted due to space convenience[5]. We are, however, interested in the limit of their difference:

[5]Their expressions are available under request in a Wolfram Mathematica file.

$$\lim_{\gamma \to 0} \frac{\partial(L_P^D - L_{DP}^D)}{\partial \beta} = -\frac{\kappa^2(\theta^H - \theta^D)\mu}{(1+\theta^D)\,(1+\theta^H)} < 0$$

Therefore, for γ small enough, $L_P^D - L_{PS}^D$ will be strictly decreasing in β, which concludes the proof. □

Proposition 4 *If $\mu = 1$ then $L_S^d \leq L_{DS}^d \iff L_P^d \geq L_{PD}^d$.*

*Prova.*Checking equations (6-5) and (6-8), we see that $\mu = 1$ gives $\pi_{2P}^c = \pi_{2S}^{ch}$. This in turn implies $\pi_{3DS}^{cd} = \pi_{3P}^{cd}$.

Note also that, with $\mu = 1$, when a Dove deviates from the separating strategy, pretending to be a Hawk, he will play $\pi_{2S}^{ch} = \pi_{2P}^c$. Similarly, deviating from the pooling strategy is to play the separating strategy. Therefore we have:

$$L_P^d = L_{DS}^d \text{ and } L_S^d = L_{DP}^d$$

$$L_S^d \leq L_{DS}^d \iff L_P^d \geq L_{PD}^d$$

This concludes the proof. If one equilibrium does not exist, the other must exist. □

Proposition 5 *If $\mu = 0$ then $L_P^d > L_{PD}^d$, i.e., there is no pooling equilibrium.*

*Prova.*First we note that regardless of whether CB_2^D plays the separating strategy or chooses anything else, agents will assign probability 0 of facing a Hawk agent in $t = 3$. Also, in either case, the inflation in $t = 3$ will be the same function of past inflation π_2.

Without benefits of pretending to be Hawk, a Dove will only pool is π_{2P}^c equals his preferred level of inflation π_{2S}^{cD}. Then it remains only to prove that $\pi_{2S}^{cd} \neq \pi_{2P}^c$. With $\mu = 0$, we have:

$$\pi_{2S}^{cD} = \frac{(1-\gamma)\{\kappa - \gamma\pi_1(1+\beta\varphi_D)\}}{(1-\gamma)^2 + \theta^D + \beta(1-\gamma)^2\varphi_D} \qquad (i)$$

$$\pi_{2P}^c = \frac{(1-\gamma)\kappa\left(1 + \frac{\beta\gamma(1-\gamma)^2(\theta^H - \theta^D)}{(\theta^H + (1-\gamma)^2)(\theta^D + (1-\gamma)^2)}\right) - \gamma(1-\gamma)\pi_1\,(1+\beta\varphi_H)}{\theta^H + (1-\gamma)^2\,(1+\beta\varphi_H)} \qquad (ii)$$

$(i) = (ii) \iff \theta^H = \theta^D$, which is false by assumption. Therefore if $\mu = 0$ a pooling equilibrium cannot not exist. □

Proposition 6 *For* $\pi_1 < \bar{\pi}_1$*,* $\Delta^S(\pi_1)$ *decreases in* π_1 *and* $\Delta^P(\pi_1)$ *increases in* π_1*.*

Prova.

.

Part I: Separating

First, note:

$$\frac{\partial L_S^D}{\partial \pi_1} = \left(\frac{\gamma^2\theta^D(1+\beta\varphi_D)}{\theta^D+(1-\gamma)^2(1+\beta\varphi_D)}\right)\pi_1 - \kappa(1-\gamma)\mu\left(\frac{\partial\left(\pi_{2S}^{cD}-\pi_{2S}^{cH}\right)}{\partial \pi_1}\right)$$

and

$$\frac{\partial L_{DS}^D}{\partial \pi_1} =$$

$$= \frac{\kappa(1-\gamma)^2\gamma\left[\theta^H(1+\beta\varphi_D)-\theta^D(1+\beta\varphi_H)\right]+\left[(\theta^H)^2(1+\beta\varphi_D)+\theta^D(1-\gamma)^2(1+\beta\varphi_H)^2\right]\gamma^2\pi_1}{\left[\theta^H+(1-\gamma)^2(1+\beta\varphi_H)\right]^2}$$
$$-\kappa(1-\gamma)(1-\mu)\left(\frac{\partial\left(\pi_{2S}^{cH}-\pi_{2S}^{cD}\right)}{\partial \pi_1}\right) - \beta\kappa(1-\gamma)\left(\frac{\partial\left(\pi_{3SD}^{cD}-\pi_{3S}^{cH}\right)}{\partial \pi_1}\right)$$

Tedious algebra shows that

$$\frac{\left[(\theta^H)^2(1+\beta\varphi_D)+\theta^D(1-\gamma)^2(1+\beta\varphi_H)^2\right]}{\left[\theta^H+(1-\gamma)^2(1+\beta\varphi_H)\right]^2} \equiv h(\theta^H)$$

increases in θ^H.

Therefore,

$$h(\theta^H) > h(\theta^D) = \frac{\theta^D(1+\beta\varphi_D)}{\theta^D+(1-\gamma)^2(1+\beta\varphi_D)}$$

Define:

$$\frac{\Delta^S(\pi_1)}{\partial \pi_1} \equiv \frac{\partial L_{DS}^D}{\partial \pi_1} - \frac{\partial L_S^D}{\partial \pi_1}$$

$$\frac{\Delta^S(\pi_1)}{\partial \pi_1} = A\pi_1 - \kappa(1-\gamma)\left(\frac{\partial\left(\pi_{2S}^{cH}-\pi_{2S}^{cD}\right)}{\partial \pi_1} + \beta\frac{\partial\left(\pi_{3SD}^{cD}-\pi_{3S}^{cH}\right)}{\partial \pi_1} - \frac{(1-\gamma)\gamma\left[\theta^H(1+\beta\varphi_D)-\theta^D(1+\beta\varphi_H)\right]}{\left[\theta^H+(1-\gamma)^2(1+\beta\varphi_H)\right]^2}\right)$$

where

$$A \equiv \gamma^2\left[\frac{\left[(\theta^H)^2(1+\beta\varphi_D)+\theta^D(1-\gamma)^2(1+\beta\varphi_H)^2\right]}{\left[\theta^H+(1-\gamma)^2(1+\beta\varphi_H)\right]^2} - \frac{\theta^D(1+\beta\varphi_D)}{\theta^D+(1-\gamma)^2(1+\beta\varphi_D)}\right] > 0$$

Also:

$$\frac{\partial\left(\pi_{2S}^{cH}-\pi_{2S}^{cD}\right)}{\partial \pi_1}=\gamma(1-\gamma)\frac{\left[\theta^H(1+\beta\varphi_D)-\theta^D(1+\beta\varphi_H)\right]}{\left[\theta^H+(1-\gamma)^2(1+\beta\varphi_H)\right]\left[\theta^D+(1-\gamma)^2(1+\beta\varphi_D)\right]}$$

$$\frac{\partial\left(\pi_{3SD}^{cD}-\pi_{3S}^{cH}\right)}{\partial \pi_1}=\gamma^2(1-\gamma)\frac{(\theta^D-\theta^H)\theta^H}{\left[\theta^H+(1-\gamma)^2(1+\beta\varphi_H)\right]\left[\theta^H+(1-\gamma)^2\right]\left[\theta^D+(1-\gamma)^2\right]}$$

Hence we can rewrite as

$$\frac{\Delta^S(\pi_1)}{\partial \pi_1}=A\pi_1-B+\beta C$$

where

$$B=\frac{\kappa(1-\gamma)^2\gamma\left[\theta^H(1+\beta\varphi_D)-\theta^D(1+\beta\varphi_H)\right]}{\left[\theta^H+(1-\gamma)^2(1+\beta\varphi_H)\right]}\times \ldots$$
$$\ldots\left[\frac{1}{\left[\theta^D+(1-\gamma)^2(1+\beta\varphi_D)\right]}-\frac{1}{\left[\theta^H+(1-\gamma)^2(1+\beta\varphi_H)\right]}\right]$$

$$C=\kappa\gamma^2(1-\gamma)^2\frac{(\theta^H-\theta^D)\theta^H}{\left[\theta^H+(1-\gamma)^2(1+\beta\varphi_H)\right]\left[\theta^H+(1-\gamma)^2\right]\left[\theta^D+(1-\gamma)^2\right]}$$

Clearly,

$$\frac{\Delta^S(\pi_1)}{\partial \pi_1}<0 \iff \pi_1<\frac{B-\beta C}{A}\equiv \bar{\pi_1}^S$$

This concludes the part of the proof referring to Separating equilibrium.

To provide a bit of intuition for this upper bound, we analyze the case where $\beta = 0$:

$$\beta=0 \Rightarrow \bar{\pi_1}^S=\frac{\kappa}{\gamma}$$

This is the upper bound to keep π_2^c positive. The reason why this is so stems from the intuition given in Figure 3.2. If π_1 is too large, the MgB line becomes negative and the grey triangle goes to the third quadrant. There, an increase in π_1 increases the triangle's area further and thus makes pooling harder.

Note, however, that in equilibrium CB_1 will never choose $\pi_1 > \frac{\kappa}{\gamma}$.

Part II: Pooling

Define:

$$\kappa^* \equiv \kappa \left(1 + \frac{\beta(1-\mu)\gamma(1-\gamma)^2(\theta^H - \theta^D)}{\left(\theta^H + (1-\gamma)^2\right)\left(\theta^D + (1-\gamma)^2\right)} \right)$$

Then:

$$\pi^c_{2P} = \frac{(1-\gamma)\kappa^* - \gamma(1-\gamma)\pi_1\left(1+\beta\varphi_H\right)}{\theta^H + (1-\gamma)^2\left(1+\beta\varphi_H\right)}$$

$$\pi_{2P} = \frac{(1-\gamma)^2\kappa^* + \gamma\pi_1\theta^H}{\theta^H + (1-\gamma)^2\left(1+\beta\varphi_H\right)}$$

It is easy to see that the difference between π^c_{2P} and π^{cH}_{2S} stems from the difference between κ and κ^*. Hence, we can write:

$$\frac{\partial L^D_P}{\partial \pi_1} = \frac{\kappa^*(1-\gamma)^2\gamma\left[\theta^H(1+\beta\varphi_D) - \theta^D(1+\beta\varphi_H)\right] + \left[(\theta^H)^2(1+\beta\varphi_D) + \theta^D(1-\gamma)^2(1+\beta\varphi_H)^2\right]\gamma^2\pi_1}{\left[\theta^H + (1-\gamma)^2(1+\beta\varphi_H)\right]^2} - \beta\kappa\mu(1-\gamma)\left(\frac{\partial\left(\pi^{cD}_{3P} - \pi^{cH}_{3P}\right)}{\partial\pi_1}\right)$$

If a Dove deviates from Pooling Equilibrium, it will choose the same inflation as his separating inflation. Note, however, that inflations expectations are different.

$$\frac{\partial L^D_{PD}}{\partial\pi_1} = \left(\frac{\gamma^2\theta^D(1+\beta\varphi_D)}{\theta^D + (1-\gamma)^2(1+\beta\varphi_D)}\right)\pi_1 - \kappa(1-\gamma)\left(\frac{\partial\left(\pi^{cD}_{2S} - \pi^c_{2P}\right)}{\partial\pi_1}\right)$$

Also:

$$\frac{\partial\left(\pi^{cD}_{2S} - \pi^c_{2P}\right)}{\partial\pi_1} = -\gamma(1-\gamma)\frac{\left[\theta^H(1+\beta\varphi_D) - \theta^D(1+\beta\varphi_H)\right]}{\left[\theta^H + (1-\gamma)^2(1+\beta\varphi_H)\right]\left[\theta^D + (1-\gamma)^2(1+\beta\varphi_D)\right]}$$

$$\frac{\partial\left(\pi^{cD}_{3P} - \pi^{cH}_{3P}\right)}{\partial\pi_1} = \gamma^2(1-\gamma)\frac{(\theta^D - \theta^H)\theta^H}{\left[\theta^H + (1-\gamma)^2(1+\beta\varphi_H)\right]\left[\theta^H + (1-\gamma)^2\right]\left[\theta^D + (1-\gamma)^2\right]}$$

Define:

$$\frac{\Delta^P(\pi_1)}{\partial\pi_1} \equiv \frac{\partial L^D_{DP}}{\partial\pi_1} - \frac{\partial L^D_P}{\partial\pi_1}$$

Hence we can rewrite as

$$\frac{\Delta^P(\pi_1)}{\partial \pi_1} = -A\pi_1 + B + \beta D$$

where A, B were defined above and:

$$D = \kappa\gamma^2(1-\gamma)^2 \frac{(\theta^H - \theta^D)}{[\theta^H + (1-\gamma)^2(1+\beta\varphi_H)]\,[\theta^H + (1-\gamma)^2]\,[\theta^D + (1-\gamma)^2]} \times \\ \times \left(\mu\theta^H - (1-\mu)(1-\gamma)^2 \frac{[\theta^H(1+\beta\varphi_D) - \theta^D(1+\beta\varphi_H)]}{[\theta^H + (1-\gamma)^2(1+\beta\varphi_H)]}\right)$$

Clearly,

$$\frac{\Delta^P(\pi_1)}{\partial \pi_1} > 0 \iff \pi_1 < \frac{B + \beta D}{A} \equiv \bar{\pi_1}^P$$

Defining

$$\bar{\pi_1} \equiv \min(\bar{\pi_1}^P, \bar{\pi_1}^s)$$

concludes the proof.□

Proposition 7 *A sufficient condition for* $\frac{\partial \pi_1}{\partial \theta_{12}} > 0$ *is* $\kappa_2 > \gamma\kappa_1$*.*

*Prova.*In order to facilitate the algebra computations, we use the software **Wolfram Mathematica**.

First we note that as $\pi_1 = \pi_1^c(1-\gamma)$, we have that $\frac{\partial \pi_1}{\partial \theta_{12}} > 0 \iff \frac{\partial \pi_1^c}{\partial \theta_{12}} > 0$. Taking derivatives, we have:

$$\frac{\partial \pi_1^c}{\partial \theta_{12}} = \frac{\Big(\kappa_2\Big((1-\gamma)^4+\theta_1\Big((1-\gamma)^2+\theta_2\Big)+(1+\beta\gamma^2)(1-\gamma)^2\theta_2\Big)-\gamma\kappa_1(1-\gamma)^2\Big((1-\gamma)^2+\theta_2\Big)\Big)}{\frac{1}{\beta(1-\gamma)^3\gamma\left((1-\gamma)^2+\theta_2\right)}\Big(\theta_1((1-\gamma)^2+\theta_2)^2+(1-\gamma)^2(\gamma(\gamma+\beta\gamma\theta_{12}-2)+1)+\theta_2^2(\beta\gamma^2+1)+2(1-\gamma)^2\theta_2\Big)^2}$$

From above we have that:

$$\frac{\partial \pi_1}{\partial \theta_{12}} > 0 \iff (\kappa_2\,(\theta_1\,((1-\gamma)^2+\theta_2) + \beta\gamma^2(1-\gamma)^2\theta_2) + (\kappa_2 - \gamma\kappa_1)\,((1-\gamma)^2+\theta_2)\,(1-\gamma)^2) > 0$$

It is easy to see that a sufficient condition is $(\kappa_2 - \gamma\kappa_1) > 0$.□

We also note the following: now that θ_{12} can be negative, we must create a lower bound so to keep the problem defined. The idea is that if θ_{12} is far too negative, the loss function will become concave and then the minimization problem no longer bound. A sufficient condition (which is quite week) for that not to happen is:

$$\theta_{12} > -\frac{\theta_2^2\Big((1-\gamma)^2\big(\beta\gamma^2+1\big)+\theta_1\Big)+2(1-\gamma)^2\theta_2\Big((1-\gamma)^2+\theta_1\Big)+(1-\gamma)^4\Big((1-\gamma)^2+\theta_1\Big)}{\beta(1-\gamma)^4\gamma^2}$$

Proposition 8 *Let* $z = \frac{\kappa_1}{\kappa_2}$*. A sufficient and necessary condition for* $\pi_1^c \geq \frac{\kappa_1(1-\gamma)}{(1-\gamma)^2+\theta_1}$ *is:*

$$\theta_1 \geq \frac{1}{2}\left(\theta_2 + (\gamma z - 1)(1-\gamma)^2 + \sqrt{(1-\gamma)^4(\gamma z - 1)^2 + 2\theta_2(1-\gamma)^2(1+\gamma z) + \theta_2^2(1+4\gamma z)}\right).$$

*Prova.*This proof is quite simple. First, since we are in the case where θ stands for an economic cost, we impose $\theta_{12} = \theta_1$.

We want a necessary and sufficient condition for

$$\pi_1^c = \frac{\kappa_1 + \beta\left((\theta_1 - \theta_2)\frac{\kappa_2\gamma(1-\gamma)^2}{(\theta_2+(1-\gamma)^2)^2}\right)}{\left(1 + \theta_1 + \beta\left(\frac{\gamma^2\theta_2^2+\gamma^2(1-\gamma)^2\theta_1}{(\theta_2+(1-\gamma)^2)^2}\right)\right)} \geq \frac{\kappa_1(1-\gamma)}{(1-\gamma)^2+\theta_1}$$

All one needs to do is to rearrange the inequality above in order to isolate θ_1. To facilitate the computations, we used the command 'Reduce' in **Mathematica** imposing the following restrictions: $\theta^H > 0$; $\theta^D > 0$; $\kappa_2 > 0$; $\kappa_1 > 0$; $\beta > 0$; $1 > \gamma > 0$. This gave us the condition stated in Proposition 8.□

Corollary 8.1 *If* $\kappa_1 = \kappa_2$ *then the sufficient and necessary condition of Proposition 2 becomes* $\theta_1 \geq \theta_2 C$ *where* C *is a constant greater than 1.*

*Prova.*Clearly $\kappa_1 = \kappa_2 \iff z = 1$. Thus the condition becomes:

$$\theta_1 \geq \frac{1}{2}\left(\theta_2 - (1-\gamma)^3 + \sqrt{(1-\gamma)^6 + 2\theta_2(1-\gamma)^2(1+\gamma) + \theta_2^2(1+4\gamma)}\right)$$

This can be rewritten as:

$$\theta_1 \geq \frac{1}{2}\left(\theta_2 - (1-\gamma)^3 + \sqrt{(\theta_2 + (1-\gamma)^3)^2 + 4\theta_2\gamma(\theta_2 + (1-\gamma)^2)}\right)$$

Since $4\theta_2\gamma(\theta_2 + (1-\gamma)^2) > 0$ we can rewrite:

$$\sqrt{(\theta_2 + (1-\gamma)^3)^2 + 4\theta_2\gamma(\theta_2 + (1-\gamma)^2)} = \theta_2 + (1-\gamma)^3 + c$$

where c is a positive constant implicitly defined by the equation above . Therefore:

$$\theta_1 \geq \frac{1}{2}\left(\theta_2 - (1-\gamma)^3 + \theta_2 + (1-\gamma)^3 + c\right)$$

Rearranging:

$$\theta_1 \geq \left(\theta_2 + \frac{c}{2}\right)$$

Defining $C \equiv 1 + \frac{c}{2\theta_2}$, we can conclude that:

$$\pi_1^c \geq \frac{\kappa_1(1-\gamma)}{(1-\gamma)^2 + \theta_1} \iff \theta_1 \geq \theta_2 C$$

.□

Corollary 8.2 *If $z \to 0$ then the condition becomes $\theta_1 \geq \theta_2$.*

*Prova.*Taking the limit of the expression as $z \to 0$:

$$\tfrac{1}{2}\left(\theta_2 + (\gamma z - 1)(1-\gamma)^2 + \sqrt{(1-\gamma)^4(\gamma z - 1)^2 + 2\theta_2(1-\gamma)^2(1+\gamma z) + \theta_2^2(1+4\gamma z)}\right) \to$$
$$\to \frac{1}{2}\left(\theta_2 - (1-\gamma)^2 + \sqrt{(\theta_2 + (1-\gamma)^2)^2}\right)$$

Clearly:

$$\frac{1}{2}\left(\theta_2 - (1-\gamma)^2 + \sqrt{(\theta_2 + (1-\gamma)^2)^2}\right) = \theta_2$$

Hence the condition becomes $\theta_1 \geq \theta_2$.□

Proposition 9 *A central bank whose successor is hawkish chooses lower inflation than one with a dovish successor in a separating equilibrium: $\pi_{1S}^H \leq \pi_{1S}^D$.*

Prova.

We will prove by contradiction: assume $\pi_{1S}^H > \pi_{1S}^D$.

First note that:

$$\frac{\partial}{\partial \theta^i}\left(\frac{\gamma^2\theta^i(1+\beta\varphi_i)}{\theta^i+(1-\gamma)^2(1+\beta\varphi_i)}\right) =$$
$$= \frac{(1-\gamma)^2\gamma^2\left((1-\gamma)^4+\theta^{i2}\left(\beta\gamma^2\left(\beta\gamma^2+3\right)+1\right)+2(1-\gamma)^2\theta^i\left(\beta\gamma^2+1\right)\right)}{\left((1-\gamma)^4+\theta^{i2}+(1-\gamma)^2\theta^i\left(\beta\gamma^2+2\right)\right)^2} > 0$$

Hence:

$$\left(\frac{\gamma^2\theta^H(1+\beta\varphi_H)}{\theta^H+(1-\gamma)^2(1+\beta\varphi_H)}\right) > \left(\frac{\gamma^2\theta^D(1+\beta\varphi_D)}{\theta^D+(1-\gamma)^2(1+\beta\varphi_D)}\right)$$

Given that the right hand side of 6-9 is the same for both $i \in \{H, D\}$, it is sufficient to show that:

$$\left[\pi_2^{ci}\frac{\partial\pi_{2S}^{ci}}{\partial\pi_{1S}^{i}} + \beta\pi_3^{ci}\frac{\partial\pi_3^{ci}}{\partial\pi_{2S}^{i}}\frac{\partial\pi_{2S}^{i}}{\partial\pi_{1S}^{i}}\right] < 0$$

After all, as the term in brackets only appears in the π_{1S}^d equation (for $\theta^H - \theta^H = 0$), π_{1S}^D would have to be greater than π_{1S}^H in order for the right hand side to equal $\kappa(1-\gamma)$.

Since $\frac{\partial\pi_{2S}^{ci}}{\partial\pi_{1S}^{i}} < 0$, $\frac{\partial\pi_3^{ci}}{\partial\pi_{2S}^{i}} < 0$ and $\frac{\partial\pi_{2S}^{i}}{\partial\pi_{1S}^{i}} > 0$, all we need is that π_2^{ci} and π_3^{ci} are positive. Before we elaborate on the parameter restrictions which assure this to be the case, it is worth noting that there is no reason why π_2^{ci} and π_3^{ci} would be negative. After all, all central banks dislike π^c far from zero and have a positive inflation bias. The only case when it would be optimal to have negative π^c is when the inherited inflation is extremely high. Indeed, $\pi_2^{cH} > 0$ as long as:

$$\pi_{1S}^h = \frac{\kappa(1-\gamma)^2}{(1-\gamma)^2+\theta^H+\beta\left(\frac{\gamma^2\theta^H(1+\beta\varphi_h)}{\theta^H+(1-\gamma)^2(1+\beta\varphi_h)}\right)} < \frac{\kappa}{\gamma(1+\beta\varphi_H)} < \frac{\kappa}{\gamma(1+\beta\varphi_D)}$$

This is always satisfied for $\theta^H > \theta^D > 0$, $\gamma \in (0,1)$, $\kappa > 0$ and $\beta \geq 0$. These are the basic restrictions of the model.

In order for $\pi_3^{cH} > 0$, it suffices that $\pi_2^H < \frac{\kappa}{\gamma}$. Algebra shows that:

$$\pi_{2S}^i < \frac{\kappa}{\gamma} \iff \pi_1 < \frac{\kappa\left((1-\gamma)^5+\theta^{i2}+(1-\gamma)^2\theta^i\left(\beta\gamma^2-\gamma+2\right)\right)}{\gamma^2\theta^i\left((1-\gamma)^2+\theta^i\right)}$$

This always holds in equilibrium (i.e. for $\pi_1 = \pi_{1S}^H$).

The case for $i = D$ will follow from our contradiction assumption.

$$\pi_{1S}^{D} < \pi_{1S}^{H} \Rightarrow \pi_{1S}^{D} < \frac{\kappa}{\gamma(1+\beta\varphi_D)} \Rightarrow \pi_2^{cD} > 0$$

Also:

$$\pi_{1S}^{D} < \pi_{1S}^{H} \Rightarrow \pi_2^{D} < \frac{\kappa}{\gamma} \Rightarrow \pi_3^{cD} > 0$$

Hence:

$$\left[\pi_2^{cD}\frac{\partial \pi_{2S}^{cD}}{\partial \pi_{1S}^{D}} + \beta\pi_3^{cD}\frac{\partial \pi_3^{cD}}{\partial \pi_{2S}^{D}}\frac{\partial \pi_{2S}^{D}}{\partial \pi_{1S}^{D}}\right] < 0$$

This contradicts $\pi_{1S}^{H} > \pi_{1S}^{D}$, concluding the proof. $\square$