US Monetary and Fiscal Policies - Conflict or Cooperation?*

Xiaoshan Chen Eric M. Leeper Campbell Leith
Durham University Indiana University Glasgow University

June 29, 2015

Abstract

Most of the literature estimating DSGE models for monetary policy analysis ignores fiscal policy and assumes that monetary policy follows a simple rule. In this paper we allow both fiscal and monetary policy to be described by rules and/or optimal policy which are subject to switches over time. We find that US monetary and fiscal policy have often been in conflict, and that it is relatively rare that we observe the benign policy combination of an conservative monetary policy paired with a debt stabilizing fiscal policy. In a series of counterfactuals, a conservative central bank following a time-consistent fiscal policy leader would come close to mimicking the cooperative Ramsey policy. However, if policy makers cannot credibly commit to such a regime, monetary accommodation of the prevailing fiscal regime may actually be welfare improving.

- JEL Codes:
- Key Words: Bayesian Estimation, interest rate rules, fiscal policy rules, optimal monetary policy, optimal fiscal policy, great moderation, commitment, discretion

1 Introduction

The 'Great Moderation' in output and inflation volatility has been the subject of much analysis, particularly for the US, where following Sims and Zha(2006) a large literature has emerged which assesses the extent to which this was simply 'good luck' - a favorable shift in shock volatilities - or 'good policy' - a desirable change in monetary policy rule parameters and/or the implicit inflation target. The improvement in policy making is typically associated with the Volcker disinflation which tends to be dated as occurring in 1979.[1] Despite the magnitude of the literature examining this issue, there is very little work examining what role fiscal policy played in the development of trend-inflation. This is somewhat surprising when one contrasts the development of inflation, real interest rates and fiscal variables including the debt to GDP ratio (see Figure 1) where the

*Address for correspondence: University of Glasgow, Economics, Adam Smith Business School, Gilbert Scott Building, Glasgow G12 8QQ, U.K. E-mail: xiaoshan.chen@durham.ac.uk, eleeper@indiana.edu and campbell.leith@glasgow.ac.uk. We would like to thank Chris Sims, Harald Uhlig, Tack Yun, Tao Zha and participants at the 2015 Tsinghua-CAEPR conference on monetary and fiscal policy in Beijing for helpful comments.

[1]See Chen et al (2013) for a discussion of the various strands of this literature.

upward trend in inflation prior to the 1980s appears to be associated with a downward trend in the debt to GDP ratio, while the moderation in inflation came at a time of a step increase in the real interest rate and rising debt to GDP ratio, at least until 1995. A notable exception to this absence of fiscal policy in explaining the development of inflation is Bianchi (2012) and Bianchi and Ilut (2015) who build on the insights of Leeper (1991) to allow for switches in the combination of monetary and fiscal policy rules over time.[2] Essentially, they find that prior to the Volcker disinflation rising inflation and falling debt to GDP levels are generated by a combination of passive monetary policy (a Taylor rule which fails to satisfy the Taylor principle requiring that real interest rates rise in response to rising inflation) and an active fiscal policy (which fails to adjust the surplus/deficit to stabilize debt). There is then a period of policy conflict following the appointment of Volcker as monetary policy turns active, while fiscal policy still fails to act to stabilize the debt before fiscal policy adjusts, turning passive in support of the post-Volcker anti-inflation policies of the Fed. This benign policy configuration then explains the fall in inflation in the 1980s, but is associated with rising debt levels.

Our current paper builds on this analysis in several ways. Firstly, we consider other types of policy making in addition to simple policy rules. Specifically, we allow monetary policy to be conducted optimally under discretion, but with fluctuations in the degree of inflation conservatism. We also allow fiscal policy to transition between active/passive fiscal rules and optimal time-consistent policy making where the fiscal authority acts as a Stackelberg leader in a game with the optimizing monetary authority. We find that this set of potential policy regimes offers a data-preferred description of monetary and fiscal policy relative to the usual rules-based approach. The paper also develops a new algorithm for solving the strategic policy game between the monetary and fiscal policy makers in the face of shifts in regime.

Secondly, we extend the policy space to allow for more potential permutations of regime. Both Bianchi (2012) and Bianchi and Ilut (2015) restrict their attention to three regimes - the benign regime where monetary policy actively targets inflation, and fiscal policy adjusts the deficit to stabilize debt (AM/PF), a less benign regime where monetary policy fails to actively target inflation possibly in order to support a fiscal policy which does not seek to stabilize debt through adjusting fiscal instruments (PM/AF) and a transitory regime where the two policies are in conflict with the monetary authority targeting inflation, and neither policy maker acting to stabilize debt (AM/AF). We consider a richer set of policy regime permutations and transitions across regimes. When we consider rules, we allow for the possible missing permutation of PM/PF. We also do not restrict the transition across regimes. Therefore it is not necessary to transition through the AM/AF regime when moving from 'bad' to 'good' regimes, PM/AF to AM/PF as in Bianchi and Ilut (2015), nor assuming a circular regime transition structure as in Bianchi (2012).This richer set of policy permutations and less restrictive transition paths adds some nuances to the description of the evolution of monetary and fiscal policies over the period and we tend to find that policy was in conflict more often than found in the analysis of Bianchi (2012) and Bianchi and Ilut(2015).

When we turn to consider time-consistent optimal policy where the fiscal authority may move

[2]Related papers which allow for regime switching in estimated fiscal policy processes prior to embedding them in a calibrated model include Davig (2004) and Davig and Leeper (2011). Traum and Yang (2010) implicitly consider switches in policy by estimating a DSGE model with fixed rules over sub-samples with priors favoring the AM/PF and PM/AF policy premutations, respectively.

between fiscal rules and acting as a Stackelberg leader which the monetary authority implements a time-consistent monetary policy with switches in the degree of conservatism, which turns out to be a data-preferred description of policy, the movement between regimes is more striking and it is rare that policy combinations conform to the usual Active/Passive pairings. There can also be substantial spillovers across regimes, with a fiscal authority behaving optimally conducting policy in a manner which takes into account the possible switches to a passive fiscal rule. This latter phenomenon is driven by the inflationary impact of alternative tax policies given that taxes are distortionary. Bianchi (2012) and Bianchi and Ilut (2015) assume lump-sum taxation.

Thirdly, by allowing for policy to be conducted optimally we obtain estimates of policy maker objective functions which allow us to construct a rich set of counterfactual analyses which include a consideration of welfare. Therefore we can distinguish between good luck and, whether for monetary or fiscal policy, good policy, improved credibility and/or avoidance of conflict are more important in re-examining the great moderation.

The plan of the paper is as follows. In Section 2 we describe our New Keynesian model augmented to include medium term government debt. Section 3 outlines the various descriptions of policy we consider before we outline our estimation approach and results in Section 4. Section 5 then undertakes a series of counterfactual exercises, before we conclude in Section 6.

2 The Model

The economy is comprised of households, a monopolistically competitive production sector, and the government. There is a continuum of goods that enter the households' consumption basket. Households form external consumption habits at the level of the consumption basket as a whole - 'superficial' habits.[3] Furthermore, we assume the economy is subject to both price and inflation inertia. Both effects have been found to be important in capturing the hump-shaped responses of output and inflation to shocks evident in VAR based studies, and are often employed in empirical applications of the New Keynesian model.[4]

On the fiscal side we allow the government to levy a tax on firms' sales revenue, which in our simple model is equivalent to a tax on all labour and profit income. These revenues are used to finance government consumption, pay for transfers to households and service/manage the outstanding stock of government debt. Unlike much theoretical analysis, we allow the government to issue a portfolio of bonds of different maturity where we will impose a geometrically declining maturity structure to tractably enable us to price bonds of different maturities within the portfolio.

[3] For a comparison of the implications for optimal policy of alternative forms of habits see Amato and Laubach (2004) and Leith et al (2012).

[4] See for example Smets and Wouters (2003), Christiano, Eichenbaum and Evans(2005) and Leith and Malley (2005).

2.1 Households

The economy is populated by a continuum of households, indexed by k and of measure 1. Households derive utility from consumption of a composite good, $C_t^k = \left(\int_0^1 \left(C_{it}^k\right)^{\frac{\eta-1}{\eta}} di \right)^{\frac{\eta}{\eta-1}}$ where η the elasticity of substitution between the goods in this basket and suffer disutility from hours spent working, N_t^k. Habits are both superficial and external implying that they are formed at the level of the aggregate consumption good, and that households fail to take account of the impact of their consumption decisions on the utility of others. To facilitate data-consistent detrending around a balanced growth path without restricting preferences to be logarithmic in form, we follow Lubik and Schorfheide (2005) and An and Schorfheide (2007) in assuming that the consumption that enters the utility function is scaled by the economy wide technology trend, implying that the household's consumption norms rise with technology as well as being affected by more familiar habits externalities. Accordingly, households derive utility from the habit-adjusted composite good,[5]

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\left(C_t^k/A_t - \theta C_{t-1}/A_{t-1}\right)^{1-\sigma} (\xi_t)^{-\sigma}}{1-\sigma} - \frac{\left(N_t^k\right)^{1+\varphi} (\xi_t)^{-\sigma}}{1+\varphi} \right]$$

where $C_{t-1} \equiv \int_0^1 C_{t-1}^k dk$ is the cross-sectional average of consumption. In other words households gain utility from consuming more than other households, are disappointed if their consumption doesn't grow in line with technical progress and are subject to a time-preference or taste-shock, ξ_t.

The process for technology is non-stationary,

$$\ln A_t = \ln \gamma + \ln A_{t-1} + \ln q_t \tag{1}$$

$$\ln q_t = \rho_q \ln q_{t-1} + \varepsilon_{q,t} \tag{2}$$

Households decide the composition of the consumption basket to minimize expenditures and the demand for individual good i is

$$C_{it}^k = \left(\frac{P_{it}}{P_t}\right)^{-\eta} C_t^k = \left(\frac{P_{it}}{P_t}\right)^{-\eta} \left(X_t^k + \theta C_{t-1}\right).$$

By aggregating across all households, we obtain the overall demand for good i as

$$C_{it} = \int_0^1 C_{it}^k dk = \left(\frac{P_{it}}{P_t}\right)^{-\eta} C_t. \tag{3}$$

Remainder of the Household's Problem The remainder of the household's problem is standard. Specifically, households choose the habit-adjusted consumption aggregate, $X_t^k =$

[5] Note that this utility specification is slightly different from that in Lubik and Shorfheide (2005) who adopt the following specification, $\frac{\left(C_t - \theta\gamma C_{t-1})/A_t\right)^{1-\sigma}(\xi_t)^{-\sigma}}{1-\sigma}$. Their specification introduces a technology shock into the definition of habits adjusted consumption which then complicates the derivation of welfare.

$C_t^k/A_t - \theta C_{t-1}/A_{t-1}$, hours worked, N_t^k, and the portfolio allocation, $B_t^{S,k}$ and $B_t^{M,k}$, to maximize expected lifetime utility

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\left(X_t^k\right)^{1-\sigma} (\xi_t)^{-\sigma}}{1-\sigma} - \frac{\left(N_t^k\right)^{1+\varphi} (\xi_t)^{-\sigma}}{1+\varphi} \right]$$

subject to the budget constraint

$$\int_0^1 P_{it} C_{it}^k di + P_t^S B_t^{S,k} + P_t^M B_t^{M,k} = B_{t-1}^{S,k} + (1+\rho P_t^M) B_{t-1}^{M,k} + W_t N_t^k + \Phi_t + Z_t \tag{4}$$

and the usual transversality condition. E_t is the mathematical expectation conditional on information available at time t, β is the discount factor $(0 < \beta < 1)$, and σ and φ are the inverses of the intertemporal elasticities of habit-adjusted consumption and work $(\sigma, \varphi > 0;\ \sigma \neq 1)$. The household's period-t income includes: wage income from providing labor services to goods producing firms, $W_t N_t^k$, a lump sum transfer from the government, Z_t, dividends from the monopolistically competitive firms, Φ_t, and payments on the portfolio of assets, $B_t^{S,k}$ and $B_t^{M,k}$. Taxes are levied on the firms' sales revenues which is equivalent to charging an income tax on all wage and profit income. Households hold two forms of government bond. The first is the familiar one period debt, B_t^S, which has a price equal to the inverse of the gross nominal interest rate, $P_t^S = R_t^{-1}$. The second type of bond is actually a portfolio of many bonds which, following Woodford (2001) pay a declining premium of ρ^j, j periods after being issued where $0 < \rho < \beta^{-1}$. The duration of the bond, assuming stable prices is $\frac{1}{1-\beta\rho}$, which means that ρ can be varied to capture changes in the maturity structure of debt. By using this simple structure we need only price a single bond, since any existing bond issued j periods ago is worth ρ^j new bonds. In the special case where $\rho = 1$ these bonds become infinitely lived consols.

There is an associated transversality condition derived as follows. Define household wealth in period t as,

$$D_t^k = (1+\rho P_t^M) B_{t-1}^{M,k} + B_{t-1}^{S,k}$$

the transversality condition can be written as,

$$\lim_{T \to \infty} E_t R_{t,T} D_T^k / P_T \geq 0$$

where $R_{t,T} = \prod_{s=t}^{T-1} \left(\frac{1+\rho P_{s+1}^M}{P_s^M} \frac{P_s}{P_{s+1}}\right)$ for $T \geq 1$ and $R_{t,t} = 1$ (see Eusepi and Preston (2010)).

We can then maximize utility subject to the budget constraint (4) to obtain the optimal allocation of consumption across time, based on the pricing of one period bonds,

$$1 = \beta E_t \left[\left(\frac{X_{t+1}^k \xi_{t+1}}{X_t^k \xi_t} \right)^{-\sigma} \frac{A_t}{A_{t+1}} \frac{P_t}{P_{t+1}} \right] R_t,$$

and the geometrically declining payoff consols,

$$P_t^M = \beta E_t \left[\left(\frac{X_{t+1}^k \xi_{t+1}}{X_t^k \xi_t} \right)^{-\sigma} \frac{A_t}{A_{t+1}} \frac{P_t}{P_{t+1}} (1+\rho P_{t+1}^M) \right],$$

Notice that when these reduce to single period bonds, $\rho = 0$, the price of these bonds is also given by $P_t^M = R_t^{-1}$. However, outside of this special case the longer term bonds introduce the term structure of interest rates to the model. It is convenient to define the stochastic discount factor for later use,

$$Q_{t,t+1} = \beta \left(\frac{X_{t+s}^k \xi_{t+1}}{X_t^k \xi_t} \right)^{-\sigma} \frac{A_t}{A_{t+1}} \frac{P_t}{P_{t+1}}, \tag{5}$$

The second foc relates to their labour supply decision and is given by, $\left(\frac{W_t}{P_t A_t}\right) = \mu_t N_t^{k\varphi} X_t^{k\sigma}$ where we have added an exogenous wage markup shock,

$$ln(\mu_t) = \rho_u \ln(\mu_{t-1}) + \varepsilon_t^{\mu}$$

which shall serve as a pure cost push shock to the New Keynesian Phillips Curve (NKPC). It should be noted that variations in tax rates, either under fiscal rules or as part of an optimal policy, will generate similar cost push effects.

2.2 Firms

We further assume that intermediate goods producers are subject to the constraints of Calvo(1983)-contracts such that, with fixed probability $(1-\alpha)$ in each period, a firm can reset its price and with probability α the firm retains the price of the previous period, but where, following Yun (1996) that price is indexed to the steady-state rate of inflation. When a firm can set the price, it can either do so in order to maximize the present discounted value of (after-tax) profits, $E_t \sum_{s=0}^{\infty} \alpha^s Q_{t,t+s} \Phi_{it+s}$, or it can follow a simple rule of thumb as in (Gali and Gertler (1999) or Leith and Malley(2005)). The constraints facing the forward looking profit maximizers are the demand for their own good (3) and the constraint that all demand be satisfied at the chosen price. Profits are discounted by the s-step ahead stochastic discount factor $Q_{t,t+s}$ and by the probability of not being able to set prices in future periods.

$$\max_{\{P_{it},\, Y_{it}\}} E_t \sum_{s=0}^{\infty} \alpha^s Q_{t,t+s} \left[\left((1-\tau_{t+s}) P_{it} \pi^s - MC_{t+s} \right) Y_{it+s} \right]$$

$$\begin{aligned} s.t. Y_{it+s} &= \left(\frac{P_{it}\pi^s}{P_{t+s}} \right)^{-\eta} Y_{t+s} \\ \text{where } Q_{t,t+s} &= \beta^s \left(\frac{X_{t+1}\xi_{t+1}}{X_t \xi_t} \right)^{-\sigma} \frac{P_t}{P_{t+s}} \end{aligned}$$

The relative price set by firms able to reset prices optimally in a forward-looking manner, satisfies the following relationship:

$$\frac{P_t^f}{P_t} = \left(\frac{\eta}{\eta - 1} \right) \frac{E_t \sum_{s=0}^{\infty} (\alpha\beta)^s \left(X_{t+s}\xi_{t+s} \right)^{-\sigma} mc_{t+s} \left(\frac{P_{t+s}\pi^{-s}}{P_t} \right)^{\eta} \frac{Y_{t+s}}{A_{t+s}}}{E_t \sum_{s=0}^{\infty} (\alpha\beta)^s \left(X_{t+s}\xi_{t+s} \right)^{-\sigma} (1-\tau_{t+s}) \left(\frac{P_{t+s}\pi^{-s}}{P_t} \right)^{\eta-1} \frac{Y_{t+s}}{A_{t+s}}} \tag{6}$$

where $mc_t = \frac{MC_t}{P_t}$ is the real marginal cost and P_t^f denotes the price set by all firms who are able to reset prices in period t and choose to do so in a profit maximizing way. Under flexible prices this implies, $mc_t = (1-\tau_t)\frac{\eta-1}{\eta}$.

However, in addition to the familiar Calvo-type price setters, we also allow for inflation inertia. To do so we assume that some firms follow simple rules of thumb when setting prices. Specifically, when a firm is given the opportunity of posting a new price, we assume that rather than posting the profit maximizing price (6), a proportion of those firms, ζ, follows a simple rule of thumb in resetting that price:

$$P_t^b = P_{t-1}^* \pi_{t-1}, \tag{7}$$

such that they update there price in line with last period's rate of inflation rather than steady-state inflation, where P_{t-1}^* denotes an index of the reset prices given by

$$\ln P_{t-1}^* = (1-\zeta)\ln P_{t-1}^f + \zeta P_{t-1}^b. \tag{8}$$

P_t represents the price level at time t. With α of firms keeping last period's price (but indexed to steady-state inflation) and $(1-\alpha)$ of firms setting a new price, the law of motion of this price index is,

$$(P_t)^{1-\eta} = \alpha\left(P_{t-1}\pi\right)^{1-\eta} + (1-\alpha)\left(P_t^*\right)^{1-\eta}.$$

Denoting the fixed share of price-setters following the rule of thumb (7) by ζ, we can derive a price inflation Phillips curve, as detailed in Leith and Malley (2005). For this we combine the rule of thumb of price setters with the optimal price setting described above, leading to the Phillips curve

$$\widehat{\pi}_t = \chi_f \beta E_t \widehat{\pi}_{t+1} + \chi_b \widehat{\pi}_{t-1} + \kappa_c(\widehat{mc}_t + \frac{\tau}{1-\tau}\widehat{\tau}_t), \tag{9}$$

where $\widehat{\pi}_t = \ln(P_t) - \ln(P_{t-1}) - \ln(\pi)$ is the deviation of inflation from its steady state value, $\widehat{mc}_t + \frac{\tau}{1-\tau}\widehat{\tau}_t = \ln(W_t/P_t) - \ln A_t + \frac{\tau}{1-\tau}\widehat{\tau}_t - \ln((\eta-1)/\eta) + \ln(1-\tau)$, are log-linearised real marginal costs adjusted for the impact of the sales revenue tax, and the reduced form parameter convolutions are defined as $\chi_f \equiv \frac{\alpha}{\Phi}$, $\chi_b \equiv \frac{\zeta}{\Phi}$, $\kappa_c \equiv \frac{(1-\alpha)(1-\zeta)(1-\alpha\beta)}{\Phi}$, with $\Phi \equiv \alpha(1+\beta\zeta)+(1-\alpha)\zeta$.

2.3 The Government

Combining the series of the representative consumer's flow budget constraints, (4), and noting the equivalence between factor incomes and national output,

$$P_t Y_t = P_t C_t + P_t G_t = W_t N_t + \Pi_t + \tau_t P_t Y_t$$

we can (after assuming the aggregate stock of one period bonds is in zero net supply, $B_t^S = 0$) rewrite the private sector's budget constraint as,

$$P_t^M B_t^M = (1+\rho P_t^M) B_{t-1}^M - P_t Y_t \tau_t + P_t G_t + P_t Z_t \tag{10}$$

Distortionary taxation and spending adjustments are required to service government debt as well as stabilize the economy. We can rewrite the federal government budget constraint in terms of debt to GDP ratio $b_t^M = \frac{P_t^M B_t^M}{P_t Y_t}$,

$$b_t^M = \frac{(1+\rho P_t^M)}{P_{t-1}^M} \frac{P_{t-1}Y_{t-1}}{P_t Y_t} b_{t-1}^M - \frac{P_t Y_t \tau_t}{P_t Y_t} + \frac{P_t G_t}{P_t Y_t} + \frac{P_t Z_t}{P_t Y_t} \tag{11}$$

where all fiscal variables are now expressed as a fraction of GDP.

2.4 The Complete Model

The complete system of non-linear equations describing the equilibrium are given in Appendix 1. Log-linearizing the equilibrium conditions (37) - (50) around the deterministic steady state detailed in the Appendix, gives the following set of equations[6]:

$$\sigma \widehat{X}_t + \varphi \widehat{N}_t = \widehat{w}_t - \widehat{\mu}_t \text{ Labor Supply} \tag{12}$$

$$\widehat{X}_t = E_t \widehat{X}_{t+1} - \frac{1}{\sigma}\left(\widehat{R}_t - E_t \widehat{\pi}_{t+1} - E_t \widehat{q}_{t+1}\right) - \widehat{\xi}_t + E_t \widehat{\xi}_{t+1} \text{Euler equation} \tag{13}$$

$$\widehat{P}_t^M = \frac{\rho\beta}{\gamma\pi} \widehat{P}_{t+1}^M - \widehat{R}_t \text{ Bond Prices} \tag{14}$$

$$\widehat{y}_t = \widehat{N}_t = \widehat{c}_t + \frac{1}{1-g} \widetilde{g}_t \text{ Resource Constraint} \tag{15}$$

$$\widehat{X}_t = (1-\theta)^{-1}(\widehat{c}_t - \theta \widehat{c}_{t-1}) \text{ Habits-Adjusted Consumption} \tag{16}$$

$$\pi_t = \chi_f \beta E_t \pi_{t+1} + \chi_b \pi_{t-1} + \kappa_c(\widehat{w}_t + \frac{1}{1-\tau}\widetilde{\tau}_t) \text{ Hybrid NKPC} \tag{17}$$

$$\begin{aligned} \widetilde{b}_t^M &= \frac{1}{\beta}\widetilde{b}_{t-1}^M + \frac{b^M}{\beta}\left(\frac{\rho\beta}{\gamma\pi}\widehat{P}_t^M - \widehat{P}_{t-1}^M + y_{t-1} - y_t - \pi_t - a_t\right) \\ & \quad -\widetilde{\tau}_t + \widetilde{g}_t + \widetilde{z}_t \text{ Govt Budget Constraint} \end{aligned} \tag{18}$$

$$\widetilde{g}_t = \rho^g \widetilde{g}_{t-1} + \phi_g(1-\rho^g)\widetilde{b}_{t-1}^M + \varepsilon_t^g \text{ Govt Spending} \tag{19}$$

$$\widetilde{z}_t = \rho_{t-1}^z \widetilde{z}_{t-1} + \varepsilon_{z,t} \text{ Transfers Shock} \tag{20}$$

$$\widehat{q}_t = \rho^z \widehat{q}_{t-1} + \varepsilon_{z,t} \text{ Technology Shock} \tag{21}$$

$$\widehat{\mu}_t = \rho^\mu \widehat{\mu}_{t-1} + \varepsilon_t^\mu \text{ Cost-Push Shock} \tag{22}$$

$$\widehat{\xi}_t = \rho^\xi \widehat{\xi}_{t-1} + \varepsilon_t^\xi \text{ Preference Shock} \tag{23}$$

[6] The fiscal variables normalized with respect to GDP (i.e. $\widetilde{b}_t^M, \widetilde{\tau}_t$, $\widetilde{g}_t$ and $\widetilde{z}_t$) are defined as the linear deviations from their steady states, while other variables are expressed as the percentage deviations from their steady states. Prior to linearization, output, consumption and real wages are all rendered stationary by scaling by technology, A_t.

where $\chi_f \equiv \frac{\alpha}{\Phi}$, $\chi_b \equiv \frac{\zeta}{\Phi}$, $\kappa_c \equiv \frac{(1-\alpha)(1-\zeta)(1-\alpha\beta)}{\Phi}$, and $\Phi \equiv \alpha(1+\beta\zeta) + (1-\alpha)\zeta$. The model is then closed through the addition of one of the descriptions of policy considered in Section 3. While the associated microfouned objective function is shown in Appendix 3 to be,

$$\Gamma_0 = -\frac{1}{2}\overline{N}^{1+\varphi} E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{c} \frac{\sigma(1-\theta)}{1-\theta\beta} \frac{c}{y} \left(\widehat{X}_t + \widehat{\xi}_t\right)^2 + \varphi \left(\widehat{y}_t - \frac{\sigma}{\varphi}\widehat{\xi}_t\right)^2 \\ + \frac{\alpha\eta}{(1-\beta\alpha)(1-\alpha)} \left(\pi_t^2 + \frac{\zeta\alpha^{-1}}{(1-\zeta)} \left[\pi_t - \pi_{t-1}\right]^2\right) \end{array} \right\} + tip + O[2] \tag{24}$$

3 Policy Making

We consider two descriptions of policy making. Firstly, where both monetary and fiscal policy follow simple rules, before allowing the monetary and fiscal authorities to behave optimally but independently.

3.1 Rules-Based Policy

When considering policy described by simple rules, we assume government spending follows a simple exogenous autoregressive process,

$$\widetilde{g}_t = \rho^g \widetilde{g}_{t-1} + \varepsilon_t^g \tag{25}$$

and tax policy follows a simple rule,

$$\begin{aligned} \widetilde{\tau}_t &= \rho^{\tau,i}\widetilde{\tau}_{t-1} + (1-\rho^{\tau,i})\left(\delta_{\tau,i}\widetilde{b}_{t-1}^M + \delta_y \widehat{y}_t\right) + \varepsilon_t^\tau \\ \text{with } i &= AM/PF, AM/AF, PM/PF, PM/PF \end{aligned} \tag{26}$$

where the coefficient on debt switches between $\delta_{\tau,i} = 0$ in the active fiscal policy regime (AF), and $\delta_{\tau,i} > 0$ in the passive fiscal policy regime (PF), and we also allow the persistence of the tax rate, $\rho^{\tau,i}$ to vary across fiscal regimes. Moreover, we allow the parameters to vary depending on which monetary policy regime the fiscal policy regime is paired with. Therefore, a passive fiscal policy operating alongside an active monetary policy may have a different coefficient on debt $\delta_{\tau,AM/PF} > 0$ than when it is paired with a passive monetary policy, $\delta_{\tau,PM/PF} > 0$

When US monetary policy is described as a generalized Taylor rule, we specify this rule following An and Schorfheide (2007),

$$\begin{aligned} R_t &= \rho^{R,j} R_{t-1} + (1-\rho^{R,j})[\psi_{1,j}\widehat{\pi}_t + \psi_{2,j}(\Delta\widehat{y}_t + \widehat{q}_t)] + \varepsilon_t^R \\ \text{with } j &= AM/PF, AM/AF, PM/PF, PM/PF \end{aligned} \tag{27}$$

where the Fed adjusts interest rates in response to movements in inflation and deviations of output growth from trend. We allow the rule parameters $(\rho^{R,j}, \psi_{1,j}, \psi_{2,j},)$ to switch between active and passive monetary policy regimes. An active monetary policy regime (AM) is characterized by $\psi_{1,i} > 1$, while the passive monetary policy regime (PM) is characterized by $0 < \psi_{1,i} < 1$. Again these parameters may vary depending on which fiscal regime the monetary policy regime is paired with.

By considering both fiscal and monetary policy changes, we can distinguish four policy regimes under rules-based policy. They are AM/PF, AM/AF, PM/PF and PM/AF. Leeper(1991) shows that, in the absence of regime switching, the existence of a unique solution to the model depends on the nature of the assumed policy regime. A unique solution can be found under both the AM/PF and PM/AF combinations of regime - in the former monetary policy actively targets inflation, while fiscal policy adjusts taxes to stabilize debt, while under the latter combination the fiscal authority does not adjust taxes to stabilize debt and the monetary authority does not actively target inflation in order to facilitate the stabilization of debt. In contrast, no stationary solution and multiple equilibria are obtained under the AM/AF and PM/AF combinations of regimes, respectively. However, when regime switching is considered, the existence and uniqueness of a solution also depends on the transition probabilities of the potential regime changes. Specifically, we allow monetary and fiscal policy rule parameters to be switched independently of each other. The transition matrices for monetary policy and fiscal policy are specified as follows:

$$P^M = \begin{bmatrix} p_{11} & 1-p_{22} \\ 1-p_{11} & p_{22} \end{bmatrix}, \tag{28}$$

$$Q^F = \begin{bmatrix} q_{11} & 1-q_{22} \\ 1-q_{11} & q_{22} \end{bmatrix} \tag{29}$$

In addition, as noted above, we allow for variations in the parameter values of active and passive monetary and fiscal rules if this is supported by the data. Indeed, we find significant variations in the AM and PF regimes depending on which policy they are combined with.

In addition to incorporating monetary and fiscal policy changes, we also account for the 'good luck' factor that is normally modelled as a decrease in the volatility of shocks hitting the economy. Therefore, we allow for independent regime switching in the variance of economics shocks. These include technology ($\widehat{q}_t$), preference ($\widehat{\xi}_t$) and cost-push ($\widehat{\mu}_t$) shocks. We assume the following transition matrix for the shock processes:

$$H = \begin{bmatrix} h_{11} & 1-h_{22} \\ 1-h_{11} & h_{22} \end{bmatrix}, \tag{30}$$

We adopt the solution algorithm proposed by Farmer et al (2011) to solve the model with Markov-switching in policy rule parameters. This model can be recast in the following system

$$\Gamma_0(S_t^j = i)X_t = \Gamma_1(S_t^j = i)X_{t-1} + \Psi(S_t^j = i)Z_t + \Pi(S_t^j = i)\eta_t. \tag{31}$$

In contrast to the standard representation of a linear model with time-invariant policy rules, Γ_0, Γ_1, Ψ and Π in (31) depend on unobserved state variables, $S_t^j = i$, where $j = M, F$ for monetary and fiscal policy, and $i = P, A$ for passive and active regimes. When a solution exists, it can be rewritten as the following AR(1) process

$$X_t = \Phi_1(S_t^j = i)X_{t-1} + \Phi_2(S_t^j = i)Z_t, Z_t \sim NID(\mathbf{0}, \mathbf{\Sigma}). \tag{32}$$

It is important to note that the law of motion of state variables not only depends on the parameters of a particular regime, but also the probability of moving across regimes. The 'spillover' from one regime to another reflects the fact that economic agents are assumed to anticipate the Markov switching between different policy rules. Equation (32) can be linked to the system of measurement equations in Section 4 for estimation.

3.2 Optimal Policy

We now turn to describe our optimal policy specifications. In Chen et al (2013, 2014) we estimate monetary policy models of the US and Euro-area respectively and find that monetary policy is best described as being optimal, but time-consistent i.e. operating under discretion rather than commitment. This dominates both a rule-based description of monetary policy and the time-inconsistent Ramsey policy. In extending this analysis to fiscal policy there are several additional factors to contend with. Firstly, the monetary and fiscal authorities should be considered to be independent policy makers with potentially different policy objectives. This, in turn, implies that we must consider the strategic interactions between the two policy makers where we allow them to play a game where either can be considered to be the Stackelberg leader (making policy decisions anticipating the reaction of the other) or a Nash equilibrium where each policy maker takes the other's policies as given in formulating their own policies. Secondly, while Chen et al (2013, 2014) find strong evidence that monetary policy has been conducted optimally, albeit with switches in the degree of conservatism over time, it is not obvious that fiscal policy can be considered to have been similarly optimal. Therefore, while we permit monetary policy to be conducted optimally, with switches in the degree of conservatism, we allow fiscal policy to switch between rule-based and optimal time-consistent policy.

In implementing optimal policy we need to describe policy objectives. In doing so, estimation with micro-founded weights is problematic. Since the micro-founded weights are functions of structural parameters, they place very tight cross-equation restrictions on the model which are generally thought to be implausible. In particular, for standard estimates of the degree of price stickiness, the micro-founded weight attached to inflation can be over 100 times that attached to the output terms (see Woodford (2003, Chapter 6). Optimal policies which were based on such a strong anti-inflation objective would clearly be inconsistent with observed inflation volatility. Therefore, for estimation, we adopt a form of the objective function for each policy maker which is consistent with the representative agents' utility, (equation 24) but allow the weights within that objective function to be freely estimated. The resulting objective function for the monetary authority is given by

$$\Gamma_0^m = -\frac{1}{2}\overline{N}^{1+\varphi}E_0\sum_{t=0}^{\infty}\beta^t\left\{\begin{array}{c}\omega_\pi\left(\pi_t^2+\frac{\zeta\alpha^{-1}}{(1-\zeta)}\left[\pi_t-\pi_{t-1}\right]^2\right)\\+\omega_1\left(\widehat{X}_t+\widehat{\xi}_t\right)^2+\omega_2\left(\widehat{y}_t-\frac{\sigma}{\varphi}\widehat{\xi}_t\right)^2\\+\omega_R(\Delta R_t)^2\end{array}\right\}+tip+O[2] \tag{33}$$

where the weight on inflation, ω_π, is normalized to 1 in the 'conservative' regime and $\omega_\pi < 1$ in the less conservative regime and we have allowed for a possible desire to smooth the policy instrument.

Similarly, the objective function for the fiscal authority is given by,

$$\Gamma_0^f = -\frac{1}{2}\overline{N}^{1+\varphi}E_0\sum_{t=0}^{\infty}\beta^t\left\{\begin{array}{c}\omega_\pi^f\left(\pi_t^2+\frac{\zeta\alpha^{-1}}{(1-\zeta)}\left[\pi_t-\pi_{t-1}\right]^2\right)\\+\omega_1\left(\widehat{X}_t+\widehat{\xi}_t\right)^2+\omega_2\left(\widehat{y}_t-\frac{\sigma}{\varphi}\widehat{\xi}_t\right)^2\\+\omega_\tau(\Delta\tau_t)^2\end{array}\right\}+tip+O[2] \tag{34}$$

where the objective functions differ in the weight attached to inflation, and the presence of an instrument smoothing objective which is specific to that policy maker. In allowing for the weight of inflation to differ across the two policy makers we are essentially allowing the government to appoint a conservative central banker as in Rogoff (1985), as well as allowing the degree of monetary policy conservatism to vary over time. Given this interpretation, we consider the fiscal policy maker's objective function (excluding the instrument smoothing term) to be a measure of social welfare when evaluating and constructing various counterfactual exercises below.[7] Otherwise the objective functions are the same and have the same form as the micro-founded welfare measure. This has the advantage of facilitating comparison across policy makers, as well as with micro-founded optimal policy studies while still retaining sufficient flexibility to enable us to explain the data.

When considering the case where the fiscal authority may fluctuate between optimal and rules-based policy, we generalize the fiscal authority's objective function as follows

$$\Omega^f = (1 - \delta^p - \delta^a)\Gamma_0^f + \delta^p(\tau_t - \tau_t^p)^2 + \delta^a(\tau_t - \tau_t^a)^2 \tag{35}$$

where $\delta^p = 1$ ($\delta^a = 1$) indicates we are in a regime where fiscal policy follows a passive (active) rule, and these indicators are zero otherwise. The setting of the policy instrument consistent with the passive (active) rule is given by equation (26).Therefore, we allow for switches in preferences which imply the policy maker may conduct policy optimally, or switch to a policy objective which implies their only goal is to implement an active or passive fiscal rule.

As with the rules-based policy, we again assume that monetary and fiscal policy switch independently of each other. When we allow fiscal policy to switch across optimal fiscal (OF), PF and AF regimes, the transition matrix of fiscal policy is modified as follows:

$$Q^F = \begin{bmatrix} q_{11} & 1 - q_{22} - q_{23} & q_{13} \\ q_{12} & q_{22} & 1 - q_{13} - q_{33} \\ 1 - q_{11} - q_{12} & q_{23} & q_{33} \end{bmatrix} \tag{36}$$

To solve the optimal policy problem outlined above, we propose a new algorithm, described in Appendices 4 and 5, with two policy makers under different structures of strategic interaction (i.e. when one policy maker can be considered a Stackelberg leader in the policy game and when they move simultaneously as part of a Nash equilibrium). Our algorithm incorporates potential changes in policy makers' preferences over time. We consider different forms of strategic interaction by allowing either policy maker to be considered the Stackelberg leader in the game or the Nash equilibrium.[8] Our estimation indicated that the case of fiscal policy leadership is preferred by the data over the other forms of the game. Therefore, we focus on the case of the fiscal leadership and consider alternative models with variations in how fiscal policy switches. By comparing the marginal data density presented in Table 1, we conclude that the model that incorporates three fiscal policy regimes: OF, PF and AF provides the best fit to the data.

It should be noted that while we allow the monetary and fiscal policy makers to play strategic games with each other, we do not assume any strategic interactions with their future selves.

[7] We can have a degree of confidence in this interpretation in that the estimated conservative monetary policy objective function is not dissimilar to that which would have been optimally chosen for a central bank by a government with such social preferences.

[8] This extends the analysis of Fragetta and Kirsanova (2010) who consider strategic interactions between policy makers when estimating simple models for the US, UK and Sweden, as they do not allow for switches in policy regime.

Therefore, for example, the fiscal policy maker today does not try to tie the hands of her future self when conducting policy today. Instead the switches in the objective function can be considered to be part of the evolution of policy maker preferences as each type of policy maker cooperates with their future selves. Particularly in the case of fiscal policy, this kind of strategic interaction over time would be an interesting area for future research.

The next section will discuss our estimation results. However, before doing so it is important to note that all model parameters are identifiable. To demonstrate this, we used the Iskrev (2010) local identification test for our models based on a simple rule as well as optimal policy under discretion.

4 Estimation

Our empirical analysis uses six US data series on real output growth (ΔGDP_t), annualized domestic inflation (INF_t), the federal funds rate (FFR_t), the annualized debt-to-GDP ratio (B_t/GDP_t), government spending to GDP ratio (G_t/GDP_t) and federal tax revenues to GDP ratio (T_t/GDP_t) from 1955Q1 to 2008Q3. All data are seasonally adjusted and at quarterly frequencies. Output growth is the log difference of real GDP, multiplied by 100. Inflation is the log difference of GDP deflator, scaled by 400. The three fiscal variables - debt, government spending and taxes - are normalized with respect to GDP and multiplied by 100. GDP, the GDP deflator, government spending and total tax revenues are obtained from the Bureau of Economic Analysis. Government debt is taken from the Federal Reserve Bank of Dallas. Finally, the federal funds rate is taken from the FRED database. Appendix 6 describes the dataset in detail.

The data are linked to the law of motion of states solved under rules-based and optimal policy through a measurement equation specified as:

$$\begin{bmatrix} \Delta GDP_t \\ INF_t \\ FFR_t \\ G_t/GDP_t \\ T_t/GDP_t \\ B_t/GDP_t \end{bmatrix} = \begin{bmatrix} \gamma^Q + \Delta\widehat{y}_t + \widehat{q}_t \\ \pi^A + 4\widehat{\pi}_t \\ r^A + \pi^A + 4\gamma^Q + 4\widehat{R}_t \\ 100g + \widetilde{g}_t \\ 100\tau + \widetilde{\tau}_t \\ \frac{100}{4}b^M + \frac{1}{4}\widetilde{b}_t^M \end{bmatrix},$$

where parameters, γ^Q, π^A, r^A, g, τ and b^M represent the steady-state values of output growth, inflation, real interest rates, government spending to GDP ratio, tax rate and debt-to-GDP at quarterly basis.

For both rules-based and optimal policy estimation, we fix the steady-state values of fiscal variables and output growth to be consistent with the sample mean values over the period of 1955Q1 - 2008Q3. The government spending to GDP ratio (g) is 8%. The federal tax revenues to GDP ratio (τ) is 17.5%. The federal debt to annualized output ratio (b^M) is 31%. The quarterly output growth (γ^Q) is 0.46%. We further assume that the steady-state of real interest rate (r^A) is 1.8% given a 2% inflation target (π^A). The average real interest rate, r^A, is linked to the discount factor, β, such that $\beta = \left(1 + r^A/400\right)^{-1}$. In addition, we fixed the average maturity of outstanding government debt at 5 years (see Leeper and Zhou (2013), Table 1).

Due to the presence of Markov-switching parameters, the likelihood function is approximated using Kim(1994)'s filter, and then combined with the prior distribution to obtain the posterior distribution. A random walk Metropolis-Hastings algorithm is then used to generate 1,500,000 draws from the posterior distribution with the first 500,000 draws being discarded and every 50th draw from the remaining draws being saved.[9]

Finally, we compute the log marginal likelihood values for each model to provide a coherent framework to compare models with different types of monetary policies. We first implement the commonly used modified harmonic mean estimator of Geweke(1999) for this task. Bayes factors are calculated based on these values. In addition, we utilize the approach of Sims et al (2008) as a robustness check. The latter is designed for models with time-varying parameters, where the posterior density may be non-Gaussian. The log marginal likelihood values are presented in Table 1, where the optimal policy model with three fiscal policy regimes (i.e. OF, PF and AF) provides the best fit to the data, whereas the rules-based policy produces significantly worse fit to the data compared to all optimal policy models.[10]

4.1 Prior Distributions

Table 2 presents the priors and posterior estimates for the rules-based policy. The priors for most of the parameters are relatively loose and broadly consistent with the literature on the estimation of New Keynesian models. Following Smets and Wouters (2003), we choose the normal distribution for inverse of the Frisch labor supply elasticity, φ, and the inverse of the intertemporal elasticity of substitution, σ, with both priors having a mean of 2.5. Habits formation, indexation and the AR(1) parameters of the technology, cost-push, taste and transfer shocks and government spending process are assumed to follow a beta distribution with a mean of 0.5 and a standard deviation of 0.15. It is important to note that the above priors are common to both the rule-based and optimal policies. The exception is the Calvo parameter where the probability of no price change, α, is set so that the average length of the contract is around one year with a fairly tight prior around that value. Allowing a looser prior on this parameter tends to result in implausibly high estimates of the degree of price stickiness.

For the interest rate rule parameters, we set symmetric priors for the AR(1) parameter of the lagged interest rate and the parameter of output growth, whereas asymmetric and truncated priors are used for the parameter of inflation to ensure that $\psi_1 > 1$ in the active monetary policy regime and $0 < \psi_1 < 1$ in the passive regime. Similarly, for the tax rule, a symmetric prior is used for the AR(1) parameter of lagged tax rate, while the parameter of debt is restricted to be zero in the active fiscal regime and positive in the passive fiscal regime. Overall, the priors of the policy rule parameters imply four distinct fiscal and monetary policy regimes: PF/AM, AF/AM, PF/PM and AF/PM. In addition, variances of shocks are chosen to be highly dispersed inverted Gamma distributions to generate realistic volatilities for the endogenous variables. To capture the good luck factor, we specifically allow for the standard deviations of the three economic shocks

[9] Geweke(1992) convergence diagnostics are available upon request.

[10] Following Jeffreys(2007), Kass and Raftery(1995) argue that values of the Bayes Factor associated with two models lying between 0 and 3.2 constitutes evidence which is "not worth more than a bare mention", between 3.2 and 10 is “substantial” evidence, between 10 and 100 is “strong” evidence and above 100 is “decisive” evidence.

(i.e. the technology, preference and cost-push shocks) in the model to subject to regime switching between high and low volatility regimes. The priors for shock variances are set to be symmetric across regimes.

Table 3 reports the priors and posterior parameter estimates for optimal policy, the relative weights (i.e. ω_1, ω_2 and ω_R) attached to the output and interest rate smoothing terms on the monetary policy objective function are assumed to follow beta distributions. In addition, asymmetric prior is chosen for the weight attached to inflation stabilization term (ω_π), where ω_π is restricted to 1 in the conservative inflation targeting regime and a value lower than 1, a beta distribution, is used for the less conservative regime. For the fiscal policy objective function, we restrict the relative weights attached to the output terms to be the same as these on the monetary policy preferences, while we estimate the weight on the inflation stabilization term (ω_π^f) placed by the fiscal authority. We assume that ω_π^f follows a Gamma distribution with prior mean of 1. This implies that the fiscal authority has a consistent preference in terms of inflation targeting with the monetary policy authority in the conservative inflation targeting regime. Furthermore, instead of stabilizing interest rates, we assume that the fiscal authority is concerned with the variation of tax rates. Therefore, we add a tax rate smoothing term to the fiscal policy objective function.

4.2 Posterior Estimates

4.2.1 Rules-Based Policy

The posterior parameter estimates of the rules-based policy are reported in Table 2. Our estimates of the structural model papers are broadly in line with other studies: an intertemporal elasticity of substitution, σ, of 2.6; a measure of price stickiness, $\alpha = 0.78$, implying that price contracts typically last for around one year; a relatively modest degree of price indexation, $\zeta = 0.21$, a significant estimate of the degree of habits, $\theta = 0.61$ and an inverse Frisch labor supply elasticity of $\varphi = 2.4$.

Under the rules-based policy, we have four alternative policy permutations: AM/PF, AM/AF, PM/PF and PM/AF. As discussed above, in order to allow for maximum flexibility in describing the policy regimes, we initially allow for variations in the parameters of active and passive monetary and fiscal rules across the four policy regimes. In other words, the active monetary (AM) rule parameters in the AM/PF regime can differ from these in the AM/AF regimes. Indeed, we find significant variations in the AM and PF regimes depending on which policy they are combined with. However, the PM and AF regimes appeared to be similar regardless of which policy they were paired with. Therefore, we restrict PM and AF to be the same across their respective paired regimes. The resultant policy regimes imply that the passive monetary policy is inertial, $\rho^R = 0.85$ only falling slightly short of the Taylor principle $\psi_1 = 0.80$ with a significant coefficient on output, $\psi_2 = 0.35$. While an active monetary policy paired with a passive fiscal policy is both inertial $\rho^R = 0.85$ and very aggressive in targeting inflation, $\psi_1 = 2.90$ with a relatively modest response to output, $\psi_2 = 0.36$. When fiscal policy is active, then an associated active monetary policy is far less aggressive as interest rate inertia falls, $\rho^R = 0.53$ along with the response to inflation, $\psi_1 = 1.33$, while the response to output increases, $\psi_2 = 0.44$. Since this latter regime is inherently unstable, it would appear that the conflict between the monetary and fiscal authority results in a moderation in the conservatism of monetary policy even while that policy remains active. Similarly, the passive fiscal policy is far more inertial $\rho^{\tau} = 0.91$ and less responsive to debt, $\delta_{\tau} = 0.03$ when it is paired with an active monetary policy than when the passive fiscal policy is paired with a passive monetary policy where tax rate inertia falls, $\rho^{\tau} = 0.42$ and the response to debt rises, $\delta_{\tau} = 0.08$. These kinds of differences in estimation across regimes could reflect the nature of the conflict between monetary and fiscal policy. In the case of the AM/AF regime the policy is unstable and only rendered determinate because of spillovers from other policy permutations, so that the moderation in monetary policy would serve to mitigate the unstable debt dynamics caused by rising debt service costs under the active monetary policy. Similarly, a passive fiscal policy which raises distortionary taxes to stabilize debt is likely to fuel inflation and lead to rising debt service costs when monetary policy is active. This is less of a danger when monetary policy is passive, so that fiscal policy can be relatively more aggressive in responding to debt in the latter case. These results suggest that the stance of one policy maker is dependent on the policies of the other. We now turn to analyze this more formally by considering optimal policy where one policy maker takes into account the actions of the other.

4.2.2 Optimal Strategic Policy

Table 3 presents the posterior parameter estimates for the optimal policy where the monetary policy authority conducts an optimal policy taking the policies of the fiscal authority as given, and where we allow that monetary authority's objective function to switch in its degree of conservatism over time. At the same time the fiscal authority acts as a Stackelberg leader in the game with the monetary authority so that the fiscal authority conducts policy anticipating the response of the Fed. Fiscal policy may switch between this leadership role and conducting policy through simple active/passive rules. Under the optimal policy model, monetary policy is always assumed to be optimal, but time-consistent with the weight attached to inflation stabilization, ω_{π}, estimated to be 0.442 in the less conservative (LC) inflation targeting regime, relative to 1 in the more conservative (MC) regime. In addition, when the fiscal authority acts as a Stackelberg leader, she appears to have a lower degree of inflation conservatism than that of monetary policy authority, with ω_{π}^{f} estimated to be 0.335. The fiscal authority also implements simple active/passive tax rules. Both active and passive fiscal rules appear to be persistent and under the latter, the response of taxes to debt is $\delta_{\tau} = 0.05$. In total, six policy regimes are identified, three for fiscal policy (i.e. OF, PF and AF) and two for monetary policy (i.e. MC and LC).

The estimates of the deep model parameters remain similar to those found under rules-based policy with a modest rise in the intertemporal elasticity of substitution to $\sigma = 2.945$ and indexation, $\zeta = 0.21$, but a slight fall in the degree of habits, $\theta = 0.51$. The other significant difference is that the estimated degree of persistence of cost-push rises dramatically as we move from the rules-based estimation to the optimal policy estimation, rising from $\rho^{\mu} = 0.28$ to $\rho^{\mu} = 0.95$. This reflects the fact that costs push shocks generate the greatest trade-off for policy makers as they raise inflation and reduce output and thereby are relied on heavily by the estimation when policy is assumed to be optimal.

4.3 Regime Switching

4.3.1 Rules-Based Policy

Bianchi (2012) considers three regimes with a circular transition matrix such that one passes from the 'good' regime of AM/PF to the 'bad' regime of PM/AF before passing through the unstable regime of AM/AF to the AM/PF regime again. Therefore, we are in the 'good' regime in the early part of the sample before transition to the PM/AF regime in the 1960s. The appointment of Volcker in 1979 leads to monetary policy turning active while fiscal policy remains active, until the early 1990s when we arrive at the good regime (AM/PF). Thinking of the transitions as reflecting a game of chicken between the monetary and fiscal authorities this set-up prevents the policy combination of PM/PF, it also prevents the direct transition from AM/PF to AM/AF as would be the case when fiscal policy turned 'bad' while monetary policy still clung to active inflation targeting.

Bianchi and Ilut (2015) use a different transition scheme allowing transitions between good and bad regimes through AM/AF in either direction where a small prior is attached to remaining in the unstable regime. Here we are in the PM/AF regime from the late 1960s to the Volcker disinflation. There are two brief attempts at disinflation in 1969 and 1973, both fail. Only the Volcker disinflation sticks with a brief period of the unstable regime AM/AF and a slight monetary policy wobble in between. Afterwards it is the 'good' regime all the way from the early 1980s onwards.

In both cases, the narrative is one of a switch from a passive to an active monetary policy around the time of Volcker with fiscal policy following suit by switching from an active to passive regime sometime later (in the early 1990s in Bianchi (2012) and early 1980s in Bianchi and Ilut (2015)).

Our assumption on policy regime transitions are less restrictive, and thus give a more nuanced picture of transitions between regimes - see Figures 2 and 3. Figure 2 details the movements across fiscal policy regimes, where the first panel describes the probability of being in the passive fiscal policy regime and the second the active fiscal policy regime. The third panel gives the probability of being in the active monetary policy regime (with its complement being the passive monetary regime). In Figure 3, we plot the estimated probabilities of being in the 'good' regime of AM/PF, the 'bad' regime of PM/AF and the regimes which imply conflict between monetary and fiscal policy (PM/PF and AM/AF). We are in the 'good' regime right up until the late 1960s. Then fiscal policy turns active in 1969, and monetary policy turns passive shortly afterwards. There is a brief attempt at disinflation in 1973, but we essentially stay in the PM/AF regime until Volcker. Afterwards monetary policy stays active, and there are brief flirtations with passive fiscal policy around 1980 and 1987, although none stick until 1992. The aftermath of the bursting of the dot com bubble around 2001 is associated with a relaxation of monetary policy (PM) while fiscal policy remains passive. This is not a regime which is possible under either Bianchi (2012) or Bianchi and Ilut(2015). Therefore, under our rules-based estimation the 1960s are not nearly as bad as these papers find and the conventional policy assignment did not re-emerge until 1992 and even then has not been uninterrupted since then.

Although much of the analysis of monetary and fiscal interactions focuses on the active/passive regimes first characterized by Leeper (1991), our estimates suggest that regimes that are only determinate because of the expectations of returning to either the AM/PF or PM/AF regime actually describe observed policy configurations for much of our sample period. Figure 3 plots the likelihood of being in the 'good' regime (AM/PF), the 'bad' regime (PM/AF) and being in a state of conflict when policy is described by simple rules. While the 'bad' regime is almost entirely focused in the 1970s, conflict is recurrent from the 1970s onwards such that 63% of the period is spent in the good regime, only 13% in the bad and 24% in conflict.

Another way of looking at this is to consider monetary and fiscal policy in isolation. Monetary policy was only passive in the 1970s (with an attempted disinflation in 1973) and following the dot-com crash. Fiscal policy was only passive prior to the 1970s and after the 1990s, although with brief periods of passivity in the 1980s.

4.3.2 Optimal Strategic Policy

When we turn to the case where we allow monetary policy to be optimal and to fluctuate between more and less conservative (MC and LC, respectively), while fiscal policy can move between OF, PF and AF, the results are even richer - see Figures 4 and 5. Looking at monetary policy alone the periods of less conservative monetary policy capture all those identified as passive in the rule-based estimation (the 1970s and following the dot-com crash). However, there are additional periods where monetary policy are identified as less conservative. The late 1950s gave way to fluctuations in conservatism throughout the first half of the 1960s, which then turned less conservative from 1967 until 1982, with only a brief bout of conservatism in 1976/7. As in Chen et al (2013) the Volcker disinflation didn't really take hold until 1982. From 1982 onwards, the more conservative regime becomes the dominant regime, with monetary policy occasionally shifting back to the less conservative regime. These include the periods after stock market crash of 1987, a boom period from 1997-1999 when the Fed chose not to tighten interest rates possibly in response to the Russian sovereign debt default of 1998 and the associated collapse of Long Term Capital Management, and after the dot com crash when the historically low federal funds rate resulted in a negative real interest rate from 2003-2005. Therefore, the sense that the Volcker disinflation was a decisive shift to a more conservative monetary policy is less clear than under rules-based estimates.

While fiscal policy is clearer. It was either passive or optimal until the late 1960s, where it turns predominantly active. The instances of non-active fiscal policy are associated with specific policy events. For example, the Nixon tax reforms of 1970 appear as an example of a passive policy, which then turned 'optimal' as fiscal policy was loosened prior to the 1972 election Here the policy was optimal in the sense that reducing tax revenues as a share of GDP reduced the inflationary impact of distortionary taxation at a time when inflation was rising sharply. Similarly, the tax rebate of President Ford in 1975 appears as a passive fiscal policy as the debt to GDP ratio has fallen below the steady state value targeted in the passive rule. It only becomes passive/optimal for a sustained period in 1994, but loses that status around 2000 for a couple of years as rising tax revenues amount to too aggressive a stabilization of debt to constitute an optimal policy. Following Clinton, the Bush tax cuts then imply a return to an active fiscal policy which then is then estimated to turn passive as the pre-2007 boom generates rising tax revenues relative to GDP despite the tax cuts. This particular boom is also associated with a relaxation in US Fed policy. The benign picture of fiscal policy turning passive in support of a conservative monetary policy shortly after the Volcker disinflation does not appear under the optimal policy model estimation. Instead it is really only with the Clinton administration that fiscal policy turns optimal/passive for any length of time in the recent past, while neither monetary nor fiscal policy appear to have permanently shifted to conservative and passive/optimal regimes, respectively.

Looking at the permutations of monetary and fiscal policy together we find that fiscal policy was initially optimal in the late 1950s before being best described by a passive rule, while monetary policy fluctuated between more or less conservative policies throughout the period to the 1970s. In the 1970s monetary policy becomes less conservative (aside from a brief bout of conservatism in 1976), while fiscal policy initially wavers between passivity/optimality and activism in the late 1960s, early 1970s before becoming solidly active until 1994. Therefore the Volcker disinflation doesn't show up as an increase in conservatism until 1982 and this is not associated with a supportive fiscal policy until the Clinton administration in 1994. Even then there is a passive monetary policy post 1997 until 1999 and from 2002/3 where the latter occurred at the same time as an active fiscal policy.

When we turn to the optimal policy estimations for comparison with the rules-based estimation we consider the combination of less (more) conservative monetary policy and passive/optimal (active) fiscal policy to represent conflict; optimal or passive fiscal policy combined with more conservative monetary policy to capture 'good' policy; and, active fiscal policy and less conservative monetary policy to indicate 'bad' policy. With these definitions 36% of the sample is spent in 'good' policy regime, 27% in the 'bad' regime, whereas the largest share of time involves conflict between monetary and fiscal policy. We can ask, to what extent does eliminating conflict improve outcomes? Is it better to avoid conflict or avoid the 'bad' regime? This is an issue to which we now turn.

4.4 Credibility and Conflict

In order to facilitate an understanding of the implications of being in the different permutations of the policy regimes we consider the following impulse response. We assume that the various endogenous states of the model are set to their observed value at the start of the sample period. This implies that the debt to GDP ratio is well above average as is government spending, while taxes are relatively low. At the same time, output is above trend, while inflation and interest rates are below average. We then consider how the economy would have evolved under the various possible regimes without being subject to any further shocks, although assuming that the estimated regime switches could still take place. How is debt stabilized from this high initial state under each of the regimes? Figure 6 plots figures of debt, inflation, the output gap, taxes and interest rates. The first column considers the estimated passive fiscal rule combined with either the more or less time-consistent monetary policy. As would be expected the passive fiscal rule succeeds in stabilizing debt, although slowly (it takes five years for the debt to GDP ratio to begin falling) given the inertia in government consumption and tax rates. The rising tax rates fuel inflation however, and the combination of a conservative monetary policy and passive fiscal policy is actually very inflationary.

Column 2 of Figure 6 then considers the active fiscal policy regime combined with the more or less conservative monetary policy. Either of these regimes fails to begin to stabilize debt within 10 years and inflation rises as debt levels rise. Column 3 then allows fiscal policy to act as a Stackelberg leader, while monetary policy follows with more or less conservative objectives.

Surprisingly, this regime also fails to stabilize debt, although inflation is moderated by the fiscal authority moderating the increases in taxes relative to those implemented under the passive fiscal rule. We shall explore why the optimal fiscal policy regime fails to stabilize debt immediately below. The final column considers the outcomes when policy acts cooperatively according to either commitment (time-inconsistent policy) or discretion (time-consistent). Here we can see the differences between commitment and discretion highlighted by Leeper and Leith (2015) as the Ramsey/commitment policy adjusts policy to very slowly stabilize debt without generating inflation. In contrast, the time-consistent cooperative policy implies a significant endogenous inflationary bias whenever debt is above its steady-state value which results in a massive increase in inflation and the desire to reduce debt back to the steady state as quickly as possible.

We now turn to re-consider the surprising result that optimal fiscal policy failed to decisively stabilize debt. To do so we replicate the IRF of Figure 6 in Figure 7, but now assume that each regime is fully credible - that is economic agents no longer expect to transition from that regime to any of the other possible regimes. Turning immediately to column 3 we see that the optimal fiscal policy now successfully stabilizes debt. It was therefore spillovers from the other regimes that resulted in the optimal policy failing to do so when transition to these regimes was still a possibility. To see which regime mattered, we can return to column 1 and consider the passive fiscal policy. When this regime is fully credible, the tax adjustment is highly inflationary and the extent to which monetary policy moderates that inflation depends on how conservative the monetary authority's objectives are - towards the end of the IRF plot inflation is highest under the combination of less conservative monetary policy and passive fiscal policy. Moreover, this level of inflation is higher than when this policy combination was not fully credible. When we consider the active fiscal policy debt remains unstabilized within 10 years even when the regime is fully credible and the path for inflation is strongly contingent on the conservatism of monetary policy, although the average level of inflation is not much changed relative to the case where the same regime lacked credibility. It is therefore the less conservative monetary policy/passive fiscal policy regime that results in spillovers to the optimal fiscal policy regime which prevent it from stabilizing debt. The optimal policy for the fiscal authority takes account of the higher taxes and therefore inflation that would emerge if policy switched to the passive fiscal rule. These alternative policies raise inflation expectations and result in the optimal fiscal authority cutting taxes today to offset the inflationary impact of expectations of moving to the high tax passive fiscal policy regime. This moderation of taxes raises debt going into the next period, and further worsens the inflation generated by a switch to the passive regime so that the fiscal authority further moderates tax increases to offset this effect. When there is no expectation of moving to the passive regime, the optimal fiscal policy is far more effective in stabilizing debt without generating inflation.[11]

The final IRF we consider follows the same format - see Figure 8 - but assumes that we eliminate the possibility that policy will be in conflict. Therefore, in the case of fiscal accommodation fiscal policy adjusts such that it becomes passive whenever monetary policy is conservative and active whenever monetary policy is less conservative. And the converse where monetary policy

[11] We can see this more formally by decomposing the extent to which inflationary expectations are driven by expected switches to the various regimes. This makes clear that raised inflationary expectations under the optimal fiscal policy regime are being driven by the expectation of switching to the passive fiscal policy rule. These results are available upon request.

adjusts to accommodate switches in fiscal policy so that a less (more) conservative monetary policy supports an active (passive/optimal) fiscal policy regime. Removing conflictual regimes does tend to reduce inflation across all the remaining regimes, particularly when fiscal policy is active. This is also associated with a lower rate of taxation, particular when fiscal policy is optimal. However, it remains the case that fiscal leadership even when associated with a conservative monetary policy has failed to begin the reduce the debt to GDP ratio within 10 years as the policy maker is still fighting against the expectations of higher taxes/inflation should the regime switch to the active tax rule, even although such a rule will always be associated with a conservative monetary policy.

4.5 Shocks

In Figure 9 we assess the extent to which the major trends in debt and inflation are driven by shifts in policy regime relative to shocks. The third column plots the estimated shock processes from our preferred model. While columns one and two plot the outcomes for the debt to GDP ratio and inflation, respectively under three scenarios. Firstly, with all estimated shocks and regime probabilities that replicate the data. Secondly, without any realized shocks, but assuming the estimated regime switches still take place, and thirdly the same scenario but with only one of the possible realized shocks. From this exercise we can see that a large part of the early fall in the debt to GDP ratio is due to the below average level of transfers estimated to be in place prior to the 1970s. Without this relatively low level of transfers inflation would have risen in the 1960s rather than 1970s. Later in the sample, the stabilization of debt in the 1990s was achievable because of the offsetting reduction in government consumption at a time of rising transfers. As might be expected, cost push shocks are needed to explain the spikes in inflation in the 1970s, but contribute surprisingly little to explaining inflation dynamics otherwise. Since transfers were not treated as being observable in the estimation it is interesting to assess to what extent our estimated transfers series matches the data. We add the data series to the relevant plot in the third column and we can see that the estimated series is a very close approximation to the actual series. From this we can conclude that our model is not utilizing improbable unobserved shock processes to explain the major developments in the debt to GDP ratio and inflation.

5 Welfare and Counterfactuals

In Table 4 we report the unconditional volatilities of key variables as well as the implied welfare cost of shocks under various policy regimes. The measure of welfare is the estimated objective function for the fiscal authority (excluding the instrument smoothing term) which we take to be a measure of social welfare. As discussed above we feel this is a natural measure of social welfare rather than the estimated objective function for the monetary authority since we assume that the government employs a 'conservative' central banker as in Rogoff (1985) to optimize the outcomes under discretion. This implies that the optimized degree of inflation conservatism that would be chosen by the government is greater than the government's underlying preference for inflation stabilization. When we compute the optimal degree of inflation aversion for a delegated central bank we find that it is remarkably similar to that estimated for the monetary authority under the more conservative regime.[12] Throughout this table we assume that the regime remains in place indefinitely, but economic agents anticipate the probability of moving to the other policy regimes. In this sense we can think of these particular policy regimes has not possessing full credibility. The results are presented in order of welfare cost, from low to high. The ranking of policy regimes is similar across low and high shock volatility regimes. The preferred policy is a passive fiscal policy, combined with a less conservative optimal monetary policy. However, this is almost indistinguishable in welfare terms from a policy in which monetary policy accommodates fiscal policy regime switches (regime MA). In this latter regime fiscal policy fluctuates between passive/active rules and optimal policy according to the estimated transition probabilities. However, whenever it does so we assume that monetary policy switches too in an accommodating manner such that an active fiscal policy is always combined with a less conservative optimal monetary policy, and a passive/optimal fiscal policy is combined with a more conservative optimal monetary policy. In other words we assume that monetary policy adjusts to switches in fiscal policy in order to avoid conflicts between the two policies. The next two preferred policies are ones which combine a passive fiscal policy with a more conservative monetary policy, closely followed by a policy where fiscal policy adjusts to accommodate switches in monetary policy so that an active monetary policy is always supported by a passive fiscal policy and a passive monetary policy implies a move to an active fiscal policy. These four permutations are relatively close in welfare terms.

The remaining four policy combinations are far less successful. Surprisingly these less successful policies include the combination of optimal monetary policy with optimal fiscal policy where the latter acts as a Stackelberg leader anticipating the reaction of the former, and the combination of an optimal fiscal policy with a less conservative monetary policy is even worse than the conventional 'bad' policy combination of an active fiscal policy with a more or less conservative monetary policy across both high and low volatility regimes. Why does the optimal fiscal policy perform so badly? We can start to see why by considering the same analysis but assuming full credibility - such that the respective policy regimes are known to stay in place with certainty - see Table 5.

When policy is fully credible, the optimal policy regime is one of full cooperative commitment - the Ramsey policy - which by definition cannot be surpassed. However, in contrast to the case when the policy regimes were subject to transition, the most preferred policy regime outside of commitment is of optimal fiscal policy (OF), with the fiscal authority acting as a Stackelberg leader with respect to a conservative monetary policy follower who takes fiscal policy variables as given. If this is not possible, then a passive fiscal policy rule in combination with an optimal conservative monetary policy is next best. The next two regimes imply the same ranking of fiscal policy (optimal policy followed by a passive rule), but where the monetary policy is now less conservative. The worst possible policies then either involve an active fiscal policy or a cooperative, but time-consistent optimal monetary and fiscal policy. This latter policy suffers from the 'debt stabilization bias' discussed in Leeper and Leith (2015). The reason why the optimal fiscal policy regime only works under full credibility is due to the spillovers across regimes discussed above. When transitions are possible, the optimal fiscal policy regime implies that the fiscal authority anticipates the rise in taxes and therefore inflation which would take place

[12] A government with such preferences would appoint a slightly more conservative banker than we estimate with a coefficient on inflation equal to 1.19 rather than 1.

whenever the economy transitions to a passive fiscal policy regime. Expecting this, especially when monetary policy is conservative, the policy maker cuts taxes today to offset the inflationary effects of anticipated increase in taxation in the future. This leads to lower inflation today, but an accumulation of government debt and greater subsequent volatility in output and inflation.

We can see more clearly the relative performance of the alternative regimes by considering the counterfactual exercise which assumes different regimes were in place, but the economy was hit by the same realization of shocks we obtain from our estimation. Figure 10 runs such a counterfactual when fiscal policy is passive. The left hand column considers the case when the passive fiscal policy is combined with a less conservative monetary policy. Within each sub-plot we plot series for the data, the outcomes that would be observed if that passive policy had been in place, but economic agents had expected the regime to switch according to the estimated probabilities and, finally, the case where the passive fiscal policy/less conservative monetary policy regime was fully credible. The right hand column considers the same series, but for the combination of passive fiscal policy and a more conservative monetary policy. Additionally, they plot the outcomes that would have emerged had either monetary or fiscal policy accommodated switches in the dominant policy. In both columns the passive fiscal policy would have resulted in debt to GDP ratios being higher in the 1970s and 80s, but would have fallen by more than observed under the Clinton administration in the mid 1990s. Debt would have been marginally higher under the more conservative monetary policy (or under any of the accommodation regimes), although inflation would have been significantly moderated in the 1970s. Credibility does not appear to be a significant issue for this regime, implying that this regime is more likely to be a cause of spillovers to other regimes, than a recipient of spillovers from those regimes.

Figure 11 repeats the exercise for an active fiscal policy. Here we see that the combination of an active fiscal policy and less conservative fiscal policy would have implied a very similar path for debt and inflation for most of the sample period, although debt would not have fallen around 1995 without the switch to an optimal fiscal policy around that time. With a more conservative monetary policy in place the extremes of inflation in the 1970s would have been avoided, although at the cost of higher debt levels from that point on. Essentially, a higher conservatism in monetary policy results in less reliance on the kinds of stabilization through inflation, bond prices and real interest rates discussed in Leeper and Leith (2015). Again the presence or absence of credibility has a relatively small impact on the counterfactual outcomes when fiscal policy is passive.

We now turn to the optimal fiscal policy regime in Figure 12. In this case the issue of credibility matters hugely. Without credibility the spillovers from the passive fiscal policy regime result in the fiscal leader moderating tax increases to mitigate the rise in inflation caused by expectations of the rise in taxes that would be adopted following a switch to a passive fiscal policy rule. Over time this moderation in tax rises from the fiscal leader results in an accumulation of government debt well above that observed in the data. This accumulation of debt would have raised inflation in the latter half of the sample quite significantly even if monetary policy have been conducted by a conservative central banker. In contrast, had the optimal fiscal policy regime been credible then debt would not have fallen quite so far in the 1970s, especially when paired with a conservative monetary policy which significantly reduces the volatility of inflation. It is interesting to note that under the optimal policy the (distortionary) tax rate plays a duel role in both stabilizing debt and inflation and we observe relatively lower (higher) tax rates in periods when inflation is high (low).

The final set of counterfactuals (see Figure 13) contrast what our welfare analysis suggested were the 'best' regimes - namely, optimal or passive fiscal policies which were fully credible and paired with a conservative monetary policy and the monetary accommodation of fiscal policy switches when such credibility is not possible - with the cooperative policies under commitment and discretion. In the second column we plot the outcomes under the cooperative policies, alongside the data. Under the Ramsey/Commitment policy we obtain a dramatic stabilization of inflation in combination with an effective tax smoothing policy which results in substantial movements in government debt. The increase in the tax rate in the mid 1980s reflects the reversal in the persistent cost push shock from positive to negative around that time, imply a desirable rise in taxation to offset the habits externality. In contrast, under the time-consistent discretionary policy there is a temptation to reduce debt through inflation surprises (whether induced by monetary policy or distortionary tax rises) which gives rise to an endogenous inflationary bias problem - what Leith and Wren-Lewis(2013) label a 'debt stabilization bias'. This, in turn, implies that the policy maker wishes to raise taxes above the tax smoothing level to mitigate this problem by returning debt to its steady state. Therefore, the time-consistent policy implies a far more rapid stabilization of debt, where the subsequent increase in debt also reflects the reversal of the cost push shock which is pushing consumption above its first best level. Contrasting the outcomes under cooperative policy with those in the first column we can see that the credible regimes of fiscal leadership or a passive fiscal rule combined with a conservative monetary follower closely mimic the outcomes under commitment, particularly in the case of fiscal leadership. It is striking that allowing the fiscal and monetary authorities to enter into a strategic game with each other is clearly welfare improving relative to the case of full cooperation when that cooperative policy is constrained to be time-consistent. In contrast we only see a relatively modest improvement in outcomes when such credibility is not achievable.

6 Conclusions

This paper stresses that the evolution of inflation dynamics in the US cannot be understood without reference to the stance of fiscal policy. Our data preferred model allows monetary policy to be optimal, but with potential switches between more or less conservative inflation aversion, and fiscal policy to switch between a passive and active fiscal rule as well as a role as a time-consistent Stackelberg leader. This model offers a more nuanced description of the evolution of monetary and fiscal policy interactions than the rules-based model. The narrative that the switch in monetary policy at the time of the Volcker disinflation was associated with a similar switch in fiscal policy making from a regime where the fiscal authorities did not act to stabilize debt to one where they did, does not fit with our estimates. Instead, we find that the Volcker disinflation occurred around 1982, but wasn't supported by a debt stabilizing fiscal policy until 1995 and even then this policy has been subject to further revisions. Moreover, there are numerous switches between the various permutations of policy regime, with policies often being in conflict in the sense that monetary policy may be strongly anti-inflationary when fiscal policy is failing to act to stabilize the debt, or the opposite case where fiscal policy is stabilizing debt, but monetary policy is not actively targeting inflation. In addition, policy changes are still ongoing in that we have not consistently been in a policy mix where fiscal policy stabilizes debt and monetary policy targets inflation even in the recent period. The implicit assumption that allows fiscal policy to be safely ignored in monetary policy models does not appear to be supported by the data.

Counterfactuals suggest that eliminating conflicts between the fiscal and monetary authorities can be just as welfare improving as adopting a 'good' policy permutation when the underlying policy regimes remain subject to switches. In addition, a more appropriate policy mix could have avoided the excesses of high inflation in the 1970s, but that the cost of this would have been a higher debt level throughout this period. The ideal time-consistent policy regime would be where the fiscal authority acts as a Stackelberg leader and the monetary authority is a conservative follower. Such a regime can come close to mimicking the outcomes that would have been observed under a cooperative Ramsey policy. However, this is contingent on the policy being fully credible in the sense of being not expected to switch to an alternative policy configuration.

It is clear from our estimation that such continuity of regime does not appear to be a feature of observed monetary and fiscal policy interactions in the US. If this policy permutation was potentially subject to switches, it would be preferable to encourage monetary policy to accommodate switches in fiscal policy such that the monetary authority only pursued a conservative anti-inflation policy when the fiscal authorities implement a passive fiscal rule or act as a time-consistent Stackelberg leader.

References

[1] Amato, J. D. and T. Laubach (2004). Implications of habit formation for optimal monetary policy. Journal of Monetary Economics 51 (2), 305-32

[2] An, S. and F. Schorfheide (2007). Bayesian Analysis of DSGE Models. Econometric Reviews 26 (2-4), 113-172.

[3] Bianchi, F. (2012). Evolving Monetary/Fiscal Policy Mix in the United States. American Economic Review 102 (3), 167-172.

[4] Bianchi, F. and C. Ilut (2015). Monetary/Fiscal Policy Mix and Agents 'Beliefs, Duke University, working paper.

[5] Calvo, G. (1983). Staggered Prices in a Utility-Maximizing Framework. Journal of Monetary Economics 12(3), 383-398.

[6] Chen, X., T. Kirsanova, and C. Leith (2013). How Optimal is US Monetary Policy? University of Glasgow Working Paper.

[7] Chen, X., T. Kirsanova, and C. Leith (2013). An Empirical Assessment of Optimal Monetary Policy Delegation in the Euro Area. University of Glasgow Working Paper.

[8] Christiano, L., M. Eichenbaum, and G. Evans (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. Journal of Political Economy 113 (1), 1-45.

[9] Davig, T. (2004), "Regime-Switching Debt and Taxation", Journal of Monetary Economics, 51 837-859.

[10] Davig, T. and E. M. Leeper (2011),. "Monetary and Fiscal Policy Interactions and Fiscal Stimulus", European Economic Review, 55, 211-227.

[11] Eser, F., C. Leith, and S. Wren-Lewis (2009). When is monetary policy all we need? University of Oxford Working Paper No.430.

[12] Eusepi, S. and B. Preston (2010). The Maturity Structure of Debt, Monetary Policy and Expectations Stabilization, mimeograph, Federal Reserve Bank of New York.

[13] Farmer, R. E. A., D. F. Waggoner, and T. Zha (2011). Minimal State Variable Solutions to Markov-switching Rational Expectations Models. Journal of Economic Dynamics and Control 35(12), 2150-2166.

[14] Fragetta, M., and T. Kirsanova. (2010). Strategic Monetary and Fiscal Policy Interactions: An Empirical Investigation. European Economic Review 54 (7): 855-879.

[15] Gali, J. and M. Gertler (1999). Inflation Dynamics: A Structural Econometric Analysis. Journal of Monetary Economics 44, 195-222.

[16] Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J. O. Berger, J. M. Bernardo, A. P. Dawid, and A. F. M. Smith (Eds.), Proceedings of the Fourth Valencia International Meetings on Bayesian Statistics, 169-194. Oxford: Oxford University Press.

[17] Geweke, J. (1999). Using simulation methods for Bayesian econometric models: inference, development, and communication. Econometric Reviews 18 (1), 1-73.

[18] Iskrev, N. (2010). Local identification in DSGE models. Journal of Monetary Economics 52(7),189-202.

[19] Kim, C. J. (1994). Dynamic linear models with Markov-switching. Journal of Econometrics, 60, 1-22.

[20] Leeper, E.M. (1991). Equilibria under 'Active' and 'Passive' monetary and fiscal policies. Journal of Monetary Economics, 27(1),129–147.

[21] Leeper, E.M. and X. Zhou. (2013). Inflation's Role in Optimal Monetary-Fiscal Policy. NBER Working Paper 19686.

[22] Leeper, E.M. and C. Leith (2015). Inflation Through the Lens of the Fiscal Theory, forthcoming in Taylor, J. and H. Uhlig, eds, Handbook of Macroeconomics, 2, Elsevier.

[23] Leith, C. and J. Malley (2005). Estimated General Equilibrium Models for the Analysis of Monetary Policy in the US and Europe. European Economic Review 49(8), 2137-2159.

[24] Leith, C., I. Moldovan, and R. Rossi (2012). Optimal Monetary Policy in a New Keynesian Model with Habits in Consumption. Review of Economic Dynamics 15 (3), 416-435.

[25] Leith, C. and S. Wren-Lewis (2013). Fiscal Sustainability in a New Keynesian Model. Journal of Money, Credit and Banking 45(8),1477-1516

[26] Lubik, T. and F. Schorfheide (2005). A Bayesian Look at the New Open Economy Macroeconomics. In NBER Macroeconomics Annual 2005, 313-379.

[27] Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. The Quarterly Journal of Economics 100 (4), 1169-1189.

[28] Sims, C. A., D. F.Waggoner, and T. Zha (2008). Methods for inference in large multiple-equation Markov-switching models. Journal of Econometrics 146, 113-144.

[29] Sims, C. A. and T. Zha (2006). Were There Regime Switches in US Monetary Policy. American Economic Review 96 (1), 54-81.

[30] Smets, F. and R. Wouters (2003). An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. Journal of the European Economic Association 1 (1), 1123-1175.

[31] Yun, T. (1996). Nominal price rigidity, money supply endogeneity, and business cycles. Journal of Monetary Economics 37 (2-3), 345-370.

[32] Traum, N. and S-C S. Yang (2010), "Monetary and Fiscal Policy Interactions in the Post-War US", European Economic Review, vol. 55(1), pp 140-164.

[33] Woodford, M. (2001). Fiscal Requirements of Price Stability, Journal of Money, Credit, and Banking. 33, 669-728.

[34] Woodford, M. (2003). Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton, NJ.: Princeton University Press.

A System of Non-linear Equations

$$\mu_t N_t^{\varphi}(X_t)^{\sigma} = \frac{W_t}{A_t P_t} \equiv w_t \tag{37}$$

$$1 = \beta E_t \left[\left(\frac{X_{t+1}^k \xi_{t+1}}{X_t^k \xi_t} \right)^{-\sigma} \frac{A_t}{A_{t+1}} \frac{P_t}{P_{t+1}} \right] R_t, \tag{38}$$

$$P_t^M = \beta E_t \left[\left(\frac{X_{t+1}^k \xi_{t+1}}{X_t^k \xi_t} \right)^{-\sigma} \frac{A_t}{A_{t+1}} \frac{P_t}{P_{t+1}} (1 + \rho P_{t+1}^M) \right], \tag{39}$$

$$N_t = (\frac{Y_t}{A_t}) \int_0^1 (\frac{P(i)_t}{P_t})^{-\eta_t} di \tag{40}$$

$$P_t Y_t = P_t C_t + P_t G_t \tag{41}$$

$$\frac{P_t^f}{P_t} = \left(\frac{\eta}{\eta - 1} \right) \frac{E_t \sum_{s=0}^{\infty} (\alpha\beta)^s \left(X_{t+s}\xi_{t+s} \right)^{-\sigma} mc_{t+s} \left(\frac{P_{t+s}\pi^{-s}}{P_t} \right)^{\eta} \frac{Y_{t+s}}{A_{t+s}}}{E_t \sum_{s=0}^{\infty} (\alpha\beta)^s \left(X_{t+s}\xi_{t+s} \right)^{-\sigma} (1 - \tau_{t+s}) \left(\frac{P_{t+s}\pi^{-s}}{P_t} \right)^{\eta - 1} \frac{Y_{t+s}}{A_{t+s}}} \tag{42}$$

$$mc_t = \frac{W_t}{A_t P_t} \tag{43}$$

$$P_t^b = P_{t-1}^* \Pi_{t-1}, \tag{44}$$

$$\ln P_{t-1}^* = (1 - \zeta) \ln P_{t-1}^f + \zeta P_{t-1}^b. \tag{45}$$

$$(P_t)^{1-\eta} = \alpha \left(P_{t-1}\pi \right)^{1-\eta} + (1 - \alpha) \left(P_t^* \right)^{1-\eta}. \tag{46}$$

$$b_t^M = \frac{(1 + \rho P_t^M)}{P_{t-1}^M} \frac{P_{t-1} Y_{t-1}}{P_t Y_t} b_{t-1}^M - \frac{P_t Y_t \tau_t}{P_t Y_t} + \frac{P_t G_t}{P_t Y_t} + \frac{P_t Z_t}{P_t Y_t} \tag{47}$$

$$\ln A_t = \ln \gamma + \ln A_{t-1} + \ln q_t \tag{48}$$

$$\ln q_t = \rho_q \ln q_{t-1} + \varepsilon_{q,t} \tag{49}$$

$$\ln(\mu_t) = \rho_u \ln(\mu_{t-1}) + \varepsilon_t^{\mu} \tag{50}$$

with an associated equation describing the evolution of price dispersion, $\int_0^1 (\frac{P(i)_t}{P_t})^{-\eta_t} di$, which is not need to tie down the equilibrium upon log-linearization. The policy variables R_t, τ_t, Z_t and G_t then need to be defined.

The Deterministic Steady State

In order to render this model stationary we need to scale certain variables by the non-stationary level of technology, A_t such that $k_t = K_t/A_t$ where $K_t = \{Y_t, C_t, W_t/P_t\}$. Fiscal variables (i.e. $P_t^M B_t^M/P_t$, G_t and Z_t) are normalized with respect to Y_t. All other real variables are naturally stationary. Applying this scaling, the steady-state equilibrium conditions reduce to:

$$N^{\varphi} X^{\sigma} = w \tag{51}$$

$$1 = \beta \left(R\pi^{-1}\right)/\gamma = \beta r/\gamma \tag{52}$$

$$P^M = \frac{\beta}{\gamma\pi - \beta\rho} \tag{53}$$

$$y = N \tag{54}$$

$$y = \frac{c}{(1-g)} \tag{55}$$

$$X = c(1-\theta) \tag{56}$$

$$mc = w \tag{57}$$

$$\frac{\eta}{\eta - 1} = \frac{1-\tau}{mc} \tag{58}$$

$$b^M = \left(\frac{\beta}{1-\beta}\right) s \tag{59}$$

To determine the steady state value of labour, we substitute for X in terms of y and then, using the aggregate production function, we obtain the following expression,

$$y^{\sigma+\varphi}\left[(1-g)(1-\theta)\right]^{\sigma} = \frac{\eta-1}{\eta}(1-\tau), \tag{60}$$

where g is the steady state share of government spending in output. We shall contrast this with the labour allocation/output that would be chosen by a social planner to obtain a measure of the steady-state distortion inherent in this economy which features distortionary taxation, monopolistic competition and the habits externality.

B The Social Planner's Problem

In order to assess the scale of the steady-state inefficiencies caused by the monopolistic competition, tax and habits externalities it is helpful to contrast the decentralized equilibrium with that which would be attained under the social planner's allocation. The social planner ignores the nominal inertia and all other inefficiencies and chooses real allocations that maximize the representative consumer's utility subject to the aggregate resource constraint, the aggregate production function, and the law of motion for habits-adjusted consumption:

$$
\begin{aligned}
&\max_{\{X_t^*,C_t^*,N_t^*\}} E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{X_t^{*1-\sigma}\xi_t^{-\sigma}}{1-\sigma} + \chi \frac{(G_t^*/A_t)^{1-\sigma}(\xi_t)^{-\sigma}}{1-\sigma} - \frac{N_t^{*1+\varphi}\xi_t^{-\sigma}}{1+\varphi} \right) \\
s.t.\ Y_t^* &= C_t^* + G_t^* \\
Y_t^* &= A_t N_t^* \\
X_t^* &= C_t^*/A_t - \theta C_{t-1}^*/A_{t-1}
\end{aligned}
$$

The optimal choice implies the following relationship between the marginal rate of substitution between labor and habit-adjusted consumption and the intertemporal marginal rate of substitution in habit-adjusted consumption

$$
\frac{(N_t^*)^{\varphi}}{(X_t^*)^{-\sigma}} = \left[1 - \theta\beta E_t \left(\frac{X_{t+1}^* \xi_{t+1}}{X_t^* \xi_t} \right)^{-\sigma} \right].
$$

The steady state equivalent of this expression can be written as,

$$
(N^*)^{\varphi+\sigma} \left[(1 - \frac{G^*}{Y^*})\,(1-\theta) \right]^{\sigma} = (1-\theta\beta)\,. \tag{61}
$$

where the optimal share of government consumption in output is given by,

$$
\frac{G_t^*}{Y_t^*} = \chi^{\frac{1}{\sigma}} (\frac{Y_t^*}{A_t})^{-\frac{\sigma+\varphi}{\sigma}}
$$

In steady state these can be combined to give the optimal share of government consumption in output,

$$
\frac{G^*}{Y^*} = (1 + (1-\theta)^{-1} \chi^{-\frac{1}{\sigma}} (1-\theta\beta)^{\frac{1}{\sigma}})^{-1}
$$

which can then used to get the steady state level of output under the social planner's allocation. [alternatively we can use the data to infer the weight on government spending in the utility function].

If we contrast this with the allocation achieved in the steady-state of our decentralized equilibrium (60), assuming that the steady state share of government consumption to GDP is the same, we can see that the two will be identical whenever the following relationship between the markup, the tax rate and the degree of habits holds,

$$
\frac{\eta}{\eta-1} = \frac{1-\tau}{1-\theta\beta} \tag{62}
$$

Notice that in the absence of habits this condition could only be supported by a negative tax rate. However, for the data given level of taxation and the estimated degree of habits this condition will define our steady-state markup, enabling us to adopt an efficient steady-state and thereby avoiding a steady-state inflationary bias problem when describing optimal policy.

C Derivation of Welfare

Individual utility in period t is

$$\frac{X_t^{1-\sigma}\xi_t^{-\sigma}}{1-\sigma}+\chi\frac{(G_t/A_t)^{1-\sigma}(\xi_t)^{-\sigma}}{1-\sigma}-\frac{N_t^{1+\varphi}\xi_t^{-\sigma}}{1+\varphi}$$

where $X_t = C_t - \theta C_{t-1}$ is the habit-adjusted aggregate consumption. Before considering the elements of the utility function, we need to note the following general result relating to second order approximations

$$\frac{Y_t - Y}{Y_t} = \widehat{Y}_t + \frac{1}{2}\widehat{Y}_t^2 + O[2]$$

where $\widehat{Y}_t = \ln\left(\frac{Y_t}{Y}\right)$ and $O[2]$ represents terms that are of order higher than 2 in the bound on the amplitude of the relevant shocks. This will be used in various places in the derivation of welfare. Now consider the second order approximation to the first term,

$$\frac{X_t^{1-\sigma}\xi_t^{-\sigma}}{1-\sigma} = \overline{X}^{1-\sigma}\left(\frac{X_t-\overline{X}}{\overline{X}}\right) - \frac{\sigma}{2}\overline{X}^{1-\sigma}\left(\frac{X_t-\overline{X}}{\overline{X}}\right)^2 - \sigma\overline{X}^{1-\sigma}\left(\frac{X_t-\overline{X}}{\overline{X}}\right)(\xi_t - 1) + tip + O[2]$$

where *tip* represents 'terms independent of policy'. Using the results above this can be rewritten in terms of hatted variables

$$\frac{X_t^{1-\sigma}\xi_t^{-\sigma}}{1-\sigma} = \overline{X}^{1-\sigma}\left\{\widehat{X}_t + \frac{1}{2}(1-\sigma)\widehat{X}_t^2 - \sigma\widehat{X}_t\widehat{\xi}_t\right\} + tip + O[2].$$

In pure consumption terms, the value of X_t can be approximated to second order by:

$$\widehat{X}_t = \frac{1}{1-\theta}\left(\widehat{c}_t + \frac{1}{2}\widehat{c}_t^2\right) - \frac{\theta}{1-\theta}\left(\widehat{c}_{t-1} + \frac{1}{2}\widehat{c}_{t-1}^2\right) - \frac{1}{2}\widehat{X}_t^2 + O[2]$$

and to a first order,

$$\widehat{X}_t = \frac{1}{1-\theta}\widehat{c}_t - \frac{\theta}{1-\theta}\widehat{c}_{t-1} + O[1]$$

which implies

$$\widehat{X}_t^2 = \frac{1}{(1-\theta)^2}\left(\widehat{c}_t - \theta\widehat{c}_{t-1}\right)^2 + O[2]$$

Therefore,

$$\frac{X_t^{1-\sigma}\xi_t^{-\sigma}}{1-\sigma} = \overline{X}^{1-\sigma}\left\{\frac{1}{1-\theta}\left(\widehat{c}_t + \frac{1}{2}\widehat{c}_t^2\right) - \frac{\theta}{1-\theta}\left(\widehat{c}_{t-1} + \frac{1}{2}\widehat{c}_{t-1}^2\right) + \frac{1}{2}(-\sigma)\widehat{X}_t^2 - \sigma\widehat{X}_t\widehat{\xi}_t\right\} + tip + O[2]$$

Summing over the future,

$$\sum_{t=0}^{\infty}\beta^t\frac{X_t^{1-\sigma}\xi_t^{-\sigma}}{1-\sigma} = \overline{X}^{1-\sigma}\sum_{t=0}^{\infty}\beta^t\left\{\frac{1-\theta\beta}{1-\theta}\left(\widehat{c}_t + \frac{1}{2}\widehat{c}_t^2\right) - \frac{1}{2}\sigma\widehat{X}_t^2 - \sigma\widehat{X}_t\widehat{\xi}_t\right\} + tip + O[2].$$

Similarly for the term in government spending,

$$\chi\frac{g_t^{1-\sigma}\xi_t^{-\sigma}}{1-\sigma} = \chi\overline{g}^{1-\sigma}\{\widehat{g}_t + \frac{1}{2}(1-\sigma)\widehat{g}_t^2 - \sigma\widehat{g}_t\widehat{\xi}_t\} + tip + O[2]$$

While the term in labour supply can be written as

$$\frac{N_t^{1+\varphi}\xi_t^{-\sigma}}{1+\varphi} = \overline{N}^{1+\varphi}\left\{\widehat{N}_t + \frac{1}{2}(1+\varphi)\widehat{N}_t^2 - \sigma\widehat{N}_t\widehat{\xi}_t\right\} + tip + O[2]$$

Now we need to relate the labour input to output and a measure of price dispersion. Aggregating the individual firms' demand for labour yields,

$$N_t = (\frac{Y_t}{A_t})\int_0^1 (\frac{P(i)_t}{P_t})^{-\eta_t} di$$

It can be shown (see Woodford, 2003, Chapter 6) that

$$\begin{aligned}\widehat{N}_t &= \widehat{y}_t + \ln[\int_0^1 (\frac{P(i)_t}{P_t})^{-\eta_t} di] \\ &= \widehat{y}_t + \frac{\eta}{2} var_i\{p(i)_t\} + O[2]\end{aligned}$$

which implies

$$\widehat{N}_t^2 = \widehat{y}_t^2$$

so we can write

$$\frac{N_t^{1+\varphi}}{1+\varphi} = \overline{N}^{1+\varphi}\left\{\widehat{y}_t + \frac{1}{2}(1+\varphi)\widehat{y}_t^2 - \sigma\widehat{y}_t\widehat{\xi}_t + \frac{\eta}{2} var_i\{p_t(i)\}\right\} + tip + O[2]$$

Welfare is then given by

$$\begin{aligned}\Gamma_0 &= \overline{X}^{1-\sigma} E_0 \sum_{t=0}^{\infty}\beta^t \left\{\frac{1-\theta\beta}{1-\theta}\left(\widehat{c}_t + \frac{1}{2}\widehat{c}_t^2\right) - \frac{1}{2}\sigma\hat{X}_t^2 - \sigma\widehat{X}_t\widehat{\xi}_t\right\} \\ &\quad + \chi\overline{g}^{1-\sigma} E_0 \sum_{t=0}^{\infty}\beta^t\{\widehat{g}_t + \frac{1}{2}(1-\sigma)\widehat{g}_t^2 - \sigma\widehat{g}_t\widehat{\xi}_t\} \\ &\quad - \overline{N}^{1+\varphi} E_0 \sum_{t=0}^{\infty}\beta^t\left\{\widehat{y}_t + \frac{1}{2}(1+\varphi)\widehat{y}_t^2 - \sigma\widehat{y}_t\widehat{\xi}_t + \frac{\eta}{2} var_i\{p_t(i)\}\right\} \\ &\quad + tip + O[2]\end{aligned}$$

From the steady-state of our model, and its comparison with the social planner's allocation we know that $\overline{X}^{1-\sigma}(1-\theta\beta) = (1-\theta)\frac{c}{y}\overline{N}^{1+\varphi}$. Similarly, assuming the same share of government spending in GDP across the social planner's and decentralized equilibrium, we also know that, $\chi\overline{g}^{1-\sigma} = \frac{g}{y}\overline{N}^{1+\varphi}$ Using the fact that,

$$\frac{c}{y}\widehat{c}_t = \widehat{y}_t - (1-\frac{c}{y})\widehat{g}_t - \frac{1}{2}\frac{c}{y}\widehat{c}_t^2 - \frac{1}{2}(1-\frac{c}{y})\widehat{g}_t^2 + \frac{1}{2}\widehat{y}_t^2 + O[2]$$

we can collect the levels terms and write the sum of discounted utilities as:

$$\Gamma_0 = -\frac{1}{2}\overline{N}^{1+\varphi}E_0\sum_{t=0}^{\infty}\beta^t\left\{\begin{array}{c}\frac{\sigma(1-\theta)}{1-\theta\beta}\frac{c}{y}\left(\widehat{X}_t+\widehat{\xi}_t\right)^2+\sigma\frac{g}{y}\left(\widehat{g}_t+\widehat{\xi}_t\right)^2\\ +(\varphi)\left(\widehat{y}_t-\frac{\sigma}{\varphi}\widehat{\xi}_t\right)^2\\ +\eta var_i\{p_t(i)\}\end{array}\right\}+tip+O[2] \tag{63}$$

Using the result from Eser et al (2009) that

$$\sum_{t=0}^{\infty}\beta^t var_i[p_t(i)] = \frac{\alpha}{(1-\beta\alpha)(1-\alpha)}\sum_{t=0}^{\infty}\beta^t\left(\pi_t^2+\frac{\zeta\alpha^{-1}}{(1-\zeta)}\left[\pi_t-\pi_{t-1}\right]^2\right)+O[2].$$

we can write the discounted sum of utility as,

$$\Gamma_0 = -\frac{1}{2}\overline{N}^{1+\varphi}E_0\sum_{t=0}^{\infty}\beta^t\left\{\begin{array}{c}\frac{\sigma(1-\theta)}{1-\theta\beta}\frac{c}{y}\left(\widehat{X}_t+\widehat{\xi}_t\right)^2+(\varphi)\left(\widehat{y}_t-\frac{\sigma}{\varphi}\widehat{\xi}_t\right)^2\\ +\frac{\alpha\eta}{(1-\beta\alpha)(1-\alpha)}\left(\pi_t^2+\frac{\zeta\alpha^{-1}}{(1-\zeta)}\left[\pi_t-\pi_{t-1}\right]^2\right)\end{array}\right\}+tip+O[2] \tag{64}$$

where we have put the terms in public consumption into tip since they are treated as an exogenous process and therefore independent of policy.

D Leadership Equilibria under Discretion with Markov Switching in Objectives

This section demonstrates how to solve non-cooperative dynamic games in the Markov jump-linear quadratic systems. Consider an economy with two policy makers: a leader (L) and a follower (F).

$$X_{t+1} = A_{11j_{t+1}}X_t + A_{12j_{t+1}}x_t + B_{11j_{t+1}}u_t^L + B_{12j_{t+1}}u_t^F + C_{j_{t+1}}\varepsilon_{t+1}, \tag{65}$$

$$E_tH_{j_{t+1}}x_{t+1} = A_{21j_t}X_t + A_{22j_t}x_t + B_{21j_t}u_t^L + B_{22j_t}u_t^F, \tag{66}$$

where $\boldsymbol{X}_t$ is a n_1 vector of predetermined variables; $\boldsymbol{x}_t$ is a n_2 vector of forward-looking variables; $\boldsymbol{u}_t = \left(u_t^{L\prime}, u_t^{F\prime}\right)'$ are the control variables, and $\boldsymbol{\varepsilon}_t$ contains a vector of zero mean *i.i.d.* shocks. Without loss of generality, the shocks are normalized so that the covariance matrix of $\boldsymbol{\varepsilon}_t$ is the identity matrix, $\boldsymbol{I}$. Therefore, the covariance matrix of the shocks to $\boldsymbol{X}_{t+1}$ is $C_{j_t}'C_{j_t}$.

The matrices A_{11j_t}, A_{21j_t}, A_{12j_t}, A_{22j_t}, H_{j_t}, B_{11j_t}, B_{12j_t}, B_{21j_t}, and B_{22j_t} can each take n different values, corresponding to the n modes $j_t = 1, 2, ...n$ in period t.The modes j_t follow a Markov process with constant transition probabilities:

$$P_{jk} = Pr\left\{j_{t+1} = k|j_t = j\right\}, \;\; j, k = 1, 2, ..., n.$$

Furthermore, P denotes the $n \times n$ transition matrix $[P_{jk}]$ and the $1 \times n$ vector $p = (p_{1t}, ..., p_{nt})$,where $p_{jt} = Pr\left\{j_t = j\right\}$, $j_t = 1, 2, ...n$ denotes the probability distribution of the modes in period t,

$$p_{t+1} = p_t P.$$

Finally, the $1 \times n$ vector $\overline{p}$ denotes the unique stationary distribution of the modes,

$$\overline{p} = \overline{p}P.$$

We assume that the intertemporal loss functions of the two policy makers are defined by the quadratic loss function

$$\mathbb{E}_t \sum_{\tau=0}^{\infty} \frac{1}{2} \beta^{\tau} \boldsymbol{L}^u_{j_{t+\tau}},$$

where $\boldsymbol{L}^u_{j_t}$ is the period loss with $u = F$ for the follower and $u = L$ for the leader, respectively. As with the structure parameter matrices in (65) and (66), $\boldsymbol{L}^u_{j_t}$ can also take different value corresponding to the n modes in period t. The period loss satisfies

$$\boldsymbol{L}^u_{j_t} = \mathbf{Y}^{u\prime}_t \boldsymbol{\Lambda}^u_{j_t} \mathbf{Y}^u_t,$$

where $\boldsymbol{\Lambda}^u_{j_t}$ is a symmetric and positive semi-definite weight matrix. $\mathbf{Y}^u_t$ are n_Y vectors of target variables for the follower and leader.

$$\mathbf{Y}^u_t = D^u \begin{bmatrix} X_t \\ x_t \\ u^L_t \\ u^F_t \end{bmatrix}.$$

It follows that the period loss function can be rewritten as

$$\boldsymbol{L}^u_{j_t} = \begin{bmatrix} X_t \\ x_t \\ u^L_t \\ u^F_t \end{bmatrix}' \mathbf{W}^u_{j_t} \begin{bmatrix} X_t \\ x_t \\ u^L_t \\ u^F_t \end{bmatrix}, \tag{67}$$

where $\boldsymbol{W}^u_{j_t} = D^{u\prime} \boldsymbol{\Lambda}^u_{j_t} D^u$ is symmetric and positive semidefinite, and

$$\mathbf{W}^u_{j_t} = \begin{bmatrix} Q^u_{11j_t} & Q^u_{12j_t} & P^u_{11j_t} & P^u_{12j_t} \\ Q^u_{21j_t} & Q^u_{22j_t} & P^u_{21j_t} & P^u_{22j_t} \\ P^{u\prime}_{11j_t} & P^{u\prime}_{21j_t} & R^u_{11j_t} & R^u_{12j_t} \\ P^{u\prime}_{12j_t} & P^{u\prime}_{22j_t} & R^{u\prime}_{12j_t} & R^u_{22j_t} \end{bmatrix}$$

is partitioned with X_t, x_t, $u_t^L$and u_t^F.

The follower and leader decide their policy u_t^F and u_t^L in period t to minimize their intertemporal loss functions defined in (67) under discretion subject to (65), (66), $\boldsymbol{X}_t$ and j_t given. The follower also observes the current decision u_t^L of the leader. Furthermore, two policy makers anticipate that they will reoptimize in period $t+1$. Reoptimization will result in the two instruments and the forward-looking variables in period $t+1$ being functions of the predetermined variables and the mode in period $t+1$ according to

$$\begin{aligned} u_{t+1}^L &= -F_{j_{t+1}}^L X_{t+1}, & (68)\\ u_{t+1}^F &= -G_{j_{t+1}}^F X_{t+1} - D_{j_{t+1}}^F u_{t+1}^L, & (69)\\ x_{t+1} &= -N_{j_{t+1}} X_{t+1}, & (70) \end{aligned}$$

where $j_{t+1} = 1, ..., n$ are the n modes at period $t+1$. The dynamics of the predetermined variables will follow

$$X_{t+1} = M_{j_t k_{t+1}} X_t + C_{j_{t+1}} \varepsilon_{t+1}, \tag{71}$$

where

$$M_{j_t k_{t+1}} = A_{11 j_{t+1}} - A_{12 j_{t+1}} N_{j_t} - B_{11 j_{t+1}} F_{j_t}^L - B_{12 j_{t+1}} G_{j_t}^F + B_{12 j_{t+1}} D_{j_t}^F F_{j_t}^L,$$

First, by (70) and (65) we have,

$$\begin{aligned} E_t H_{j_{t+1}} x_{t+1} &= -E_t H_{j_{t+1}} N_{j_{t+1}} X_{t+1} \\ &= -E_t H_{j_{t+1}} N_{j_{t+1}} \left(A_{11 j_{t+1}} X_t + A_{12 j_{t+1}} x_t + B_{11 j_{t+1}} u_t^L + B_{12 j_{t+1}} u_t^F \right) \end{aligned}$$

(where $E_t H_{j_{t+1}} N_{j_{t+1}} = \sum_{k=1}^{n} P_{j_t k_{t+1}} H_{k_{t+1}} N_{k_{t+1}}$, conditional on $j_t = 1, 2, ...n$ at the period t.) Combining this with (66) gives

$$\begin{aligned} & E_t H_{j_{t+1}} N_{j_{t+1}} \left(A_{11 j_{t+1}} X_t + A_{12 j_{t+1}} x_t + B_{11 j_{t+1}} u_t^L + B_{12 j_{t+1}} u_t^F \right) \\ = \; & A_{21 j_t} X_t + A_{22 j_t} x_t + B_{21 j_t} u_t^L + B_{22 j_t} u_t^F. \end{aligned}$$

Solving for x_t we obtain

$$x_t = -J_{j_t} X_t - K_{j_t}^L u_t^L - K_{j_t}^F u_t^F, \tag{72}$$

where

$$\begin{aligned} J_{j_t} &= \left(A_{22j} + \sum_{k=1}^{n} P_{j_t k_{t+1}} H_{k_{t+1}} N_{k_{t+1}} A_{12k} \right)^{-1} \left(A_{21j} + \sum_{k=1}^{n} P_{j_t k_{t+1}} H_{k_{t+1}} N_{k_{t+1}} A_{11k} \right), \\ K_{j_t}^L &= \left(A_{22j} + \sum_{k=1}^{n} P_{j_t k_{t+1}} H_{k_{t+1}} N_{k_{t+1}} A_{12k} \right)^{-1} \left(B_{21j} + \sum_{k=1}^{n} P_{j_t k_{t+1}} H_{k_{t+1}} N_{k_{t+1}} B_{11k} \right), \\ K_{j_t}^F &= \left(A_{22j} + \sum_{k=1}^{n} P_{j_t k_{t+1}} H_{k_{t+1}} N_{k_{t+1}} A_{12k} \right)^{-1} \left(B_{22j} + \sum_{k=1}^{n} P_{j_t k_{t+1}} H_{k_{t+1}} N_{k_{t+1}} B_{12k} \right). \end{aligned}$$

We assume that $A_{22j_t} + \sum_{k=1}^{n} P_{j_t k_{t+1}} H_{k_{t+1}} N_{k_{t+1}} A_{12k}$ is invertible.

Second, substituting x_t from (65) using (72) gives

$$X_{t+1} = \widetilde{A}_{j_t k_{t+1}} X_t + \widetilde{B}^L_{j_t k_{t+1}} u^L_t + \widetilde{B}^F_{j_t k_{t+1}} u^F_t + C_{j_{t+1}} \varepsilon_{t+1}, \tag{73}$$

where

$$\begin{aligned}
\widetilde{A}_{j_t k_{t+1}} &= A_{11k_{t+1}} - A_{12k_{t+1}} J_{j_t}, \\
\widetilde{B}^L_{j_t k_{t+1}} &= B_{11k_{t+1}} - A_{12k_{t+1}} K^L_{j_t}, \\
\widetilde{B}^F_{j_t k_{t+1}} &= B_{12k_{t+1}} - A_{12k_{t+1}} K^F_{j_t}.
\end{aligned}$$

D.1 Policy of the Follower

Using (72) in the follower's loss function (67) gives

$$\begin{aligned}
\boldsymbol{L}^F_{j_t} &= \begin{bmatrix} X_t \\ x_t \\ u^L_t \\ u^F_t \end{bmatrix}' \begin{bmatrix} Q^F_{11j_t} & Q^F_{12j_t} & P^F_{11j_t} & P^F_{12j_t} \\ Q^F_{21j_t} & Q^F_{22j_t} & P^F_{21j_t} & P^F_{22j_t} \\ P^{F\prime}_{11j_t} & P^{F\prime}_{21j_t} & R^F_{11j_t} & R^F_{12j_t} \\ P^{F\prime}_{12j_t} & P^{F\prime}_{22j_t} & R^{F\prime}_{12j_t} & R^F_{22j_t} \end{bmatrix} \begin{bmatrix} X_t \\ x_t \\ u^L_t \\ u^F_t \end{bmatrix} \\
&= \begin{bmatrix} X_t \\ u^L_t \\ u^F_t \end{bmatrix}' \begin{bmatrix} \widetilde{Q}^F_{j_t} & \widetilde{P}^F_{1j_t} & \widetilde{P}^F_{2j_t} \\ \widetilde{P}^{F\prime}_{1j_t} & \widetilde{R}^F_{11j_t} & \widetilde{R}^F_{12j_t} \\ \widetilde{P}^{F\prime}_{2j_t} & \widetilde{R}^{F\prime}_{12j_t} & \widetilde{R}^F_{22j_t} \end{bmatrix} \begin{bmatrix} X_t \\ u^L_t \\ u^F_t \end{bmatrix}
\end{aligned} \tag{74}$$

where

$$\begin{aligned}
\widetilde{Q}^F_{j_t} &= Q^F_{11j_t} - Q^F_{12j_t} J_{j_t} - J'_{j_t} Q^F_{21j_t} + J'_{j_t} Q^F_{22j_t} J_{j_t} && (75) \\
\widetilde{P}^F_{1j_t} &= P^F_{11j_t} - Q^F_{12j_t} K^L_{j_t} + J'_{j_t} Q^F_{22j_t} K^L_{j_t} - J'_{j_t} P^F_{21j_t} && (76) \\
\widetilde{P}^F_{2j_t} &= P^F_{12j_t} - Q^F_{12j_t} K^F_{j_t} + J'_{j_t} Q^F_{22j_t} K^F_{j_t} - J'_{j_t} P^F_{22j_t} && (77) \\
\widetilde{R}^F_{11j_t} &= K^{L\prime}_{j_t} Q^F_{22j_t} K^L_{j_t} - K^{L\prime}_{j_t} P^F_{21j_t} - P^{F\prime}_{21j_t} K^L_{j_t} + R^F_{11j_t} && (78) \\
\widetilde{R}^F_{12j_t} &= K^{L\prime}_{j_t} Q^F_{22j_t} K^F_{j_t} - K^{L\prime}_{j_t} P^F_{22j_t} - R^F_{12j_t} K^F_{j_t} + R^F_{12j_t} && (79) \\
\widetilde{R}^F_{22j_t} &= K^{F\prime}_{j_t} Q^F_{22j_t} K^F_{j_t} - K^{F\prime}_{j_t} P^F_{22j_t} - P^{F\prime}_{22j_t} K^F_{j_t} + R^F_{22j_t} && (80)
\end{aligned}$$

The optimal value of the problem in period t is associated with the symmetric positive semi-definite matrix $V^F_{k_{t+1}}$ and it satisfies the Bellman equation:

$$X_t V^F_{j_t} X_t = \min_{u^F_{j_t}} \left\{ \boldsymbol{L}^F_{j_t} + \beta E_t \left[X'_{t+1} V^F_{k_{t+1}} X_{t+1} \right] \right\} \tag{81}$$

subject to (73) and (74). The first-order condition with respect to u_t^F is

$$\begin{aligned} 0 &= X_t'\widetilde{P}_{2j_t}^F + u_t^{L\prime}\widetilde{R}_{12j_t}^F + u_t^{F\prime}\widetilde{R}_{22j_t}^F + \beta E_t X_t'\widetilde{A}_{j_t k_{t+1}}' V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F \\ &\quad + \beta E_t u_t^{L\prime}\widetilde{B}_{j_t k_{t+1}}^{L\prime} V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F + \beta E_t u_t^{F\prime}\widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F. \end{aligned}$$

This leads to the optimal policy function u_t^F of the follower

$$u_t^F = -G_{j_{t+1}}^F X_{t+1} - D_{j_{t+1}}^F u_{t+1}^L, \tag{82}$$

where

$$\begin{aligned} G_{j_{t+1}}^F &= \left(\widetilde{R}_{22j_t}^{F\prime} + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F\right)^{-1} \left(\widetilde{P}_{2j_t}^{F\prime} + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \widetilde{A}_{j_t k_{t+1}}\right), \\ D_{j_{t+1}}^F &= \left(\widetilde{R}_{22j_t}^{F\prime} + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F\right)^{-1} \left(\widetilde{R}_{12j_t}^{F\prime} + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k+1}^F \widetilde{B}_{j_t k_{t+1}}^L\right). \end{aligned}$$

Furthermore, using (82) and (68) in (72) gives

$$x_t = -N_{j_t} X_t,$$

where

$$N_{j_t} = J_{j_t} - K_{j_t}^L F_{j_t}^L - K_{j_t}^F G_{j_t}^F + K_{j_t}^F D_{j_t}^F F_{j_t}^L,$$

and using (82) and (68) and (**??**) in (65) gives

$$X_{t+1} = M_{j_t k_{t+1}} X_t + C_{j_{t+1}} \varepsilon_{t+1},$$

where

$$M_{j_t k_{t+1}} = A_{11j_{t+1}} - A_{12j_{t+1}} N_{j_t} - B_{11j_{t+1}} F_{j_t}^L - B_{12j_{t+1}} G_{j_t}^F + B_{12j_{t+1}} D_{j_t}^F F_{j_t}^L$$

Finally, using (68) and (82) in (81) results in

$$\begin{aligned} V_{j_t}^F &= \widetilde{Q}_{j_t}^F - \widetilde{P}_{1j_t}^F F_{j_t}^L - F_{j_t}^{L\prime} \widetilde{P}_{1j_t}^{F\prime} + F_{j_t}^{L\prime} \widetilde{R}_{11j_t}^F F_{j_t}^L \qquad (83) \\ &\quad + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L\right)' V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L\right) \\ &\quad - \left[\widetilde{P}_{2j_t}^{F\prime} - \widetilde{R}_{12j_t}^{F\prime} F_{j_t}^L + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L\right)\right]' \\ &\quad \left(\widetilde{R}_{22j_t}^{F\prime} + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F\right)^{-1} \qquad (84) \\ &\quad \left[\widetilde{P}_{2j_t}^{F\prime} - \widetilde{R}_{12j_t}^{F\prime} F_{j_t}^L + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L\right)\right]. \end{aligned}$$

D.2 Policy of the Leader

Using (72) and (82) in the leader's loss function (67) gives

$$
\begin{aligned}
\boldsymbol{L}_{j_t}^L &= \begin{bmatrix} X_t \\ x_t \\ u_t^L \\ u_t^F \end{bmatrix}' \begin{bmatrix} Q_{11j_t}^L & Q_{12j_t}^L & P_{11j_t}^L & P_{12j_t}^L \\ Q_{21j_t}^L & Q_{22j_t}^L & P_{21j_t}^L & P_{22j_t}^L \\ P_{11j_t}^{L\prime} & P_{21j_t}^{L\prime} & R_{11j_t}^L & R_{12j_t}^L \\ P_{12j_t}^{L\prime} & P_{22j_t}^{L\prime} & R_{12j_t}^{L\prime} & R_{22j_t}^L \end{bmatrix} \begin{bmatrix} X_t \\ x_t \\ u_t^L \\ u_t^F \end{bmatrix} \\
&= \begin{bmatrix} X_t \\ u_t^L \end{bmatrix}' \begin{bmatrix} \widetilde{Q}_{j_t}^L & \widetilde{P}_{j_t}^L \\ \widetilde{P}_{j_t}^{L\prime} & \widetilde{R}_{j_t}^L \end{bmatrix} \begin{bmatrix} X_t \\ u_t^L \end{bmatrix}, \qquad (85)
\end{aligned}
$$

where

$$
\begin{aligned}
\widetilde{Q}_{j_t}^L &= Q_{11j_t}^L - P_{12j_t}^L G_{j_t}^F - G_{j_t}^{F\prime} P_{12j_t}^{L\prime} + G_{j_t}^{F\prime} R_{22j_t}^L G_{j_t}^F - Q_{12j_t}^L \widetilde{J}_{j_t} \qquad (86) \\
&\quad - \widetilde{J}_{j_t}' Q_{21j_t}^L + \widetilde{J}_{j_t}' Q_{22j_t}^L \widetilde{J}_{j_t} + \widetilde{J}_{j_t}' P_{22j_t}^L G_{j_t}^F + G_{j_t}^{F\prime} P_{22j_t}^{L\prime} \widetilde{J}_{j_t}, \\
\widetilde{P}_{j_t}^L &= P_{11j_t}^L - Q_{12j_t}^L \widetilde{K}_{j_t} - P_{12j_t}^L D_{j_t}^F + \widetilde{J}_{j_t}' Q_{22j_t}^L \widetilde{K}_{j_t} - \widetilde{J}_{j_t}' P_{21j_t}^L \qquad (87) \\
&\quad + \widetilde{J}_{j_t}' P_{22j_t}^L D_{j_t}^F + G_{j_t}^{F\prime} P_{22j_t}^{L\prime} \widetilde{K}_{j_t} - G_{j_t}^{F\prime} R_{12j_t}^{L\prime} + G_{j_t}^{F\prime} R_{22j_t}^L D_{j_t}^F, \\
\widetilde{R}_{j_t}^L &= R_{11j_t}^L + \widetilde{K}_{j_t}' Q_{22j_t}^L \widetilde{K}_{j_t} - R_{12j_t}^L D_{j_t}^F - D_{j_t}^{F\prime} R_{12j_t}^{L\prime} + D_{j_t}^{F\prime} R_{22j_t}^L D_{j_t}^F \qquad (88) \\
&\quad - \widetilde{K}_{j_t}' P_{21j_t}^L + \widetilde{K}_{j_t}' P_{22j_t}^L D_{j_t}^F - P_{21j_t}^{L\prime} \widetilde{K}_{j_t} + D_{j_t}^{F\prime} P_{22j_t}^{L\prime} \widetilde{K}_{j_t}.
\end{aligned}
$$

and $\widetilde{J}_{j_t} = \left(J_{j_t} - K_{j_t}^F G_{j_t}^F\right)$ and $\widetilde{K}_{j_t} = \left(K_{j_t}^L - K_{j_t}^F D_{j_t}^F\right)$

The value of the problem in period t is associated with the symmetric positive semidefinite matrix $V_{k_{t+1}}^L$ and it satisfies the Bellman equation

$$
X_t V_{j_t}^L X_t = \min_{u_{j_t}^L} \left\{ \boldsymbol{L}_{j_t}^L + \beta E_t \left[X_{t+1}' V_{k_{t+1}}^L X_{t+1} \right] \right\}, \qquad (89)
$$

subject to (85), (73) and (82) . The first-order condition with respect to u_t^L is

$$
\begin{aligned}
0 &= X_t' \widetilde{P}_{j_t}^L + u_t^{L\prime} \widetilde{R}_{j_t}^L + \beta E_t X_t' \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^F G_{j_t}^F \right)' V_{k_{t+1}}^L \left(\widetilde{B}_{j_t k_{t+1}}^L - \widetilde{B}_{j_t k_{t+1}}^F D_{j_t}^F \right) \\
&\quad + \beta E_t u_t^{L\prime} \left(\widetilde{B}_{j_t k_{t+1}}^L - \widetilde{B}_{j_t k_{t+1}}^F D_{j_t}^F \right)' V_{k_{t+1}}^L \left(\widetilde{B}_{j_t k_{t+1}}^L - \widetilde{B}_{j_t k_{t+1}}^F D_{j_t}^F \right),
\end{aligned}
$$

This leads to the optimal policy function of the leader

$$
u_t^L = -F_j^L X_t, \qquad (90)
$$

where

$$
\begin{aligned}
F_j^L &= \left[\widetilde{R}_{j_t}^{L\prime} + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \left(\widetilde{B}_{j_t k_{t+1}}^L - \widetilde{B}_{j_t k_{t+1}}^F D_{j_t}^F \right)' V_{k_{t+1}}^L \left(\widetilde{B}_{j_t k_{t+1}}^L - \widetilde{B}_{j_t k_{t+1}}^F D_{j_t}^F \right) \right]^{-1} \\
&\quad \left[\widetilde{P}_{j_t}^{L\prime} + \beta \sum_{k=1}^n P_{j_t k_{t+1}} \left(\widetilde{B}_{j_t k_{t+1}}^L - \widetilde{B}_{j_t k_{t+1}}^F D_{j_t}^F \right)' V_{k_{t+1}}^L \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^F G_{j_t}^F \right) \right]
\end{aligned}
$$

Furthermore, using (82) and (90) in (72) gives

$$x_t = -N_{j_t} X_t, \tag{91}$$

where

$$N_{j_t} = J_{j_t} - K^L_{j_t} F^L_{j_t} - K^F_{j_t} G^F_{j_t} + K^F_{j_t} D^F_{j_t} F^L_{j_t},$$

and using (82), (90) and (91) in (65) gives

$$X_{t+1} = M_{j_t k_{t+1}} X_t + C_{j_{t+1}} \varepsilon_{t+1},$$

where

$$M_{j_t k_{t+1}} = A_{11 j_{t+1}} - A_{12 j_{t+1}} N_{j_t} - B_{11 j_{t+1}} F^L_{j_t} - B_{12 j_{t+1}} G^F_{j_t} + B_{12 j_{t+1}} D^F_{j_t} F^L_{j_t}$$

Finally, using (90) and (73) in (89) results in

$$\begin{aligned}
V^L_{j_t} \quad = \quad & \widetilde{Q}^L_{j_t} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} G^F_{j_t} \right)' V^L_k \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} G^F_{j_t} \right) \\
& - \left[\widetilde{P}^{L\prime}_{j_t} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{B}^L_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} D^F_{j_t} \right)' V^L_{k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} G^F_{j_t} \right) \right] \\
& \left[\widetilde{R}^{L\prime}_{j_t} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{B}^L_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} D^F_{j_t} \right)' V^L_{k_{t+1}} \left(\widetilde{B}^L_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} D^F_{j_t} \right) \right]^{-1} \\
& \left[\widetilde{P}^{L\prime}_{j_t} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{B}^L_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} D^F_{j_t} \right)' V^L_{k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} G^F_{j_t} \right) \right]
\end{aligned} \tag{92}$$

To sum up, the first order conditions to the optimization problem (65), (66) and (67) can be written in the following form:

$$N_{j_t} = J_{j_t} - K^L_{j_t} F^L_{j_t} - K^F_{j_t} G^F_{j_t} + K^F_{j_t} D^F_{j_t} F^L_{j_t}, \tag{93}$$

$$\begin{aligned}
V^F_{j_t} \quad \equiv \quad & \widetilde{Q}^F_{j_t} - \widetilde{P}^F_{1 j_t} F^L_{j_t} - F^{L\prime}_{j_t} \widetilde{P}^{F\prime}_{1 j_t} + F^{L\prime}_{j_t} \widetilde{R}^F_{11 j_t} F^L_{j_t} \\
& + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}^L_{j_t k_{t+1}} F^L_{j_t} \right)' V^F_{k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}^L_{j_t k_{t+1}} F^L_{j_t} \right) \\
& - \left[\widetilde{P}^{F\prime}_{2 j_t} - \widetilde{R}^{F\prime}_{12 j_t} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \widetilde{B}^{F\prime}_{j_t k_{t+1}} V^F_{k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} + \widetilde{B}^L_{j_t k_{t+1}} F^L_{j_t} \right) \right]' \\
& \left(\widetilde{R}^{F\prime}_{22 j_t} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \widetilde{B}^{F\prime}_{j_t k_{t+1}} V^F_{k_{t+1}} \widetilde{B}^F_{j_t k_{t+1}} \right)^{-1} \\
& \left[\widetilde{P}^{F\prime}_{2 j_t} - \widetilde{R}^{F\prime}_{12 j_t} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \widetilde{B}^{F\prime}_{j_t k_{t+1}} V^F_{k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} + \widetilde{B}^L_{j_t k_{t+1}} F^L_{j_t} \right) \right],
\end{aligned} \tag{94}$$

$$
\begin{aligned}
V_{j_t}^{L} \equiv\ & \widetilde{Q}_{j_t}^{L} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^{F} G_{j_t}^{F} \right)' V_k^{L} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^{F} G_{j_t}^{F} \right) \\
& - \left[\widetilde{P}_{j_t}^{L\prime} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{B}_{j_t k_{t+1}}^{L} - \widetilde{B}_{j_t k_{t+1}}^{F} D_{j_t}^{F} \right)' V_{k_{t+1}}^{L} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^{F} G_{j_t}^{F} \right) \right]' \\
& \left[\widetilde{R}_{j_t}^{L\prime} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{B}_{j_t k_{t+1}}^{L} - \widetilde{B}_{j_t k_{t+1}}^{F} D_{j_t}^{F} \right)' V_{k_{t+1}}^{L} \left(\widetilde{B}_{j_t k_{t+1}}^{L} - \widetilde{B}_{j_t k_{t+1}}^{F} D_{j_t}^{F} \right) \right]^{-1} \\
& \left[\widetilde{P}_{j_t}^{L\prime} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{B}_{j_t k_{t+1}}^{L} - \widetilde{B}_{j_t k_{t+1}}^{F} D_{j_t}^{F} \right)' V_{k_{t+1}}^{L} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^{F} G_{j_t}^{F} \right) \right],
\end{aligned} \tag{95}
$$

$$
\begin{aligned}
F_j^{L} =\ & \left[\widetilde{R}_{j_t}^{L\prime} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{B}_{j_t k_{t+1}}^{L} - \widetilde{B}_{j_t k_{t+1}}^{F} D_{j_t}^{F} \right)' V_{k_{t+1}}^{L} \left(\widetilde{B}_{j_t k_{t+1}}^{L} - \widetilde{B}_{j_t k_{t+1}}^{F} D_{j_t}^{F} \right) \right]^{-1} \\
& \left[\widetilde{P}_{j_t}^{L\prime} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{B}_{j_t k_{t+1}}^{L} - \widetilde{B}_{j_t k_{t+1}}^{F} D_{j_t}^{F} \right)' V_{k_{t+1}}^{L} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^{F} G_{j_t}^{F} \right) \right],
\end{aligned} \tag{96}
$$

$$
G_{j_{t+1}}^{F} = \left(\widetilde{R}_{22j_t}^{F\prime} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^{F} \widetilde{B}_{j_t k_{t+1}}^{F} \right)^{-1} \left(\widetilde{P}_{2j_t}^{F\prime} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^{F} \widetilde{A}_{j_t k_{t+1}} \right), \tag{97}
$$

$$
D_{j_{t+1}}^{F} = \left(\widetilde{R}_{22j_t}^{F\prime} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^{F} \widetilde{B}_{j_t k_{t+1}}^{F} \right)^{-1} \left(\widetilde{R}_{12j_t}^{F\prime} + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k+1}^{F} \widetilde{B}_{j_t k_{t+1}}^{L} \right) \tag{98}
$$

The discretion equilibrium is a fixed point $\left(N_j, V_j^L, V_j^F\right) \equiv \left\{N_{j_t}, V_{j_t}^L, V_{j_t}^F\right\}_{j_t=1}^{n}$ of the mapping and a corresponding $\left(F_j^L, G_j^F, D_j^F\right) \equiv \left\{F_{j_t}^L, G_{j_t}^F, D_{j_t}^F\right\}_{j_t=1}^{n}$. The fixed point can be obtained as the limit of $\left(N_{jt}, V_{jt}^L, V_{jt}^F\right)$ when $t \longrightarrow -\infty$.

E Nash Equilibrium under Discretion with Markov Switching in Objectives

Under Nash, the two policy makers decide their policy u_t^L and u_t^F simultaneously. It is therefore arbitrary which we label leader and follower. The policy reaction of the follower (F) can not include the leader (L)'s contemporary policy instrument, therefore, $D_{j_{t+1}}^F = 0$ in (69). Reoptimization in period $t+1$ result in the two instruments and the forward-looking variables being functions of the predetermined variables and the mode as follows

$$
\begin{aligned}
u^L_{t+1} &= -F^L_{j_{t+1}} X_{t+1}, & (99)\\
u^F_{t+1} &= -F^F_{j_{t+1}} X_{t+1}, & (100)\\
x_{t+1} &= -N_{j_{t+1}} X_{t+1}, & (101)
\end{aligned}
$$

E.1 Policy of the Follower

Using (72) and (99) in the follower's loss function (67) gives

$$
\begin{aligned}
\boldsymbol{L}^F_{j_t} &= \begin{bmatrix} X_t \\ x_t \\ u^L_t \\ u^F_t \end{bmatrix}' \begin{bmatrix} Q^F_{11j_t} & Q^F_{12j_t} & P^F_{11j_t} & P^F_{12j_t} \\ Q^F_{21j_t} & Q^F_{22j_t} & P^F_{21j_t} & P^F_{22j_t} \\ P^{F\prime}_{11j_t} & P^{F\prime}_{21j_t} & R^F_{11j_t} & R^F_{12j_t} \\ P^{F\prime}_{12j_t} & P^{F\prime}_{22j_t} & R^{F\prime}_{12j_t} & R^F_{22j_t} \end{bmatrix} \begin{bmatrix} X_t \\ x_t \\ u^L_t \\ u^F_t \end{bmatrix} \\
&= \begin{bmatrix} X_t \\ u^F_t \end{bmatrix}' \begin{bmatrix} \widetilde{Q}^F_{j_t} & \widetilde{P}^F_{j_t} \\ \widetilde{P}^{F\prime}_{j_t} & \widetilde{R}^F_{j_t} \end{bmatrix} \begin{bmatrix} X_t \\ u^F_t \end{bmatrix} \qquad (102)
\end{aligned}
$$

where

$$
\begin{aligned}
\widetilde{Q}^F_{j_t} &= Q^F_{11j_t} - Q^F_{12j_t}\widetilde{J}^L_{j_t} - P^F_{11j_t}F^L_{j_t} - \widetilde{J}^{L\prime}_{j_t}Q^F_{21j_t} + \widetilde{J}^{L\prime}_{j_t}Q^F_{22j_t}\widetilde{J}^L_{j_t} + \widetilde{J}^{L\prime}_{j_t}P^F_{21j_t}F^L_{j_t} & (103)\\
&\quad -F^{L\prime}_{j_t}P^{F\prime}_{11j_t} + F^{L\prime}_{j_t}P^{F\prime}_{21j_t}\widetilde{J}^L_{j_t} + F^{L\prime}_{j_t}R^F_{11j_t}F^L_{j_t}, \\
\widetilde{P}^F_{j_t} &= -Q^F_{12j_t}K^F_{j_t} + P^F_{12j_t} + \widetilde{J}^{L\prime}_{j_t}Q^F_{22j_t}K^F_{j_t} - \widetilde{J}^{L\prime}_{j_t}P^F_{22j_t} + F^{L\prime}_{j_t}P^{F\prime}_{21j_t}K^F_{j_t} - F^{L\prime}_{j_t}R^F_{12j_t}, & (104)\\
\widetilde{R}^F_{j_t} &= K^F_{j_t}Q^F_{22j_t}K^F_{j_t} - K^F_{j_t}P^F_{22j_t} - P^{F\prime}_{22j_t}K^F_{j_t} + R^F_{22j_t}, & (105)
\end{aligned}
$$

and $\widetilde{J}^L_{j_t} = \left(J_{j_t} - K^L_{j_t}F^L_{j_t}\right)$.

The optimal value of the problem in period t is associated with the symmetric positive semi-definite matrix $V^F_{k_{t+1}}$ and it satisfies the Bellman equation:

$$
X_t V^F_{j_t} X_t = \min_{u^F_{j_t}} \left\{ \boldsymbol{L}^F_{j_t} + \beta E_t \left[X'_{t+1} V^F_{k_{t+1}} X_{t+1} \right] \right\} \qquad (106)
$$

subject to (73), (99) and (102). The first-order condition with respect to u^F_t is

$$
0 = X'_t\widetilde{P}^F_{j_t} + u^{F\prime}_t\widetilde{R}^F_{j_t} + \beta E_t X'_t\left(\widetilde{A}_{j_tk_{t+1}} - \widetilde{B}^L_{j_tk_{t+1}}F^L_{j_t}\right)' V^F_{k_{t+1}}\widetilde{B}^F_{j_tk_{t+1}} + \beta E_t u^{F\prime}_t \widetilde{B}^{F\prime}_{j_tk_{t+1}} V^F_{k_{t+1}} \widetilde{B}^F_{j_tk_{t+1}},
$$

The optimal policy function of the follower is given by

$$
u^F_t = -F^F_{j_t} X_t, \qquad (107)
$$

where

$$F_{j_t}^F = \left(\widetilde{R}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F\right)^{-1} \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L\right)\right]$$

Furthermore, using (107) and (99) in (72) gives

$$x_t = -N_{j_t} X_t, \tag{108}$$

where

$$N_{j_t} = J_{j_t} - K_{j_t}^L F_{j_t}^L - K_{j_t}^F F_{j_t}^F,$$

and using (107),(108) and (99) in (65) gives

$$X_{t+1} = M_{j_t k_{t+1}} X_t + C_{j_{t+1}} \varepsilon_{t+1},$$

where

$$M_{j_t k_{t+1}} = A_{11 j_{t+1}} - A_{12 j_{t+1}} N_{j_t} - B_{11 j_{t+1}} F_{j_{t+1}}^L - B_{12 j_{t+1}} F_{j_t}^F$$

Finally, using (99) and (107) in (106) results in

$$\begin{aligned} V_{j_t}^F \;=\; & \widetilde{Q}_{j_t}^F + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}}' - F_{j_{t+1}}^{L\prime} \widetilde{B}_{j_t k_{t+1}}^{L\prime}\right) V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_{t+1}}^L\right) \\ & - \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L\right)\right]' \\ & \left[\widetilde{R}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F\right]^{-1} \\ & \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L\right)\right] \end{aligned} \tag{109}$$

E.2 Policy of the Leader

Using (72) and (100) in the follower's loss function (67) gives

$$\begin{aligned} \boldsymbol{L}_{j_t}^F \;&=\; \begin{bmatrix} X_t \\ x_t \\ u_t^L \\ u_t^F \end{bmatrix}' \begin{bmatrix} Q_{11j_t}^L & Q_{12j_t}^L & P_{11j_t}^L & P_{12j_t}^L \\ Q_{21j_t}^L & Q_{22j_t}^L & P_{21j_t}^L & P_{22j_t}^L \\ P_{11j_t}^{L\prime} & P_{21j_t}^{L\prime} & R_{11j_t}^L & R_{12j_t}^L \\ P_{12j_t}^{L\prime} & P_{22j_t}^{L\prime} & R_{12j_t}^{L\prime} & R_{22j_t}^L \end{bmatrix} \begin{bmatrix} X_t \\ x_t \\ u_t^L \\ u_t^F \end{bmatrix} \\ &=\; \begin{bmatrix} X_t \\ u_t^L \end{bmatrix}' \begin{bmatrix} \widetilde{Q}_{j_t}^L & \widetilde{P}_{j_t}^L \\ \widetilde{P}_{j_t}^{L\prime} & \widetilde{R}_{j_t}^L \end{bmatrix} \begin{bmatrix} X_t \\ u_t^L \end{bmatrix} \end{aligned} \tag{110}$$

where $\widetilde{J}_{j_t}^F = J_{j_t} - K_{j_t}^F F_{j_t}^F$ and

$$
\begin{aligned}
\widetilde{Q}^L_{j_t} &= Q^L_{11j_t} - Q^L_{12j_t}\widetilde{J}^F_{j_t} - P^L_{12j_t}F^F_{j_t} - \widetilde{J}^{F\prime}_{j_t}Q^L_{21j_t} + \widetilde{J}^{F\prime}_{j_t}Q^L_{22j_t}\widetilde{J}^F_{j_t} + \widetilde{J}^{F\prime}_{j_t}P^L_{22j_t}F^F_{j_t} \qquad (111)\\
&\quad -F^{F\prime}_{j_t}P^{L\prime}_{12j_t} + F^{F\prime}_{j_t}P^{L\prime}_{22j_t}\widetilde{J}^F_{j_t} + F^{F\prime}_{j_t}R^L_{22j_t}F^F_{j_t},\\
\widetilde{P}^L_{j_t} &= -Q^L_{12j_t}K^L_{j_t} + P^L_{11j_t} + \widetilde{J}^{F\prime}_{j_t}Q^L_{22j_t}K^L_{j_t} - \widetilde{J}^{F\prime}_{j_t}P^L_{21j_t} + F^{F\prime}_{j_t}P^{L\prime}_{22j_t}K^L_{j_t} - F^{F\prime}_{j_t}R^{L\prime}_{12j_t}, \qquad (112)\\
\widetilde{R}^L_{j_t} &= K^{L\prime}_{j_t}Q^L_{22j_t}K^L_{j_t} - K^{L\prime}_{j_t}P^L_{21j_t} - P^{L\prime}_{21j_t}K^L_{j_t} + R^L_{11j_t} \qquad (113)
\end{aligned}
$$

The optimal value of the problem in period t is associated with the symmetric positive semi-definite matrix $V^L_{k_{t+1}}$ and it satisfies the Bellman equation:

$$
X_t V^L_{j_t} X_t = \min_{u^L_{j_t}} \left\{ \boldsymbol{L}^L_{j_t} + \beta E_t \left[X'_{t+1} V^L_{k_{t+1}} X_{t+1} \right] \right\} \qquad (114)
$$

subject to (73), (100) and (110). The first-order condition with respect to u^L_t is

$$
0 = X'_t \widetilde{P}^F_{j_t} + u^{L\prime}_t \widetilde{R}^F_{j_t} + \beta E_t X'_t \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} F^F_{j_t} \right)' V^L_{k_{t+1}} \widetilde{B}^L_{j_t k_{t+1}} + \beta E_t u^{L\prime}_t \widetilde{B}^{L\prime}_{j_t k_{t+1}} V^L_{k_{t+1}} \widetilde{B}^L_{j_t k_{t+1}}
$$

The optimal policy function of the leader is given by

$$
u^L_t = -F^L_{j_t} X_t, \qquad (115)
$$

where

$$
\begin{aligned}
F^L_{j_t} &= \left(\widetilde{R}^{F\prime}_{j_t} + \beta E_t \widetilde{B}^{L\prime}_{j_t k_{t+1}} V^L_{k_{t+1}} \widetilde{B}^L_{j_t k_{t+1}} \right)^{-1} \\
&\quad \left[\widetilde{P}^{F\prime}_{j_t} + \beta E_t \widetilde{B}^{L\prime}_{j_t k_{t+1}} V^L_{k_{t+1}} \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}^F_{j_t k_{t+1}} F^F_{j_t} \right) \right].
\end{aligned}
$$

Furthermore, using (107) and (115) in (72) gives

$$
x_t = -N_{j_t} X_t, \qquad (116)
$$

where

$$
N_{j_t} = J_{j_t} - K^L_{j_t} F^L_{j_t} - K^F_{j_t} F^F_{j_t}
$$

and using (107) and (115) and (116) in (65) gives

$$
X_{t+1} = M_{j_t k_{t+1}} X_t + C_{j_{t+1}} \varepsilon_{t+1},
$$

where

$$
M_{j_t k_{t+1}} = A_{11j_{t+1}} - A_{12j_{t+1}} N_{j_t} - B_{11j_{t+1}} F^L_{j_{t+1}} - B_{12j_{t+1}} F^F_{j_t}.
$$

Finally, using (115) and (107) in (114) results in

$$
\begin{aligned}
V_{j_t}^L &= \widetilde{Q}_{j_t}^L + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{A}'_{j_t k_{t+1}} - F_{j_t}^{F\prime} \widetilde{B}_{j_t k_{t+1}}^{F\prime} \right) V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^F F_{j_t}^F \right) \\
&\quad - \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{L\prime} V_{k_{t+1}}^L \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^F F_{j_t}^F \right) \right]' \\
&\quad \left(\widetilde{R}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{L\prime} V_{k_{t+1}}^L \widetilde{B}_{j_t k_{t+1}}^L \right)^{-1} \\
&\quad \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{L\prime} V_{k_{t+1}}^L \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^F F_{j_t}^F \right) \right]
\end{aligned} \tag{117}
$$

To sum up, the first order conditions to the optimization problem (65), (66) and (67) can be written in the following form:

$$
N_{j_t} = J_{j_t} - K_{j_t}^L F_{j_t}^L - K_{j_t}^F F_{j_t}^F,
$$

$$
\begin{aligned}
V_{j_t}^F &= \widetilde{Q}_{j_t}^F + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{A}'_{j_t k_{t+1}} - F_{j_{t+1}}^{L\prime} \widetilde{B}_{j_t k_{t+1}}^{L\prime} \right) V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_{t+1}}^L \right) \\
&\quad - \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L \right) \right]' \\
&\quad \left[\widetilde{R}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F \right]^{-1} \\
&\quad \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L \right) \right],
\end{aligned} \tag{118}
$$

$$
\begin{aligned}
V_{j_t}^L &= \widetilde{Q}_{j_t}^L + \beta \sum_{k=1}^{n} P_{j_t k_{t+1}} \left(\widetilde{A}'_{j_t k_{t+1}} - F_{j_t}^{F\prime} \widetilde{B}_{j_t k_{t+1}}^{F\prime} \right) V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^F F_{j_t}^F \right) \\
&\quad - \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{L\prime} V_{k_{t+1}}^L \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^F F_{j_t}^F \right) \right]' \\
&\quad \left(\widetilde{R}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{L\prime} V_{k_{t+1}}^L \widetilde{B}_{j_t k_{t+1}}^L \right)^{-1} \\
&\quad \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{L\prime} V_{k_{t+1}}^L \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^F F_{j_t}^F \right) \right],
\end{aligned} \tag{119}
$$

$$
F_{j_t}^F = \left(\widetilde{R}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \widetilde{B}_{j_t k_{t+1}}^F \right)^{-1} \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{F\prime} V_{k_{t+1}}^F \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^L F_{j_t}^L \right) \right], \tag{120}
$$

$$
F_{j_t}^L = \left(\widetilde{R}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{L\prime} V_{k_{t+1}}^L \widetilde{B}_{j_t k_{t+1}}^L \right)^{-1} \left[\widetilde{P}_{j_t}^{F\prime} + \beta E_t \widetilde{B}_{j_t k_{t+1}}^{L\prime} V_{k_{t+1}}^L \left(\widetilde{A}_{j_t k_{t+1}} - \widetilde{B}_{j_t k_{t+1}}^F F_{j_t}^F \right) \right]. \tag{121}
$$

The discretion equilibrium is a fixed point $\left(N_j, V_j^L, V_j^F\right) \equiv \left\{N_{j_t}, V_{j_t}^L, V_{j_t}^F\right\}_{j_t=1}^n$ of the mapping and a corresponding $\left(F_j^L, F_j^F\right) \equiv \left\{F_{j_t}^L, F_{j_t}^F\right\}_{j_t=1}^n$. The fixed point can be obtained as the limit of $\left(N_{jt}, V_{jt}^L, V_{jt}^F\right)$ when $t \longrightarrow -\infty$.

F Data Appendix

We followed Bianchi and Ilut (2015) when construct our fiscal variables. The fiscal variables, such as government spending and tax revenues, are taken from National Income and Product Accounts (NIPA) Table 3.2 (Federal Government Current Receipts and Expenditures) released by the Bureau of Economics Analysis. These data series are nominal and in levels.

Government Spending. Government spending is defined as the sum of consumption expenditure (line 21), gross government investment (line 42), net purchases of nonproduced assets (line 44), minus consumption of fixed capital (line 45), minus wage accruals less disbursements (line 33).

Total tax revenues. Total tax revenues are constructed as the difference between current receipts (line 38) and current transfer receipts (line 16).

Federal government debt. Federal government debt is the market value of privately held gross Federal debt, which is downloaded from Dallas Fed web-site

The above three fiscal variables are normalized with respect to Nominal GDP. **Nominal GDP** is taken from NIPA Table 1.1.5 (Gross Domestic Product).

Real GDP. Real GDP is take download from NIPA Table 1.1.6 (Real Gross Domestic Product, Chained Dollars)

The GDP deflator. The GDP deflator is obtained from NIPA Table 1.1.5 (Gross Domestic Product).

Effective Federal Funds Rate. Effective Federal Funds Rate is taken from the St. Louis Fed website.

Finally, we calculate the total government transfers to contrast with the exogenous AR(1) transfer shock in Figure 6. All data used to construct government transfers are taken from NIPA Table 3.2. Total government transfers is defined as current transfer payments (line 22) minus current transfer receipts (line 16) plus capital transfers payments (line 43) minus capital transfer receipts (line 39) plus subsidies (line 32).

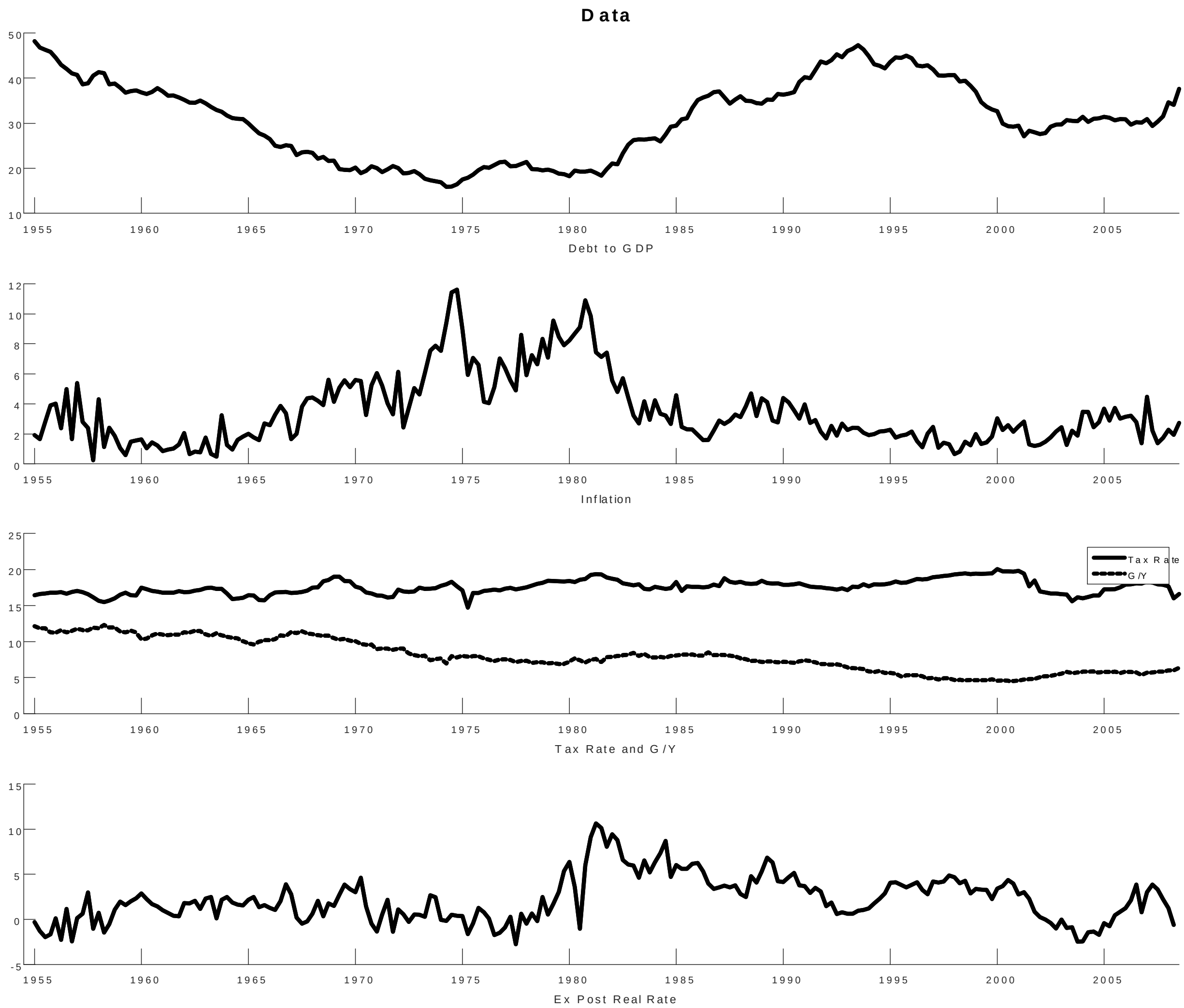

Figure 1: Data

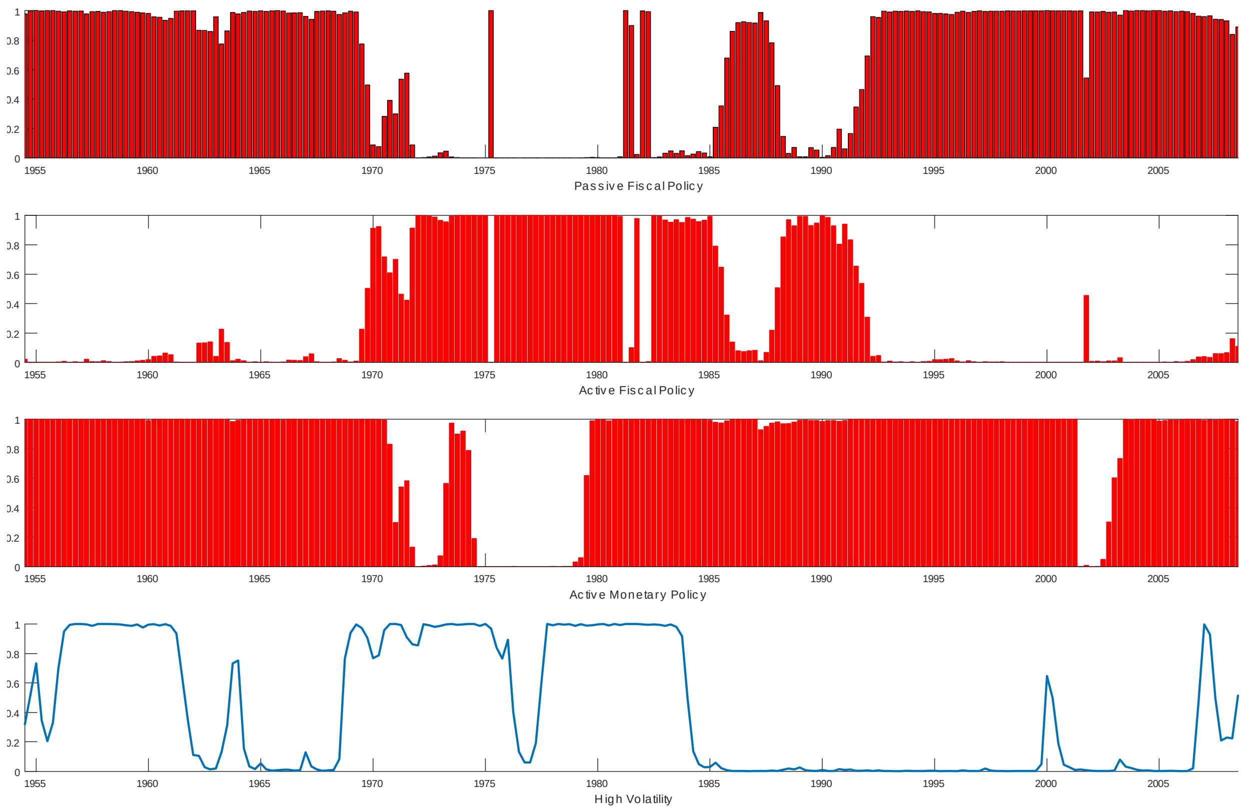

Figure 2: Regime Switching Rules I

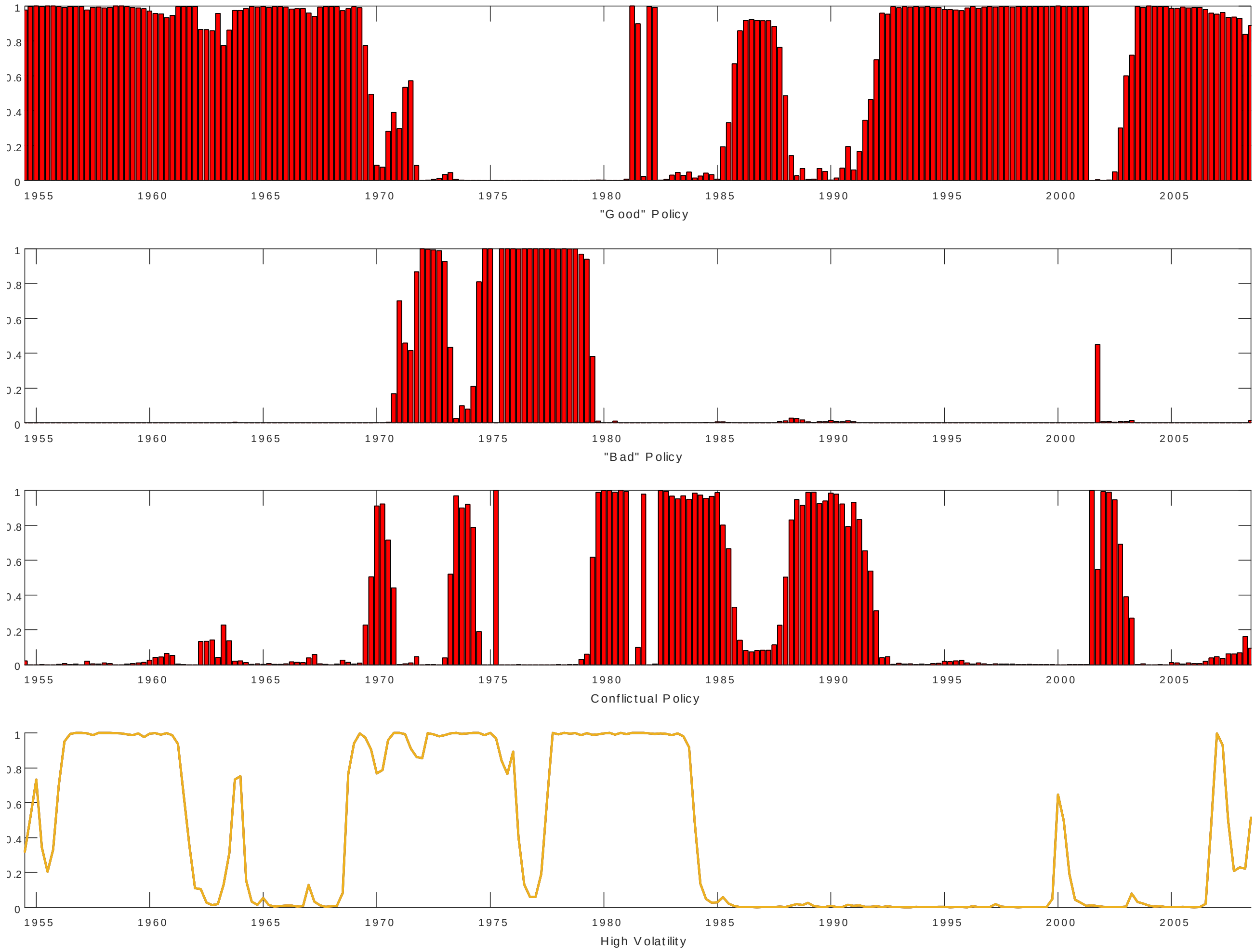

Figure 3: Regime Switching Rules II

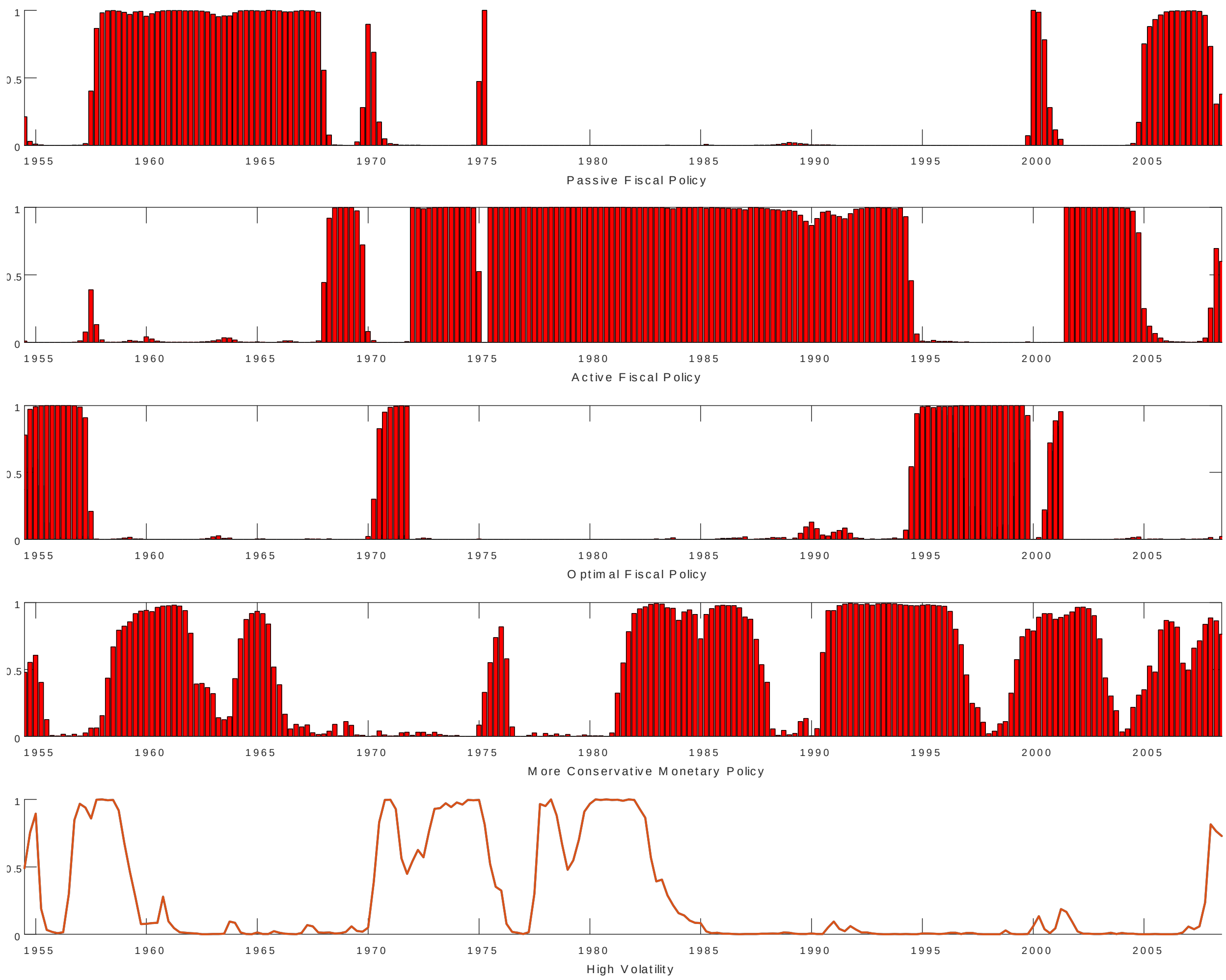

Figure 4: Regime Switching Optimal Policy I

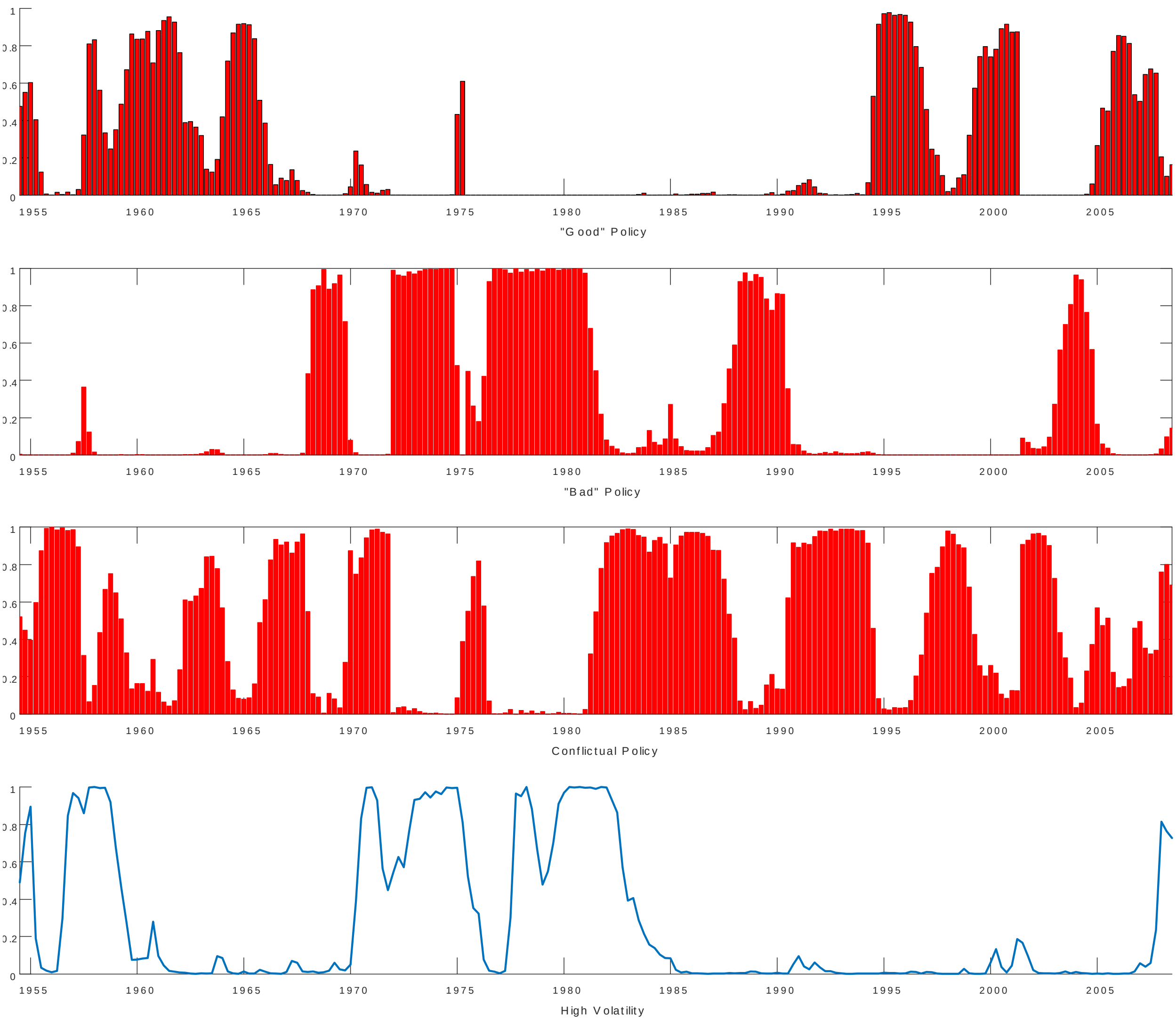

Figure 5: Regime Switching Optimal Policy II

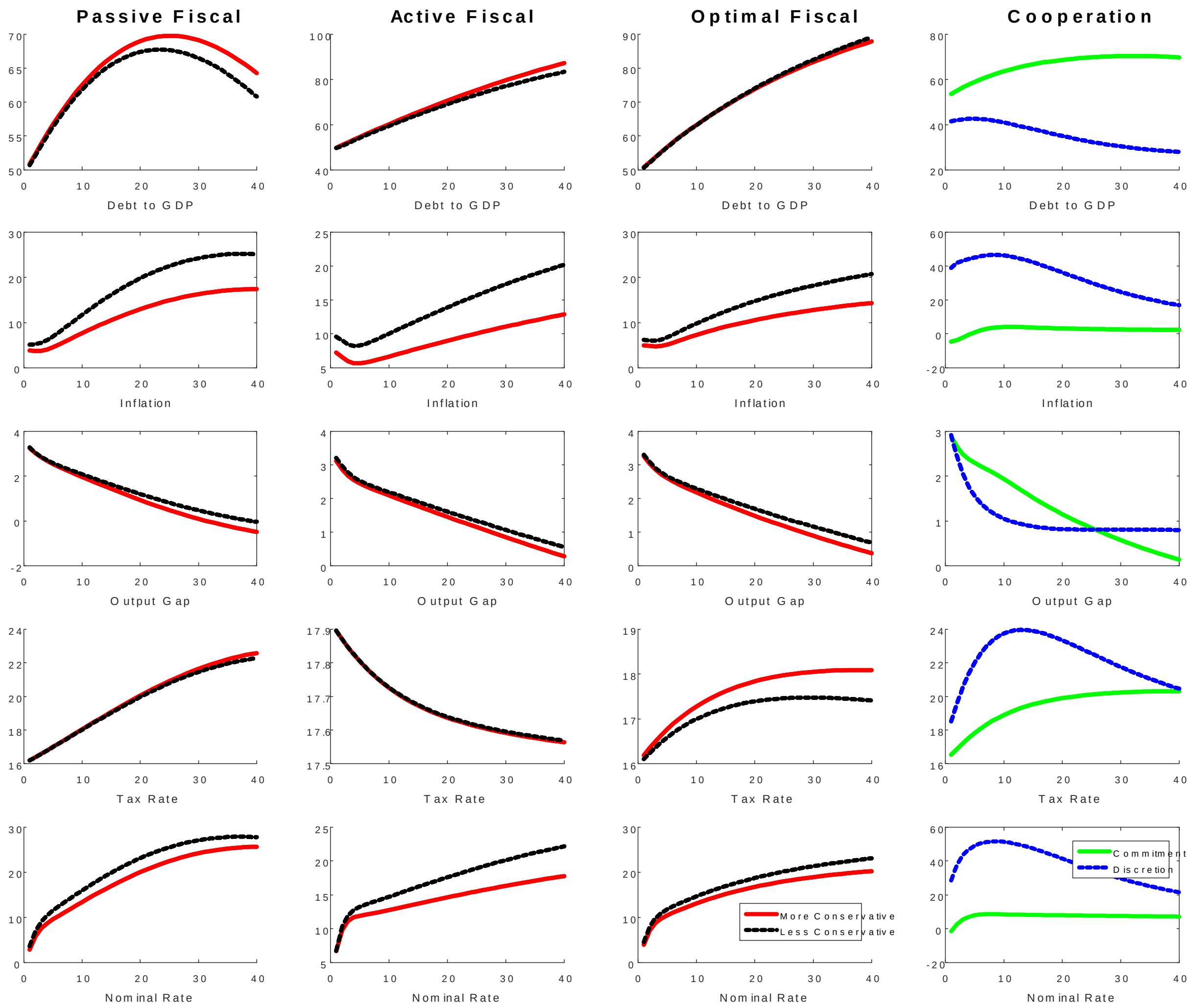

Figure 6: Estimated Regimes IRF

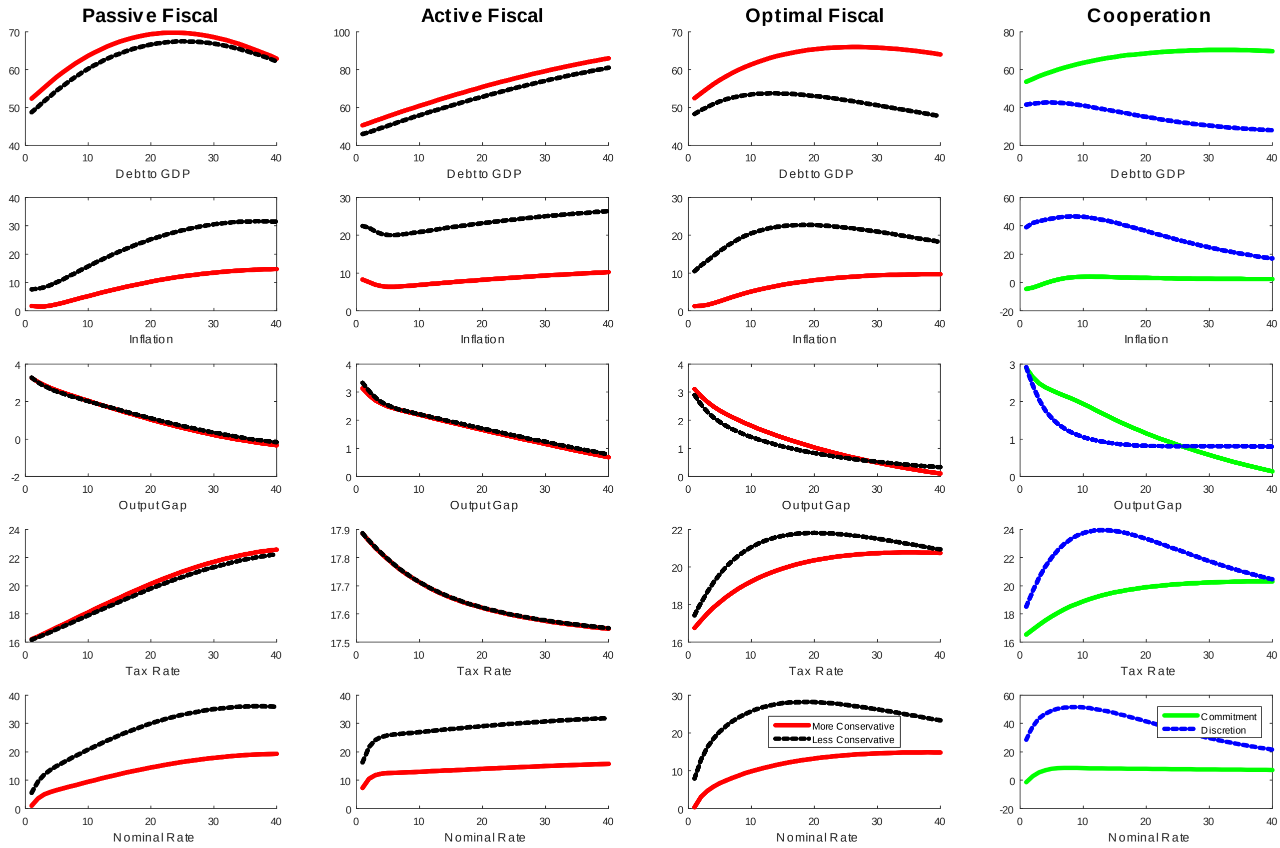

Figure 7: Credible Regimes IRF

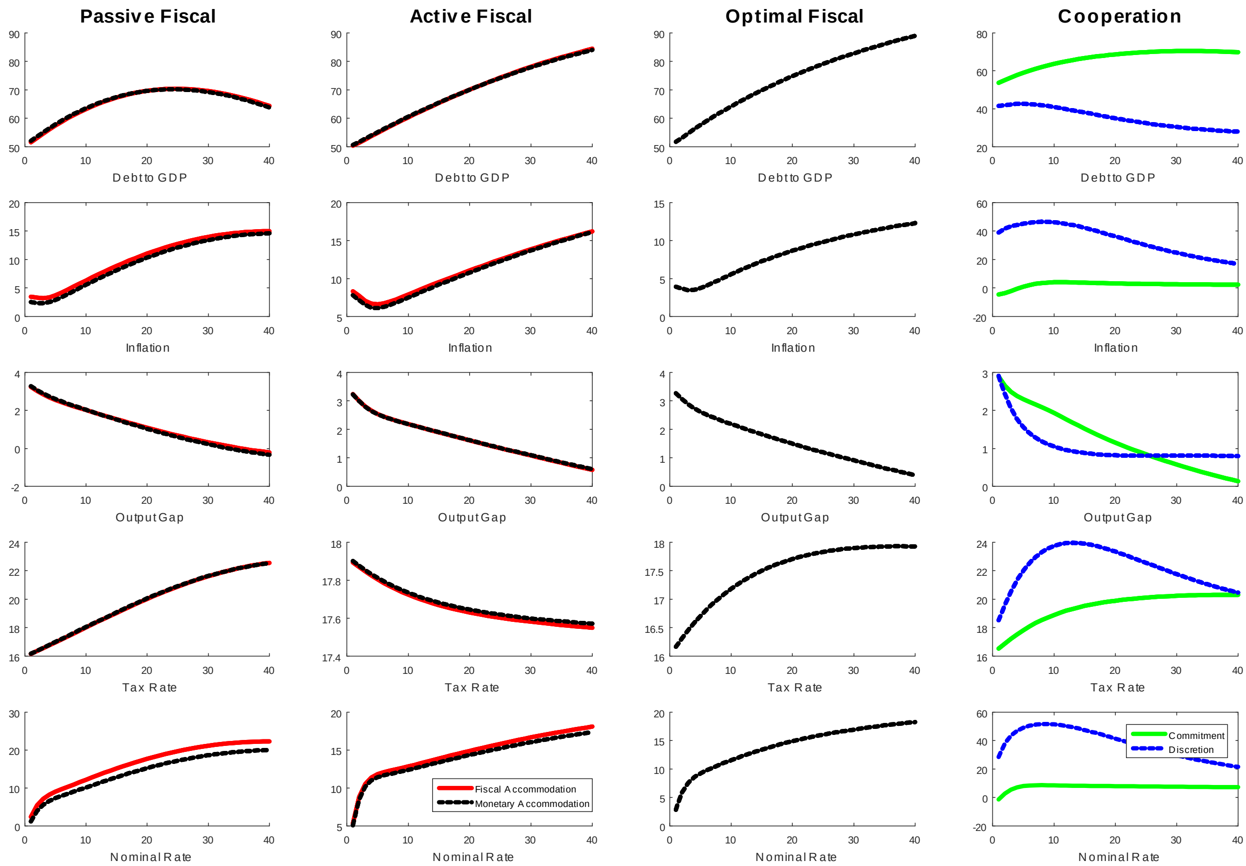

Figure 8: Eliminating Conflict IRF

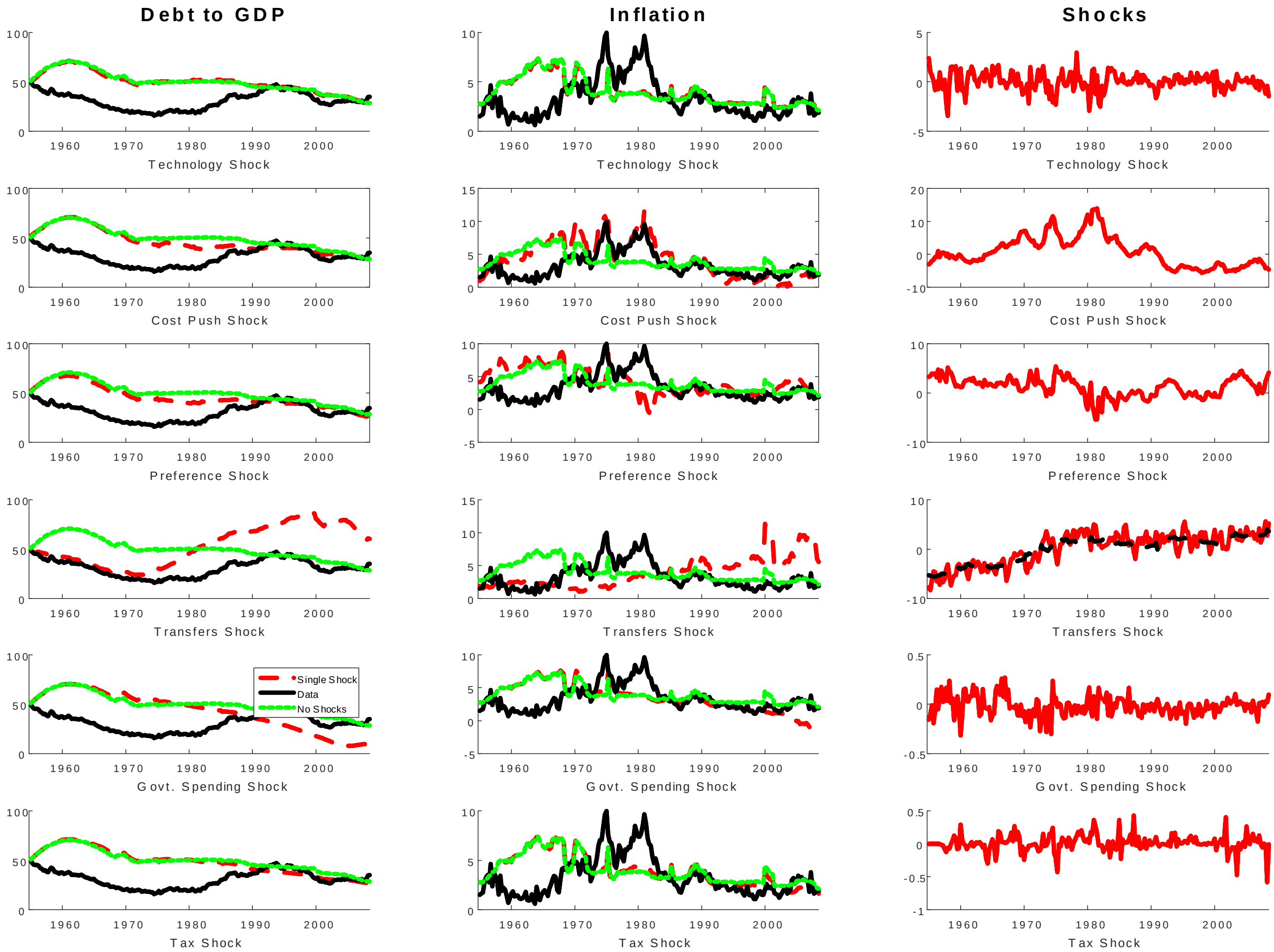

Figure 9: Contribution of Shocks to Trends in Debt to GDP and Inflation.

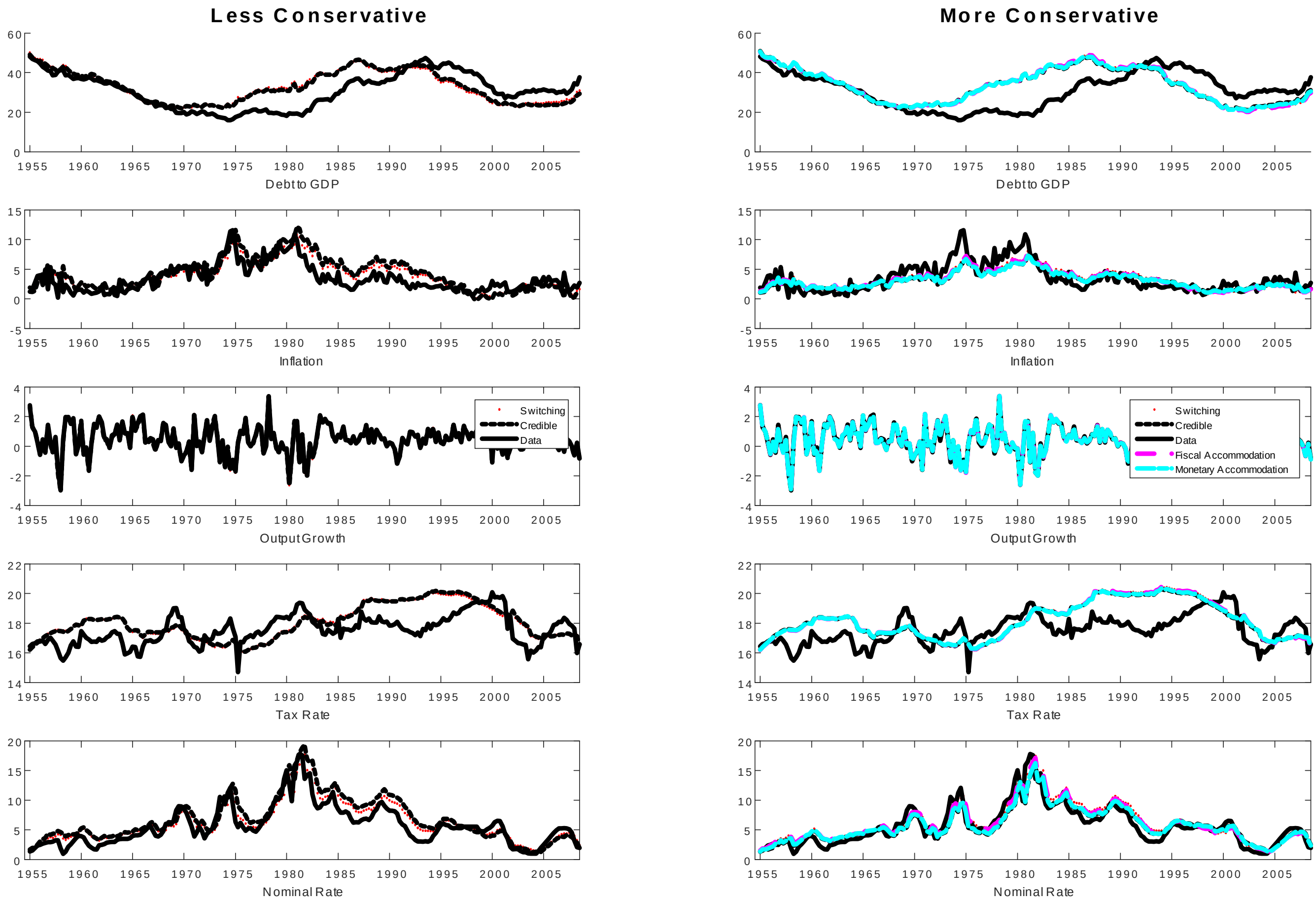

Figure 10: Passive Fiscal Policy Counterfactual

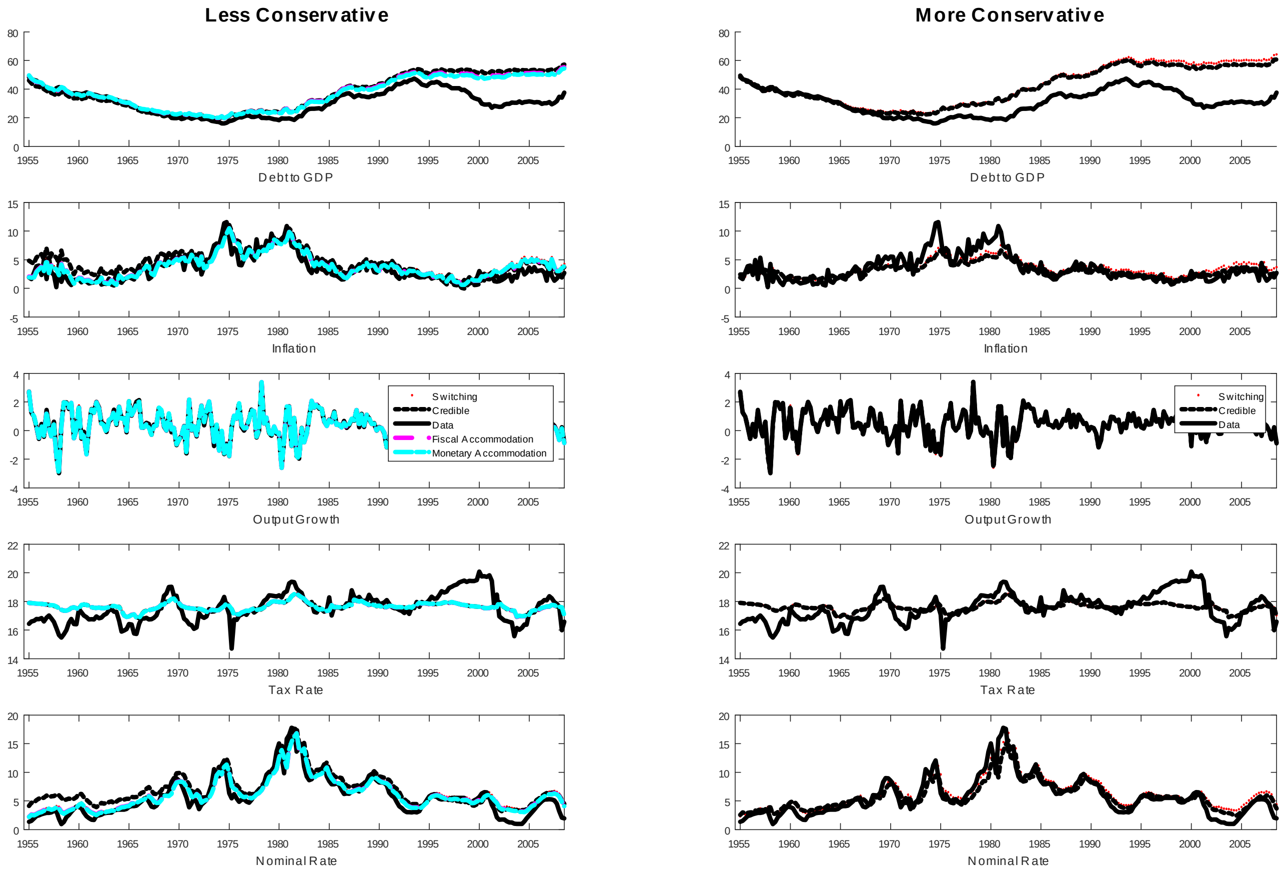

Figure 11: Active Fiscal Policy Counterfactual

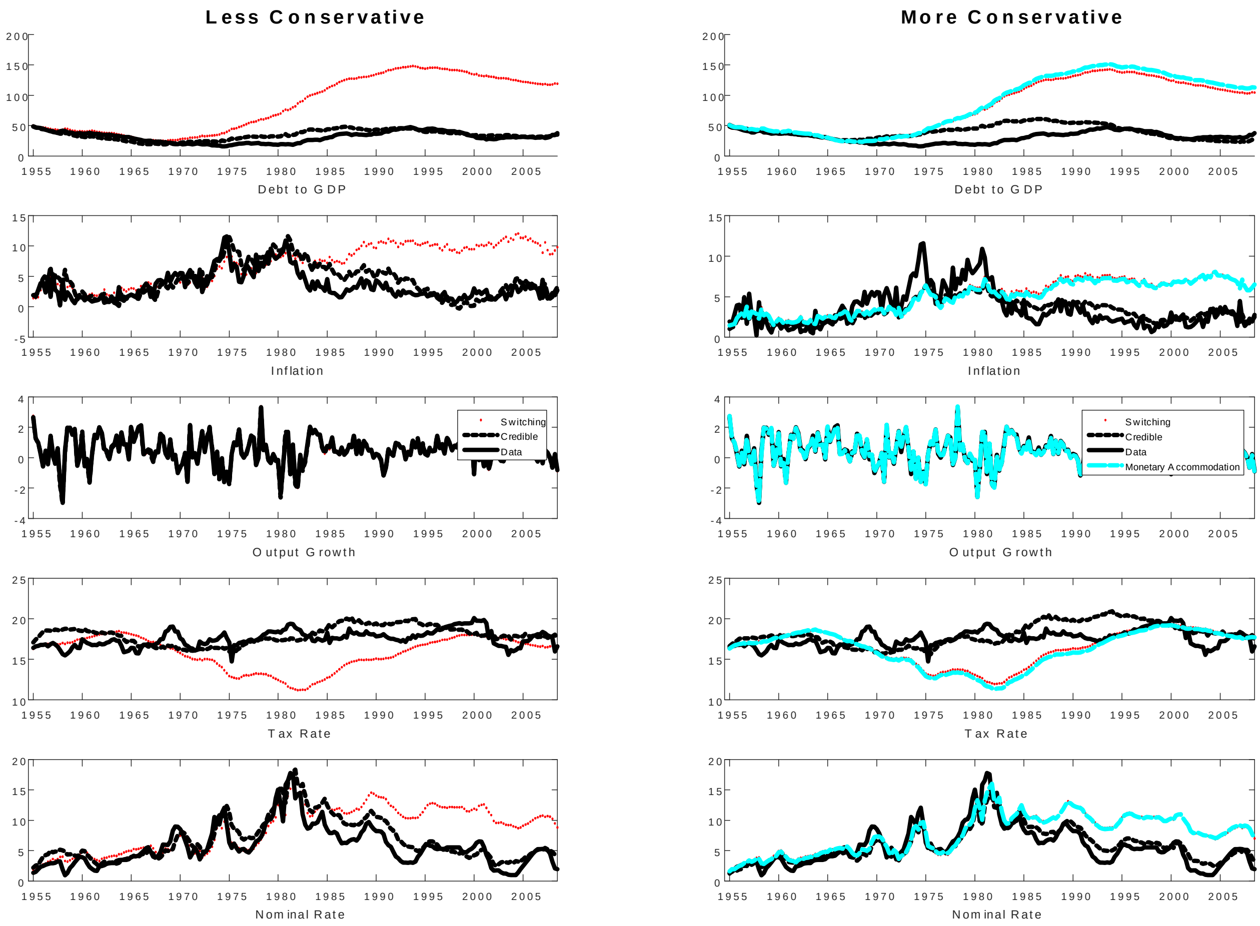

Figure 12: Optimal Fiscal Policy Counterfactual

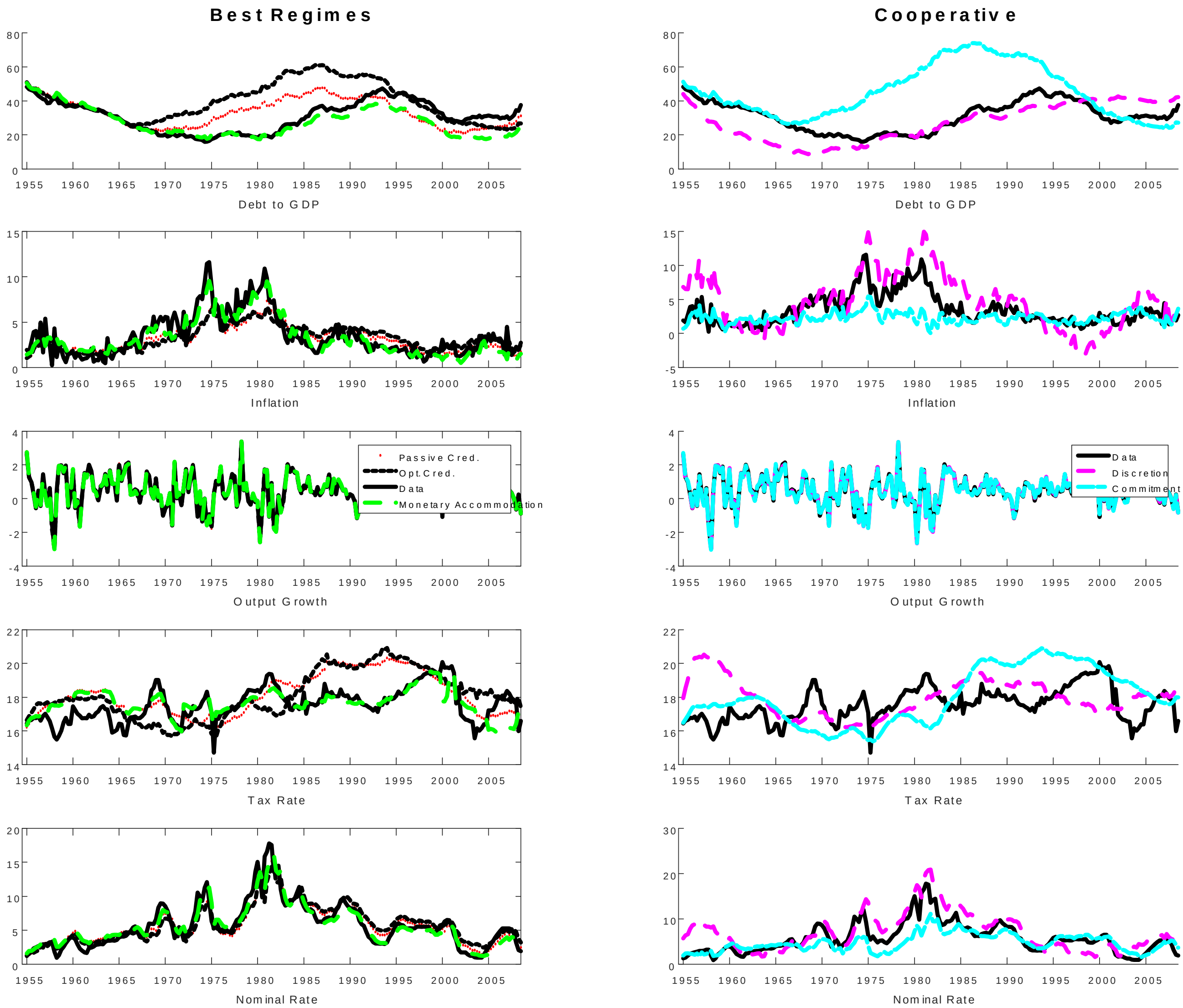

Figure 13: 'Best' Policy Regimes Counterfactual

Table 1 Model Comparison

Models	Geweke(1999)	Bayes factor rel. to of&pf&af	Sims et al.(2008) (q_L statistic)
op&pf&af	-1341.2083	1	-1341.074 (0.178)
pf&af	-1348.4564	$exp\,[7.248]$	-1347.2503 (0.138)
op&pf	-1350.0912	$exp\,[8.883]$	-1349.36 (0.155)
Rules-Based policy	-1417.001	$exp\,[75.793]$	-1417.4422 (0.1007)

Table 2: Rules-Based Policy (Monetary-Fiscal Mix)

Parameters	Mode	5%	95%	Type	Mean	Std Dev
			Active Monetary-Passive Fiscal			
ρ^R,lagged interest rate	0.859	0.833	0.887	Beta	0.50	0.25
ψ_1.interest rate resp. to inflation	2.898	2.583	3.214	Gamma	2.00	0.50
ψ_2,interest rate resp. to output	0.360	0.175	0.550	Gamma	0.50	0.25
ρ^τ,lagged tax rate	0.909	0.874	0.943	Beta	0.50	0.25
δ_τ,tax rate resp. to debt	0.031	0.017	0.045	Gamma	0.07	0.02
δ_y,tax rate resp. to output	0.055	0.003	0.109	Gamma	0.10	0.10
			Active Monetary-Active Fiscal			
ρ^R,lagged interest rate	0.534	0.417	0.638	Beta	0.50	0.25
ψ_1.interest rate resp. to inflation	1.351	1.199	1.506	Gamma	2.00	0.50
ψ_2,interest rate resp. to output	0.460	0.293	0.633	Gamma	0.50	0.25
ρ^τ,lagged tax rate	0.666	0.563	0.761	Beta	0.50	0.25
δ_τ,tax rate resp. to debt	0	-	-	F	-	-
δ_y,tax rate resp. to output	0.055	0.003	0.109	Gamma	0.10	0.10
			Passive Monetary-Passive Fiscal			
ρ^R,lagged interest rate	0.847	0.800	0.892	Beta	0.50	0.25
ψ_1.interest rate resp. to inflation	0.805	0.697	0.917	Gamma	0.80	0.30
ψ_2,interest rate resp. to output	0.347	0.141	0.551	Gamma	0.50	0.25
ρ^τ,lagged tax rate	0.422	0.293	0.553	Beta	0.50	0.25
δ_τ,tax rate resp. to debt	0.079	0.056	0.105	Gamma	0.07	0.02
δ_y,tax rate resp. to output	0.055	0.003	0.109	Gamma	0.10	0.10
			Passive Monetary-Active Fiscal			
ρ^R,lagged interest rate	0.847	0.800	0.892	Beta	0.50	0.25
ψ_1.interest rate resp. to inflation	0.805	0.697	0.917	Gamma	0.80	0.30
ψ_2,interest rate resp. to output	0.347	0.141	0.551	Gamma	0.50	0.25
ρ^τ,lagged tax rate	0.666	0.563	0.761	Beta	0.50	0.25
δ_τ,tax rate resp. to debt	0	-	-	F	-	-
δ_y,tax rate resp. to output	0.055	0.003	0.109	Gamma	0.10	0.10

Table 2: Rules-Based Policy (continued)

Parameters	Mode	5%	95%	Type	Mean	Std Dev
σ,Inv. of intertemp. elas. of subst.	2.577	2.337	2.817	Normal	2.50	0.25
α,Calvo parameter	0.780	0.752	0.810	Beta	0.75	0.02
ζ,inflation inertia	0.214	0.135	0.295	Beta	0.50	0.15
θ,habit persistence	0.605	0.448	0.762	Beta	0.50	0.15
φ,Inverse of Frisch elasticity	2.377	2.132	2.625	Normal	2.50	0.25
Serial correl. and std. of shocks						
ρ^{ξ},AR coeff., taste shock	0.877	0.832	0.925	Beta	0.50	0.15
ρ^{μ},AR coeff., cost-push shock	0.278	0.096	0.456	Beta	0.50	0.15
ρ^{z},AR coeff., productivity shock	0.345	0.250	0.447	Beta	0.50	0.15
ρ^{tr},AR coeff., transfer shock	0.558	0.473	0.648	Beta	0.50	0.15
ρ^{g},AR coeff., government spending	0.981	0.971	0.991	Beta	0.50	0.15
$\sigma_{\xi(s=1)}$,taste shock (L)	0.404	0.280	0.513	Inv. Gamma	0.50	2
$\sigma_{\xi(s=2)}$,taste shock (H)	1.010	0.721	1.284	Inv. Gamma	0.50	2
$\sigma_{\mu(s=1)}$,cost-push shock (L)	2.481	1.425	3.606	Inv. Gamma	0.50	2
$\sigma_{\mu(s=2)}$,cost-push shock (H)	4.603	4.131	5.000	Inv. Gamma	0.50	2
$\sigma_{z(s=1)}$,productivity shock (L)	0.549	0.447	0.651	Inv. Gamma	0.50	2
$\sigma_{z(s=2)}$,productivity shock (H)	1.221	1.045	1.402	Inv. Gamma	0.50	2
σ_{g},government shock	0.248	0.229	0.268	Inv. Gamma	0.50	2
σ_{tr},transfer shock	4.066	3.673	4.449	Inv. Gamma	0.50	2
σ_{τ},tax rate shock	0.356	0.325	0.388	Inv. Gamma	0.50	2
σ_{R},interest rate shock	0.213	0.192	0.234	Inv. Gamma	0.50	2
Transition probs						
p_{11},monetary policy: remaining active	0.972	0.956	0.988	Beta	0.90	0.05
p_{22},monetary policy: remaining passive	0.867	0.833	0.901	Beta	0.90	0.05
q_{11},fiscal policy: remaining passive	0.910	0.883	0.937	Beta	0.90	0.05
q_{22},fiscal policy: remaining active	0.880	0.849	0.914	Beta	0.90	0.05
z_{11},volatility: remaining with low volatility	0.933	0.888	0.978	Beta	0.90	0.05
z_{22},volatility: remaining with high volatility	0.924	0.877	0.973	Beta	0.90	0.05

Table 3: Optimal Policy (Monetary-Fiscal Mix)

Parameters	Mode	5%	95%	Type	Mean	Std Dev
ω_1,gap term, $\hat{X}_t - \hat{\xi}_t$	0.164	0.126	0.210	Beta	0.50	0.15
ω_2,gap term, $\hat{y}_t - \frac{\sigma}{\varphi}\hat{\xi}_t$	0.183	0.141	0.216	Beta	0.50	0.15
Conservative Monetary-Optimal Fiscal						
ω_R,change in interest rate	0.734	0.604	0.879	Beta	0.50	0.15
ω_π,inflation	1	-	-	F	-	-
ω_τ,change in tax rate	0.728	0.564	0.855	Beta	0.50	0.15
ω_π^f,inflation	0.335	0.232	0.445	Gamma	1.00	0.30
Less Conservative Monetary-Optimal Fiscal						
ω_R,change in interest rate	0.734	0.604	0.879	Beta	0.50	0.15
ω_π,inflation	0.442	0.347	0.539	Beta	0.50	0.15
ω_τ,change in tax rate	0.728	0.564	0.855	Beta	0.50	0.15
ω_π^f,inflation	0.335	0.232	0.445	Gamma	1.00	0.30
Conservative Monetary-Passive Fiscal						
ω_R,change in interest rate	0.734	0.604	0.879	Beta	0.50	0.15
ω_π,inflation	1	-	-	F	-	-
ρ^τ,lagged tax rate	0.965	0.956	0.974	Beta	0.50	0.25
δ_τ,tax rate resp. to debt	0.050	0.047	0.053	Gamma	0.07	0.02
δ_y,tax rate resp. to output	0.038	0.001	0.074	Gamma	0.10	0.10
Less Conservative Monetary-Passive Fiscal						
ω_R,change in interest rate	0.734	0.604	0.879	Beta	0.50	0.15
ω_π,inflation	0.442	0.347	0.539	Beta	0.50	0.15
ρ^τ,lagged tax rate	0.965	0.956	0.974	Beta	0.50	0.25
δ_τ,tax rate resp. to debt	0.050	0.047	0.053	Gamma	0.07	0.02
δ_y,tax rate resp. to output	0.038	0.001	0.074	Gamma	0.10	0.10
Conservative Monetary-Active Fiscal						
ω_R,change in interest rate	0.734	0.604	0.879	Beta	0.50	0.15
ω_π,inflation	1	-	-	F	-	-
ρ^τ,lagged tax rate	0.912	0.888	0.937	Beta	0.50	0.25
δ_τ,tax rate resp. to debt	0	-	-	F	-	-
δ_y,tax rate resp. to output	0.038	0.001	0.074	Gamma	0.10	0.10
Less Conservative Monetary-Active Fiscal						
ω_R,change in interest rate	0.734	0.604	0.879	Beta	0.50	0.15
ω_π,inflation	0.442	0.347	0.539	Beta	0.50	0.15
ρ^τ,lagged tax rate	0.912	0.888	0.937	Beta	0.50	0.25
δ_τ,tax rate resp. to debt	0	-	-	F	-	-
δ_y,tax rate resp. to output	0.038	0.001	0.074	Gamma	0.10	0.10

Table 3: Optimal Policy (continued)

Parameters	Mode	5%	95%	Type	Mean	Std Dev
σ,Inv. of intertemp. elas. of subst.	2.913	2.736	3.076	Normal	2.50	0.25
α,Calvo parameter	0.795	0.773	0.817	Beta	0.75	0.02
ζ,inflation inertia	0.270	0.182	0.367	Beta	0.50	0.15
θ,habit persistence	0.563	0.421	0.743	Beta	0.50	0.15
φ,Inverse of Frisch elasticity	2.265	2.128	2.385	Normal	2.50	0.25
ϕ_g,govt.spending rule resp. to debt	-0.015	-0.022	-0.007	Normal	0.00	0.10
Serial correl. and std. of shocks						
ρ^{ξ},AR coeff., taste shock	0.931	0.915	0.947	Beta	0.50	0.15
ρ^{μ},AR coeff., cost-push shock	0.946	0.930	0.962	Beta	0.50	0.15
ρ^{z},AR coeff., productivity shock	0.263	0.193	0.332	Beta	0.50	0.15
ρ^{tr},AR coeff., transfer shock	0.871	0.840	0.901	Beta	0.50	0.15
ρ^{g},AR coeff., government spending	0.983	0.977	0.989	Beta	0.50	0.15
$\sigma_{\xi(s=1)}$,taste shock (L)	0.731	0.127	0.524	Inv. Gamma	0.50	2
$\sigma_{\xi(s=2),}$taste shock (H)	2.545	1.840	3.213	Inv. Gamma	0.50	2
$\sigma_{\mu(s=1),}$cost-push shock (L)	0.467	0.386	0.545	Inv. Gamma	0.50	2
$\sigma_{\mu(s=2),}$cost-push shock (H)	1.408	1.233	1.589	Inv. Gamma	0.50	2
$\sigma_{z(s=1),}$productivity shock (L)	0.685	0.613	0.765	Inv. Gamma	0.50	2
$\sigma_{z(s=2)}$,productivity shock (H)	1.117	0.975	1.246	Inv. Gamma	0.50	2
σ_g,government shock	0.163	0.150	0.176	Inv. Gamma	0.50	2
σ_{tr},transfer shock	1.732	1.612	1.844	Inv. Gamma	0.50	2
σ_{τ},tax rate shock	0.244	0.220	0.266	Inv. Gamma	0.50	2
Transition probs						
p_{11},monetary policy: remaining conservative	0.909	0.886	0.931	Beta	0.90	0.05
p_{22},monetary policy: remaining less conservative	0.921	0.895	0.944	Beta	0.90	0.05
q_{11},fiscal policy: remaining optimal	0.882	0.852	0.914	Dirichlet	0.90	0.05
q_{12},optimal to passive fiscal policy	0.009	0.0004	0.017	Dirichlet	0.05	0.05
q_{22},fiscal policy: remaining passive	0.962	0.946	0.979	Dirichlet	0.90	0.05
q_{23},passive to active fiscal policy	0.006	0.0001	0.012	Dirichlet	0.05	0.05
q_{33},fiscal policy: remaining active	0.922	0.898	0.945	Dirichlet	0.90	0.05
q_{31},active to optimal fiscal policy	0.005	0.0001	0.010	Dirichlet	0.05	0.05
z_{11},volatility: remaining with low volatility	0.956	0.935	0.978	Beta	0.90	0.05
z_{22},volatility: remaining with high volatility	0.893	0.865	0.925	Beta	0.90	0.05

Table 4: Unconditional Variances and Welfare With Regime Switching

Regime	Output	Inflation	Interest Rate	Tax Rate	Welfare Cost
			Low Volatility		
PF/LC	1.22	0.75	0.66	6.23	2.60
MA	1.40	0.33	0.40	5.31	2.60
PF/MC	1.4	0.34	0.60	7.51	2.61
FA	1.51	0.45	0.50	4.00	2.62
AF/LC	3.92	1.95	1.36	0.004	3.00
OF/MC	4.89	0.78	0.87	2.06	3.06
AF/MC	7.64	0.97	0.93	0.01	3.23
OF/LC	5.21	2.28	1.56	3.58	3.43
			High Volatility		
PF/LC	1.40	1.25	1.03	6.73	24.68
MA	1.63	0.58	0.72	5.56	24.69
PF/MC	1.69	0.56	0.92	7.84	24.75
FA	1.75	0.79	0.84	4.20	24.75
AF/LC	4.24	2.50	1.76	0.004	25.24
AF/MC	8.08	1.20	1.29	0.01	25.49
OF/MC	6.85	1.19	1.38	6.01	26.20
OF/LC	7.15	3.41	2.35	9.03	27.16

Key to the Table: AF - Active Fiscal, PF - Passive Fiscal, OF - Optimal Fiscal Policy (Stackelberg Leadership), MC - More Conservative Optimal Monetary, LC - Less Conservative Optimal Monetary, MA - Monetary Policy Accommodates Fiscal Regime, FA - Fiscal Policy Accommodates Monetary Regime.

Table 5: Unconditional Variances and Welfare with Full Credibility

Regime	Output	Inflation	Interest Rate	Tax Rate	Welfare Cost
			Low Volatility		
Commitment	1.37	0.06	0.16	3.53	2.45
OF/MC	1.23	0.16	0.24	4.00	2.48
PF/MC	1.34	0.24	0.33	7.08	2.56
OF/LC	0.93	0.70	0.69	4.71	2.58
PF/LC	1.23	1.17	1.15	6.59	2.70
Discretion	1.03	1.52	1.37	6.46	2.75
AF/MC	3,64	0.36	0.49	0.0004	2.84
AF/LC	3.36	2.10	2.23	0.0004	3.00
			High Volatility		
Commitment	1.63	0.16	0.50	4.58	24.28
OF/MC	1.43	0.31	0.52	4.66	24.46
PF/MC	1.60	0.42	0.64	7.40	24.62
OF/LC	1.09	1.31	1.11	5.20	24.71
PF/LC	1.43	1.81	1.64	7.13	24.91
AF/MC	3.95	0.52	0.80	0.0004	24.96
Discretion	1.20	2.71	2.14	6.92	25.24
AF/LC	3.66	2.63	2.69	0.0004	25.28

Key to the Table: AF - Active Fiscal, PF - Passive Fiscal, OF - Optimal Fiscal Policy (Stackelberg Leadership), MC - More Conservative Optimal Monetary, LC - Less Conservative Optimal Monetary, Commitment - Cooperative Ramsey Policy, Discretion - Cooperative Time-Consistent Optimal Policy.