Online Appendix

Self-Fulfilling Debt Crises:
Can Monetary Policy Really Help?

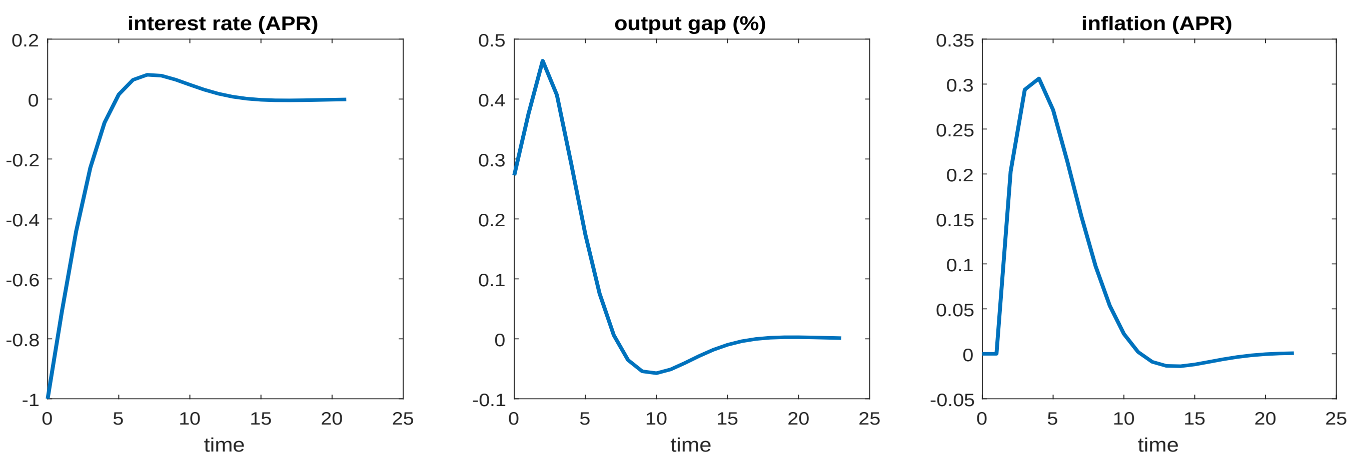
Philippe Bacchetta, Elena Perazzi and Eric van Wincoop

May 2017

This online Appendix has 5 sections. Section I shows that the effect of a standard monetary shock in the NK model with our benchmark calibration. Section II shows that for s_{high} large enough, there will not be default when $\tilde{s} = s_{high}$, as assumed in section 2.1 of the paper. Sections III derives the NK Phillips curve. Section IV discusses the case of uncertainty about T, briefly mentioned in section 2.1 of the paper. Finally, Section V discusses the relationship between equilibrium output and expected interest rates, discussed in section 4.1 of the paper.

I Effect of a Standard Monetary Shock

The plots below show the response of the interest rate, output and inflation to a 1% negative monetary shock at $t = 0$.



II Minimum Level of s_{high}

As mentioned in Section 2.1, we assume that the primary surplus s_{high} in the good state is sufficiently high such that default never happens in that state. We derive a condition for s_{high} under which this is the case. We only do so in the non-monetary LW model of Section 2.1. If $b_T > s_{high}/(R-1)$, there is default even when $\tilde{s} = s_{high}$. When $b_T = s_{high}/(R-1)$ the price schedule then drops down a second time, to

$$Q_{T-1} = \frac{\zeta}{(R-1)b_T}\left(\psi s_{low} + (1-\psi)s_{high}\right) \tag{1}$$

We need to show that there can be a level of s_{high} such that there is no equilibrium with $b_T > s_{high}/(R-1)$. This is the case if the pricing schedule is always above the debt accumulation schedule. For a given $b_T > s_{high}/(R-1)$, Q_{T-1} from the pricing schedule must be higher than from the debt accumulation schedule. This is the case when

$$\frac{\zeta}{(R-1)b_T}\left(\psi s_{low} + (1-\psi)s_{high}\right) > \frac{\chi^{\kappa}\kappa b_0 - \chi^{s}\overline{s}}{b_T - (1-\delta)^T b_0} \tag{2}$$

If this is the case for $b_T = s_{high}/(R-1)$, it also holds for larger b_T, leading to the condition

$$\frac{\zeta}{(R-1)}\left(\psi s_{low} + (1-\psi)s_{high}\right) > \frac{\chi^{\kappa}\kappa b_0 - \chi^{s}\overline{s}}{1 - (R-1)(1-\delta)^T b_0/s_{high}} \tag{3}$$

Since the left-hand side of this expression depends positively on s_{high} and goes to infinity when $s_{high} \to \infty$, while the right hand side depends negatively on s_{high} and goes to a constant when $s_{high} \to \infty$, it follows that for s_{high} above some cutoff level this condition is always satisfied.

III Derivation of the NK Phillips Curve

We derive the optimal price set by firms, taking into account all the features of our model, in particular: i) habit formation by households; ii) price indexation; iii) lagged response in price adjustment. Another interpretation of the latter feature is that firms base they pricing decision on information that is d periods old.

The consumption-leisure tradeoff equation for households, derived from the utility function (26) in the paper, is

$$\frac{N_t(i)^{\phi}}{(C_t - \eta C_{t-1})^{-\sigma} - \eta\beta E_t(C_{t+1} - \eta C_t)^{-\sigma}} = \frac{W_t(i)}{P_t} \tag{4}$$

Notice that the labor supply $N_t(i)$ and the wage $W_t(i)$ are firm-specific.

Our assumptions about information delay imply that at time $t-d$ firms choose the relative price $h_t^*(i) \equiv P_t^*(i)/P_t$. Due to our assumption of indexation, if the firm optimizes its prices at time t and does not re-optimize again in subsequent periods, the price at $t+k$ is

$$P_{t+k}(i) = h_t^*(i) P_t \left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{\gamma} \tag{5}$$

Profits at $t+k$ for firm i are

$$\Pi_{t+k}(i) = h_t^*(i) P_t \left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{\gamma} Y_{t+k}(i) - W_{t+k}(i)\left(\frac{Y_{t+k}(i)}{A}\right)^{\frac{1}{1-\alpha}} \tag{6}$$

Demand for firm i's goods is equal to

$$Y_{t+k}(i) = \left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{-\gamma\varepsilon} \left(\frac{h_t^*(i) P_t}{P_{t+k}}\right)^{-\varepsilon} Y_{t+k} \tag{7}$$

where Y_{t+k} is aggregate output.

Maximizing the expected present discounted value of profits

$$E_{t-d} \sum_{k=0}^{\infty} \theta^k Q_{t,t+k} \Pi_{t+k}(i)$$

with respect to $h_t^*(i)$ we obtain

$$E_{t-d} \sum_{k=0}^{\infty} \theta^k Q_{t,t+k} \frac{P_t^{-\varepsilon} Y_{t+k}}{P_{t+k}^{-\varepsilon}} \left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{\gamma(1-\varepsilon)} \left[1 - \mu \frac{1}{h_t^*(i)} \frac{P_{t+k}}{P_t} \left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{-\gamma} MC_{t+k}(i)\right] = 0 \tag{8}$$

where $\mu = \varepsilon/(\varepsilon - 1)$ is the steady state markup, $MC_{t+k}(i)$ is the real marginal cost of production, $Q_{t,t+k}$ is the stochastic discount factor from t to $t+k$, and d is the information processing delay. The real marginal cost of production $M_{t+k}(i)$ is defined as the derivative of the real labor cost $W_{t+k}(i) N_{t+k}(i)/P_{t+k}$ with respect to $Y_{t+k}(i)$.

It follows from (8) that in a zero-inflation steady state the marginal cost is $1/\mu$. Linearizing around the steady state, and using $Q_{t,t+k} = \beta^k$ in steady state, we obtain

$$E_{t-d} \sum_{k=0}^{\infty} (\beta\theta)^k \left[h_t^*(i) - (p_{t+k} - p_t) + \gamma(p_{t+k-1} - p_{t-1}) - mc_{t+k}(i)\right] = 0 \qquad (9)$$

where lower case variables denote logs in deviation from their steady state. (9) can be rewritten as

$$h_t^*(i) = (1 - \beta\theta) E_{t-d} \sum_{k=0}^{\infty} (\beta\theta)^k \left[(p_{t+k} - p_t) - \gamma(p_{t+k-1} - p_{t-1}) + mc_{t+k}(i)\right] \qquad (10)$$

Using the production function, we can rewrite (4) as

$$\frac{W_{t+k}(i)}{P_{t+k}} = \frac{N_{t+k}(i)^{\phi}}{\tilde{C}_{t+k}} = \frac{(Y_{t+k}(i)/A)^{\phi/(1-\alpha)}}{(Y_{t+k} - \eta Y_{t+k-1})^{-\sigma} - \eta\beta E_{t+k}(Y_{t+k+1} - \eta Y_{t+k})^{-\sigma}} \qquad (11)$$

The total real cost of production, $W_{t+k}(i)N_{t+k}(i)/P_{t+k}$, is

$$\frac{(Y_{t+k}(i)/A)^{(\phi+1)/(1-\alpha)}}{(Y_{t+k} - \eta Y_{t+k-1})^{-\sigma} - \eta\beta E_{t+k}(Y_{t+k+1} - \eta Y_{t+k})^{-\sigma}} \qquad (12)$$

and the marginal cost is

$$MC_{t+k} = \frac{\phi + 1}{1 - \alpha} \frac{1}{A} \frac{(Y_{t+k}(i)/A)^{(\phi+\alpha)/(1-\alpha)}}{(Y_{t+k} - \eta Y_{t+k-1})^{-\sigma} - \eta\beta E_{t+k}(Y_{t+k+1} - \eta Y_{t+k})^{-\sigma}} \qquad (13)$$

When prices are all the same and constant, $Y_{t+k} = Y_{t+k}(i)$ is equal to the natural rate Y_{t+k}^n, and the marginal cost is equal to $1/\mu$. From (13) we can then solve for the natural rate of output, which is constant as productivity is constant.

Using

$$y_{t+k}(i) = y_{t+k} - \varepsilon h_t^* + \varepsilon(p_{t+k} - p_t) - \gamma\varepsilon(p_{t+k-1} - p_{t-1})$$

which is the log-linearized version of(7), we obtain the log-linearized version of (13) in the form

$$\begin{aligned} mc_{t+k} \quad = \quad & \frac{\phi + \alpha}{1 - \alpha} x_{t+k} + \frac{\sigma}{(1 - \beta\eta)(1 - \eta)} \tilde{x}_{t+k} - \frac{(\phi + \alpha)\varepsilon}{1 - \alpha} h_t^* \\ & + \frac{(\phi + \alpha)\varepsilon}{1 - \alpha}(p_{t+k} - p_t) - \frac{(\phi + \alpha)\varepsilon}{1 - \alpha}\gamma(p_{t+k-1} - p_{t-1}) \qquad (14) \end{aligned}$$

where, as in the paper, $x_t \equiv y_y - y_t^n$ and $\tilde{x}_t \equiv x_t - \eta x_{t-1} - \beta\eta E_t(x_{t+1} - \eta x_t)$. Substituting (14) in (10) we obtain

$$h_t^* = (1-\beta\theta)E_{t-d}\sum_{k=0}^{\infty}(\beta\theta)^k\left[(p_{t+k}-p_t) - \gamma(p_{t+k-1}-p_{t-1}) + \zeta_1 x_{t+k} + \zeta_2\tilde{x}_{t+k}\right] \tag{15}$$

where

$$\zeta_1 = \frac{\phi+\alpha}{1-\alpha}\frac{1-\alpha}{1-\alpha+(\alpha+\phi)\varepsilon}$$
$$\zeta_2 = \frac{\sigma}{(1-\eta\beta)(1-\eta)}\frac{1-\alpha}{1-\alpha+(\alpha+\phi)\varepsilon}$$

We can write (15) as

$$h_t^* = \theta\beta E_{t-d}h_{t+1}^* + \theta\beta E_{t-d}\pi_{t+1} - \gamma\theta\beta E_{t-d}\pi_t + (1-\theta\beta)\zeta_1 E_{t-d}x_t + (1-\theta\beta)\zeta_2 E_{t-d}\tilde{x}_t) \tag{16}$$

Let $p_t^* = h_t^* + p_t$ be the log price at t of the firms that re-optimize. (16) implies

$$\begin{aligned} p_t^* - p_t &= \theta\beta E_{t-d}(p_{t+1}^* - p_{t+1}) + \theta\beta E_{t-d}\pi_{t+1} - \gamma\theta\beta E_{t-d}\pi_t \\ &+ (1-\theta\beta)\zeta_1 E_{t-d}x_t + (1-\theta\beta)\zeta_2 E_{t-d}\tilde{x}_t \end{aligned} \tag{17}$$

which can also be written as

$$\begin{aligned} p_t^* - p_{t-1} &= \theta\beta E_{t-d}(p_{t+1}^* - p_t) + \pi_t - \gamma\theta\beta E_{t-d}\pi_t \\ &+ (1-\theta\beta)\zeta_1 E_{t-d}x_t + (1-\theta\beta)\zeta_2 E_{t-d}\tilde{x}_t \end{aligned} \tag{18}$$

If $\bar{p}_t$ denotes the price level of firms that do not re-optimize, we have

$$p_t = \theta\bar{p}_t + (1-\theta)p_t^* \tag{19}$$

Subtracting p_{t-1}

$$\pi_t = \theta(\bar{p}_t - p_{t-1}) + (1-\theta)(p_t^* - p_{t-1}) \tag{20}$$

Firms that do not re-optimize index to past inflation, so that $\bar{p}_t - p_{t-1} = \gamma\pi_{t-1}$. This gives

$$\pi_t = \theta\gamma\pi_{t-1} + (1-\theta)(p_t^* - p_{t-1}) \tag{21}$$

Taking the expectation at $t-d$ of the same expression one period later, multiplied by $\theta\beta$, we obtain

$$\theta\beta E_{t-d}\pi_{t+1} = \theta^2\gamma\beta E_{t-d}\pi_t + (1-\theta)\theta\beta E_{t-d}(p_{t+1}^* - p_t) \tag{22}$$

Taking the difference between the last two equations and substituting (18) we obtain

$$\pi_t = \gamma\pi_{t-1} + \beta E_{t-d}\pi_{t+1} - \gamma\beta E_{t-d}\pi_t + \omega_1 E_{t-d}x_t + \omega_2 E_{t-d}\tilde{x}_t \tag{23}$$

where

$$\omega_1 = \frac{1-\theta}{\theta}(1-\theta\beta)\frac{\phi+\alpha}{1-\alpha+(\alpha+\phi)\varepsilon} \tag{24}$$

$$\omega_2 = \frac{1-\theta}{\theta}(1-\theta\beta)\frac{1-\alpha}{1-\alpha+(\alpha+\phi)\varepsilon}\frac{\sigma}{(1-\eta\beta)(1-\eta)} \tag{25}$$

This is the NK Phillips curve.

IV Uncertainty about the Date of Default Decision

We have assumed that the only uncertainty in the model is about the level of primary surpluses that can be generated from T onward. In other words, there is uncertainty about whether the government is able to enact reforms that raise the primary surplus. But this uncertainty is resolved at a known date and the default decision is then made at that time. We will now discuss an extension whereby there is uncertainty about T itself.

In general there can be uncertainty about both the date that we find out if reforms will be enacted and about the reforms themselves. We now abstract from the latter by setting $\psi = 1$. In this case the agents know that there will be no reform that raises primary surpluses, but they do not know at what time a decision will be made to default or not. We further simplify by considering only two possible dates for the default decision. The default decision will take place at T_1 with probability p and at T_2 with probability $1-p$, with $T_1 < T_2$. The asset price prior to T_1 now takes into account the possibility of default at either T_1 or T_2.

Monetary policy now takes the following form. The central bank chooses interest rates R_0 to R_{T_1-1}. After that, if a default decision is made at time T_1, the central bank chooses interest rates $\tilde{R}_{T_1}, ..., \tilde{R}_H$ and after that the Taylor rule applies. If no default decision is made at time T_1, the central bank chooses interest rates $R_{T_1}, ..., R_H$ and after that the Taylor rule applies. Central bank policy starting at T_1 will therefore depend on whether a default decision is made at time T_1.

Without uncertainty about T, the debt accumulation and pricing schedules took the form of two relationships between b_T and Q_{T-1}. They now take the form of two relationships between b_{T_1} and Q_{T_1-1}. The debt accumulation schedule is the same as before, replacing T with T_1:

$$b_{T_1} = (1-\delta)^{T_1}\frac{B_0}{P_{T_1}} + \frac{P_{T_1-1}}{P_{T_1}}\frac{\chi^\kappa \kappa B_0/P_0 - \chi^s \bar{s}}{Q_{T_1-1}} \tag{26}$$

where

$$\begin{aligned} \chi^\kappa &= \left[r_{T_1-2}...r_1 r_0 + (1-\delta) r_{T_1-2}...r_1 \frac{P_0}{P_1} + ... + (1-\delta)^{T_1-1}\frac{P_0}{P_{T_1-1}} \right] \\ \chi^s &= 1 + r_{T_1-2} + r_{T_1-2} r_{T_1-3} + ... + r_{T_1-2}...r_1 r_0 \end{aligned}$$

The pricing schedule is now more complex as we need to take into account possible default at T_2. Starting at date T_1, there may be multiple equilibria if the debt is in an intermediate range. In this case there is either a self-fulfilling default at T_2 or no default at T_2. But agents now need to make an assumption prior to T_1 about which of these two equilibria will be picked if both equilibria exist and there is no default decision at T_1. We will consider the worst case scenario under which agents belief that the default equilibrium will be picked if it exists.

Starting at date T_1, the problem is like the one discussed in Section 2 as there can only be default at T_2 if no default decision is made at T_1. Using the results from section 2, the lowest the lowest value of debt for which there is a default equilibrium is

$$b_{low} = \frac{\zeta s^{pdv}(T_2) + r_{T_2-1}\bar{\chi}^s \bar{s}}{(P_{T_1}/P_{T_2})(1-\delta)^{T_2-T_1}((1-\delta)Q_{T_2}+\kappa)\zeta + r_{T_2-1}\bar{\chi}^\kappa \kappa} \tag{27}$$

where

$$\begin{aligned} \bar{\chi}^\kappa &= \left[r_{T_2-2}...r_{T_1} + (1-\delta) r_{T_2-2}...r_{T_1+1}\frac{P_{T_1}}{P_{T_1+1}} + ... + (1-\delta)^{T_2-1}\frac{P_{T_1}}{P_{T_2-1}} \right] \\ \bar{\chi}^s &= 1 + r_{T_2-2} + r_{T_2-2} r_{T_2-3} + ... + r_{T_2-2}...r_{T_1+1} r_{T_1} \\ s^{pdv}(T_2) &= \left[1 + \frac{1}{r_{T_2}} + \frac{1}{r_{T_2} r_{T_2+1}} + ... \right] s_{low} \\ Q_{T_2} &= \frac{\kappa}{R_{T_2}} + \frac{(1-\delta)\kappa}{R_{T_2} R_{T_2+1}} + \frac{(1-\delta)^2\kappa}{R_{T_2} R_{T_2+1} R_{T_2+2}} + ... \end{aligned}$$

The price at $T_1 - 1$ will be equal to

$$Q_{T-1} = \frac{(1-\delta)(p\tilde{H}_{T_1} + (1-p)\bar{Q}_{T_1}) + \kappa}{R_{T_1-1}} \tag{28}$$

where $\tilde{H}_{T_1}$ is the payoff at time T_1 when a default decision is made at T_1 and $\bar{Q}_{T_1}$ is the asset price at T_1 if no default decision is made at T_1. In what follows we derive expressions for $\tilde{H}_{T_1}$ and $\bar{Q}_{T_1}$. These will be functions of b_{T_1-1}, which therefore delivers the price schedule.

First consider $\tilde{H}_{T_1}$. If the default decision takes place at T_1, we have

$$\begin{aligned} s^{pdv}(T_1) &= \left[1 + \frac{1}{\tilde{r}_{T_1}} + \frac{1}{\tilde{r}_{T_1}\tilde{r}_{T_1+1}} + ...\right] s_{low} \\ \tilde{Q}_{T_1} &= \frac{\kappa}{\tilde{R}_{T_1}} + \frac{(1-\delta)\kappa}{\tilde{R}_{T_1}\tilde{R}_{T_1+1}} + \frac{(1-\delta)^2\kappa}{\tilde{R}_{T_1}\tilde{R}_{T_1+1}\tilde{R}_{T_1+2}} + ... \end{aligned}$$

We have

$$\begin{aligned} \tilde{H}_{T_1} &= \tilde{Q}_{T_1} \qquad \text{if } b_{T_1} \leq \frac{s^{pdv}(T_1)}{(1-\delta)\tilde{Q}_{T_1} + \kappa} && (29) \\ &= \frac{\zeta s^{pdv}(T_1)}{b_{T_1}} \quad \text{if } b_{T_1} > \frac{s^{pdv}(T_1)}{(1-\delta)\tilde{Q}_{T_1} + \kappa} && (30) \end{aligned}$$

Next consider $\bar{Q}_{T_1}$. This is the asset price at T_1 if there is no default decision at T_1. When $b_{T_1} \leq b_{low}$, there will be no default at T_2 and

$$\bar{Q}_{T_1} = \frac{\kappa}{R_{T_1}} + \frac{(1-\delta)\kappa}{R_{T_1}R_{T_1+1}} + \frac{(1-\delta)^2\kappa}{R_{T_1}R_{T_1+1}R_{T_1+2}} + ... \tag{31}$$

When $b_{T_1} > b_{low}$, there will be default at T_2. The first step is to compute b_{T_2} based on debt accumulation from T_1 to T_2. Using the results from section 2 of the paper we have

$$b_{T_2} = (1-\delta)^{T_2-T_1}\frac{P_{T_1}}{P_{T_2}}b_{T_1} + \frac{P_{T_2-1}}{P_{T_2}}\frac{\chi^\kappa \kappa b_{T_1} - \chi^s \bar{s}}{Q_{T_2-1}} \tag{32}$$

where

$$\begin{aligned} \chi^\kappa &= \left[r_{T_2-2}...r_{T_1+1}r_{T_1} + (1-\delta)r_{T_2-2}...r_{T_1+1}\frac{P_{T_1}}{P_{T_1+1}} + ... + (1-\delta)^{T_2-T_1-1}\frac{P_{T_1}}{P_{T_2-1}}\right] \\ \chi^s &= 1 + r_{T_2-2} + r_{T_2-2}r_{T_2-3} + ... + r_{T_2-2}...r_{T_1+1}r_{T_1} \end{aligned}$$

When then have

$$Q_{T_2-1} = \frac{\zeta s^{pdv}(T_2)}{b_{T_2}} \tag{33}$$

Using

$$Q_t = \frac{(1-\delta)Q_{t+1} + \kappa}{R_t}$$

we can then work backward to obtain Q_{T_1}, which gives us $\bar{Q}_{T_1}$.

The pricing schedule drops down twice, at $b_{T_1} = s^{pdv}(T_1)/[(1-\delta)\tilde{Q}_{T_1} + \kappa]$ and at $b_{T_1} = b_{low}$. In order to avoid an equilibrium with default, the debt accumulation schedule must cross below the pricing schedule and therefore below these two points where the pricing schedule drops down. We cannot tell a priori which of these two points will be binding and therefore follow the following approach. We first maximize utility subject to the constraint that at $b_{T_1} = s^{pdv}(T_1)/[(1-\delta)\tilde{Q}_{T_1} + \kappa]$ the asset price from the pricing schedule is equal to that from the debt accumulation schedule. We check that at the same time the asset price from the debt accumulation schedule is no larger than from the pricing schedule at $b_{T_1} = b_{low}$. If this is not the case, it does not represent optimal policy as there will exist a default equilibrium. We next maximize utility subject to the constraint that at $b_{T_1} = b_{low}$ the asset price from both schedules is the same. We then check that at $b_{T_1} = s^{pdv}(T_1)/[(1-\delta)\tilde{Q}_{T_1} + \kappa]$ the asset price from the debt accumulation schedule is no larger than from the pricing schedule. If this is not the case, there exists a default equilibrium. There is always an optimal policy solution that avoids default under one of these two constraints, which represents the binding constraint.

Figure A1 provides an illustration for the case where $T_1 = 10$ and $T_2 = 20$. Except for $\psi = 1$, all other parameters are the same as in the benchmark parameterization. The chart on the left shows the maximum inflation rate under optimal policy for different values of B_0, while the chart on the right shows the ultimate price level as a result of the optimal policy. The charts also show results for the case where $T = 10$ and $T = 20$ without uncertainty, and $\psi = 1$. The thick section on the horizontal axis represents the range of B_0 for which there are multiple equilibria under uncertainty about T in the absence of monetary policy, which is 0.84 to 1.47.

The case of uncertainty lies between the two cases without uncertainty. The range of B_0 for which there are multiple equilibria is shifted to somewhere in between the two cases without uncertainty, but otherwise the results remain very similar to those without uncertainty. Unless B_0 is very close to the lowest value for which there are multiple equilibria, it remains the case that very significant inflation is needed to avoid multiple equilibria. For example, when B_0 is 1.42

times B_{low}, which is 1.19 and again near the middle of the multiplicity range, the maximum inflation rate is 20%, while ultimately the price level will increase by a factor 3.9.

V Interest Rates and Output

There is a limit to how much real interest rates can change and therefore contribute to help avoid self-fulfilling crises. In the standard model without habit formation the consumption Euler equation, together with $c_t = y_t$, implies

$$y_t = y_{t+1} - \frac{1}{\sigma} r_t \tag{34}$$

where $r_t = i_t - E_t \pi_{t+1} - r^n$ is the real interest rate minus the natural rate. This implies

$$y_0 = -\frac{1}{\sigma} E_0 \sum_{t=0}^{\infty} r_t \tag{35}$$

If there are large changes in the real interest rate, they will need to be reversed as the sum of all the changes in the real interest rate is equal to $-\sigma y_t$ and y_t cannot change by too much in the first quarter of a change in policy (if it does, the model should not be taken seriously).

With habit formation the last expression needs to be modified a bit. In this section we derive the modified expression. The consumption Euler equation, together with $c_t = y_t$, implies

$$\tilde{y}_t = E_t \tilde{y}_{t+1} - \frac{1-\beta\eta}{\sigma} r_t \tag{36}$$

where

$$\tilde{y}_t = y_t - \eta y_{t-1} - \beta\eta E_t (y_{t+1} - \eta y_t) \tag{37}$$

Collecting terms, and removing the expectation operator for future variables in what follows to save on notation, we have

$$(1 + \eta + \beta\eta^2) y_t - \eta y_{t-1} - (1 + \beta\eta + \beta\eta^2) y_{t+1} + \beta\eta y_{t+2} = -\frac{1-\beta\eta}{\sigma} r_t \tag{38}$$

or

$$y_{t-1} = \frac{1 + \eta + \beta\eta^2}{\eta} y_t - \frac{1 + \beta\eta + \beta\eta^2}{\eta} y_{t+1} + \beta y_{t+2} + \frac{1-\beta\eta}{\eta\sigma} r_t \tag{39}$$

We can write this as

$$\begin{pmatrix} y_{t+1} \\ y_t \\ y_{t-1} \end{pmatrix} = A \begin{pmatrix} y_{t+2} \\ y_{t+1} \\ y_t \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{1-\beta\eta}{\eta\sigma} r_t \end{pmatrix} \tag{40}$$

where

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \beta & -\frac{1+\beta\eta+\beta\eta^2}{\eta} & \frac{1+\eta+\beta\eta^2}{\eta} \end{pmatrix} \tag{41}$$

Define

$$a = -A^{-1} \begin{pmatrix} 0 \\ 0 \\ \frac{1-\beta\eta}{\eta\sigma} \end{pmatrix} \tag{42}$$

Then

$$\begin{pmatrix} y_{t+2} \\ y_{t+1} \\ y_t \end{pmatrix} = A^{-1} \begin{pmatrix} y_{t+1} \\ y_t \\ y_{t-1} \end{pmatrix} + a r_t \tag{43}$$

We can diagonalize the matrix $(A^{-1})'$:

$$(A^{-1})' = P \Delta P^{-1} \tag{44}$$

Then

$$A^{-1} = (P')^{-1} \Delta P' \tag{45}$$

Define $\tilde{a} = P'a$. Then

$$P' \begin{pmatrix} y_{t+2} \\ y_{t+1} \\ y_t \end{pmatrix} = \Delta P' \begin{pmatrix} y_{t+1} \\ y_t \\ y_{t-1} \end{pmatrix} + \tilde{a} r_t \tag{46}$$

Define

$$x_t = P' \begin{pmatrix} y_{t+1} \\ y_t \\ y_{t-1} \end{pmatrix} \tag{47}$$

Then we have

$$x_{t+1} = \Delta x_t + \tilde{a} r_t \tag{48}$$

It turns out that the first and second eigenvalues are explosive (the first one is larger than 1 and the second one is 1). Call these λ_1 and λ_2. Then the first and second elements of the difference equation above are

$$x_{t+1}(1) = \lambda_1 x_t(1) + \tilde{a}(1) r_t \tag{49}$$

$$x_{t+1}(2) = \lambda_2 x_t(2) + \tilde{a}(2) r_t \tag{50}$$

Solving these gives

$$x_t(1) = -\tilde{a}(1) \sum_{i=0}^{\infty} \frac{1}{\lambda_1^{i+1}} r_{t+i} \tag{51}$$

$$x_t(2) = -\tilde{a}(2) \sum_{i=0}^{\infty} \frac{1}{\lambda_2^{i+1}} r_{t+i} \tag{52}$$

At time zero we can write

$$x_0(1) = -\tilde{a}(1) \sum_{t=0}^{\infty} \frac{1}{\lambda_1^{t+1}} r_t \tag{53}$$

$$x_0(2) = -\tilde{a}(2) \sum_{t=0}^{\infty} \frac{1}{\lambda_2^{t+1}} r_t \tag{54}$$

Assume that y is zero at time -1 (the shock happens at time 0). Define $Z = (P'[1:2,1:2])^{-1}$. It then follows that

$$y_0 = -\sum_{t=0}^{\infty} \left(Z[2,1]\tilde{a}(1) \frac{1}{\lambda_1^{t+1}} + Z[2,2]\tilde{a}(2) \frac{1}{\lambda_2^{t+1}} \right) r_t \tag{55}$$

Numerically this is equal to (for benchmark parameters):

$$y_0 = -0.58r_0 - 0.73r_1 - 0.83r_2 - 0.89r_3 - 0.93r_4 - 0.95r_5 - 0.97r_6 - 0.98r_7 - 0.99r_8 - \ldots \tag{56}$$

Further coefficients are very close to -1. Except for the first two or three coefficients, there is not a lot of difference relative to the case without habit formation. The first couple of coefficients are smaller as consumption responds less with habit formation. This allows for larger changes in the equilibrium interest rate for a given change in y_0. But it remains the case that the sum of the interest rates cannot change much. For the benchmark parameterization $y_0 = 0.0157$ corresponds to the sum above when we plug in the solution for the real interest rates (everything is in deviation from steady state). Much larger real interest changes imply that output in the first quarter rises by much more than 1.57% (6.3% on an annualized basis), which is implausible.