Part 5: Monetary and Fiscal Policy

Chapters

17 Fiscal Policy I: Public Debt and the Effectiveness of Fiscal Policy 261
18 Fiscal Policy II: The Long-Run Determinants of Fiscal Policy 279
19 Monetary Policy: An Introduction 295
20 Rules vs Discretion 315
21 Recent Debates in Monetary Policy 323
22 New Developments in Monetary and Fiscal Policy 345

LSE Press
ADVANCED MACRO-ECONOMICS
an easy guide
Filipe Campante, Federico Sturzenegger and Andrés Velasco

Filipe Campante,
Federico Sturzenegger
and Andrés Velasco

Advanced Macro-economics

LSE Press

Suggested citation: Campante, Filipe; Sturzenegger, Federico; and Velasco, Andrés. (2021) 'Part 5: Monetary and Fiscal Policy', in Filipe Campante, Federico Sturzenegger and Andrés Velasco, *Advanced Macro-Economics: An Easy Guide*. London: LSE Press, 2021. Chapters 17–22. https://doi.org/10.31389/lsepress.ame.5

Filipe Campante, Federico Sturzenegger and Andrés Velasco 2021. Published under a Creative Commons Attribution NonCommercial licence allowing users to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license does not allow for commercial use.

Monetary and Fiscal Policy

CHAPTER 17

Fiscal policy I: Public debt and the effectiveness of fiscal policy

We have now reached the point at which we can use our hard-earned theoretical tools to study in depth the main policy instruments of macroeconomics, and eventually their long-run and short-run effects. We will start by considering fiscal policy, which means we will now address some of the most pressing policy questions in both developing and developed countries: does an increase in government expenditure increase output, or does it reduce it? What are the determinants of fiscal policy? What are the views?

Fiscal policy can always be thought of under two lights. On one level, it is a component of aggregate demand and will interest us as a tool of macroeconomic stabilisation; on the other hand, it also has a role in terms of determining the provision of public goods which brings in a whole host of other long-run considerations associated with the quality of that spending.

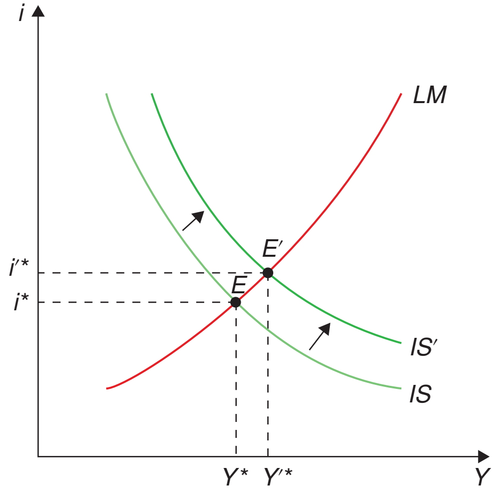
As a tool of macroeconomic stabilisation it is fundamentally different from monetary policy because fiscal policy requires resources, in other words, it needs to be financed. The result is that whatever expansion is obtained by fiscal spending will be diluted by the negative effect produced by the fact that it needs to obtain resources from the economy to finance itself. As a result it will always unavoidably put in motion a countervailing effect that must be taken into account.

This chapter will deal with the first role of fiscal policy, its role for demand management, which happens to be where we typically have our first encounter with the topic at the undergraduate level. In fact, our treatment of fiscal policy illustrates how the things we learn in undergraduate-level macro can be misleading. For instance, in the traditional Keynesian rendition, fiscal policy helps stabilise output – you may remember that an increase in G moves the IS to the right (see Figure 17.1). But this "undergraduate level" analysis is incomplete because it assumes that private consumption or investment are not affected by the increased expenditure. But how is the expenditure financed? Does this financing affect other components of aggregate demand? Imagine a permanent increase in expenditures financed with taxes. Our model of permanent income would anticipate a one to one decrease in private consumption, quite the opposite of a consumption function that is a rigid function of income.

In fact, if we ignore the financing side of the expenditure (be it through taxes or debt) it is evident that some budget constraint will be violated. Either the government is spending what it does not have, or consumers are spending what they cannot afford. Once we include the complete picture we realise

How to cite this book chapter:
Campante, F., Sturzenegger, F. and Velasco, A. 2021. *Advanced Macroeconomics: An Easy Guide.*
Ch. 17. 'Fiscal policy I: Public debt and the effectiveness of fiscal policy', pp. 261–278. London: LSE Press.
DOI: https://doi.org/10.31389/lsepress.ame.q License: CC-BY-NC 4.0.

Figure 17.1 A fiscal expansion in the IS-LM model



that some variables need to adjust – but when does this occur? And how? And by how much? It is these issues that make the analysis of fiscal policy so complex.

We will thus start our analysis by considering the government's intertemporal budget constraint. We have also explored the role of intertemporal budget constraints in analysing consumption and current account dynamics, the same analytical tools we used then will come in handy here. We will see the results that flow from this analysis in terms of the scope of fiscal policy and public debt dynamics.

17.1 | The government budget constraint

We must recognise that the government cannot create resources out of nothing, so it must respect an intertemporal budget constraint. This entails that the present value of government spending is limited by the present value of taxes (minus initial debt).

Let us start by looking carefully at the government budget constraint. Let g_t and τ_t denote the government's real purchases and tax revenues at time t, and d_0, its initial real debt outstanding. The simplest definition of the budget deficit is that it is the rate of change of the stock of debt. The rate of change in the stock of real debt equals the difference between the government's purchases and revenues, plus the real interest on its debt. That is,

$$\dot{d}_t = \left(g_t - \tau_t\right) + rd_t, \tag{17.1}$$

where r is the real interest rate.

The term in parentheses on the right-hand side of (17.1) is referred to as the primary deficit. It is the deficit excluding interest payments of pre-existing debt, and it is often a better way of gauging how fiscal policy at a given time is contributing to the government's budget constraint, since it leaves aside the effects of what was inherited from previous policies.

The government is also constrained by the standard solvency (no-Ponzi game) condition

$$\lim_{T\rightarrow\infty}\left(d_T e^{-rT}\right) \leq 0. \tag{17.2}$$

The government's budget constraint does not prevent it from staying permanently in debt, or even from always increasing the amount of its debt. Recall that the household's budget constraint in the Ramsey model implies that in the limit the present value of its wealth cannot be negative. Similarly, the restriction the budget constraint places on the government is that in the limit of the present value of its debt cannot be positive. In other words, the government cannot run a Ponzi scheme in which it raises new debt to pay for the ballooning interest on pre-existing debt. This condition is at the heart of the discussions on the solvency or sustainability of fiscal policy, and is the reason why sustainability exercises simply test that dynamics do not have government debt increase over time relative to GDP.

How do we impose this solvency condition on the government? We follow our standard procedure of solving the differential equation that captures the flow budget constraint. We can solve (17.1) by applying our familiar method of multiplying it by the integrating factor e^{-rt}, and integrating between 0 and T:

$$\begin{aligned} \dot{d}_t e^{-rt} - r d_t e^{-rt} &= \left(g_t - \tau_t\right) e^{-rt} \Rightarrow \\ d_T e^{-rT} &= d_0 + \int_0^T \left(g_t - \tau_t\right) e^{-rt} dt. \end{aligned} \tag{17.3}$$

$$\lim_{T\to\infty} \left(d_T e^{-rT}\right) = d_0 + \int_0^\infty \left(g_t - \tau_t\right) e^{-rt} dt \leq 0, \tag{17.4}$$

Applying the solvency condition (17.2) and rearranging, this becomes

$$\int_0^\infty g_t e^{-rt} dt \leq -d_0 + \int_0^\infty \tau_t e^{-rt} dt. \tag{17.5}$$

A household's budget constraint is that the present value of its consumption must be less than or equal to its initial wealth plus the present value of its labour income. A government faces an analogous constraint: the present value of its purchases of goods and services must be less than or equal to its initial wealth plus the present value of its tax receipts. Note that because d_0 represents debt rather than wealth, it enters negatively into the budget constraint.

An alternative way of writing this constraint is

$$\int_0^\infty \left(\tau_t - g_t\right) e^{-rt} dt \geq d_0. \tag{17.6}$$

Expressed this way, the budget constraint states that the government must run primary surpluses large enough in present value to offset its initial debt.

17.2 | Ricardian equivalence

We add households to our problem, and derive a very important result in the theory of fiscal policy, namely Ricardian equivalence: given a path for government spending, it doesn't matter whether it is financed via taxes or debt. We discuss the conditions under which this result holds, and the caveats to its application in practice.

Let us now add households to the story, and ask questions about the effects of fiscal policy decisions on private actions and welfare. In particular, for a given path of spending, a government can choose to finance it with taxes or bonds. Does this choice make any difference? This question is actually at the heart of the issue of countercyclical fiscal policy. If you remember the standard Keynesian exercise, a fiscal stimulus is an increase in government spending that is not matched by an increase in current taxes. In other words, it is a debt-financed increase in spending. Is this any different from a policy that would raise taxes to finance the increased spending?

A natural starting point is the neoclassical growth model we have grown to know and love. To fix ideas, consider the case in which the economy is small and open, domestic agents have access to an international bond b which pays an interest rate r (the same as the interest rate on domestic government debt), and in which all government debt is held by domestic residents.

When there are taxes, the representative household's budget constraint is that the present value of its consumption cannot exceed its initial wealth plus the present value of its after-tax labour income:

$$\int_0^\infty c_t e^{-rt} dt = b_0 + d_0 + \int_0^\infty \left(y_t - \tau_t\right) e^{-rt} dt. \tag{17.7}$$

Notice that initial wealth now apparently has two components: international bond-holdings b_0 and domestic public debt holdings d_0.

17.2.1 | The effects of debt vs tax financing

Breaking the integral on the right-hand side of (17.7) in two gives us

$$\int_0^\infty c_t e^{-rt} dt = b_0 + d_0 + \int_0^\infty y_t e^{-rt} dt - \int_0^\infty \tau_t e^{-rt} dt. \tag{17.8}$$

It is reasonable to assume that the government satisfies its budget constraint, (17.5), with equality. If it did not, its wealth would be growing forever, which does not seem realistic. With that assumption, (17.5) implies that the present value of taxes, $\int_{t=0}^\infty \tau_t e^{-rt} dt$, equals initial debt d_0 plus the present value of government purchases, $\int_{t=0}^\infty g_t e^{-rt} dt$. Substituting this fact into (17.8) gives us

$$\int_0^\infty c_t e^{-rt} dt = b_0 + \int_0^\infty y_t e^{-rt} dt - \int_0^\infty g_t e^{-rt} dt. \tag{17.9}$$

Equation (17.9) shows that we can express the households' budget constraint in terms of the present value of government purchases without any reference to the division of the financing of those purchases at any point in time between taxes and bonds. Since the path of taxes does not enter either households' budget constraint or their preferences, it does not affect consumption. Thus, we have a key result: only the quantity of government purchases, not the division of the financing of those purchases between taxes and bonds, affects the economy. This was first pointed out by British economist David Ricardo, back in the 19th century, which is why it is called *Ricardian equivalence*. Barro (1974) revived this result, proving it in the context of the NGM.

To see the point more starkly, focus on the case in which the agent maximises

$$\int_0^\infty u(c_t) e^{-\rho t} dt, \tag{17.10}$$

where $\rho > 0$ is the rate of time preference. Assume moreover that $r = \rho$, as we have done in the past. This implies that the optimal consumption path is flat:

$$c_t = c \text{ for all } t \geq 0. \tag{17.11}$$

Applying this to (17.9) yields

$$c = rb_0 + r\int_0^{\infty} y_t e^{-rt}dt - r\int_0^{\infty} g_t e^{-rt}dt \tag{17.12}$$

so that consumption is equal to permanent income. Again, since the path of taxes does not enter the RHS of (17.12), it does not affect consumption.

What is the intuition for this remarkable result? If a certain amount of spending is financed by issuing debt, the solvency condition implies that this debt must be repaid at some point. In order to do so, the government will have to raise taxes in the exact present value of that spending. This means that government bonds are not net wealth: for a given path of government expenditures, a higher stock of debt outstanding means a higher present value of taxes to be paid by households. In other words, the government cannot create resources out of nothing, it can only transfer them over time; rational, forward-looking consumers recognise this, and do not change their behaviour from what it would have been had that spending been financed with taxes immediately.

17.2.2 | Caveats to Ricardian equivalence

This very strong result was obtained under rather stringent assumptions, which we now underscore:

- Consumers live forever, so that a change in taxes even very far away in time matters to them. But note, if consumers had finite lives but acted altruistically in regards their sons and daughters, the result would still hold. This was the insight of Barro (1974).
- Taxes are lump-sum and therefore non-distortionary. If the present value of private income $\int_0^{\infty} y_t e^{-rt}dt$, for instance, fell if taxes increased (due to distortionary effects on investment or labour supply), then Ricardian equivalence would not hold.
- Consumers face no borrowing constraints so that all they care about is the present value of taxes. Consider, by contrast, an agent who cannot borrow and who has no initial wealth. Then, a tax increase today that was perfectly offset (in present value terms) by a tax break T periods later would reduce his income today – and therefore his consumption today– regardless of what happened in the future.
- Agents and the government can borrow at the same rate of interest r. If government could borrow more cheaply than consumers, for instance, by cutting taxes and running a larger deficit today, it would increase the wealth of consumers, enabling them to increase their consumption.

17.3 | Effects of changes in government spending

We show that changes in government spending have real effects, and that they are very different depending on whether they are temporary or permanent. Permanent increases are matched one-for-one by decreases in consumption. Temporary increases are matched less than one-for-one, with a current account deficit emerging as a result.

Now ask the opposite question. What is the effect of changes in government spending on consumption, savings, and the current account of the economy?

To get started, notice that the private budget constraint is

$$\dot{b}_t + \dot{d}_t = r\left(b_t + d_t\right) + y_t - \tau_t - c_t. \tag{17.13}$$

Combining this with the government budget constraint (17.1) we have

$$\dot{b}_t = rb_t + y_t - g_t - c_t. \tag{17.14}$$

Suppose now that income is constant throughout, but government consumption can vary. In particular, suppose that g_t can take two values: g^H and g^L, with $g^H > g^L$. To gain further insights, we study the effects of permanent and temporary changes in government spending that take place at time 0.

17.3.1 | The initial steady state

Assume that just prior to the shock, spending is expected to remain constant forever; that is, $g_t = g^L$ for all t. The economy is thus in a steady state in which consumption is given by

$$c = rb_0 + y - g^L. \tag{17.15}$$

In the initial steady state, the current account is zero.

17.3.2 | Permanent increase in government spending

Suppose now that at time 0 there is an unanticipated and permanent increase in spending from g^L to g^H. From (17.12) it follows that consumption adjusts instantaneously to its new (and lower) value:

$$c' = rb_0 + y - g^H, \qquad t \geq 0. \tag{17.16}$$

Since consumption falls one-to-one with government spending, the trade and current accounts do not change. Hence, an unanticipated and permanent increase in spending has no impact on the current account.

17.3.3 | Temporary increase in spending

Suppose that the economy is in the initial steady state described above, with consumption given by (17.15). At time 0, there is an unanticipated and *temporary increase* in spending:

$$g_t = \begin{cases} g^H, & 0 \leq t < T \\ g^L, & t \geq T, \end{cases} \tag{17.17}$$

for some $T > 0$.

First compute the consumption path. From (17.12) it follows that consumption falls immediately to the level given by

$$c'' = rb_0 + y - g^H\left(1 - e^{-rT}\right) - g^L e^{-rT}, \qquad t \geq 0. \tag{17.18}$$

Next compute the current account path. Plugging (17.18) into (17.14) we have

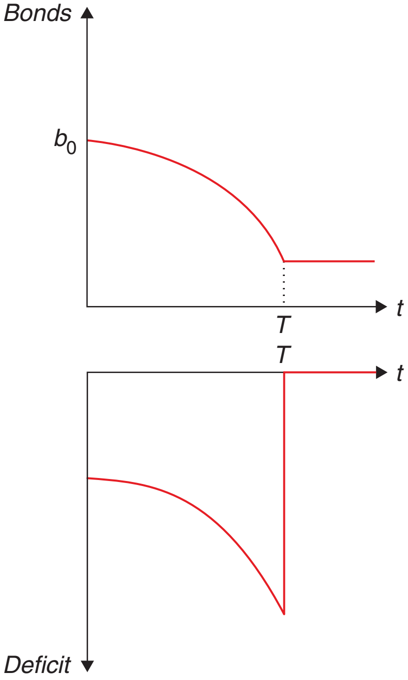
$$\dot{b}_t = r\left(b_t - b_0\right) + \left(g^L - g^H\right)e^{-rT}, \qquad 0 \leq t < T. \tag{17.19}$$

Notice that at time $t = 0$ this implies

$$\dot{b}_0 = \left(g^L - g^H\right)e^{-rT} < 0. \tag{17.20}$$

The current account is negative ($\dot{b}_0 < 0$) from the start. It follows that $b_t - b_0 < 0$, and, from (17.19), that $\dot{b}_t < 0$ for all times between 0 and T. The current account worsens over time and then jumps back

Figure 17.2 Response to a transitory increase in government spending



to zero at time T. Figure 17.2 shows the trajectory of the deficit and foreign assets in response to this temporary shock.

How do we know that the current account goes to 0 at time T? Recall the current account is given by

$$\dot{b}_t = rb_t + y_t - g_t - c_t. \tag{17.21}$$

Solving this for a constant path of income, spending, and consumption we have

$$b_T e^{-rT} = b_0 + \left(1 - e^{-rT}\right)\left(\frac{y - g^H - c''}{r}\right). \tag{17.22}$$

Plugging (17.18) into this expression we have

$$rb_T = rb_0 + \left(1 - e^{-rT}\right)\left(g^L - g^H\right). \tag{17.23}$$

Evaluating (17.14) at time T we have

$$\dot{b}_T = rb_T + y - g^L - c''. \tag{17.24}$$

Finally, using (17.18) and (17.23) in (17.24), we obtain

$$\begin{aligned} \dot{b}_T &= rb_0 + \left(1 - e^{-rT}\right)\left(g^L - g^H\right) + y - g^L \\ &\quad -rb_0 - y + g^H\left(1 - e^{-rT}\right) + g^L e^{-rT} \qquad (17.25) \\ &= 0. \qquad (17.26) \end{aligned}$$

That is, the current account is zero at time T.

The key policy lesson of this discussion is that consumption does not fully offset the changes in government consumption when those changes are transitory. So, if government expenditure aims at changing aggregate demand, it has a much higher chance of achieving that objective if it moves in a transitory fashion. Here the excess spending affects the current account which can easily be governed by temporary changes in fiscal policy. This is why you may often hear mentions of twin deficits in policy debates, in reference to the idea that fiscal deficits can be accompanied by current account deficits.

17.4 | Fiscal policy in a Keynesian world

We go over how fiscal policy fits into the (New) Keynesian framework. When there is the possibility of increasing output via aggregate demand (that is, in the short run), fiscal policy can affect output. That depends, however, on the consumer and firm responses, which take into account the budget constraints.

We have seen that, once we consider the government's budget constraint, and its interaction with the households' budget constraint, the effects of fiscal policy are radically modified, compared to what you may have been used to in undergraduate discussions.

As a starter, the effectiveness of fiscal policy as a tool for managing aggregate demand, in a full employment economy, is clearly dubious. For example, think about the NGM in a closed economy. It should be immediately clear that government spending *cannot* have an effect on output from the demand side: the government cannot through aggregate management increase the productive capacity of the economy, and this means its increased spending must crowd out something else. (In a small open economy model, as we have seen, it crowds out mostly consumption; in a closed or large open economy, it would also crowd out investment.) This has the flavour of Ricardian equivalence.

Having said that, we are still left with the question: where is the role for Keynesian fiscal policy? Is the conclusion that such policy is impossible? Not quite. Recall that the models so far in this chapter had no room for aggregate demand policy by assumption: the economies we've seen were not demand constrained and output was exogenous. In other words, one might think of them as descriptions of long-term effects. However, in a recession, the government can try to engage those that are sitting idle because of insufficient aggregate demand. This is the view that lies behind the thinking of proponents of fiscal stimulus.

To think more systematically about this, let us look at the role of government in the context of the canonical (discrete-time) New Keynesian model we saw in Chapter 15. The key change that we incorporate is to include government spending, which (in a closed economy) entails the following resource constraint:

$$Y_t = C_t + G_t. \tag{17.27}$$

Expressing this in terms of (percentage) deviations from the steady state, it is easy to see that the log-linearised resource constraint is[1]

$$\hat{y}_t = (1-\gamma)\hat{c}_t + \gamma\hat{g}_t, \tag{17.28}$$

where $\gamma = \frac{\bar{G}}{\bar{Y}}$ is the steady state share of government spending, and the hat above a variable z represents log deviations from the steady state: $\hat{z}_t = z_t - \bar{z}_t$.

Now take the Euler equation coming from consumer optimisation. Log-linearising around the (non-stochastic) steady state of the economy (and leaving aside the term in the discount rate ρ, for notational simplicity) yields our familiar

$$\hat{c}_t = E_t\hat{c}_{t+1} - \sigma r_t. \tag{17.29}$$

Now, we can't simply substitute output for consumption, as we did back in Chapter 15. Instead, substituting the new linearised resource constraint into (17.29) yields, after some simple manipulation.[2]

$$\hat{y}_t = E_t\hat{y}_{t+1} - \sigma(1-\gamma)\left(i_t - E_t\pi_{t+1} - r^n_g\right), \tag{17.30}$$

where $r^n_g = -\frac{\gamma}{\sigma(1-\gamma)}\Delta\hat{g}_{t+1}$. This term r^n_g is the modified version of the natural interest rate – that is to say, the rate that would keep output stable – that we encountered in Chapter 15.[3] It is easy to see that a *transitory* increase in $\hat{g}_t$ corresponds to an increase in that natural interest rate. In other words, a transitory fiscal expansion increases aggregate demand.

Here's where you may ask: but isn't the resource constraint being violated? Well, no: we have used (17.28) explicitly in our derivation, after all. The trick is that the resource constraint embodies the possibility that output itself is not fixed (in the short run), but can increase. That is the key distinction that allows for a potential effect of fiscal policy on aggregate demand.

Does that mean that an increase in government spending will lead to an increase in output? Not necessarily.

First, consumption will change as well. Note that the Euler equation that underpins the NKIS in (17.29) does not pin down the full consumption path: for that we must incorporate the intertemporal budget constraint. And we know that an increase in government spending must correspond to an increase in taxes at some point, and consumers will respond to that. In particular, in a Ricardian world, that response will be a one-for-one decrease in consumption if the increase is permanent, thereby negating the impact of government spending on aggregate demand. That a permanent increase in g has no effect on aggregate demand can be seen by noticing that a permanent increase in g will not change r^n_g in (17.30). That permanent changes in g have no effect may come as a surprise, given what it typically portrayed in intermediate texts, but on a simple inspection is obvious. Countries differ radically in the amount of government expenditures, but there is no relation between aggregate demand, recessions, or macroeconomic performance between them. Permanent changes in government spending just crowd out other components of aggregate demand.

More generally, the impact of government spending on aggregate demand is thus predicated on it *being temporary*, or on departures from Ricardian assumptions, as we have discussed.

On top of that, there is also a supply-side effect, via the behaviour of firms: the New Keynesian Phillips Curve (NKPC) will also be affected. You will recall that the NKPC looks like this:

$$\pi_t = \beta E_t\pi_{t+1} + \kappa\hat{y}_t. \tag{17.31}$$

A shock to government spending will also have a supply-side effect via an increase in the output gap, thus affecting inflation expectations, which will feed back into the NKIS. In short, the fiscal multiplier – that is, the change in output in response to a change in government spending – will depend on the specification and parameters of the model.

Finally, the NKIS makes it clear that this will depend on the response of monetary policy – and remember, the full specification of the canonical New Keynesian DSGE model requires the specification of a monetary policy rule. For instance, if monetary policy responds by increasing the nominal interest rate, it undoes the impact of fiscal expansion.

Summing up, our point is that to assess the role of fiscal policy it's important to keep in mind that people will react to policy, and will consider (to a smaller or greater extent) its dynamic effects. Part of a fiscal stimulus will end up being saved to face future taxes, and the extent to which this will crowd out other sources of demand will depend on how demand-constrained the economy is. It must also take into account its effects on the perceived sustainability of current fiscal policies, as well as the response of monetary policy. All of these are empirical debates that take place, at least to some extent, within the context of this framework we've discussed. Let's finish our discussion by considering what happens in practice.

17.4.1 | The current (empirical) debate: Fiscal stimulus and fiscal adjustment

With all of these different effects in mind, we can look at some of the evidence regarding the effectiveness of fiscal policy and, closely related, on the implications of the kind of fiscal adjustment that might be needed in the face of soaring debt ratios.

The question on the effectiveness of fiscal policy as a tool for aggregate demand management is often subsumed in the discussion on the size of the multiplier: if I increase government spending a given amount, by what multiple of this amount will it increase output? The undergraduate text provides an unambiguous answer to this question: it goes up. But our model has led us to think of a number of things that might affect that multiplier: whether the increase is permanent or temporary and how agents react to it, the context in which it takes place, and what it entails for the future path of spending. There is a wide range of beliefs about that in the policy debate – as illustrated by the many countries in which proponents of fiscal expansion or austerity are pitched against one another. In practice, it is very hard to isolate that impact since there can obviously be reverse causality, and there are always lots of other things, besides the change in fiscal policy, that are going on at the same time and can also affect output. In sum, we have a classic identification challenge.

There are two main approaches to overcome that challenge. First, some people (e.g. Blanchard and Perotti 2002 and Ilzetzki et al. 2013) use what is called a structural vector autoregression (SVAR) econometric technique, which basically relies on the assumption that it takes time for fiscal policy to respond to news about the state of the economy. It uses VAR (regressing a vector of variables on their lags, essentially) to tease out predictable responses of fiscal policy to output and vice-versa, and attributes any correlation between the unpredicted components to the impact of government spending on output. The problem, of course, is that what is unpredicted by the econometrician need not be so for the agents themselves, so we can never be sure we're estimating the true impact. The second approach is to use natural experiments where government spending increases due to some credibly exogenous factor. A popular example is wars and associated military buildups (e.g. Ramey and Shapiro 1999, Ramey 2009, Barro and Redlick 2011), and, on a similar note, a recent paper by Nakamura and Steinsson (2014) looks at the differential effect of military buildup shocks across U.S. regions. (Another example, from Shoag (2013), uses variation induced by changes in state pension fund returns.) The problem is that exercises rely on specific sources of spending that may not be typical. In other words, it's hard to assess the external validity of these natural experiments.

The variation in the estimates is enormous. In the natural experiment literature, they range from 0.5 to 1.5, from military buildups, and reach above 2 in other approaches. The SVAR literature has wildly divergent numbers across time and countries and specific structural assumptions, ranging from −2.3 to 3.7! One state-of-the-art attempt, Ilzetzki et al. (2013), who use better (quarterly) data for a number of countries, also reached nuanced conclusions: (i) Short-run multiplier is negative in developing countries; (ii) Multiplier is zero or negative when debt level is high; (iii) Multiplier is zero under flexible exchange rates, but positive under fixed exchange rates, and is higher in closed economies than in open economies. Note that (i) and (ii) are quite consistent with the effects we have been talking about: if people believe that higher spending is unsustainable, they would expect higher taxes in the near future, dampening the effect of the expansion. By the same token, (iii) is consistent with a simple Mundell-Fleming model, which we haven't talked about in the book, but which you may recall from undergraduate macro: under flexible exchange rates, more government spending leads to a currency appreciation, which leads to a reduction in net exports that undoes the fiscal stimulus. These results share the flavour with our previous discussion of Ricardian equivalence: when deficits and debt levels become large, consumer behaviour seems to more clearly undo the effects of fiscal policy.

Of course, in practice, policymakers have to act even with all this uncertainty. Here's an example, Table 17.1, shows the estimates generated by the non-partisan Congressional Budget Office (CBO) for the impact of the 2009 stimulus in the U.S. One can see patterns that are broadly in line with our discussion: (temporary) spending has a larger effect than tax cuts, tax cuts to people who are more likely to be credit-constrained have a larger effect. (Note that this short-term estimate keeps at a minimum the possible supply-side effects that might come from distortionary taxation. We will get to this in the next chapter.)

This discussion on the effectiveness of fiscal policy also colours the debate on the desirability and implications of fiscal adjustments aimed at putting the debt situation under control. As discussed in Alesina et al. (2019), the conventional wisdom that these adjustments would necessarily be contractionary is too simplistic. The inherent intertemporal nature of fiscal policy sees to that: maybe a fiscal contraction, by convincing people that a previously unsustainable debt trajectory is now sustainable, will "crowd in" private spending, as people anticipate lower taxes (or no tax hikes) in the future. It may also have the supply-side effects that will be the focus of our next chapter.

The evidence in Alesina et al. (2019) suggests that, in practice, many fiscal contractions have indeed been expansionary! It seems that this is particularly likely if the adjustment is about cutting spending rather than increasing taxes – perhaps because the former are deemed to be more sustainable in the long run. While there is controversy about the evidence – who would have thought? (see Ball et al. (2013), particularly with respect to what is selected as an episode of major fiscal contraction) – the point remains that the practice is more complex than what the simple undergraduate textbook view would have suggested.

17.5 | What have we learned?

The key lesson from this chapter is that evaluating the macroeconomic impact of fiscal policy requires considering its dynamic implications and the existence of budget constraints. The fact that economic agents understand that means that any change in fiscal policy will meet with a response from the private sector, which in turn immediately affects the impact of that change.

The first expression of this logic is the Ricardian equivalence result: it does not matter whether a certain path of spending is financed via debt or taxes, because in the end it all has to come from taxes at some point. As a result, changes in government spending create a counteracting adjustment in private spending, because of future taxes. The size of the adjustment depends on whether the change is permanent or temporary – in the later case, the adjustment by consumption is less than one-for-one, and aggregate demand moves accordingly.

Table 17.1 Estimating the multiplier for the 2009 U.S. fiscal stimulus

Estimated Output Multipliers and Budgetary Costs of Major Provisions of the American Recovery and Reinvestment Act				
	Estimated Output Multipliers[a]			
Type of Activity	Low Estimate	High Estimate	Major Provisions of ARRA	Total Budgetary Cost of Provisions, 2009–2019[b]
Purchases of Goods and Services by the Federal Government	1.0	2.5	Division A, Title II: Other; Title IV: Energy Efficiency and Renewable Energy; Title IV: Innovative Technology Loan Guarantee Program; Title IV: Other Energy Programs; Title V: Federal Buildings Fund; Title VIII: National Institutes of Health; Title VIII: Other Department of Health and Human Services	$88 billion
Transfer Payments to State and Local Governments for Infrastructure	1.0	2.5	Division A, Title VII: Clean Water and Drinking Water State Revolving Funds; Title XI: Other Housing Assistance; Title XII: Highway Construction; Title XII: Other Transportation	$44 billion
Transfer Payments to State and Local Governments for Other Purposes	0.7	1.9	Division A, Title VIII: Education for the Disadvantaged; Title VIII: Special Education; Title IX: State Fiscal Stabilization Fund; Division B, Title V: State Fiscal Relief Fund	$215 billion
Transfer Payments to Individuals	0.8	2.2	Division A, Title I: Supplemental Nutrition Assistance Program; Title VIII: Student Financial Assistance; Division B, Title II: Unemployment Compensation; Title III: Health Insurance Assistance	$100 billion
One-Time Payments to Retirees	0.2	1.2	Division B, Title II: Economic Recovery Payments, Temporary Aid to Needy Families, and Child Support[c]	$18 billion
Two-Year Tax Cuts for Lower- and Middle-Income People	0.5	1.7	Division B, Title I: Refundable Tax Credits; Making Work Pay Credit;[d] American Opportunity Tax Credit[d]	$168 billion
One-Year Tax Cut for Higher-Income People	0.1	0.5	Increase in Individual AMT Exemption Amount[d]	$70 billion
Extension of First-Time Homebuyer Credit	0.2	1.0	Extension of First-Time Homebuyer Credit[d]	$7 billion

Still, we saw that fiscal policy can affect output in the short run, in a Keynesian world. But even in this case, understanding the impact of fiscal policy requires thinking about the private sector adjustments – driven by the shadow of the intertemporal budget constraints – as well as the response of monetary policy.

17.6 | What next?

If you are interested in learning more about the debate on the impact of fiscal policy, a great starting point is the book *Fiscal Policy After The Financial Crisis*, edited by Alesina and Giavazzi (2013). It contains a thorough overview of different perspectives in that debate.

17.7 | Appendix

17.7.1 | Debt sustainability

Let us now turn to some practical issues when analysing fiscal policy from a dynamic perspective. Our discussion in this chapter has talked about how fiscal policy has to fit into an intertemporal budget constraint. But that raises the question: how do we know, in practice, whether a certain fiscal trajectory fits into the intertemporal budget constraint? If it does not we say the debt level is unsustainable, and we must expect some sort of change. We deal with this issue first. We then turn to some problems related to measurement, which are particularly important given the inherently dynamic nature of fiscal policy.

17.7.2 | A simplified framework

An important issue in macroeconomic analysis is to figure out if the debt dynamics are sustainable. Strictly speaking, sustainability means satisfying the NPG condition: the present discounted value of debt in the long run has to be zero. This restriction is relatively imprecise as to short term debt dynamics. Debt can grow, and grow very fast, and still satisfy the NPG condition. There is an ample literature on this, but a very practical, though primitive, way of answering this question is by asking whether the debt-to-GDP ratio is increasing or not. According to this stark definition debt sustainability is achieved when the debt-to-GDP ratio is constant or declining over time.

With this in mind a typical debt sustainability analysis would start with the dynamic equation for debt:

$$D_t = \left(1 + r_t\right) D_{t-1} + \left(\overbrace{G_t - T_t}^{\text{primary deficit}}\right). \tag{17.32}$$

(Note that this is the discrete-time equivalent of (17.1).) Divide this by GDP_t (we also multiply and divide the first term to the right of the equal sign by GDP_{t-1}) to obtain

$$\frac{D_t}{GDP_t} = \left(1 + r_t\right) \frac{D_{t-1}}{GDP_{t-1}} \frac{GDP_{t-1}}{GDP_t} + \left(\frac{G_t}{GDP_t} - \frac{T_t}{GDP_t}\right). \tag{17.33}$$

Denoting with lower cases the variables relative to GDP, the above becomes

$$d_t = \frac{(1+r_t)}{(1+\gamma)} d_{t-1} + \Big(\underbrace{g_t - t_t}_{pd_t} \Big) \tag{17.34}$$

or

$$d_t = \frac{(1+r_t)}{(1+\gamma)} d_{t-1} - \Big(\underbrace{t_t - g_t}_{ps_t} \Big),$$

where *pd* and *ps* stand for primary deficit and surplus (again, the difference between government expenditure and income, and vice versa, respectively, prior to interest payments), γ is the growth rate of GDP. Assume that we are in a steady state such that everything is constant. In this case

$$d = \frac{(1+r)}{(1+\gamma)} d - ps, \tag{17.35}$$

which can be solved for the primary surplus needed for debt to GDP ratios to be stable:

$$ps = d\left[\frac{(1+r)}{(1+\gamma)} - 1\right] = d\left[\frac{1+r-1-\gamma}{1+\gamma}\right] = d\left[\frac{r-\gamma}{1+\gamma}\right]. \tag{17.36}$$

Any surplus larger than that in (17.36) will also do. In order to make this equation workable, practitioners typically try to estimate expected growth in the medium term (growth has to be assessed in the same denomination of the debt, so if debt is, say, in U.S. dollars, you should use the growth rate of GDP in current U.S. dollars), the interest rate can be estimated from the average interest payments, and debt to GDP ratios are typically available. Table 17.2 makes this computation for a series of values of g and d given a specific r (in this table we use 7%).

The table shows the primary surplus required to stabilise the debt to GDP ratio. It can also be used to estimate restructuring scenarios. Imagine a country (Greece?) with a debt to GDP ratio of 150% and an expected growth rate of 2%. The table suggests it needs to run a budget surplus of 7.6% of GDP. If only a 2.5% surplus is feasible debt, sustainability obtains (at this interest rate) if debt is 50% of GDP. To obtain debt sustainability the country needs a haircut of 66%.

Recent models have emphasized the stochastic nature of fiscal revenues and expenditures. Simulating these processes, rather than estimating a single number, they compute a distribution function for fiscal results thus allowing to compute the value at risk in public debt.

17.8 | Measurement issues

The government budget constraint (17.5) involves the present values of the entire paths of purchases and revenues, and not the deficit at a point in time. As a result, conventional measures of either the primary or total deficit can be misleading about the real evolution of fiscal accounts. We illustrate this with three examples.

Table 17.2 Required primary surpluses

Public debt to GDP (%)	GDP growth rate (%)					
	1	2	3	4	5	6
35	2.1	1.7	1.4	1.0	0.7	0.3
40	2.4	2.0	1.6	1.2	0.8	0.4
45	2.7	2.2	1.7	1.3	0.9	0.4
50	3.0	2.5	1.9	1.4	1.0	0.5
55	3.3	2.7	2.1	1.6	1.0	0.5
60	3.6	2.9	2.3	1.7	1.1	0.6
65	3.9	3.2	2.5	1.9	1.2	0.6
70	4.2	3.4	2.7	2.0	1.3	0.7
75	4.5	3.7	2.9	2.2	1.4	0.7
80	4.8	3.9	3.1	2.3	1.5	0.8
85	5.0	4.2	3.3	2.5	1.6	0.8
90	5.3	4.4	3.5	2.6	1.7	0.8
95	5.6	4.7	3.7	2.7	1.8	0.9
100	5.9	4.9	3.9	2.9	1.9	0.9
110	6.5	5.4	4.3	3.2	2.1	1.0
120	7.1	5.9	4.7	3.5	2.3	1.1
130	7.7	6.4	5.0	3.8	2.5	1.2
140	8.3	6.9	5.4	4.0	2.7	1.3
150	8.9	7.4	5.8	4.3	2.9	1.4
160	9.5	7.8	6.2	4.6	3.0	1.5

17.8.1 | The role of inflation

The first example is the effect of inflation on the measured deficit. The change in nominal debt outstanding – that is, the conventionally measured budget deficit – equals the difference between nominal purchases and revenues, plus the nominal interest on the debt. If we let D denote the nominal debt, the nominal deficit is thus

$$\dot{D}_t = P_t\left(g_t - \tau_t\right) + i_t D_t, \tag{17.37}$$

where P is the price level and i is the nominal interest rate. When inflation rises, the nominal interest rate rises for a given real rate. Thus, interest payments and the deficit increase. Yet the higher interest payments are just offsetting the fact that the higher inflation is eroding the value of debt. The behaviour of the real stock of debt, and thus the government's budget constraint is not affected.

To see this formally, we use the fact that, by definition, the nominal interest rate equals the real rate plus inflation. This allows us to rewrite our expression for the nominal deficit as

$$\dot{D}_t = P_t\left(g_t - \tau_t\right) + \left(r + \pi_t\right) D_t. \tag{17.38}$$

Dividing through by P_t this yields

$$\frac{\dot{D}_t}{P_t} = g_t - \tau_t + \left(r + \pi_t\right)\frac{D_t}{P_t}. \tag{17.39}$$

Next define real debt as

$$d_t = \frac{D_t}{P_t}, \tag{17.40}$$

so that (17.39) becomes

$$\frac{\dot{D}_t}{P_t} = g_t - \tau_t + \left(r + \pi_t\right) d_t. \tag{17.41}$$

But recall next from (17.1) that

$$\dot{d}_t = g_t - \tau_t + rd_t. \tag{17.42}$$

Using this in (17.41) we finally have

$$\frac{\dot{D}_t}{P_t} = \dot{d}_t + \pi_t d_t. \tag{17.43}$$

Higher inflation raises the conventional (nominal) measure of the deficit even when it is deflated by the price level. This was Brazil's problem for many years.

17.8.2 | Asset sales

The second example is the sale of an asset. If the government sells an asset, it increases current revenue and thus reduces the current deficit. But it also forgoes the revenue the asset would have generated in the future. In the natural case where the value of the asset equals the present value of the revenue it will produce, the sale has no effect on the present value of the government's revenue. Thus, the sale affects the current deficit but does not affect the government's net worth. It follows that the effect of privatisation on the fiscal position of the government has to be analysed carefully, as it is not given by the bump in measured current revenue. If there is a positive impact it should be predicated on the idea that the present value of the revenues to the private sector is greater than what would be the case for the government (say, because the government runs it inefficiently), so that the buyer would be willing to pay more than the present value of the revenues the government would obtain from it. But this argument, not to speak of the computation, is seldom done.

17.8.3 | Contingent liabilities

The third example is a contingent liability. A contingent liability is a government commitment to incur expenses in the future that is made without provision for corresponding revenues. (Did anyone say Social Security?) In contrast to an asset sale, a contingent liability affects the budget constraint without affecting the current deficit. If the government sells an asset, the set of policies that satisfy the budget constraint is unchanged. If it incurs a contingent liability, on the other hand, satisfying the budget constraint requires higher future taxes or lower future purchases. In industrialised countries, the largest contingent or unfunded liabilities are entitlement programs, particularly social security and health insurance. These unfunded liabilities are typically larger than the conventionally measured stock of government debt; they are the main reason that fiscal policies in these countries do not appear to be on sustainable paths.

One way to compute these contingent liabilities is to do a debt decomposition exercise, decomposing the evolution of the debt-to-GDP ratio into its components: the primary deficit, economic growth, etc. The residual in this computation is the recognition of contingent liabilities over the years.

17.8.4 | The balance sheet approach

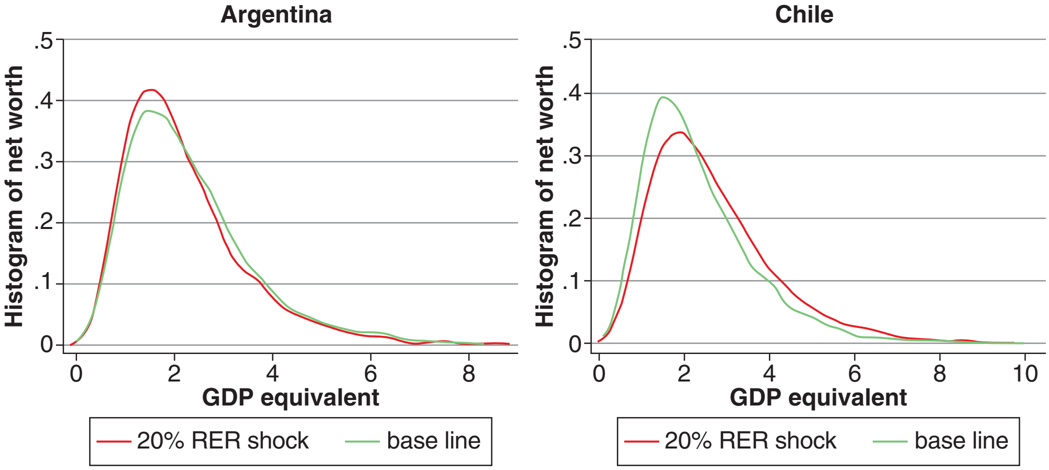
Another alternative to look into fiscal accounts is to work out the whole balance sheet of the government. This requires understanding that the main asset of the government is the NPV of its taxes, and that the main liability is not explicit debt, but the NPV of future wages and pensions. If so, the budget constraint can be written as

Assets	Liabilities
NPV of future taxes	Explicit debt
Liquid assets	NPV of future pensions and wages
Other assets	Contingent liabilities

This so-called balance sheet approach has been shown to qualify quite dramatically the role played by explicit liabilities in determining the vulnerability or currency exposure of governments. For example, for many countries that issue debt in foreign currency, it is believed that a real exchange rate depreciation compromises fiscal sustainability because it increases the real value of debt. However, this statement should consider what other liabilities it has in domestic currency (say, wages), and how a real depreciation may increase its tax base. If the government collects a sizable fraction of its income, say, from export taxes, then its net worth may actually increase with a devaluation. An example is provided in Figure 17.3, by looking at the histogram of the NPV of government and its reaction to a real depreciation for Argentina and Chile.

Argentina, for example, shows an improvement in the net worth of the government as a result of a devaluation, in contrast with the results from the literature that focuses on explicit debt.

Figure 17.3 Net worth: Argentina vs Chile, from Levy-Yeyati and Sturzenegger (2007)



Notes

[1] Is it really that easy? Let $\bar{Y}$ be the steady state level of output, and, similarly, for C and G. Now subtract $\bar{Y}$ from both sides of the equation, and divide both sides by $\bar{Y}$. Recognising that $\bar{Y} = \bar{C} + \bar{G}$, this gives you $\frac{Y_t - \bar{Y}}{\bar{Y}} = \frac{C_t - \bar{C}}{\bar{Y}} + \frac{G_t - \bar{G}}{\bar{Y}}$. Now multiply and divide the first term on the RHS by $\bar{C}$, and the second term by $\bar{G}$, and you will get the result.

[2] Here's how that one goes: note from the resource constraint that $\hat{c}_t = \frac{1}{1-\gamma}\hat{y}_t - \frac{\gamma}{1-\gamma}\hat{g}_t$, and plug this into (17.29). Multiplying both sides by $(1-\gamma)$ yields $\hat{y}_t = E_t\hat{y}_{t+1} - \gamma\left(E_t\hat{g}_{t+1} - \hat{g}_t\right) - \sigma(1-\gamma)r_t$. Define $\Delta\hat{g}_{t+1} = \left(E_t\hat{g}_{t+1} - \hat{g}_t\right)$ and collect terms, and the result follows.

[3] Recall that, for notational simplicity, we are leaving aside the terms in the discount rate and the productivity shock that shifts the natural rate of output, which is implicit within the log deviation term.

References

Alesina, A., Favero, C., & Giavazzi, F. (2019). Effects of austerity: Expenditure-and tax-based approaches. *Journal of Economic Perspectives*, *33*(2), 141–62.

Alesina, A. & Giavazzi, F. (2013). *Fiscal policy after the financial crisis*. University of Chicago Press.

Ball, L., Furceri, D., Leigh, D., & Loungani, P. (2013). *The distributional effects of fiscal austerity* (tech. rep.) DESA Working Paper No. 129.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy*, *82*(6), 1095–1117.

Barro, R. J. & Redlick, C. J. (2011). Macroeconomic effects from government purchases and taxes. *The Quarterly Journal of Economics*, *126*(1), 51–102.

Blanchard, O. & Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics*, *117*(4), 1329–1368.

Ilzetzki, E., Mendoza, E. G., & Végh, C. A. (2013). How big (small?) are fiscal multipliers? *Journal of Monetary Economics*, *60*(2), 239–254.

Levy-Yeyati, E. & Sturzenegger, F. (2007). A balance-sheet approach to fiscal sustainability. *CID Working Paper Series*.

Nakamura, E. & Steinsson, J. (2014). Fiscal stimulus in a monetary union: Evidence from U.S. regions. *American Economic Review*, *104*(3), 753–92.

Ramey, V. A. (2009). Defense news shocks, 1939–2008: Estimates based on news sources. *Unpublished paper, University of California, San Diego*.

Ramey, V. A. & Shapiro, M. D. (1999). *Costly capital reallocation and the effects of government spending*. National Bureau of Economic Research.

Shoag, D. (2013). Using state pension shocks to estimate fiscal multipliers since the Great Recession. *American Economic Review*, *103*(3), 121–24.

CHAPTER 18

Fiscal policy II: The long-run determinants of fiscal policy

In our previous chapter we learned that it may be quite misleading to think about fiscal policy without thinking about the budget constraint. Our overarching question was whether fiscal policy can actually affect aggregate demand, and the central message was that the existence of budget constraints limits substantially what fiscal policy can do in that domain. In particular, all fiscal policy interventions will have to be eventually paid for, and rational individuals will understand that and adjust their consumption in ways that might undo the effects of those interventions – completely or partially, depending on the nature of the latter. Fiscal policy is more constrained than our old IS-LM world had led us to believe.

This relatively muted effect on aggregate demand, led to skepticism with respect to the prospects of government spending as a tool for macroeconomic stabilisation, with monetary policy taking over as the go-to aggregate demand management tool. In fact, over the years, Central Bankers have taken a much more prominent role in managing stabilisation policies than fiscal policy. (The very low interest rates of recent times changed more than a few minds on this point, particularly because of the zero-rate lower bound issue. We will turn to that in due course.)

In any event, this leads us to think of fiscal policy as it relates to the needs for public good provision, the second angle we highlighted at the start of the last chapter. We can think of this as a long-run view: that of setting up a system aimed at providing these public goods (though keeping in mind that some of these may depend on the cycle, such as unemployment insurance). With this view in mind, how should the level of government spending and taxes move over time? How stable should they be?

In order to think that optimal long-run fiscal policy we will initially consider a government that faces an exogenous stream of government expenditures (public goods and entitlements, which may be shocked by natural disasters, wars, pandemics, etc.) How should it finance them? What should the path of taxes and borrowing be?

As it turns out, with non-distortionary taxation, long-lived individuals, and perfect capital markets, we know that this choice is inconsequential: it's the world of Ricardian equivalence! But assume now that taxes are in fact distortionary, so that a high tax rate is costly – for instance, because it reduces incentives to work and/or to invest. What should the government do? After analysing this problem,

How to cite this book chapter:
Campante, F., Sturzenegger, F. and Velasco, A. 2021. *Advanced Macroeconomics: An Easy Guide.*
Ch. 18. 'Fiscal policy II: The long-run determinants of fiscal policy', pp. 279–294. London: LSE Press.
DOI: https://doi.org/10.31389/lsepress.ame.r License: CC-BY-NC 4.0.

we will discuss the decision of how government spending itself may react to shocks. If income falls, should government expenditure remain the same?

We then go over the question of whether fiscal policy, in practice, behaves in the way that theory would prescribe. The answer is that, of course, there are significant departures, because policy makers are often not behaving in the way our simple theory suggests – in particular, they are not benevolent planners, but rather part of an often messy political process. We need to get into the political economy of fiscal policy to think through these issues.

Finally, we briefly address the issue of what type of taxes should be raised to finance government expenditure, in particular whether it makes sense to finance expenditures with taxes on labor or taxes on capital. While this is a topic that overlaps with issues discussed in the field of public finance, it delivers a remarkable result with implications for capital accumulation, and therefore for the macroeconomics discussion.

But let's start at the beginning by looking at the optimal path for aggregate taxes.

18.1 | Tax smoothing

We establish the tax smoothing principle, that suggests a countercyclical pattern for fiscal policy: one should run deficits when spending needs are higher than normal, and surpluses when they are below normal. By the same token, temporary expenditures should be financed by deficits.

Let g_t and τ_t denote the government's real purchases and tax revenues at time t, and d_0 is the initial real debt outstanding. The budget deficit is

$$\dot{d}_t = \left(g_t - \tau_t\right) + rd_t, \tag{18.1}$$

where r is the real interest rate.

The government is also constrained by the standard solvency condition:

$$\lim_{T\to\infty}\left(d_T e^{-rT}\right) \leq 0, \tag{18.2}$$

In the limit the present value of its debt cannot be positive.

18.1.1 | The government objective function

The government wants to minimise tax distortions:

$$L = \int_0^\infty y_t \ell\left(\frac{\tau_t}{y_t}\right) e^{-rt} dt, \tag{18.3}$$

where $\ell(0) = 0$, $\ell(.)' > 0$ and $\ell''(.) > 0$. This function is a shorthand to capture the fact that taxes usually distort investment and labour-supply decisions, such that the economy (the private sector) suffers a loss, which is higher the higher is the ratio of tax revenues to income. Notice that we can think of the ratio $\frac{\tau_t}{y_t}$ as the tax rate. The loss function is also convex: the cost increases at an increasing rate as the tax rate $\frac{\tau_t}{y_t}$ rises. Notice also that these are pecuniary losses, which the government discounts at the rate of interest.

18.1.2 | Solving the government's problem

Assume that the government controls tax revenue τ_t (which becomes the control variable) while government spending g_t is given. The Hamiltonian for the problem is

$$H = y_t \ell \left(\frac{\tau_t}{y_t} \right) + \lambda_t \left(g_t - \tau_t + r d_t \right), \tag{18.4}$$

where debt d_t is the state variable and λ_t the costate. The FOCs are

$$\ell' \left(\frac{\tau_t}{y_t} \right) = \lambda_t, \tag{18.5}$$

$$\dot{\lambda}_t = \lambda_t (r - r) = 0, \tag{18.6}$$

$$\lim_{T\to\infty} \left(\lambda_T d_T e^{-rT} \right) = 0. \tag{18.7}$$

18.1.3 | The time profile of tax distortions

The combination of (18.5) and (18.6) implies that tax revenue as a share of output should be constant along a perfect foresight path:

$$\ell' \left(\frac{\tau_t}{y_t} \right) = \lambda \text{ for all } t \geq 0. \tag{18.8}$$

We call this tax smoothing. The intuition is the same as in consumption smoothing. With the rate of interest equal to the rate at which loss is discounted, there is no incentive to have losses be higher in one moment than in another. The intuition is that, because the marginal distortion cost per unit of revenue raised is increasing in the tax rate (the ratio $\frac{\tau_t}{y_t}$), a smooth tax rate minimises distortion costs.

Denote the implicit tax rate by

$$\phi_t = \frac{\tau_t}{y_t}. \tag{18.9}$$

Expression (18.8) says that along a perfect foresight path, the tax rate should be constant:

$$\phi_t = \frac{\tau_t}{y_t} = \phi \text{ for all } t. \tag{18.10}$$

This is known as the "tax smoothing" principle, and is the key result of the paper by Barro (1979).[1]

Notice also that, if λ is constant and non-zero, the TVC (18.7) implies that

$$\lim_{T\to\infty} \left(d_T e^{-rT} \right) = 0. \tag{18.11}$$

That is, the solvency condition will hold with equality. Since the shadow value of debt is always positive, the government will choose to leave behind as much debt as possible (in present value) – that is to say, zero.

18.1.4 | The level of tax distortions

Solving (18.1) forward starting from some time 0 yields

$$d_T e^{-rT} = d_0 + \int_0^T \left(g_t - \tau_t\right) e^{-rt} dt. \tag{18.12}$$

Next, apply to this last equation the transversality/solvency condition (18.11) to get

$$\lim_{T\to\infty} \left(d_T e^{-rT}\right) = d_0 + \int_0^\infty \left(g_t - \tau_t\right) e^{-rt} dt = 0. \tag{18.13}$$

Rearranging, this becomes

$$\int_0^\infty \tau_t e^{-rt} dt = d_0 + \int_0^\infty g_t e^{-rt} dt. \tag{18.14}$$

But, given (18.10), this can be rewritten as

$$\phi \int_0^\infty y_t e^{-rt} dt = d_0 + \int_0^\infty g_t e^{-rt} dt, \tag{18.15}$$

or

$$\phi = \frac{d_0 + \int_0^\infty g_t e^{-rt} dt}{\int_0^\infty y_t e^{-rt} dt}. \tag{18.16}$$

That is to say, the optimal flat tax rate equals the ratio of the present value of the revenue the government must raise to the present value of output.

18.1.5 | The steady state

Imagine that initially both output and expenditures are expected to remain constant, at levels $g_t = g^L$ and $y_t = y$. Then, (18.16) implies

$$\tau = \phi y = r d_0 + g^L, \tag{18.17}$$

so that the chosen tax revenue is equal to the permanent expenditures of government. Using this result in budget constraint (18.1) we have

$$\dot{d}_t = r d_t + g^L - \tau_t = r d_t + g^L - r d_0 - g^L = r\left(d_t - d_0\right). \tag{18.18}$$

Evaluating this expression at time 0 we obtain

$$\dot{d}_0 = r\left(d_0 - d_0\right) = 0. \tag{18.19}$$

Hence, the stock of debt is constant as well.

18.1.6 | Changes in government expenditures

Suppose now that at time 0 there is an unanticipated and permanent increase in spending from g^L to g^H. From (18.17) it follows that tax revenue adjusts instantaneously to its new (and higher) value:

$$\tau' = \phi' y = rd_0 + g^H, \qquad t \geq 0. \tag{18.20}$$

The adjustment takes place via an increase in the tax rate ϕ, to a higher level ϕ'. Since revenues increases one-to-one with government spending, fiscal deficit does not change. Hence, an unanticipated and permanent increase in spending has no impact on the deficit nor on government debt.

How about temporary shocks? Suppose that the economy is in the initial steady-state described above, with revenue given by (18.17). At time 0, there is an unanticipated and *temporary increase* in spending:

$$g_t = \begin{cases} g^H, & 0 \leq t < T \\ g^L, & t \geq T, \end{cases} \tag{18.21}$$

for some $T > 0$.

First compute the revenue path. Expression (18.16) becomes

$$\phi = \frac{rd_0 + r\int_0^\infty g_t e^{-rt} dt}{y}. \tag{18.22}$$

Combining (18.21) and (18.22) we have that revenue rises immediately to the level given by:

$$\tau'' = \phi'' y = rd_0 + g^H \left(1 - e^{-rT}\right) + g^L e^{-rT}, \qquad t \geq 0. \tag{18.23}$$

where $\phi'' > \phi$ is now the new and constant tax rate.

Note that, quite naturally, the increase in the tax rate is lower under the temporary increase in spending than under the permanent increase:

$$\phi' - \phi'' = \frac{\left(g^H - g^L\right) e^{-rT}}{y} > 0. \tag{18.24}$$

Next, compute the path for the fiscal deficit. Plugging (18.23) into (18.1) we have

$$\dot{d}_t = r\left(d_t - d_0\right) + \left(g^H - g^L\right) e^{-rT}, \qquad 0 \leq t < T. \tag{18.25}$$

Notice that at time $t = 0$ this implies

$$\dot{d}_0 = \left(g^H - g^L\right) e^{-rT} > 0. \tag{18.26}$$

There is a fiscal deficit ($\dot{d}_0 > 0$) from the start. From (18.25), this means that $d_t > d_0$ for all times between 0 and T. The fiscal deficit worsens over time and then jumps back to zero at time T. Figure 18.1 shows the evolution of the deficit and government debt in response to this temporary spending shock.

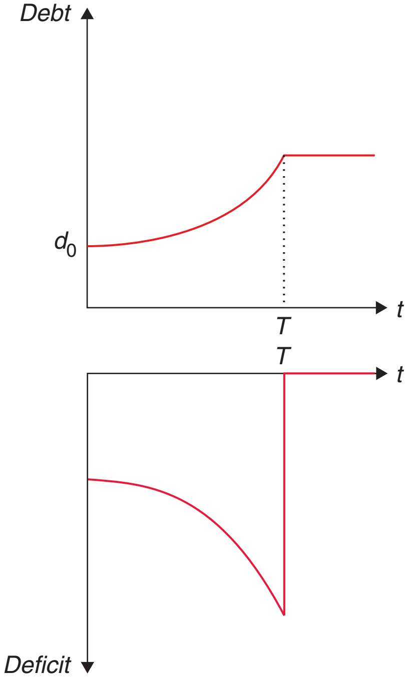
How do we know that the fiscal deficit goes to 0 at time T? Recall from (18.12) that

$$d_T e^{-rT} = d_0 + \left(1 - e^{-rT}\right)\left(\frac{g^H - \tau''}{r}\right). \tag{18.27}$$

Plugging (18.23) into this expression we have

$$rd_T = rd_0 + \left(1 - e^{-rT}\right)\left(g^H - g^L\right). \tag{18.28}$$

Figure 18.1 Response to an increase in government spending



Evaluating (18.1) at time T we have

$$\dot{d}_T = rd_T + g^L - \tau''. \tag{18.29}$$

Finally, using (18.23) and (18.28) in (18.29) we obtain

$$\begin{aligned} \dot{d}_T &= rd_0 + \left(1 - e^{-rT}\right)\left(g^H - g^L\right) + g^L \\ &\quad -rd_0 - g^H\left(1 - e^{-rT}\right) - g^L e^{-rT} \\ &= 0. \end{aligned} \tag{18.30}$$

Hence, debt is constant at time T and thereafter.

18.1.7 | Countercyclical fiscal policy

The pattern we have just established gives us a standard framework for thinking about how fiscal policy should respond to fluctuations: you should run deficits when expenditure needs are unusually high, and compensate with surpluses when they are relatively low. In short, the logic of tax smoothing provides a justification for running a *countercyclical* fiscal policy, based on long-run intertemporal optimisation.

This basic principle can be even stronger under plausible alternative assumptions, relative to what we have imposed so far. Consider, for instance, the case of a loss function for the deadweight loss of taxes that depends uniquely on the tax rate (i.e. eliminating the factor y_t that multiplies $\ell(\cdot)$ in (18.3) above). Then it is easy to see that the FOC is

$$\ell'\left(\frac{\tau_t}{y_t}\right)\frac{1}{y_t} = \lambda, \tag{18.31}$$

which means that the tax rate should be higher in booms and lower in recessions. The same happens if, plausibly, distortions are higher in recessions.[2] Or if government expenditure has a higher value in recessions than in booms – say, because of unemployment insurance. All of these changes further strengthen the countercyclical nature of fiscal policy.

18.1.8 | Smoothing government spending

The opposite happens if there is a desire to smooth government spending over time. Consider, for example, a case where tax revenues are now exogenous, though maybe hit by exogenous shocks. This may capture important cases, such as those in which economies are heavily reliant on the proceeds from natural resources, on whose prices they may not have much influence.[3] How should such an economy plan its spending profile?

To discuss this question imagine the case of a country where the government maximises

$$\int_0^\infty \left(\frac{\sigma}{1-\sigma}\right) g_t^{\frac{\sigma-1}{\sigma}} e^{-\rho t} dt, \tag{18.32}$$

subject to the budget constraint:

$$\dot{b}_t = rb_t + \tau_t + \epsilon_t - g_t, \tag{18.33}$$

where b stands for government assets, plus the NPG condition that we omit for brevity. Notice that this is identical to our standard consumption optimisation in an open economy, and, therefore, we know that the government has a desire to smooth government expenditures over time. We assume, though, that the shock to income ϵ_t follows:

$$\epsilon_t = \epsilon_0 e^{-\delta t}. \tag{18.34}$$

This embeds a full range of cases – for instance, if $\delta \to \infty$ then we have a purely transitory shock, but if $\delta \to 0$, the shock would be permanent.

Assuming, for simplicity, as we've done before, that $\rho = r$, the FOC for this problem is $\dot{g} = 0$: it is now government spending that should remain constant over time. This means that

$$g = r\left[b_0 + \frac{\tau}{r} + \frac{\epsilon_0}{r+\delta}\right]. \tag{18.35}$$

Notice that the final two terms give the present discounted value of taxes and of the income shock. Using (18.33) in (18.35) (and allowing $b_0 = 0$):

$$\dot{b}_t = \frac{\delta}{r+\delta}\epsilon_t. \tag{18.36}$$

Notice that if the shock is permanent ($\delta = 0$), the change in debt is zero, and government spending adjusts immediately to its new level. If shocks are transitory and positive then the government accumulates assets along the converging path. What is the implication of this result? That if the government wants to smooth its consumption, it will actually decrease its expenditures when hit by a negative shock. How persistent the shock is determines the impact on government expenditures. In other words, the desire to smooth this response somewhat weakens the countercyclical results we have discussed above.

18.1.9 | Summing up

If taxation is distortionary, then a government should endeavor to smooth the path of taxes as much as possible. That means that taxes respond fully to permanent changes in government expenditure, but only partially to transitory changes. The government borrows in bad times (when its expenditure is unusually high) and repays in good times (when its expenditure is unusually low).

If you think about it, this is exactly the same logic of consumption-smoothing, for the exact same reasons. But it has very important policy consequences: any extraordinary expenditure – e.g. a war, or a big infrastructure project – should be financed with debt, and not with taxes. The path followed by debt, however, ought to satisfy the solvency constraint; in other words, it must be sustainable.

This implies a *countercyclical* pattern for fiscal policy: you should run deficits when expenditure needs are unusually high, and compensate with surpluses when they are relatively low. This may interact with the business cycle as well: to the extent that spending needs go up in cyclical downturns (e.g. because of unemployment insurance payments), and the revenue base goes up when income goes up, the tax smoothing principle will suggest that deficits increase during recessions, and decrease when times are good. It doesn't make sense to increase tax rates to deal with higher unemployment insurance needs.

Yet this is different from using fiscal policy as an aggregate demand management tool, which is often the sense in which one may hear the term countercyclical fiscal policy being deployed. That said, there is a clear complementarity between the two angles: running the optimal fiscal policy, from a tax smoothing perspective, will also operate in the right direction when needed for aggregate demand management purposes, to the extent that it has an effect on aggregate demand, as per the previous chapter.

18.2 | Other determinants of fiscal policy

We discuss how political considerations may explain why, in practice, we see departures from the tax smoothing prescriptions. We also talk about how rules and institutions might address that.

While tax smoothing is a good starting point to understand and plan fiscal policy, tax smoothing can't explain many of the observed movements in fiscal policy. What are we missing? Alesina and Passalacqua (2016) provide a comprehensive literature review that you should read to understand the political subtleties of how fiscal policy gets determined in practice. They start by acknowledging that Barro's tax smoothing theory works very well when it comes to explaining U.S. and UK's last 200 years of debt dynamics. For these two countries, we see, generally speaking, debt increase during wars and then decline in the aftermath.[4]

More generally, tax smoothing is less descriptively useful when thinking about the short term. First of all, the prescriptions of tax smoothing depend strongly on expectations. When is an increase in spending temporary, thereby requiring smoothing? For a practitioner, it becomes an informed guess at best. Beyond that, however, there are many fiscal episodes that cannot be reconciled with tax smoothing: burgeoning debt-to-GDP ratios, protracted delays in fiscal adjustments, differences in how countries respond to the same shocks, etc.

Such episodes require a political economy story – namely, the realisation that policy makers in practice are not benevolent social planners.[5] The literature has come up with a menu of alternative possibilities.

18.2.1 | The political economy approach

The first of these possibilities relies on the concept of *fiscal illusion*, which claims that people don't fully understand the implications of fiscal policy. Voters overestimate the benefits of current expenditure and underestimate the future tax burden. Opportunistic politicians (you?) may take advantage of this. If so, there is a bias towards deficits. A derivation of this theory is the so-called political business cycle literature, which looks at the timing of spending around elections. Even with rational voters, we can still have cycles related to the different preferences of politicians from different parties or ideological backgrounds. There is evidence of electoral budget cycles across countries, but that evidence suggests that they are associated with uninformed voters, and, hence, that they tend to disappear as a democracy consolidates or as transparency increases. The main conclusion is that these factors may explain relatively small and short-lived departures from optimal fiscal policy, but not large and long-lasting excessive debt accumulation.

Another issue has to do with *intergenerational redistribution*. The idea is that debt redistributes income across generations. Obviously, if there is Ricardian equivalence (for example, because through bequests there are intertemporal links into the infinite future), this is inconsequential. But there are models where this can easily not be the case. For example, imagine a society with poor and rich people. The rich leave positive bequests, but the poor would like to have negative bequests. Because this is not possible, running budget deficits is a way of borrowing on future generations. In this case, the poor vote or push for expansionary fiscal policies. This effect will be stronger the more polarised society is, where the median voter has a lower income relative to average income.

The distributional conflicts may not be across time but actually at the same time. This gives rise to the theory of deficits as the result of distributive conflicts. Alesina and Drazen (1991) kicked off the literature with their model on deficits and *wars of attrition*. The idea is that whichever group gives in (throws the towel) will pay a higher burden of stabilisation costs. Groups then wait it out, trying to signal their toughness in the hope the other groups will give in sooner. Notice that appointing extremists to fight for particular interest groups can be convenient, though this increases the polarisation of the political system. In this setup, a crisis triggers a stabilisation and fiscal adjustment by increasing the costs of waiting. Drazen and Grilli (1993) show the surprising result that, in fact, a crisis can be welfare enhancing! Laban and Sturzenegger (1994) argue that delays occur because adjustment generates uncertainty, and therefore, there is value in waiting, even if delaying entails costs. Imagine a sick person who fears the risk of an operation to cure his ailment. Under a broad range of parameters, he may choose to wait in the certain state of poor health and put off the chances of an unsuccessful operation. This literature has opened the room to analyse other issues. Signalling, for example, is an important issue. Cukierman and Tommasi (1998) explain that it takes a Nixon to go to China: political preferences opposite to the policies implemented convey more credibly the message of the need for reform (Sharon or Rabin, Lula, etc. are all examples of this phenomenon).

Yet another story related to conflicting preferences about fiscal policy refers to *debt as a commitment*, or *strategic debt*. (See Persson and Svensson (1989) and Alesina and Tabellini (1990)). The idea is that a government who disagrees with the spending priorities of a possible successor chooses to overburden it with debt so it restricts its spending alternatives. Then, the more polarised a society, the higher its debt levels. If the disagreement is on the size of government, the low spender will reduce taxes and increase debt. (Does this ring any bells?)

Another version has to do with externalities associated with the provision of local public goods, a version of what is called the tragedy of the commons. Battaglini and Coate (2008) consider a story where a legislature decides on public good provision, but with two kinds of public goods: one that benefits all citizens equally, and another that is local (i.e. benefits only those who live in the locality where it is provided). If decisions are made by a legislature where members are elected in local districts, an externality arises: a member's constituents get all the benefits of the bridge their representative got built, but everyone in the country shares in the cost. They show that, in this case, debt will be above the efficient level, and tax rates will be too volatile.

Departures from optimal fiscal policy can also be due to rent-seeking. Acemoglu et al. (2011) or Yared (2010) consider scenarios where there is an agency problem between citizens and self-interested politicians, and the latter can use their informational advantage to extract rents. The tension between the need to provide these incentives to politicians (pushing to higher revenue and lower spending on public goods), on the one hand, and the needs of citizens (lower taxes and higher spending) on the other – and how they fluctuate in response to shocks – can lead to volatility in taxes and over-accumulation of debt.

Finally, an interesting failure of the tax smoothing prediction that is widespread in developing countries is the issue of procyclical fiscal policy. As we have noted, the intuition suggests that fiscal policy should be countercyclical, saving in good times to smooth out spending in bad times. However, Gavin and Perotti (1997) noted that Latin American governments, in particular, have historically tended to do the exact opposite, and the evidence was later extended to show that most developing countries do the same. One explanation is that these countries are credit-constrained: when times are bad, they lose all access to credit, and this prevents them from smoothing. This explanation again begs the question of why these governments don't save enough so that they eventually don't need to resort to capital markets for smoothing. Another explanation, proposed by Alesina et al. (2008), builds on a political economy story: in countries with weak institutions and corruption, voters will rationally want the government to spend during booms, because any savings would likely be stolen away. This predicts that the procyclical behaviour will be more prevalent in democracies, where voters exert greater influence over policy, a prediction that seems to hold in the data.

18.2.2 | Fiscal rules and institutions

Since there are so many reasons why politicians may choose to depart from optimal fiscal policy prescriptions – and impose costs on society as a result – it is natural to ask whether it might be possible to constrain their behaviour in the direction of optimal policy.

One type of approach in that direction is the adoption of specific rules – the most typical example of which is in the form of balanced-budget requirements. These impose an obvious cost in terms of flexibility: you do want to run deficits and surpluses, as a matter of optimal fiscal policy! The question is whether the political distortions are so large that this may be preferable. One key insight from the literature on such rules is that the costs of adopting them tend to arise in the short run, while the benefits accrue in the longer term. This in itself raises interesting questions about their political sustainability.

You could imagine a variety of rules and institutions, among different dimensions: do you want to decentralise your fiscal decisions? If so, how do you deal with tax sharing? Do you want to delegate decisions to technocrats? Does it all make a difference? The general conclusions is that, even taking into account the endogeneity of institutional arrangements, institutions matter to some extent. Yet there is room for skepticism. After all, many countries have passed balanced-budget rules and yet have never mastered fiscal balance. Chile, on the other hand, had no law, but run a structural, full employment, 1% surplus for many years to compensate for its depletion of copper reserves. In the end, the policy question of how you build credibility is still there: is it by passing a law that promises you will behave, or simply by behaving?

All in all, we can safely conclude that fiscal policy is heavily affected by political factors, which is another reason why the reach of the fiscal instrument as a tool for macroeconomic policy may be blunted in most circumstances. This is particularly so in developing countries, which tend to be more politically (and credit) constrained. As a result, our hunch is that you want transparent, simple, stable (yet *flexible*), predictable rules for governing fiscal policy. Your success will hinge on convincing societies to stick to these simple rules!

18.3 | Optimal taxation of capital in the NGM

We end with a public finance detour: what should be the optimal way of taxing capital in the context of the NGM? We show that even a planner who cares only about workers may choose to tax capital very little. The optimal tax on capital can grow over time, or converge to zero, depending on the elasticity of intertemporal substitution.

Let's end our discussion of fiscal policy with an application of the NGM. Our exercise will have more the flavour of public finance (what kinds of taxes should we use to fund a given path of government expenditure) rather than the macroeconomic perspective we have focused on so far. We develop it here to show the wealth of questions the framework we have put so much effort to develop throughout the book can help us understand. It will also allow us to present an additional methodological tool: how to solve for an optimal taxation rule.

Imagine that you need to finance a certain amount of public spending within the framework of the NGM. What would be the optimal mix of taxes? Should the central planner tax labour income, or should it tax capital? The key intuition to solving these problems is that the tax will affect behaviour, so the planner has to include this response in his planning problem. For example, if a tax on capital is imposed, you know it will affect the Euler equation for the choice of consumption, so the planner needs to *add* this modified Euler equation as an additional constraint in the problem. Adding this equation will allow the planner to internalise the response of agents to taxes, which needs to be done when computing the optimal level of taxation.

To make things more interesting, let's imagine the following setup which follows Judd (1985) and Straub and Werning (2020). There are two types of agents in the economy: capitalists that own capital, earn a return from it, and consume C_t; and workers that only have as income their labour, which they consume (c_t).[6] We will assume the central planner cares only about workers (it is a revolutionary government!), which is a simplification for exposition. However, we will see that even in this lopsided case some interesting things can happen, including having paths where, asymptotically, workers decide not to tax capitalists!

So let's get to work. The planner's problem (in discrete time) would be to maximise

$$\sum_{t=0}^{\infty} \left(\frac{1}{1+\rho} \right)^{t} u\left(c_t\right), \tag{18.37}$$

subject to

$$u'(C_t) = \frac{(1 + r_{t+1})}{1 + \rho}(1 - \tau_{t+1})u'(C_{t+1}), \tag{18.38}$$

$$C_t + b_{t+1} = (1 - \tau_t)(1 + r_t)b_t, \tag{18.39}$$

$$c_t = w_t + T_t, \tag{18.40}$$

$$C_t + c_t + k_{t+1} = f(k_t) + (1 - \delta)k_t, \tag{18.41}$$

$$k_t = b_t. \tag{18.42}$$

The first equation (18.38) is the Euler condition for capitalists. As we said above, this equation needs to hold and, therefore, has to be considered by the planner as a constraint. The following two equations (18.39) and (18.40) define the budget constraint of capitalists and workers. Capitalist earn income from their capital, though this may be taxed at rate τ. Workers are just hand-to-mouth consumers: they eat all their income. They receive, however, all the proceeds of the capital tax $T_t = \tau_t k_t$. Equation (18.41) is the resource constraint of the economy where we assume capital depreciates at rate δ. Finally equation (18.42) states that in equilibrium the assets of the capitalists is just the capital stock. We have omitted the transversality conditions and the fact that wages and interest rates are their marginal products – all conditions that are quite familiar by now.

Before proceeding you may think that, if the planner cares about workers, and not at all about capitalists, the solution to this problem is trivial: just expropriate the capital and give it to workers. This intuition is actually correct, but the problem becomes relatively uninteresting.[7] To rule that out, we will add the extra assumption that workers would not know what to do with capital. You actually have to live with the capitalists so they can run the capital.

Something else you may notice is that we have not allowed for labour taxation. Labour taxation creates no distortion here, but would be neutral for workers (what you tax is what you return to them), so the interesting question is what is the optimal tax on capital.

Typically, you would just set out your Bellman equation (or Hamiltonian, if we were in continuous time) to solve this problem, simply imposing the Euler equation of capitalists as an additional constraint. However, if their utility is log, the solution is simpler because we already know what the optimal consumption of capitalists is going to be: their problem is identical to the consumption problem we solved in Chapter 12! You may recall that the solution was that consumption was a constant share of assets $C_t = \frac{\rho}{(1+\rho)}(1 + r_t)b_t = (1 - s)(1 + r_t)b_t$, where the second equality defines s. Capitalists consume what they don't save. In our case, this simplifies the problem quite a bit. We can substitute for the capitalists' consumption in the resource constraint, knowing that this captures the response function of this group to the taxes.[8] Using the fact that

$$C_t = (1 - s)(1 - \tau_t)(1 + r_t)b_t, \tag{18.43}$$

and that

$$b_{t+1} = s(1 - \tau_t)(1 + r_t)b_t, \tag{18.44}$$

it is easy to show (using the fact that $b_t = k_t$) that

$$C_t = \frac{(1 - s)}{s}k_{t+1}. \tag{18.45}$$

We can replace this in the resource constraint to get

$$\frac{k_{t+1}}{s} + c_t = f(k_t) + (1 - \delta)k_t. \tag{18.46}$$

Notice that this is equivalent to the standard growth model except that the cost of capital is now increased by $\frac{1}{s}$. Capital has to be intermediated by capitalists, who consume part of the return, thus making accumulation more expensive from the perspective of the working class.

Solving (18.37) subject to (18.46) entails the first order condition

$$u'(c_t) = \frac{s}{1+\rho}u'(c_{t+1})(f'(k_{t+1}) + 1 - \delta). \tag{18.47}$$

In steady state, this equation determines the optimal capital stock:

$$1 = \frac{s}{1+\rho}(f'(k^*) + 1 - \delta) = s\frac{1}{1+\rho}R^*, \tag{18.48}$$

where R^* is the interest rate before taxes. Notice that in steady state we must also have that the savings are sufficient to keep the capital stock constant,

$$sRk = k, \tag{18.49}$$

where R is the after-tax return for capitalists. Using (18.48) and (18.49), it is easy to see that

$$\frac{R}{R^*} = \frac{1}{1+\rho} = (1-\tau), \tag{18.50}$$

or simply that $\tau = \frac{\rho}{1+\rho}$.

In short: the solution is a constant tax on capital, benchmarked by the discount rate. In particular, the less workers (or the planner) discount the future, the smaller the tax: keeping the incentives for capitalists to accumulate pays of in the future and therefore is more valuable the less workers discount future payoffs. In the limit, with very patient workers/planner, the tax on capital approaches zero, and this happens even though the planner does not care about the welfare of capitalists! This is a powerful result.

Yet, this result relies heavily on the log utility framework. In fact, Straub and Werning (2020) solve for other cases, and show that things can be quite different. As you may have suspected by now, this is all about income versus substitution effects. Consider a more general version of the utility function of capitalists: $U(C_t) = \frac{C_t^{1-\sigma}}{(1-\sigma)}$. If $\sigma > 1$, the elasticity of intertemporal substitution is low, and income effects dominate. If $\sigma < 1$, in contrast, substitution effects dominate. Does the optimal tax policy change? It does – in fact, Straub and Werning show that, in the first case, optimal taxes on capital grow over time! In the second case, on the other hand, they converge to zero.

The intuition is pretty straightforward: workers want capitalists to save more since, the larger the capital stock, the larger the tax they collect from it. If taxes are expected to increase, and income effects prevail, capitalists will save more in expectation of the tax hike, which makes it optimal from the perspective of the workers to increase taxes over time. The opposite occurs when substitution effects prevail. As we can see, even this simple specification provides interesting and complex implications for fiscal policy.

18.4 | What have we learned?

We have seen that, from the standpoint of financing a given path of government spending with a minimum of tax distortions, there is a basic principle for optimal fiscal policy: tax smoothing. Optimal fiscal policy will keep tax rates constant, and finance temporarily high expenditures via deficits, while running surpluses when spending is relatively low. This countercyclical fiscal policy arises not because of aggregate demand management, but from a principle akin to consumption smoothing. Because high taxes entail severe distortions, you don't want to have to raise them too high when spending needs arise. Instead, you want to tax at the level you need to finance the permanent level of expenditures, and use deficits and surpluses to absorb the fluctuations.

As it turns out, in practice there are many important deviations from this optimal principle – more often than not in the direction of over-accumulation of debt. We thus went over a number of possible political economy explanations for why these deviations may arise. We can safely conclude that fiscal policy is heavily affected by political factors, which gives rise to the question of whether rules and institutions can be devised to counteract the distortions.

Finally, we briefly went over the public finance question of optimal taxation of capital in the context of the NGM model. We obtained some surprising results – a social planner concerned only with workers may still refrain from taxing capital to induce more capital accumulation, which pays off in the long run. Yet, these results are very sensitive to the specification of preferences, particularly in the elasticity of intertemporal substitution, further illustrating the power of our modelling tools in illuminating policy choices.

18.5 | What next?

The survey by Alesina and Passalacqua (2016) is a great starting point for the literature on the determinants of fiscal policy.

Notes

[1] In a model with uncertainty the equivalent to equation (18.10) would be $\phi_t = E(\phi_{t+1})$, that is, that the tax rates follow a random walk.

[2] A good application of this is to think about the optimal response to the Covid-19 pandemic of 2020. In a forced recession, taxes became almost impossible to pay, in some cases leading to bankruptcies and thus promoting policies of tax breaks during the outbreak.

[3] Take the case of Guyana, which one day found oil and saw its revenues grow by 50% in one year – this actually happened in 2020.

[4] Even in these countries, however, there are anomalies, such as the U.S. accumulating debt over the relatively peaceful 1980s – more on this later...

[5] In fact, the need for a political economy story strikes even deeper. Suppose we had a truly benevolent and farsighted government, what should it do in the face of distortionary taxation? Well, Aiyagari et al. (2002) have the answer: it should accumulate assets gradually, so that, eventually, it would have such a huge pile of bonds that it could finance any level of spending just with the interest it earned from them – no need for any distortionary taxes! In that sense, even the tax smoothing logic ultimately hinges on there being some (binding) upper bound on the level of assets that can be held by the government, often referred to as an ad hoc asset limit.

[6] The problem for capitalists is to maximise $\sum_{t=0}^{\infty}\left(\frac{1}{1+\rho}\right)^t u\left(C_t\right)$, subject to $C_t + b_{t+1} = (1+r_t)b_t$. Notice that this problem is virtually identical, but not exactly the same, to the consumer problem we solved in Chapter 12. The difference is a timing convention. Before savings at time t did not generate interest income in period t, here they do. Thus what before was b_t will now become $b_t(1+r_t)$.

[7] Though in the real world this option sometimes is implemented. Can you think of reasons why we don't see it more often?

[8] The log case is particularly special because the tax rate does not affect the capitalist's consumption path – more on this in a bit.

References

Acemoglu, D., Golosov, M., & Tsyvinski, A. (2011). Political economy of Ramsey taxation. *Journal of Public Economics*, *95*(7-8), 467–475.

Aiyagari, S. R., Marcet, A., Sargent, T. J., & Seppälä, J. (2002). Optimal taxation without state-contingent debt. *Journal of Political Economy*, *110*(6), 1220–1254.

Alesina, A. & Drazen, A. (1991). Why are stabilizations delayed? *American Economic Review*, *81*(5), 1170–1188.

Alesina, A., Campante, F. R., & Tabellini, G. (2008). Why is fiscal policy often procyclical? *Journal of the European Economic Association*, *6*(5), 1006–1036.

Alesina, A. & Passalacqua, A. (2016). The political economy of government debt. *Handbook of Macroeconomics*, *2*, 2599–2651.

Alesina, A. & Tabellini, G. (1990). A positive theory of fiscal deficits and government debt. *The Review of Economic Studies*, *57*(3), 403–414.

Barro, R. J. (1979). On the determination of the public debt. *Journal of Political Economy*, *87*(5, Part 1), 940–971.

Battaglini, M. & Coate, S. (2008). A dynamic theory of public spending, taxation, and debt. *American Economic Review*, *98*(1), 201–36.

Cukierman, A. & Tommasi, M. (1998). When does it take a Nixon to go to China? *American Economic Review*, 180–197.

Drazen, A. & Grilli, V. (1993). The benefit of crises for economic reforms. *The American Economic Review*, *83*(3), 598–607.

Gavin, M. & Perotti, R. (1997). Fiscal policy in Latin America. *NBER Macroeconomics Annual*, *12*, 11–61. https://www.journals.uchicago.edu/doi/pdf/10.1086/654320.

Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, *28*(1), 59–83.

Laban, R. & Sturzenegger, F. (1994). Distributional conflict, financial adaptation and delayed stabilizations. *Economics & Politics*, *6*(3), 257–276.

Persson, T. & Svensson, L. E. (1989). Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences. *The Quarterly Journal of Economics*, *104*(2), 325–345.

Straub, L. & Werning, I. (2020). Positive long-run capital taxation: Chamley-Judd revisited. *American Economic Review*, *110*(1), 86–119.

Yared, P. (2010). Politicians, taxes and debt. *The Review of Economic Studies*, *77*(2), 806–840.

CHAPTER 19

Monetary policy: An introduction

19.1 | The conundrum of money

We have finally reached our last topic: monetary policy (MP), one of the most important topics in macroeconomic policy, and perhaps the most effective tool of macroeconomic management. While among practitioners there is a great deal of consensus over the way monetary policy should be implemented, it always remains a topic where new ideas flourish and raise heated debates. Paul Krugman tweeted,

> Nothing gets people angrier than monetary theory. Say that Trump is a traitor and they yawn; say that fiat money works and they scream incoherently.

Our goal in these final chapters is to try to sketch the consensus, its shortcomings, and the ongoing attempts to rethink MP for the future, even if people scream!

We will tackle our analysis of monetary policy in three steps. In this chapter we will start with the basics: the relation of money and prices, and the optimal choice of inflation. This will be developed first, in a context where output is exogenous. This simplifies relative to the New Keynesian approach we discussed in Chapter 15, but will provide some of the basic intuitions of monetary policy. The interaction of money and output creates a whole new wealth of issues. Is monetary policy inconsistent? Should it be conducted through rules or with discretion? Why is inflation targeting so popular among central banks? We will discuss these questions in the next chapter. Finally, in the last two chapters we will discuss new frontiers in monetary policy, with new challenges that have become more evident in the new period of very low interest rates. In Chapter 21 we discuss monetary policy when constrained by the lower bound, and the new approach of quantitative easing. In Chapter 22 we discuss a series of topics: secular stagnation, the fiscal theory of the price level, and bubbles. Because these last two chapters are more prolific in referencing this recent work, we do not add the what next section at the end of the chapter, as the references for future exploration are already plenty within the text.

But before we jump on to this task, let us briefly note that monetary economics rests on a fairly shaky foundation: the role of money – why people hold it, and what are its effects on the economy – is one of the most important issues in macroeconomics, and yet it is one of the least understood. Why is this? For starters, in typical micro models – and pretty much in all of our macro models as well – we did not deal with money: the relevant issues were always discussed in terms of relative prices,

How to cite this book chapter:
Campante, F., Sturzenegger, F. and Velasco, A. 2021. *Advanced Macroeconomics: An Easy Guide.* Ch. 19. 'Monetary policy: An introduction', pp. 295–314. London: LSE Press.
DOI: https://doi.org/10.31389/lsepress.ame.s License: CC-BY-NC 4.0.

not nominal prices. There was no obvious (or, at least, essential) role for money in the NGM that we used throughout this book. In fact, the *non plus ultra* of micro models, the general equilibrium Arrow-Debreu framework, not only does not need money, it also does not have trading! (Talk about outrageous assumptions.) In that model, all trades are consummated at the beginning of time, and then you have the realisation of these trades, but no new trading going on over time. Of course, the world is not that complete, so we need to update our trading all the time. We use money as an insurance for these new trades. However, it is much easier to say people use money for transactions than to model it, because we need to step into the world of incomplete markets, and we do not know how to handle that universe well.

The literature has thus taken different paths for introducing money into general equilibrium models. The first is to build a demand for money from micro-foundations. The question here is whether one commodity (maybe gold, shells, salt?) may become a vehicle that people may naturally choose for transactions, i.e. what we usually refer to as money. Kiyotaki and Wright (1989), for example, go this way. While nice, by starting from first principles, this approach is intractable and did not deliver models which are sufficiently flexible to discuss other issues, so this research has only produced a plausible story for the existence of money but not a workable model for monetary policy.

The other alternative is to introduce money in our typical overlapping generations model. Money serves the role of social security, and captures the attractive feature that money has value because you believe someone down the road will take it. Unfortunately, the model is not robust. Any asset that dominates money in rate of return will simply crowd money out of the system, thus making it impossible to use this model to justify the use of money in cases in which the inflation rate is minimally positive when money is clearly dominated in rate of return.

A third approach is to just assume that money needs to be used for purchases, the so-called cash in advance constraints. In this framework the consumer splits itself at the beginning of each period into a consumer self and a producer self. Given that the consumer does not interact with the producer, she needs to bring cash from the previous period, thus the denomination of cash in advance. This is quite tractable, but has the drawback that gives a very rigid money demand function (in fact, money demand is simply equal to consumption).

A more flexible version is to think that the consumer has to devote some time to shopping, and that shopping time is reduced by the holdings of money. This provides more flexibility about thinking in the demand for money.

Finally, a take-it-all-in alternative is just to add money in the utility function. While this is a reduced form, it provides a flexible money demand framework, and, therefore, has been used extensively in the literature. At any rate, it is obvious that people demand money, so we just postulate that it provides utility. An additional benefit is that it can easily be accommodated into the basic framework we have been using in this book, for example, by tacking it to an optimisation problem akin to that of the NGM.

Thus, we will go this way in this chapter. As you will see, it provides good insights into the workings of money in the economy.

19.1.1 | Introducing money into the model

Let's start with the simplest possible model. Output exogenous, and a government that prints money and rebates the proceeds to the consumer. We will lift many of these assumptions as we go along. But before we start we need to discuss the budget constraints.

Assume there is only one good the price of which in terms of money is given by P_t. The agent can hold one of two assets: money, whose nominal stock is M_t, and a real bond, whose real value is given, as in previous chapters, by b_t. Note that we now adopt the convention that real variables take on small-case letters, and nominal variables are denoted by capital letters. The representative agent's budget constraint is given by

$$\frac{\dot{M}_t}{P_t} + \dot{b}_t = rb_t + y_t - \tau_t - c_t, \tag{19.1}$$

where τ_t is real taxes paid to the government and, as usual, y_t is income and c_t consumption. Define the real quantity of money as

$$m_t = \frac{M_t}{P_t}. \tag{19.2}$$

Taking logs of both sides, and then time derivatives, we arrive at:

$$\dot{m}_t = m_t \frac{\dot{M}_t}{M_t} - m_t \frac{\dot{P}_t}{P_t} = \frac{M_t}{P_t}\frac{\dot{M}_t}{M_t} - m_t \frac{\dot{P}_t}{P_t}. \tag{19.3}$$

Defining $\pi_t \equiv \frac{\dot{P}_t}{P_t}$ as the rate of inflation and rearranging, we have:

$$\frac{\dot{M}_t}{P_t} = \dot{m}_t + \pi_t m_t. \tag{19.4}$$

The LHS of (19.4) is the real value of the money the government injects into the system. We call this total revenue from money creation, or seigniorage. Notice from the RHS of (19.4) that this has two components:

- The term $\dot{m}_t$ is the increase in real money holdings by the public. (It is sometimes referred to as seigniorage as well; we'll keep our use consistent).
- The term $m_t \pi_t$ is the inflation tax: the erosion, because of inflation, of the real value of the money balances held by the public. We can think of m_t as the tax base, and π_t as the tax rate.

Using (19.4) in (19.1) we have that

$$\dot{m}_t + \dot{b}_t = rb_t + y_t - \tau_t - c_t - \pi_t m_t. \tag{19.5}$$

On the LHS we have accumulation by the agent of the two available financial assets: money and bonds. The last term on the RHS is an additional expense: taxes paid on the real balances held.

Let us consider a steady state in which all variables are constant, then (19.5) becomes

$$rb + y = \tau + c + \pi m. \tag{19.6}$$

Hence, total income on the LHS must be enough to finance total expenditures (including regular taxes τ and the inflation tax πm).

A useful transformation involves adding and subtracting the term rm_t to the RHS of (19.5):

$$\dot{m}_t + \dot{b}_t = r\left(m_t + b_t\right) + y_t - \tau_t - c_t - \left(r + \pi_t\right) m_t. \tag{19.7}$$

Now define

$$a_t = m_t + b_t \tag{19.8}$$

as total financial assets held by the agent, and

$$i_t = r + \pi_t \tag{19.9}$$

as the nominal rate of interest. Using these two relationships in (19.7) we get

$$\dot{a}_t = ra_t + y_t - \tau_t - c_t - i_t m_t. \tag{19.10}$$

The last term on the RHS shows that the cost of holding money, in an inflationary environment, is the nominal rate of interest i_t.

19.2 | The Sidrauski model

Following Sidrauski (1967), we assume now the representative agent's utility function is

$$\int_0^{\infty} [u\left(c_t\right) + v(m_t)]e^{-\rho t}dt. \tag{19.11}$$

Here $v(m_t)$ is utility from holdings of real money balances. Assume $v'(m_t) \geq 0$, $v''(m_t) < 0$ and that Inada conditions hold. The agent maximises (19.11) subject to (19.10), which we repeat here for clarity, though assuming, without loss of generality, that output remains constant

$$\dot{a} = ra_t + y - \tau_t - c_t - i_t m_t,$$

plus the standard solvency condition

$$\lim_{T\to\infty} \left[a_T e^{-rT}\right] \geq 0, \tag{19.12}$$

and the initial condition a_0. The Hamiltonian is

$$H = [u(c_t) + v(m_t)] + \lambda_t \left(ra_t + y - \tau_t - c_t - i_t m_t\right), \tag{19.13}$$

where m_t and c_t are control variables, a_t is the state variable and λ_t is the co-state. First order conditions for a maximum are

$$u'(c_t) = \lambda_t, \tag{19.14}$$

$$v'(m_t) = \lambda_t i_t, \tag{19.15}$$

$$\dot{\lambda}_t = \lambda_t(\rho - r) = 0, \tag{19.16}$$

where the last equality comes from assuming $r = \rho$ as usual. Equations (19.14) and (19.16) together imply that c_t is constant and equal to c for all t. Using this fact and combining (19.14) and (19.15) we have

$$v'(m_t) = i_t u'(c). \tag{19.17}$$

We can think of equation (19.17) as defining money demand: demand for real balances is decreasing in the nominal interest rate i_t and increasing in steady state consumption c. This is a way to microfound the traditional money demand functions you all have seen before, where demand would be a positive function of income (because of transactions) and a negative function of the nominal interest rate, which is the opportunity cost of holding money.

19.2.1 | Finding the rate of inflation

What would the rate of inflation be in this model? In order to close the model, notice that

$$\frac{\dot{m}_t}{m_t} = \sigma - \pi_t, \tag{19.18}$$

where σ is the rate of money growth. We will also assume that the money printing proceeds are rebated to the consumer, which means that

$$\tau = -\sigma m_t. \tag{19.19}$$

Replacing (19.18) and (19.19) into (19.10), using $\rho = r$, and realizing the agent has no incentive to hold debt, gives that $c = y$, so that marginal utility is also constant and can be normalised to 1. Using (19.9), equation (19.17) becomes

$$v'(m_t) = \rho + \pi_t, \tag{19.20}$$

which substituting in (19.18) gives

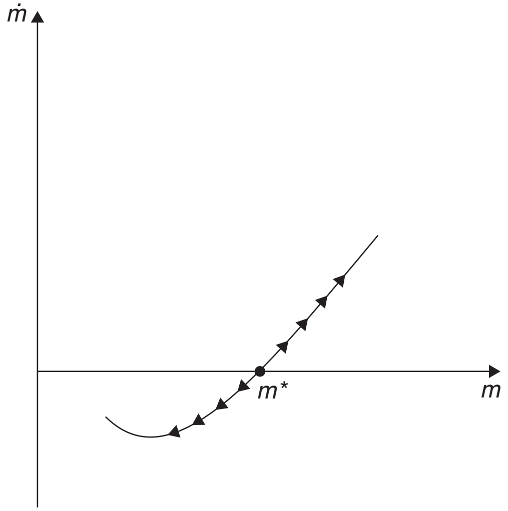
$$\dot{m}_t = (\rho + \sigma)m_t - v'(m_t)m_t. \tag{19.21}$$

Equation (19.21) is a differential equation that defines the equilibrium. Notice that because $v''(m) < 0$, this is an unstable differential equation. As the initial price level determines the initial point (m is a jump variable in our definitions of Chapter 3), the equilibrium is unique at the point where $\dot{m}_t = 0$. The dynamics are shown in Figure 19.1.

This simple model provides some of the basic intuitions of monetary theory.

- An increase in the quantity of nominal money will leave m unchanged and just lead to a jump in the price level. This is the quantitative theory of money that states that any increase in the stock of money will just result in an equivalent increase in prices.

Figure 19.1 The Sidrauski model

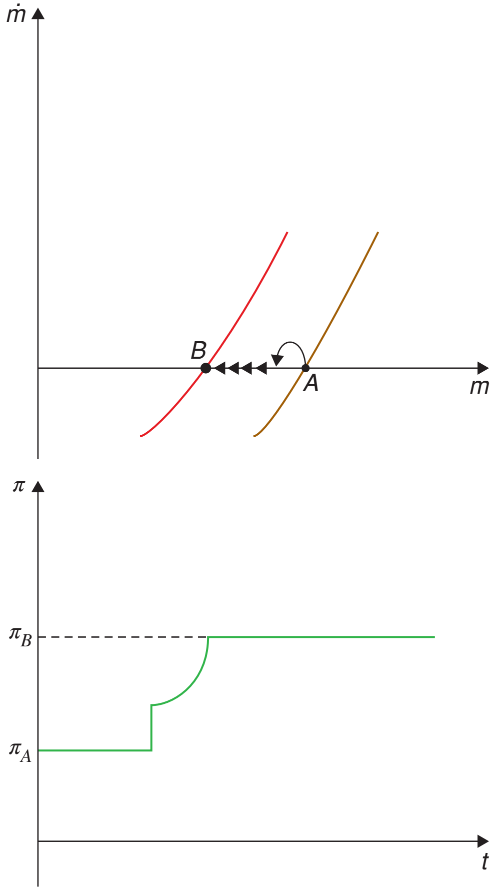


- The rate of inflation is the rate of growth of money (see equation (19.18)). Inflation is a monetary phenomenon.
- What happens if, suddenly, the rate of growth of money is *expected* to grow in the future? The dynamics entail a jump in the price level today and a divergent path which places the economy at its new equilibrium when the rate of growth finally increases. In short, increases in future money affect the price and inflation levels today. The evolution of m and π are shown in Figure 19.2.
- Does the introduction of money affect the equilibrium? It doesn't. Consumption is equal to income in all states of nature. This result is called the neutrality of money.

19.2.2 | The optimal rate of inflation

Let's assume now that we ask a central planner to choose the inflation rate in order to maximise welfare. What σ, and, therefore, what inflation rate would be chosen?

Figure 19.2 An anticipated increase in the money growth



We know from (19.20) and (19.17) that the steady-state stock of money held by individuals solves the equation

$$v'(m) = (\rho + \pi) = (\rho + \sigma). \tag{19.22}$$

This means that the central bank can choose σ to maximise utility from money-holdings. This implies choosing

$$\pi^{best} = \sigma^{best} = -\rho < 0, \tag{19.23}$$

so that

$$v'(m^{best}) = 0. \tag{19.24}$$

This means that m^{best} is the satiation stock of real balances and you achieve it by choosing a *negative* inflation rate. This is the famous Friedman rule for optimal monetary policy. What's the intuition? You should equate the marginal cost of holding money from an individual perspective (the nominal interest rate) to the social cost of printing money, which is essentially zero. A zero nominal rate implies an inflation rate that is equal to minus the real interest rate.

In practice, we don't see a lot of central banks implementing deflationary policy. Why is it so? Probably because deflation has a lot of costs that are left out of this model: its effect on debtors, on aggregate demand, etc., likely in the case when prices and wages tend to be sticky downwards.

We should thus interpret our result as meaning that policy makers should aim for low levels of inflation, so as to keep social and private costs close. In any case, there is a huge literature on the costs of inflation that strengthens the message of this result, we will come back to this at the end of the chapter.

19.2.3 | Multiple equilibria in the Sidrauski model

In the previous section we analysed the steady state of the model, but, in general, we have always been cautious as to check if other equilibria are possible. In this monetary model, as it happens, they are.

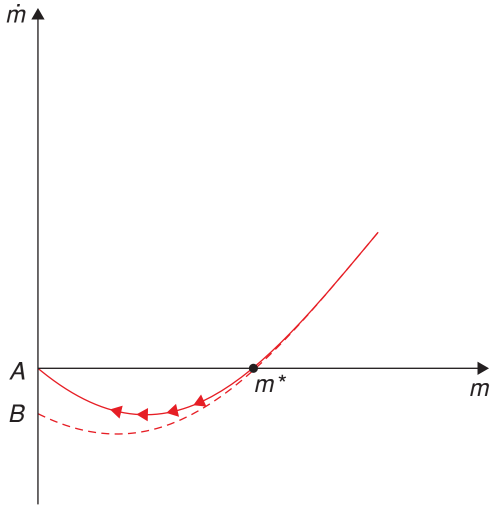
Figure 19.3 shows the possible configurations for equation (19.21), for all m. We know that

$$\left.\frac{\partial \dot{m}}{\partial m}\right|_{SS} = -v''(m) > 0, \tag{19.25}$$

so that the curve crosses the steady state with a positive slope. But what happens to the left of the steady state? Figure 19.3, shows two paths depending on whether the value of the term $v'(m)m$ approaches zero or a positive number as m approaches zero. If money is very essential and its marginal utility is very high as you reduce your holdings of money, then $v'(m)m > 0$ as m approaches zero. This case corresponds to the path denoted by the letter B. If $v'(m)m \rightarrow 0$, as $m \rightarrow 0$ then the configuration is of the path leading to A.

With this we can now study other equilibria. The paths to the right are deflationary paths, where inflation is negative and real balances increase without bound. We do not see these increasing deflationary paths, so, from an empirical point of view, they do not seem very relevant (mathematically they are feasible, and some people resorted to these equilibria to explain the low inflation rates in the U.S. in recent years, see Sims (2016)). The paths to the left of the steady state are inflationary paths. Paths along the B curve are inconsistent, as they require $\dot{m} < 0$ when m hits zero, which is unfeasible. However, paths that do end up at zero, denoted A in Figure 19.3, are feasible. In these cases money is not so essential, so it is wiped out by a hiperinflationary process. In a classical paper, Cagan (1956)

Figure 19.3 Multiple equilibria in the Sidrauski model



speculated on the possibility of these self-sustaining inflationary dynamics in which the expectation of higher inflation leads to lower money demand, fuelling even higher inflation. So these feasible paths to the left of the steady state could be called Cagan equilibria. The general equilibrium version of the Cagan equilibria described here was first introduced by Obstfeld and Rogoff (1983).

19.2.4 | Currency substitution

The model is amenable to discussing the role of currency substitution, that is, the possibility of phasing out the currency and being replaced by a sounder alternative.

The issue of understanding how different currencies interact, has a long tradition in monetary economics. Not only because, in antiquity, many objects operated as monies, but also because, prior to the emergence of the Fed, currency in the U.S. were issued by commercial banks, so there was an innumerable number of currencies circulating at each time. A popular way to think this issue is Gresham's Law; faced with a low quality currency and a high quality currency, Gresham's Law argues that people will try to get rid of the low quality currency while hoarding the high quality currency, *bad money displaces good money*. Of course while this may be true at the individual level, it may not be so at the aggregate level because prices may increase faster when denominated in units of the bad-quality currency debasing its value. Sturzenegger (1994) discusses this issue and makes two points.

- When there are two or more currencies, it is more likely that the condition $v'(m)m=0$ is satisfied (particularly for the low quality currency). Thus, the hyperinflation paths are more likely.
- If the dynamics of money continue are described by an analogous to (19.21) such as

$$\dot{m}_1 = (\rho + \sigma_1)m_1 - v'(m_1, m_2)m_1, \tag{19.26}$$

  notice that if the second currency m_2 reduces the marginal utility of the first one, then the inflation rate on the equilibrium path is lower: less inflation is needed to wipe out the currency.

This pattern seems to have occurred in a series of hyperinflations in Argentina in the late 80s, each new wave coming faster but with *lower* inflation. Similarly, at the end of the 2000s, also in Argentina, very tight monetary conditions during the fixed exchange regime led to the development of multiple private currencies. Once the exchange rate regime was removed, these currencies suffered hyperinflations and disappeared in a wink (see Colacelli and Blackburn 2009).

19.2.5 | Superneutrality

How do these results extend to a model with capital accumulation? We can see this easily also in the context of the Sidrauski model (we assume no population growth), but where we give away the assumption of exogenous output and allow for capital accumulation. Consider now the utility function

$$\int_0^\infty u\left(c_t, m_t\right) e^{-\rho t} dt, \tag{19.27}$$

where $u_c, u_m > 0$ and $u_{cc}, u_{mm} < 0$. However, we'll allow the consumer to accumulate capital now. Defining again $a = k + m$, the resource constraint can be written as

$$\dot{a}_t = r_t a_t + w_t - \tau_t - c_t - i_t m_t. \tag{19.28}$$

The Hamiltonian is

$$H = u\left(c_t, m_t\right) + \lambda_t \left[r a_t + w_t - \tau_t - c_t - i_t m_t\right]. \tag{19.29}$$

The FOC are, as usual,

$$u_c(c_t, m_t) = \lambda_t, \tag{19.30}$$

$$u_m(c_t, m_t) = \lambda_t i_t, \tag{19.31}$$

$$\dot{\lambda}_t = (\rho - r)\lambda_t. \tag{19.32}$$

The first two equations give, once again, a money demand function $u_m = u_c i$, but the important result is that because the interest rate now is the marginal product of capital, in steady state $r = \rho = f'(k^*)$, where we use the $*$ superscript to denote the steady state. We leave the computations to you, assuming $\tau = -\sigma m$, and using the fact that w is the marginal product of labour, replacing in (19.28) we find that

$$c = f(k^*). \tag{19.33}$$

But this is the level of income that we would have had in the model with no money! This result is known as superneutrality: not only does the introduction of money not affect the equilibrium, neither does the inflation rate.

Later, we will see the motives for why we believe this is not a good description of the effects of inflation, which we believe in the real world are harmful for the economy.

19.3 | The relation between fiscal and monetary policy

If inflation originates in money printing, the question is, what originates money printing? One possible explanation for inflation lies in the need of resources to finance public spending. This is called the public finance approach to inflation and follows the logic of our tax smoothing discussion in the previous chapter. According to this view, taxes generate distortions, and the optimal taxation mix entails equating these distortions across all goods. If we think of money holdings as another good, it is optimal to tax money holdings – that is, with inflation – if other goods are taxed as well. Thus, the higher the cost of collecting other taxes (the weaker your tax system), the more you should rely on inflation as a form of collecting income. If the marginal cost of taxes increases with recessions, then you should use more inflation in downturns.

Another reason for inflation is to compensate the natural tendency towards deflation. If prices were constant, we would probably have deflation, because we know that price indexes suffer from an upward bias. As new products come along and relative prices move, people change their consumption mix looking for cheaper alternatives, so their actual basket is always "cheaper" than the measured basket. For the U.S., this bias is allegedly around 1% per year, but it has been found larger for emerging economies.[1] Thus an inflation target of 1 or 2% in fact aims, basically, at price stability.

However, the main culprit for inflation, is, obviously, fiscal needs regardless of any optimisation consideration. The treasury needs resources, does not want to put with the political pain of raising taxes, and simply asks the central bank to print some money which eventually becomes inflation.

19.3.1 | The inflation-tax Laffer curve

The tax collected is the combination of the inflation rate and the money demand that pays that inflation tax. Thus, a question arises as to whether countries may choose too high an inflation rate. May the inflation rate be so high that discouraging money demand actually reduces the amount collected through the inflation tax? In other words are we on the wrong side of the Laffer curve?[2]

To explore this question let's start with the budget constraint for the government,

$$\dot{m}_t = rd_0 + g - \tau - \pi_t m_t, \tag{19.34}$$

which, in steady state, becomes

$$rd_0 + g - \tau = \pi m. \tag{19.35}$$

Assuming a typical demand function for money

$$m = ye^{-\gamma i}, \tag{19.36}$$

we can rewrite this as

$$rd_0 + g - \tau = \pi ye^{-\gamma(r+\pi)}. \tag{19.37}$$

Note that

$$\frac{\partial \left(\pi e^{-\gamma(r+\pi)}\right)}{\partial \pi} = ye^{-\gamma(r+\pi)}(1 - \gamma\pi), \tag{19.38}$$

so that revenue is increasing in π for $\pi < \gamma^{-1}$, and decreasing for $\pi > \gamma^{-1}$. It follows that $\pi = \gamma^{-1}$ is the revenue maximising rate of inflation. Empirical work, however, has found, fortunately, that government typically place themselves on the correct side of the Laffer curve.[3]

19.3.2 | The inflation-tax and inflation dynamics

What are the dynamics of this fiscally motivated inflation? Using (19.36), we can write,

$$\pi_t = \gamma^{-1}(\log(y) - \log(m_t)) - r. \tag{19.39}$$

This in (19.34) implies

$$\dot{m}_t = rd_0 + g - \tau - \gamma^{-1}\left(\log(y) - \log(m_t)\right) m_t + rm_t. \tag{19.40}$$

Notice that,

$$\left.\frac{\partial \dot{m}_t}{\partial m_t}\right|_{SS} = -\gamma^{-1}(\log(y) - \log(m)) + \gamma^{-1} + r, \tag{19.41}$$

which using (19.39)

$$\pi_t = \gamma^{-1}(\log(y) - \log(m_t)) - r. \tag{19.42}$$

simplifies to,

$$\left.\frac{\partial \dot{m}_t}{\partial m_t}\right|_{SS} = \gamma^{-1} - \pi_t. \tag{19.43}$$

Hence, $\left.\frac{\partial \dot{m}_t}{\partial m_t}\right|_{SS} > 0$ for the steady state inflation below γ^{-1}, and $\left.\frac{\partial \dot{m}_t}{\partial m_t}\right|_{SS} < 0$ for the steady state inflation rate above γ^{-1}.

This means that the high inflation equilibrium is stable. As m is a jumpy variable, this means that, in addition to the well-defined equilibrium at low inflation, there are infinite equilibria in which inflation converges to the high inflation equilibria.

Most practitioners disregard this high inflation equilibria and focus on the one on the good side of the Laffer curve, mostly because, as we said, it is difficult to come up with evidence that countries are on the wrong side. However, the dynamics should be a reminder of the challenges posed by stabilisation.

19.3.3 | Unpleasant monetary arithmetic

In this section we will review one of the most celebrated results in monetary theory, the unpleasant monetarist arithmetic presented initially by Sargent and Wallace (1981). The result states that a monetary contraction may lead to higher inflation in the future. Why? Because, if the amount of government spending is exogenous and is not financed with seigniorage, it has to be financed with bonds. If eventually seigniorage is the only source of revenue, the higher amount of bonds will require more seigniorage and, therefore, more inflation. Of course, seigniorage is not the only financing mechanism, so you may interpret the result as applying to situations when, eventually, the increased cost of debt is not financed, at least entirely, by other revenue sources. Can it be the case that the expected future inflation leads to higher inflation now? If that were the case, the contractionary monetary policy would be ineffective even in the short run! This section discusses if that can be the case.

The tools to discuss this issue are all laid out in the Sidrauski model discussed in section 19.2, even though the presentation here follows Drazen (1985).

Consider the evolution of assets being explicit about the components of a,

$$\dot{b}_t + \dot{m}_t = -\pi_t m_t + y + \rho b_t - c_t. \tag{19.44}$$

Where we assume $r = \rho$ as we've done before. The evolution of real money follows

$$\dot{m}_t = (\sigma - \pi_t)m_t. \tag{19.45}$$

Replacing (19.45) into (19.44), we get

$$\dot{b}_t = -\sigma m_t + y - c_t + \rho b_t, \tag{19.46}$$

where the term $y - c$ can be interpreted as the fiscal deficit.[4] Call this expression D. Replacing (19.20) in (19.45) we get

$$\dot{m}_t = (\sigma + \rho - v'(m_t))m_t. \tag{19.47}$$

Equations (19.46) and (19.47) will be the dynamic system, which we will use to discuss our results. It is easy to see that the $\dot{b}$ equation slopes upwards and that the $\dot{m}$ is an horizontal line. The dynamics are represented in Figure 19.4. A reduction in σ shifts both curves upwards.

Notice that the system is unstable. But b is not a jump variable. The system reaches stability only if the rate of money growth is such that it can finance the deficit stabilising the debt dynamics. It is the choice of money growth that will take us to the equilibrium. b here is not the decision variable.

Our exercise considers the case where the rate of growth of money falls for a certain period of time after which it moves to the value needed to keep variables at their steady state. This exercise represents well the case studied by Sargent and Wallace.

To analyse this we first compute all the steady state combinations of m and b for different values of σ. Making $\dot{b}$ and $\dot{m}$ equal to zero in (19.46) and (19.47) and substituting σ in (19.46) using (19.47), we get

$$b = \frac{mv'(m)}{\rho} - m - \frac{D}{\rho}. \tag{19.48}$$

This is the SS locus in Figure 19.5. We know that eventually the economy reverts to a steady state along this line. To finalize the analysis, show that the equation for the accumulation of assets can be written as

$$\dot{a}_t = \rho a_t - v'(m_t)m_t + D. \tag{19.49}$$

Figure 19.4 The dynamics of *m* and *b*

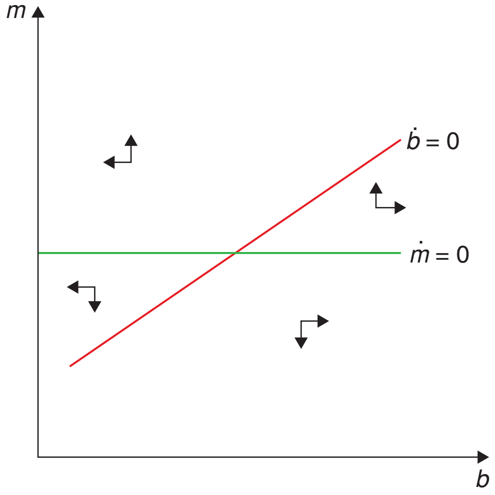
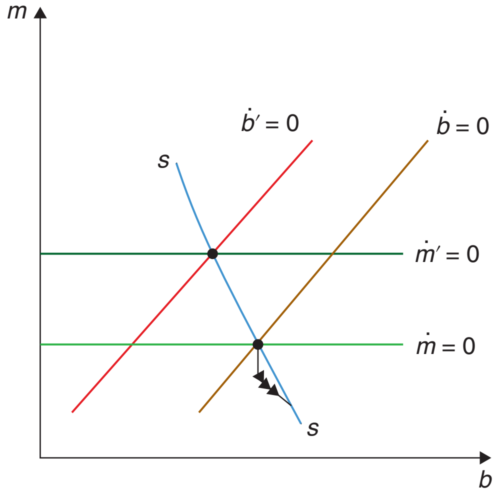


Figure 19.5 Unpleasant monetarist arithmetic



notice, however, that if $\dot{a} = 0$ this equation coincides with (19.48). This means that above the steady states locus the dynamic paths have a slope that is less than one (so that the sum of m and b grows as you move) and steeper than one below it (so that the total value of assets falls).

We have now the elements to discuss our results. Consider first the case where $v(m) = \log(m)$. In this case the inflation tax is constant and independent of the inflation rate. Notice that this implies from (19.44) that the $\dot{b} = 0$ line is vertical. In this case, the reduction in the growth rate of money implies a jump to the lower inflation rate, but the system remains there and there is no unpleasant monetary arithmetic. A lowering of the rate of growth of money, does not affect the collection of the inflation tax and thus does not require more debt financing, so the new lower inflation equilibrium can sustain itself, and simply jumps back to the original point when the growth rate of money reverts to its initial value.

Now consider that case where the demand for money is relatively inelastic, which implies that, in order to increase seigniorage, a higher inflation rate is required and the slope of the SS curve is negative.[5] Now the policy of reducing seigniorage collection for some time will increase inflation in the long run as a result of the higher level of debt. This is the soft version of the Sargent-Wallace result.

But the interesting question is whether it may actually increase inflation even in the short run, something we call the hard version of the unpleasant monetarist arithmetic, or, in Drazen's words, the spectacular version.

Whether this is the case will depend on the slope of the SS curve. If the curve is flat then a jump in m is required to put the economy on a path to a new steady state. In this case, only the soft, and not the hard, version of the result holds (an upwards jump in m happens only if inflation falls). However, if the SS curve is steeper than negative one (the case drawn in (19.5), only a downwards jump in m can get us to the equilibrium. Now we have Sargent and Wallace's spectacular, unpleasant monetary result: lowering the rate of money growth can actually increase the inflation rate in the short run! The more inelastic money demand, then the more likely this is to be in this case.

Of course these results do not carry to all bond issues. If, for example, a central bank sells bonds B_t to buy foreign reserves Re_t (where e_t is the foreign currency price in domestic currency units), the central bank income statement changes by adding an interest cost $i\Delta B_t$ but also adds a revenue equal to i^*Re_t where i and i^* stand for the local and foreign interest rates. If $\Delta B = Re$, to the extent that $i = i^* + \frac{\dot{e}}{e}$ (uncovered interest parity), there is no change in net income, and therefore no change in the equilibrium inflation rate.

This illustrates that the Sargent-Wallace result applies to bond sales that compensate money printed to finance the government (i.e. with no backing). In fact, in Chapter 21 we will discuss the policy of quantitative easing, a policy in which Central Banks issue substantial amount of liquidity in exchange for real assets, such as corporate bonds, and other financial instruments, finance with interest bearing reserves. To the extent that these resources deliver an equilibrium return, they do not change the monetary equilibrium.

19.3.4 | Pleasant monetary arithmetic

Let's imagine now that the government needs to finance a certain level of government expenditure, but can choose the inflation rates over time. What would be the optimal path for the inflation tax? To find out, we assume a Ramsey planner that maximises consumer utility, internalising the optimal behaviour of the consumer to the inflation tax itself, much in the same way we did in the previous chapter in our discussion of optimal taxation; and, of course, subject to its own budget constraint.[6] The problem is then to maximise

$$\int_0^\infty [u\left(y\right) + v(L(i_t, y))]e^{-\rho t}dt, \tag{19.50}$$

where we replace c for y and m_t for $L(i_t, y)$, as per the results of the Sidrauski model. The government's budget constraint is

$$\dot{a}_t = \rho a_t - i_t m_t + d_t, \tag{19.51}$$

where $a_t = \frac{B_t+M_t}{P_t}$ is the real amount of liabilities of the government, d_t is the government deficit and we've replaced $r = \rho$. The Ramsey planner has to find the optimal sequence of interest rates, that is, of the inflation rate. The FOCs are

$$v_m L_i + \lambda_t \left[L(i_t, y) + i_t L_i\right] = 0, \tag{19.52}$$

plus

$$\dot{\lambda}_t = \rho\lambda_t - \rho\lambda_t. \tag{19.53}$$

The second FOC show that λ is constant. Given this the first FOC shows the nominal interest is constant as well. Optimal policy smooths the inflation tax across periods, a result akin to our tax smoothing result in the previous chapter (if we include a distortion from taxation, we would get that the marginal cost of inflation should equal the marginal cost of taxation, delivering the result that inflation be countercyclical).

Consider what happens if the government faces a decreasing path for government expenditures, that is

$$d_t = d_0 e^{-\delta t}. \tag{19.54}$$

The solution still requires a constant inflation rate but now the seigniorage needs to satisfy

$$\pi^* m^* = \rho b_0 + \rho\frac{d_0}{\rho + \delta}. \tag{19.55}$$

Integrating (19.51) gives the solution for a_t

$$b_t = \frac{\pi^* m^*}{\rho} - \frac{d_0}{\rho + \delta} e^{-\delta t}. \tag{19.56}$$

Notice that debt increases over time: the government smoothes the inflation tax by running up debt during the high deficit period. This debt level is higher, of course, relative to a policy of financing the deficit with inflation in every period (this would entail a decreasing inflation path pari passu with the deficit). At the end, the level of debt is higher under the smoothing equilibrium than under the policy of full inflation financing, leading to higher steady state inflation. This is the monetaristic arithmetic at work. However, far from being unpleasant, this is the result of an optimal program. The higher long run inflation is the cost of smoothing the inflation in other periods.

19.4 | The costs of inflation

The Sidrauski model shows that inflation does not affect the equilibrium. But somehow we do not believe this result to be correct. On the contrary, we believe inflation is harmful to the economy. In their celebrated paper, Bruno and Easterly (1996) found that, beyond a certain threshold inflation was negatively correlated with growth, a view that is well established among practicioners of monetary policy. This result is confirmed by the literature on growth regressions. Inflation always has a negative and significant effect on growth. In these regressions it may very well be that inflation is capturing a more fundamental weakness as to how the political system works, which may suggest that for these countries it is not as simple as "choosing a better rate of inflation".

However, to make the point on the costs of inflation more strongly, we notice that even disinflation programs are expansionary. This means that the positive effects of lowering inflation are strong, so much so that they even undo whatever potential costs a disinflation may have. Figures 19.6 and 19.7 show all recent disinflation programs for countries that had reached an inflation rate equal to or higher than 20% in recent years. The figure is split in two panels, those countries that implemented disinflation with a floating regime and those that used some kind of nominal anchor (typically the exchange rate), and shows the evolution of inflation (monthly) in 19.6 and GDP (quarterly) in 19.7 since the last time they reached 20% anual inflation. The evidence is conclusive: disinflations are associated with higher growth.

So what are these costs of inflation that did not show up in the Sidrauski model? There has been a large literature on the costs of inflation. Initially, these costs were associated with what were dubbed shoe-leather costs: the cost of going to the bank to get cash (the idea is that the higher the inflation, the lower your demand for cash, and the more times you needed to go to the bank to get your cash). This was never a thrilling story (to say the least), but today, with electronic money and credit cards, simply no longer makes any sense. On a more benign note we can grant it tries to capture all the increased transaction costs associated with running out (or low) of cash.

Other stories are equally disappointing. Menu costs (the idea that there are real costs of changing prices) is as uneventful as the shoe-leather story. We know inflation distorts tax structures and redistributes incomes across people (typically against the poorest in the population), but while these are undesirable consequences they on their own do not build a good explanation for the negative impact of inflation on growth.

Figure 19.6 Recent disinflations (floating and fixed exchange rate regimes)

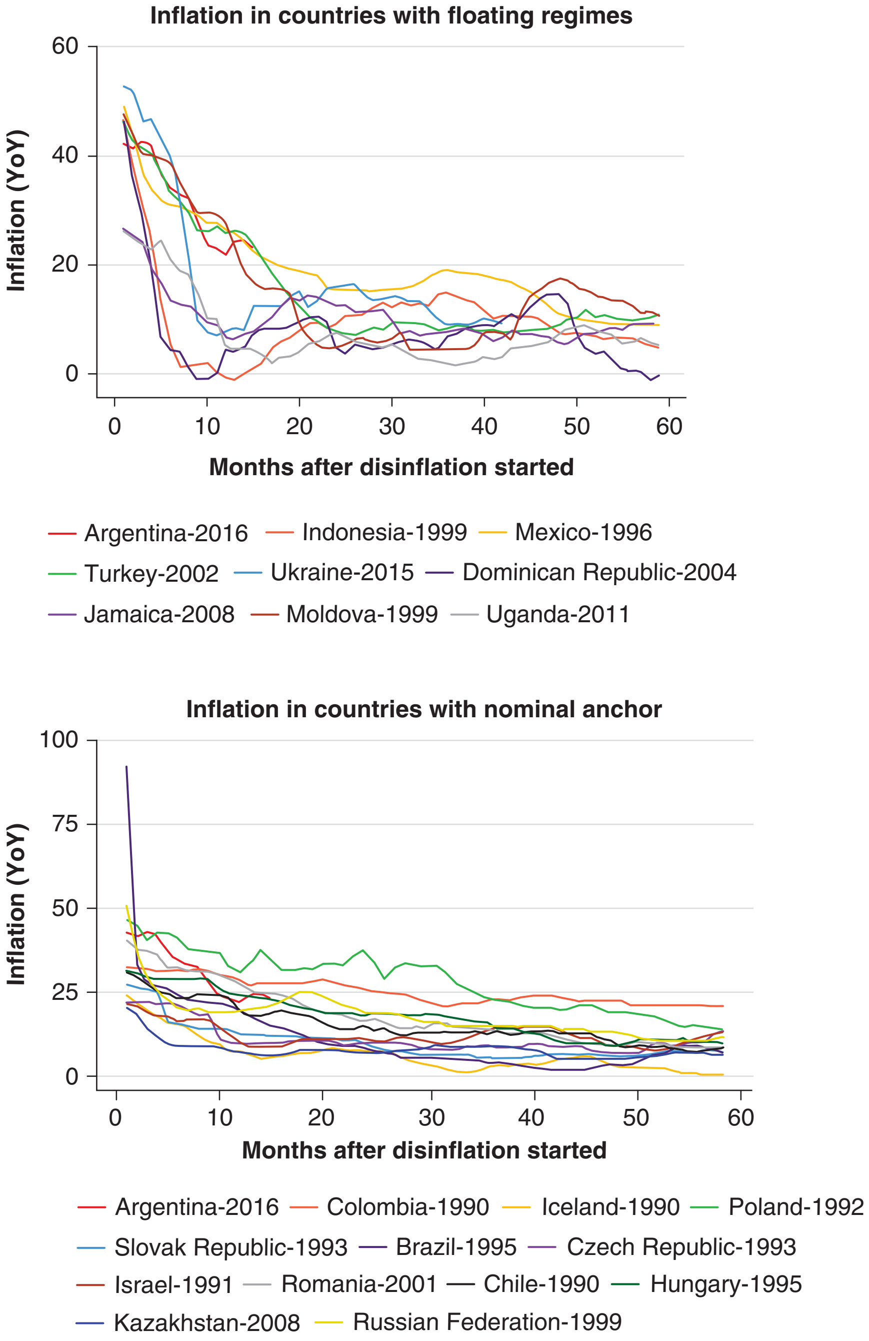
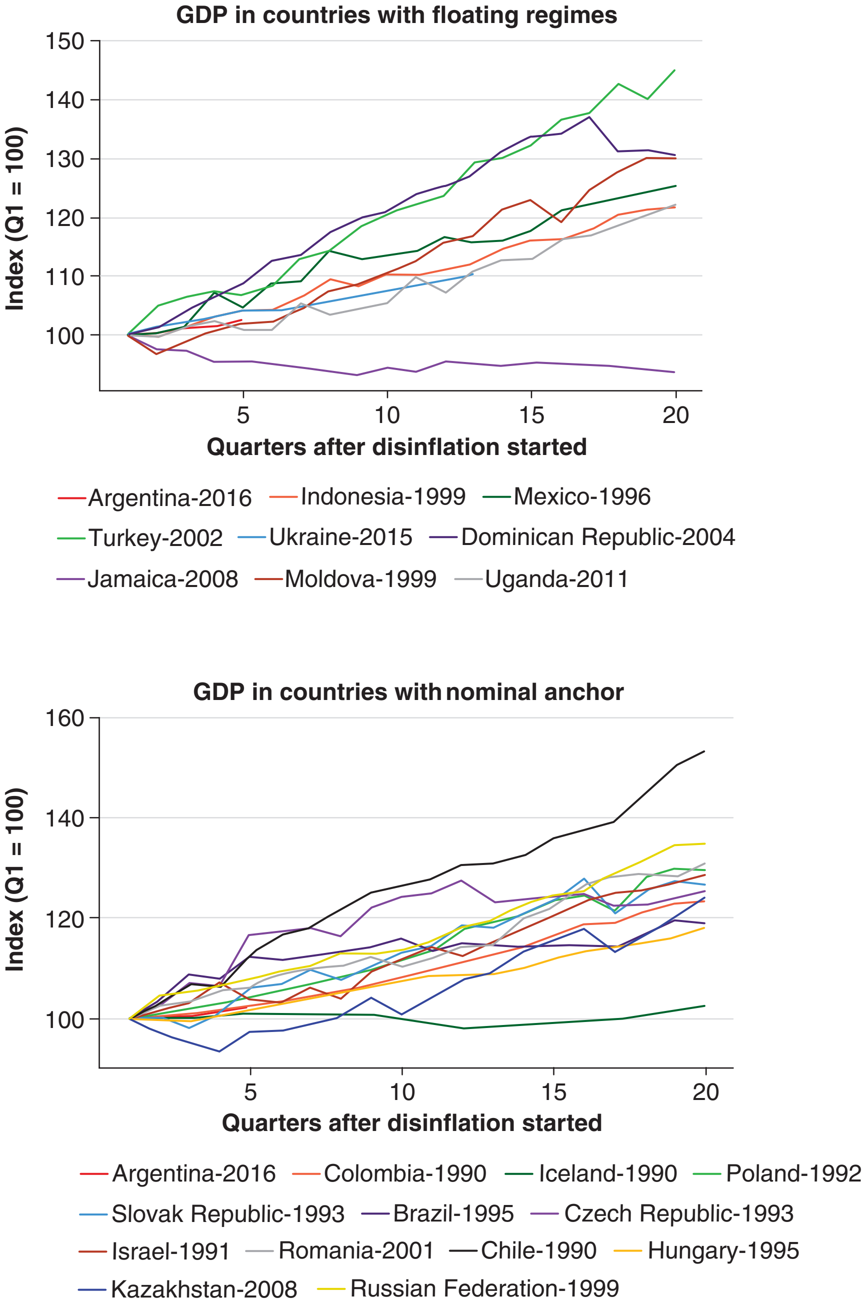


Figure 19.7 GDP growth during disinflations



19.4.1 | The Tommasi model: Inflation and competition

So the problem with inflation has to be significant and deep. An elephant in the room that seems difficult to see. Tommasi (1994) provides what we believe is a more plausible story based on the role of inflation in messing up the price system. Tommasi focuses on a well-known fact: increases in inflation increase the volatility in relative prices (this occurs naturally in any model where prices adjust at different times or speed). Tommasi argues that relative prices changes, not only generate economic inefficiencies but also change the relative power of sellers and purchasers pushing the economy away from its competitive equilibrium. To see this, let's draw from our analysis of search discussed in Chapter 16.

Imagine a consumer that is searching for a low price. Going to a store implies finding a price, the value of which can be described by

$$rW(p) = (\bar{x} - p) + \rho[U - W(p)]. \tag{19.57}$$

Having a price implies obtaining a utility $\bar{x} - p$. If relative prices were stable, the consumer could go back to this store and repurchase, but if relative prices change, then this price is lost. This occurs with probability ρ. If this event occurs, the consumer is left with no offer (value U). The ρ parameter will change with inflation and will be our object of interest. If the consumer has no price, he needs to search for a price with cost C and value U as in

$$rU = -C + \alpha \int_0^\infty max(0, W(p) - U)dF(p). \tag{19.58}$$

Working analogously as we did in the case of job search, remember that the optimal policy will be determined by a reservation price p_R. As this reservation price is the one that makes the customer indifferent between accepting or not accepting the price offered, we have that $rW(p_R) = \bar{x} - p_R = rU$, which will be handy later on. Rewrite (19.57) as

$$W(p) = \frac{\bar{x} - p + \rho U}{r + \rho}. \tag{19.59}$$

Subtracting U from both sides (and using $rU = \bar{x} - p_R$), we have

$$W(p) - U = \frac{\bar{x} - p + \rho U}{r + \rho} - U = \frac{\bar{x} - p - rU}{r + \rho} = \frac{p_R - p}{r + \rho}. \tag{19.60}$$

We can now replace rU and $W(p) - U$ in (19.58) to obtain

$$\bar{x} - p_R = -C + \frac{\alpha}{r + \rho} \int_0^{p_R} (p_R - p)dF(p), \tag{19.61}$$

or, finally,

$$p_R = C + \bar{x} - \frac{\alpha}{r + \rho} \underbrace{\int_0^{p_R} (p_R - p)\, dF(p)}_{(+)}. \tag{19.62}$$

The intuition is simple. The consumer is willing to pay up to his valuation of the good $\bar{x}$ plus the search cost C that can be saved by purchasing this unit. However, the reservation price falls if there is expectation of a better price in a new draw.

The equation delivers the result that if higher inflation implies a the higher ρ, then the higher is the reservation price. With inflation, consumers search less thus deviating the economy from its competitive equilibrium.

Other stories have discussed possible other side effects of inflation. There is a well documented negative relation between inflation and the size of the financial sector (see for example Levine and Renelt (1991) and Levine and Renelt (1992)). Another critical feature is the fact that high inflation implies that long term nominal contracts disappear, a point which becomes most clear if inflation may change abruptly. Imagine a budget with an investment that yields a positive or negative return x or $-x$, in a nominal contract this may happen if inflation moves strongly. Imagine that markets are incomplete and agents cannot run negative net worth (any contract which may run into negative wealth is not feasible). The probability of eventually running into negative wealth increases with the length of the contract. [7] The disappearance of long term contracts has a negative impact on productivity.

19.4.2 | Taking stock

We have seen how money and inflation are linked in the long run, and that a simple monetary model can help account for why central banks would want to set inflation at a low level. We haven't really talked about the short run, in fact, in our model there are no real effects of money or monetary policy. However, as you anticipate by now, this is due to the fact that there are no price rigidities. To the extent that prices are flexible in the long run, the main concern of monetary policy becomes dealing with inflation, and this is how the practice has evolved in recent decades. If there are rigidities, as we have seen previously, part of the effect of monetary policy will translate into output, and not just into the price dynamics. It is to these concerns that we turn in the next chapter.

Notes

[1] de Carvalho Filho and Chamon (2012) find a 4.5% annual bias for Brazil in the 80s. Gluzmann and Sturzenegger (2018) find a whopping 7% bias for 85–95 in Argentina, and 1% for the period 95–2005.

[2] You may know this already, but the Laffer curve describes the evolution of tax income as you increase the tax rate. Starts at zero when the tax rate is zero, and goes back to zero when the tax rate is 100%, as probably at this high rates the taxable good has disappeared. Thus, there is a range of tax rates where increasing the tax rate decreases tax collection income.

[3] See Kiguel and Neumeyer (1995).

[4] If $y = c + g$ then $y - c = g$, and as there are no tax resources, it indicates the value of the deficit.

[5] We disregard the equilibria where the elasticity is so high that reducing the rate of money growth increases the collection of the inflation tax. As in the previous section, we disregard these cases because we typically find the inflation tax to operate on the correct side of the Laffer curve.

[6] This section follows Uribe (2016).

[7] For a contract delivering a positive or negative return x with equal probabilities in each period, the possibility of the contract eventually hitting a negative return is .5 if it lasts one period and $.5 + \sum_{3,5,\ldots}^{\infty} \frac{1}{n+1}\frac{(n-2)!!}{(n-1)!!}$ if it lasts n periods. This probability is bigger than 75% after nine periods, so, quickly long term contracts become unfeasible. See Neumeyer (1998) for a model along these lines.

References

Bruno, M. & Easterly, W. (1996). Inflation and growth: In search of a stable relationship. *Federal Reserve Bank of St. Louis Review*, *78*(May/June 1996).

Cagan, P. (1956). The monetary dynamics of hyperinflation. *Studies in the Quantity Theory of Money*.

Colacelli, M. & Blackburn, D. J. (2009). Secondary currency: An empirical analysis. *Journal of Monetary Economics*, *56*(3), 295–308.

de Carvalho Filho, I. & Chamon, M. (2012). The myth of post-reform income stagnation: Evidence from Brazil and Mexico. *Journal of Development Economics*, *97*(2), 368–386.

Drazen, A. (1985). Tight money and inflation: Further results. *Journal of Monetary Economics*, *15*(1), 113–120.

Gluzmann, P. & Sturzenegger, F. (2018). An estimation of CPI biases in Argentina 1985–2005 and its implications on real income growth and income distribution. *Latin American Economic Review*, *27*(1), 8.

Kiguel, M. A. & Neumeyer, P. A. (1995). Seigniorage and inflation: The case of Argentina. *Journal of Money, Credit and Banking*, *27*(3), 672–682.

Kiyotaki, N. & Wright, R. (1989). On money as a medium of exchange. *Journal of Political Economy*, *97*(4), 927–954.

Levine, R. & Renelt, D. (1991). *Cross-country studies of growth and policy: Methodological, conceptual, and statistical problems* (Vol. 608). World Bank Publications.

Levine, R. & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *The American Economic Review*, 942–963.

Neumeyer, P. A. (1998). Inflation-stabilization risk in economies with incomplete asset markets. *Journal of Economic Dynamics and Control*, *23*(3), 371–391.

Obstfeld, M. & Rogoff, K. (1983). Speculative hyperinflations in maximizing models: Can we rule them out? *Journal of Political Economy*, *91*(4), 675–687.

Sargent, T. J. & Wallace, N. (1981). Some unpleasant monetarist arithmetic. *Federal Reserve Bank of Minneapolis Quarterly Review*, *5*(3), 1–17.

Sidrauski, M. (1967). Rational choice and patterns of growth in a monetary economy. *The American Economic Review*, *57*(2), 534–544.

Sims, C. A. (2016). Fiscal policy, monetary policy and central bank independence. https://bit.ly/3BmMJYE.

Sturzenegger, F. A. (1994). Hyperinflation with currency substitution: Introducing an indexed currency. *Journal of Money, Credit and Banking*, *26*(3), 377–395.

Tommasi, M. (1994). The consequences of price instability on search markets: Toward understanding the effects of inflation. *The American Economic Review*, 1385–1396.

Uribe, M. (2016). *Is the monetarist arithmetic unpleasant? (tech. rep.).* National Bureau of Economic Research.

CHAPTER 20

Rules vs Discretion

Now, let's move back to the New Keynesian world of Chapter 15 where the existence of nominal rigidities implies that monetary policy (MP) can have real effects. Most central banks believe this a more realistic description of the environment, at least in the short run. In such a world, MP has to assess the trade-offs when it comes to stabilising inflation versus stabilising output. In this chapter we develop a framework that will let us analyse this.

20.1 | A basic framework

Fortunately, we have already developed most of the ingredients of such framework: it's the canonical New Keynesian model! As you may recall, it is founded on two basic equations, the New Keynesian IS curve (NKIS), and the New Keynesian Phillips curve (NKPC), which we rewrite here for your convenience. First, the NKIS:

$$y_t = E_t[y_{t+1}] - \sigma \left(i_t - E_t[\pi_{t+1}] - \rho\right) + u_t^{IS}. \tag{20.1}$$

This is exactly as we had before, with u_t^{IS} corresponding to an (aggregate demand) shock. We specify shocks being a random, white-noise disturbance.

Now, the NKPC:

$$\pi_t = \kappa(y_t - y_t^n) + \beta E_t[\pi_{t+1}] + u_t^{\pi}, \tag{20.2}$$

with u_t^{π} corresponding to an (aggregate supply) shock. If you check this against the specification of previous chapters, the main difference you will notice is the existence of these demand and supply shocks.

You will recall that, when we discussed the canonical NK model, we talked about an interest rate rule, namely the celebrated Taylor rule. Now is the time to think about the nature of monetary policy rules more broadly.

20.1.1 | Time inconsistency

The first thing we have to do is to think about what the central bank/policy-maker (CB, for shorthand) wants to do. We assume that, when it comes to inflation, it wants to minimise departures from the optimal level, which we normalize to zero. (Again, it could be positive, could be negative – it's just

---

How to cite this book chapter:
Campante, F., Sturzenegger, F. and Velasco, A. 2021. *Advanced Macroeconomics: An Easy Guide.*
Ch. 20. 'Rules vs Discretion', pp. 315–322. London: LSE Press.
DOI: https://doi.org/10.31389/lsepress.ame.t License: CC-BY-NC 4.0.

a normalization.) When it comes to output, we will introduce a more consequential assumption: we take that the CB wants to minimise deviations not from the natural rate (y), but rather from what we may call the Walrasian rate of output, which we call y^*. Think of this as the output level that would prevail in the absence of any market distortions, such as monopoly power or distortionary taxation. The idea is that it is almost surely the case that $y^* > y$ – monopolies produce suboptimal quantities, distortionary taxes lead to suboptimal effort, etc.

In order to capture this idea, we will usually think of the CB as minimising a loss function like this:

$$L = \frac{1}{2}\left[\alpha\pi_t^2 + \left(y_t - y^*\right)^2\right], \tag{20.3}$$

where $\alpha > 0$ denotes the relative importance of inflation as compared to output deviations.

To discuss the implications, let's develop a model to deal with this issue in the spirit of Rogoff (1985). The details follow Velasco (1996) which uses a simpler Phillips curve, but which captures the spirit of (20.2). In this simplified version the economy is fully characterised by the expectational Phillips curve

$$y_t - y = \theta\left(\pi_t - \pi_t^e\right) + z_t, \quad \theta > 0, \tag{20.4}$$

where π is the actual rate of inflation, π^e is the expected rate, y_t is actual output, y is steady state (or natural rate) output, and z_t is a random shock (which should be interpreted here as a supply shock) with mean zero and variance σ^2. The term $\theta\left(\pi_t - \pi_t^e\right)$ implies that whenever actual inflation is below expected inflation, output falls. Notice that the supply shock is the only shock here (we assume away demand shocks, whether of the nominal or real kind).

The social loss function is

$$L = \left(\frac{1}{2}\right)\left(\alpha\pi_t^2 + \left(y_t - \gamma y\right)^2\right), \quad \alpha > 0, \gamma > 1. \tag{20.5}$$

The function (20.5) indicates that society dislikes fluctuations in both inflation and output. Notice that the bliss output rate is $y^* = \gamma y$, is above the natural rate of y. This will be a source of problems.

The timing of actions is as follows. The economy has a natural output rate y which is known by all players. The public moves first, setting its expectations of inflation. The shock z_t is then realised. The policymaker moves next, setting π to minimise (20.5) subject to (20.4), the realisation of the shock (known to the policymaker) and the public's expectations of inflation. Notice this timing implies the policymaker has an informational advantage over the public.

By assuming that the policymaker can control π_t directly, we are finessing the issue of whether that control is exercised via a money rule (and, therefore, flexible exchange rates), an interest rate rule, or an exchange rate rule. What is key is that the authorities can set whatever policy tool is at their disposal once expectations have been set.

The policy maker, acting with discretion sets, π_t optimally, taking π_t^e (which has been already set) as given. Substituting (20.4) into (20.5) the objective function becomes

$$L = \left(\frac{1}{2}\right)\alpha\pi_t^2 + \left(\frac{1}{2}\right)\left[\theta\left(\pi_t - \pi_t^e\right) + z_t - y(\gamma - 1)\right]^2. \tag{20.6}$$

Minimising with respect to π_t yields

$$\alpha\pi_t + \theta\left[\theta\left(\pi_t - \pi_t^e\right) + z_t - y(\gamma - 1)\right] = 0. \tag{20.7}$$

Rearranging we arrive at

$$\theta\pi_t = (1-\lambda)\left[\theta\pi_t^e - z_t + y(\gamma - 1)\right], \tag{20.8}$$

where $\lambda \equiv \frac{\alpha}{\alpha+\theta^2} < 1$.

If, in addition, we impose the rational expectations condition that $\pi_t^e = E\pi_t$, we have from (20.8) that

$$\theta\pi_t^e = \left(\frac{1-\lambda}{\lambda}\right)(\gamma - 1)y. \tag{20.9}$$

Hence, under discretion, inflation expectations are positive as long as $(\gamma - 1)y$ is positive. Since $(\gamma - 1)y$ is the difference between the natural rate of output and the target rate in the policymaker's loss function, we conclude that, as long as this difference is positive, the economy exhibits an inflation bias: expected inflation is positive.

Using (20.9) in (20.8) yields

$$\theta\pi_t = \left(\frac{1-\lambda}{\lambda}\right)(\gamma - 1)y - (1-\lambda)z_t, \tag{20.10}$$

or, more simply,

$$\pi_t = \left(\frac{1-\lambda}{\theta\lambda}\right)(\gamma - 1)y - \frac{(1-\lambda)}{\theta}z_t, \tag{20.11}$$

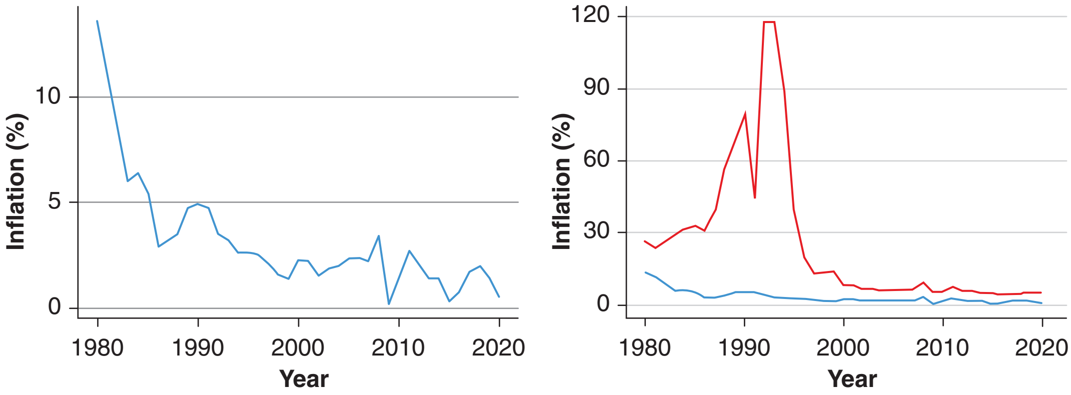
so that actual inflation depends on the shock as well as on the fixed inflation bias term. The fact that the CB wants to boost output above its natural level leads to a problem of dynamic inconsistency and inflationary bias that was originally pointed out by Kydland and Prescott (1977), and Barro and Gordon (1983). This is one of the most important and influential results for modern macroeconomic policy-making, and its intuition points squarely at the limits of systematic policy in a world where people are rational and forward-looking: they will figure out the CB's incentives, and, because of that, the tradeoff that the CB would like to exploit vanishes. Rational expectations implies that the equilibrium will occur at an inflation rate sufficiently high so that the cost of increasing inflation further would not be desirable to the CB. Once this anticipation is included in the model, discretion does not help expand output. In fact, if all could agree to a lower inflation, everybody would be better off.

The main takeaway is that *credibility* is a key element of monetary policy practice: if people believe the CB's commitment to fight inflation and not to exploit the inflation-output tradeoff systematically, the terms of the tradeoff in the short run become more favourable. This idea has been encapsulated in the mantra of *rules vs discretion*: policy-makers are better off in the long run if they are able to commit to rules, rather than trying to make policy in discretionary fashion.

20.1.2 | A brief history of monetary policy

In common policy parlance, the lesson is that being subject to time inconsistency, the CB needs to find an anchor for monetary policy. This anchor helps keep inflation expectations in check, and ameliorate the time inconsistency problem. The drawback is that the anchor may be too rigid, and make monetary policy less effective or have other side effects. Therefore the key issue is how to find an anchor that delivers credibility while not jeopardising the ability to react to shocks. One such mechanism is to appoint conservative central bankers,[1] who would have a low γ; or insuring the independence of the CB and having it focus squarely on inflation. These two policies, now widely used, have helped to reduce the inflation bias as shown in Figure 20.1. But in addition to these obvious solutions, the quest to

Figure 20.1 Inflation: advanced economies (blue line) and emerging markets (red line)



build a monetary framework that provides credibility and flexibility has gone on for decades. Mishkin (1999)) provides a nice narrative that we summarize as follows:

- The age of discretion lasted until the early 70s when there was a belief that there was a long term tradeoff between inflation and output. During this period there were no major objections to the use of monetary policy. The Keynesian/monetarist debate focused on the relative merits of fiscal vs. monetary policy.
- The rise of inflation in the 1970s led to increased skepticism on the role of monetary policy, and led to the acknowledgement that a nominal anchor was required. The discussion took place mostly in the U.S., as most other countries still had a fixed exchange rate that they carried over from the Bretton Woods system (and therefore no monetary policy of their own). But once countries started recovering their monetary policies by floating the exchange rate, monetary aggregates became the prime nominal anchor. Central banks committed to a certain growth in monetary aggregates over the medium term, while retaining flexibility in the short run.
- By the 1980s, it was clear that monetary aggregate targeting was not working very well, mostly due to instability in the demand for money. Gerald Bouey, then governor of the Bank of Canada, described the situation in his famous quote “We didn’t abandon monetary aggregates, they abandoned us.”
- Starting in the 1990s, central banks have increasingly used inflation itself as the nominal target. This is the so called inflation targeting regime. Other central banks (the Fed in the U.S.) have remained committed to low inflation, but without adopting an explicit target (though Fed governors embrace openly the idea of a 2% target for annual inflation recently updated to “an average of 2% over time”). Other countries remained using fixed exchange rates, while monetary targeting went in disuse.
- Inflation targeting, however, has a drawback: it magnifies output volatility when the economy is subject to substantial supply shocks. As a response many central bankers do not run a strict inflation targeting but a flexible inflation targeting, where the target is a long run objective retaining substantial flexibility in the short run.[2]
- More recently, some central banks have veered away from targeting inflation and started targeting inflation expectations instead (see Adrian et al. (2018)).

20.2 | The emergence of inflation targeting

Given its increasing popularity, let's spend some time analysing the monetary framework of inflation targeting. We laid the framework above which gave us a solution for the inflation rate.

Recall that using (20.9) in (20.8) yields

$$\theta\pi_t = \left(\frac{1-\lambda}{\lambda}\right)(\gamma-1)\,y-(1-\lambda)\,z_t, \tag{20.12}$$

so that actual inflation depends on the shock as well as on the fixed inflation bias term. Subtracting (20.9) from (20.12) yields

$$\theta\left(\pi_t-\pi_t^e\right)=-(1-\lambda)\,z_t, \tag{20.13}$$

or

$$y_t = y+\lambda z_t. \tag{20.14}$$

That is, deviations of output from the natural rate are random and depend on the shock and on the parameter λ.

Finally, using (20.12) and (20.14) in (20.5) yields

$$L=\left(\frac{1}{2}\right)\left(\frac{1-\lambda}{\lambda}\right)\left[(\gamma-1)\,y-\lambda z_t\right]^2+\left(\frac{1}{2}\right)\left(y\,(1-\gamma)+\lambda z_t\right)^2, \tag{20.15}$$

and taking expectations we have

$$L=\left(\frac{1}{2}\right)\left(\frac{1}{\lambda}\right)\left[(\gamma-1)^2\,y^2+\lambda^2 Ez_t^2\right], \tag{20.16}$$

or

$$EL^{disc}=\left(\frac{1}{2}\right)\left[\frac{(\gamma-1)^2\,y^2}{\lambda}+\lambda\sigma^2\right], \tag{20.17}$$

where σ^2 is the variance of z_t and the expectation is unconditional – that is, taken without knowing the realisation of z_t. Hence, expected social loss is increasing in the natural rate y, in the difference between γ and 1, and in the variance of the shock.

20.2.1 | A rigid inflation rule

Consider what happens, on the other hand, if the policymaker has precommitted not to manipulate inflation, therefore setting $\pi_t = 0$. The Phillips curve dictates that

$$y_t = y-\theta\pi_t^e+z_t. \tag{20.18}$$

If, in addition, the rule is credible, so that $\pi_t^e = 0$, we have

$$y_t = y+z_t. \tag{20.19}$$

Notice that, unlike the case of discretionary policy (see expression (20.14)), here output absorbs the full impact of the shock (the coefficient λ is missing).

The corresponding loss is

$$L^{rule}=\left(\frac{1}{2}\right)\left[-y\,(\gamma-1)+z_t\right]^2. \tag{20.20}$$

The unconditional expectation of (20.20) is

$$EL^{rule}=\left(\frac{1}{2}\right)\left[y^2\,(\gamma-1)^2+\sigma^2\right]. \tag{20.21}$$

20.2.2 | Which regime is better?

If the unconditional expectation of the loss is the welfare criterion, then deciding which regime is better depends on parameter values. Expressions (20.17) and (20.21) reveal that $EL^{rule} < EL^{disc}$ if and only if $(\gamma - 1)\, y > \sigma\sqrt{\lambda}$. The LHS is a proxy for the inflation bias under discretion; the RHS is a proxy for the variability cost under a rigid rule. The rigid rule is better when the former is larger, and vice versa. In short, you prefer a fixed rule if your inflation bias is large and the supply shocks small.

20.2.3 | The argument for inflation targeting

Suppose now that the social objective function is still given by (20.5), but that now the policymaker is given the objective function

$$L^{p} = \left(\frac{1}{2}\right)\alpha\left(\pi_t + \pi\right)^2 + \left(\frac{1}{2}\right)\left(y_t - \gamma y\right)^2, \tag{20.22}$$

where $-\pi$ is the bliss rate of inflation for the policymaker. We can interpret this as the target assigned to the policymaker by society.

Substituting (20.4) into (20.22), one gets

$$L^{p} = \left(\frac{1}{2}\right)\alpha\left(\pi_t + \pi\right)^2 + \left(\frac{1}{2}\right)\left[\theta\left(\pi_t - \pi_t^e\right) + z_t - (\gamma - 1)\, y\right]^2. \tag{20.23}$$

Minimising with respect to π_t yields

$$\alpha\left(\pi_t + \pi\right) + \theta\left[\theta\left(\pi_t - \pi_t^e\right) + z_t - (\gamma - 1)\, y\right] = 0. \tag{20.24}$$

Rearranging we arrive at

$$\theta\pi_t = (1 - \lambda)\left[\theta\pi_t^e - z_t + y(\gamma - 1)\right] - \lambda\theta\pi. \tag{20.25}$$

Taking expectations we have

$$\theta\pi_t^e = \left(\frac{1 - \lambda}{\lambda}\right)(\gamma - 1)\, y - \theta\pi, \tag{20.26}$$

so the inflation bias is positive or negative depending on the setting of π. Suppose the target is set so that the inflation bias is zero. Having $\theta\pi_t^e = 0$ implies

$$\lambda\theta\pi = (1 - \lambda)(\gamma - 1)\, y. \tag{20.27}$$

Using this in (20.25) yields

$$\theta\pi_t = -(1 - \lambda)\, z_t. \tag{20.28}$$

Using this and $\pi_t^e = 0$ in (20.4) yields

$$y_t - y = \lambda z_t, \tag{20.29}$$

so that deviations of output from its long run level are the same as under discretion.

Finally, using (20.28) and (20.29) into the public's loss function (20.5) yields

$$L = \left(\frac{1}{2}\right)\lambda(1-\lambda)z_t^2 + \left(\frac{1}{2}\right)\left[\lambda z_t - (\gamma - 1)y\right]^2. \tag{20.30}$$

Taking expectations and rearranging

$$EL^{target} = \left(\frac{1}{2}\right)\left[\lambda\sigma^2 + (\gamma-1)^2 y^2\right]. \tag{20.31}$$

It is easy to check that EL^{target} is smaller than either EL^{disc} or EL^{rule}. That is, inflation targeting is better for welfare than fully discretionary policy and a rigid rule. The intuition should be simple: targeting enjoys the flexibility benefits of discretion and the credibility benefits of a rule (the inflation bias is zero).

20.2.4 | In sum

As inflation in the world decreased, monetary policy entered into a happy consensus by the 2000s. Credibility had been restored, and even those central banks that did not explicitly target inflation were widely understood to be essentially doing the same. The short-term interest rate was the policy tool of choice. Enhanced credibility, or the so called "flattening of the Phillips curve" made monetary policy more powerful as a stabilisation mechanism, and as a result became the tool of choice for steering the business cycle. Some central bankers even acquired heroic, pop-culture status.

But then, the 2008/2009 crisis hit. The consensus was revealed inadequate to deal with the crisis at its worst, and questions were raised as to the extent to which monetary policy had failed to prevent (and perhaps contributed to) the Great Recession and, later on, the European Crisis. Perhaps with the benefit of hindsight, the challenge proved to be central bank's finest hour: the recoveries were relatively swift and inflation remained low. The hard-gained credibility provided the room for massive increases in liquidity, that shattered not a bit the credibility of the central banks and allowed to counteract the drainage of liquidity during the crises. This was perhaps best epitomised in a celebrated quote by Mario Draghi, then governor of the European Central Bank who, in July 2012, announced that the Central Bank would do "whatever it takes". This phrase, as of today, is the symbol of the coming of age of modern CB when full discretion can be pursued without rising an eyebrow or affecting expectations!

Notes

[1] One way of illustrating this debate is to remember the discussion surrounding the creation of the European Central Bank. As a novel institution whose governance was in the hands of a number of countries, it was not clear how it would build its credibility. Someone suggested to locate it in Frankfurt, so it could absorb (by proximity?) Germany's conservative approach to monetary policy. The french wanted to control the presidency, but this was considered not sufficiently strong at least at the beginning, so they compromised on a two year presidency with a Dutch. However, after two years, when French Jean Marie Trichet took over, he still had to be overly conservative to build his, and the institution's, credibility.

[2] We should also keep in mind that inflation targeting does not mean that the central bank or policy maker does not care about anything other than inflation. As we show in the model in the next section, the central bank's objective function may take deviations of output into account – the relative weight of output will affect the tolerance of the central bank to deviations of inflation from the target as a result of shocks.

References

Adrian, T., Laxton, M. D., & Obstfeld, M. M. (2018). *Advancing the frontiers of monetary policy*. International Monetary Fund.

Barro, R. J. & Gordon, D. B. (1983). Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics*, *12*(1), 101–121.

Kydland, F. E. & Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, *85*(3), 473–491.

Mishkin, F. S. (1999). International experiences with different monetary policy regimes. *Journal of Monetary Economics*, *43*(3), 579–605.

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

Velasco, A. (1996). Fixed exchange rates: Credibility, flexibility and multiplicity. *European Economic Review, 40*, 1023–1035.

CHAPTER 21

Recent debates in monetary policy

In the last two chapters we presented the basic analytics of monetary policy in the long and in the short run. For the short run, we developed a simple New Keynesian model that can parsimoniously make sense of policy as it has been understood and practised over the last few decades.

Before the 2008 financial crisis, most advanced-country central banks, and quite a few emerging-market central banks as well, carried out monetary policy by targeting a short-term interest rate. In turn, movements in this interest rate were typically guided by the desire to keep inflation close to a pre-defined target — this was the popular policy of inflation targeting. This consensus led to a dramatic decrease in inflation, to the point of near extinction in most economies, over the last two or three decades.

But this benign consensus was shaken by the Great Financial Crisis of 2008-2009. First, there was criticism that policy had failed to prevent (and perhaps contributed to unleashing) the crisis. Soon, all of the world's major central banks were moving fast and courageously into uncharted terrain, cutting interest rates sharply and all the way to zero. A first and key issue, therefore, was whether the conventional tools of policy had been rendered ineffective by the zero lower bound.

In response to the crisis, and in a change that persists until today, central banks adopted all kinds of unconventional or unorthodox monetary policies. They have used central bank reserves to buy Treasury bonds and flood markets with liquidity, in a policy typically called quantitative easing. And they have also used their own reserves to buy private sector credit instruments (in effect lending directly to the private sector) in a policy often referred to as credit easing.

Interest rate policy has also become more complex. Central banks have gone beyond controlling the contemporary short rate, and to announcing the future path of short rates (for a period of time that could last months or years), in an attempt at influencing expectations a policy known as forward guidance. Last but not least, monetary authorities have also begun paying interest on their own reserves — which, to the extent that there is a gap between this rate and the short-term market rate of interest (say, on bonds), gives central bankers an additional policy tool.

These policies can be justified on several grounds. One is the traditional control of inflation — updated in recent years to include avoidance of deflation as well. Another is control of aggregate demand and output, especially when the zero lower bound on the nominal interest limits the effectiveness of traditional monetary policy. A third reason for unconventional policies is financial stability: if

---

How to cite this book chapter:
Campante, F., Sturzenegger, F. and Velasco, A. 2021. *Advanced Macroeconomics: An Easy Guide.* Ch. 21. 'Recent debates in monetary policy', pp. 323–344. London: LSE Press.
DOI: https://doi.org/10.31389/lsepress.ame.u License: CC-BY-NC 4.0.

spikes in spreads, for instance, threaten the health of banks and other financial intermediaries (this is exactly what happened in 2007-09), then monetary policy may need to act directly on those spreads to guarantee stability and avoid runs and the risk of bankruptcy.

Do these policies work, in the sense of attaining some or all of these objectives? How do they work? Why do they work? What does their effectiveness (or lack of effectiveness) hinge on?

A massive academic literature on these questions has emerged during the last decade. Approaches vary, but the most common line of attack has been to append a financial sector to the standard New Keynesian model (yes, hard to believe, but, until the crisis, finance was largely absent from most widely-used macro models), and then explore the implications.

This change brings at least two benefits. First, finance can itself be a source of disturbances, as it occurred in 2007-09 and had also occurred in many earlier financial crises in emerging markets. Second, the enlarged model can be used to study how monetary policy can respond to both financial and conventional disturbances, with the financial sector also playing the role of potential amplifier of those shocks.

Here we cannot summarise that literature in any detail (but do look at Eggertsson and Woodford (2003), Gertler and Karadi (2011), and the survey by Brunnermeier et al. (2013) for a taste). What we do is extend our standard NK model of earlier sections and chapters to include a role for liquidity and finance, and we use the resulting model to study a few (not all) varieties of unconventional monetary policy.

The issues surrounding conventional and unconventional monetary policies have taken on new urgency because of the Covid-19 crisis. In the course of 2020, central banks again resorted to interest, cutting it all the way to the zero lower bound, coupled with quantitative easing and credit easing policies that are even more massive than those used over a decade ago. And in contrast to the Great Financial Crisis, when only advanced-country central banks experimented with unconventional policies, this time around many emerging-economy central banks have dabbled as well. So understanding how those policies work has key and urgent policy relevance — and that is the purpose of this chapter.

21.1 | The liquidity trap and the zero lower bound

John Hicks, in the famous paper where he introduced the IS-LM model (1937), showed how monetary policy on occasion might become ineffective. These "liquidity traps" as he called them, occurred when the interest rate fell to zero and could not be pushed further down. In this section we model this liquidity trap in our New Keynesian framework.

Until not too long ago, economists viewed the liquidity trap as the stuff of textbooks, not reality. But then in the 1990s Japan got stuck in a situation of very low or negative inflation and no growth. No matter what the Japanese authorities tried, nothing seemed to work. In 1998, Paul Krugman pointed out that "here we are with what surely looks a lot like a liquidity trap in the world's second-largest economy". And then he proceeded to show that such a trap could happen not just in the static IS-LM model, but in a more sophisticated, dynamic New Keynesian model.

Of course, the experience of Japan was not the only one in which a liquidity trap took center stage. During the world financial crisis of 2008-09, the world's major central banks cut their interests to zero or thereabouts, and found that policy alone was not sufficient to contain the collapse of output. The same, perhaps with greater intensity and speed, has occurred during the Covid-19 crisis of 2020-21, with monetary authorities cutting rates to zero and searching for other policy tools to contain the destruction of jobs and the drop in activity. So, the issues surrounding the zero lower bound and liquidity traps are a central concern of macroeconomists today[1].

To study such traps formally, let us return to the two-equation canonical New Keynesian model of Chapter 15

$$\dot{\pi}_t = \rho\pi_{\mathrm{t}} - \kappa x_t, \tag{21.1}$$

$$\dot{x}_{\mathrm{t}} = \sigma\left(i_t - \pi_{\mathrm{f}} - r^n\right), \tag{21.2}$$

where, recall, π_t is inflation, x_t is the output gap, i_t is the policy-determined nominal interest rate, $r^n \equiv \rho + \sigma^{-1}g$ is the natural or Wicksellian interest rate, which depends on both preferences (the discount rate ρ and the elasticity σ) and trend productivity growth (g).

To close the model, instead of simply assuming a mechanic policy rule (of the Taylor type or some other type, as we did in Chapter 15), we consider alternative paths for the interest rate in response to an exogenous shock. Werning (2011), in an influential and elegant analysis of the liquidity trap, studies formal optimisation by the policymaker, both under rules and under discretion. Here we take a somewhat more informal approach, which draws from his analysis and delivers some of the same policy insights.[2]

Define a liquidity trap as a situation in which the zero lower bound is binding and monetary policy is powerless to stabilise inflation and output. To fix ideas, consider the following shock:

$$r^n_{\mathrm{t}} = \begin{cases} \underline{r}^n < 0 & \text{for} \quad 0 \le t < T \\ r^n > 0 & \text{for} \quad t \ge T. \end{cases} \tag{21.3}$$

Starting from r^n, at time 0 the natural rate of interest unexpectedly goes down to $\underline{r}^n$, and it remains there until time T, when it returns to r^n and stays there forever. The key difference between this shock and that studied in Chapter 15 in the context of the same model, is that now the natural rate of interest is *negative* for an interval of time. Recall that this rate depends on preferences and on trend growth in the natural rate of output. So if this productivity growth becomes sufficiently negative, r^n_t could be negative as well.

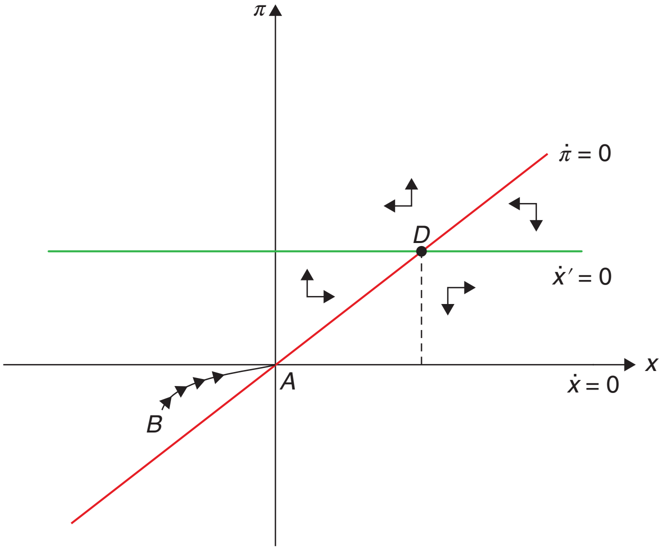
Notice that the combination of flagging productivity and a negative natural rate of interest corresponds to what Summers (2018) has labelled secular stagnation. The point is important, because, if secular stagnation, defined by Summers precisely as a situation in which the natural rate of interest falls below zero for a very long time (secular comes from the Latin *soeculum*, meaning *century*), then economies will often find themselves in a liquidity trap.

The other novel component of the analysis here, compared to Chapter 15, is that now we explicitly impose the zero lower bound on the nominal interest rate, and require that $i_t \ge 0 \forall t$.

If the central bank acts with discretion, choosing its preferred action at each instant, the zero lower bound will become binding as it responds to the shock. To see this, let us first ask what the central bank will optimally do once the shock is over at time T. Recall the canonical New Keynesian model displays, what Blanchard and Galí (2007) called the divine coincidence: there is no conflict between keeping inflation low and stabilising output. If $i = r^n$, then $\pi_t = x_t = 0$ is an equilibrium. So starting at time T, any central bank that is happiest when both inflation and the output gap are at zero will engineer exactly that outcome, ensuring $\pi_t = x_t = 0 \quad \forall t \ge T$.

In terms of the phase diagram in Figure 21.1, we assume that initially (before the shock) $i = r^n$, so that $\pi_t = 0 \forall t < 0$. Therefore, the initial steady state was at point A, and to that point exactly the system must return at time T. What happens between dates 0 and T ? Trying to prevent a recession and the

Figure 21.1 Monetary policy in the ZLB



corresponding deflation, the central bank will cut the nominal interest all the way to zero. That will mean that between dates 0 and *T*, dynamics correspond to the system with steady state at point D, but because of the zero lower bound, policy cannot take the economy all the way back to the pre-shock situation and keep $\pi_t = x_t = 0$ always. So, on impact the system jumps to point B, and both inflation and the output gap remain negative (deflation and depression or at least recession take hold) in the aftermath of the shock and until date *T*[3].

Both Krugman (1998) and Werning (2011) emphasise that the problem is the central bank's lack of credibility: keeping the economy at $\pi_t = x_t = 0$ is optimal starting at time *T*, and so people in this economy will pay no attention to announcements by the central bank that claim something else. In technical language, the monetary authority suffers from a time inconsistency problem of the kind identified by Kydland and Prescott (1977) and Calvo (1978) (see Chapter 20): from the point of view of any time before time *T*, engineering some inflation after *T* looks optimal. But when time *T* arrives, zero inflation and a zero output gap become optimal.

What is to be done? This is Krugman's (1998) answer: The way to make monetary policy effective, then, is for the Central Bank to credibly promise to be irresponsible – to make a persuasive case that it will permit inflation to occur, thereby producing the negative real interest rates the economy needs. In fact, there are simple paths for the nominal interest rate that, if the central bank could commit to them, would deliver a better result. Consider a plan, for instance, that keeps inflation and the output gap constant at

$$\pi_t = -\underline{r}^n > 0 \ \ and \ \ x_t = -\frac{\underline{r}^n}{\kappa} > 0 \quad \forall t \geq 0. \tag{21.4}$$

Since $i_t = r_t^n + \pi_t$, it follows that $i_t = 0 \quad \forall t < T$, and $i_t = r^n - \underline{r}^n > 0 \quad \forall t \geq T$. Although this policy is not fully optimal, it may well (depending on the social welfare function and on parameter values) deliver higher welfare than the policy of $i_t = 0$ forever, which causes recession and deflation between

0 and T. And note that as prices become less sticky (in the limit, as κ goes to infinity), the output gap goes to zero, so this policy ensures no recession (and no boom either)[4].

Notice, strikingly, that this policy - just like the one described in the phase diagram above - also involves keeping the nominal interest stuck against the zero lower bound during the whole duration of the adverse shock, between times 0 and T. So if the policy is the same over that time interval, why are results different? Why is there no recession as a result of the shock? Crucially, the difference arises because now people expect there will be inflation and a positive output gap after time T, and this pushes up inflation before T (recall from Chapter 15 that inflation today increases with the present discounted value of the output gaps into the infinite future), reducing the real interest rate and pushing up consumption demand and economic activity.

Of course, the alternative policy path just considered is just one such path that avoids recession, but not necessarily the optimal path. Werning (2011) and, before that, Eggertsson and Woodford (2003) characterised the fully optimal policies needed to get out of a liquidity trap. Details vary, but the main message is clear: during the shock, the central bank needs to be able to persuade people (to pre-commit, in the language of theory) it will create inflation *after* the shock is over.

What can central banks do to acquire the much-needed credibility to become "irresponsible"? One possibility is that they try to influence expectations through what has become known as "forward guidance". One example, is the Fed's repeated assertion that it anticipates that "weak economic conditions are likely to warrant exceptionally low levels of the federal funds rate for some time". Alternatively, central bankers can stress that they will remain vigilant and do whatever it takes to avoid a deep recession. For instance, on 28 February 2020, when the Covid 19 pandemic was breaking out, Fed Chairman Jerome Powell issued this brief statement:

> The fundamentals of the U.S. economy remain strong. However, the coronavirus poses evolving risks to economic activity. The Federal Reserve is closely monitoring developments and their implications for the economic outlook. We will use our tools and act as appropriate to support the economy.

When put this way, the problem seems relatively simple to solve: the CB needs only to use these additional tools to obtain a similar result to what it would obtain by simply playing around with the short-term nominal interest rate, as in normal times. Unfortunately, this is not that easy precisely because of the crucial role played by expectations and credibility. The crucial point is that the central bankers need to convince the public that it will pursue expansionary policies in the future, even if inflation runs above target, and this runs counter to their accumulated credibility as hawkish inflation-fighters and committed inflation-targeters.

Recent thinking on these issues - and on other policy alternatives available to policymakers when against the zero lower bound - is summarised in Woodford (2016). He argues that, when it comes to forward guidance, what is needed are explicit criteria or rules about what would lead the central bank to change policy in the future - criteria that would facilitate commitment to being irresponsible.

One way to do that is to make policy history-dependent: the central bank commits to keep a certain path for interest rates unless certain criteria, in terms of a certain target for the output gap or unemployment or nominal GDP, for instance, are met. The Fed has actually moved recently towards that approach, stating that current low rates will be maintained unless unemployment falls below a certain level, or inflation rises above a certain level. The recent inflation targeting shift by the Bank of Japan can also be interpreted in line with this approach.

Another way forward is to move from an inflation target to a price level target (see Eggertsson and Woodford (2003) and Eggertsson and Woodford (2004)). The benefit of a price-level target over an inflation target to fight deflation is that it meets enhanced deflationary pressure with an intensified commitment to pursue expansionary policy in the future (even if the target price level is unchanged). An inflation target, on the other hand, lets bygones be bygones: a drop in prices today does not affect the course of policy in the future, since, under inflation targeting, the central bank is focused only on the current rate of change in prices. Thus, inflation targeting does not induce the same kind of stabilising adjustment of expectations about the future course of policy as does price-level targeting[5].

And if a rethinking of the traditional inflation targeting framework is called for, another rule that has gained adherents recently is the so-called NGDP or nominal GDP level targeting (see Sumner (2014) and Beckworth (2019)). In targeting nominal GDP the central bank could commit to compensate for falls in output by allowing for higher inflation. The underlying point is that NGDP would provide a better indicator, compared to inflation alone, of the kind of policy intervention that is needed.

21.2 | Reserves and the central bank balance sheet

As we mentioned, the Great Financial Crisis introduced a wealth of new considerations for monetary policy. In this section we develop a model of quantitative easing where the Central Bank pays money on its reserves, adding a new variable to the policy tool which was not present in our traditional monetary models where the rate of return on all Central Bank liabilities was fixed at zero. We will see this introduces a number of new issues. While the modelling does not make this necessarily explicit, underlying the new paradigm is the understanding that there is a financial sector that intermediates liquidity. Thus, before going into the full fledged optimisation problem, we lay out a more pedestrian approach to illustrate some of the issues.

21.2.1 | Introducing the financial sector

To introduce these new issues we can start from a simple IS-LM type of model, as in the lower panel of Figure 21.2.[6]

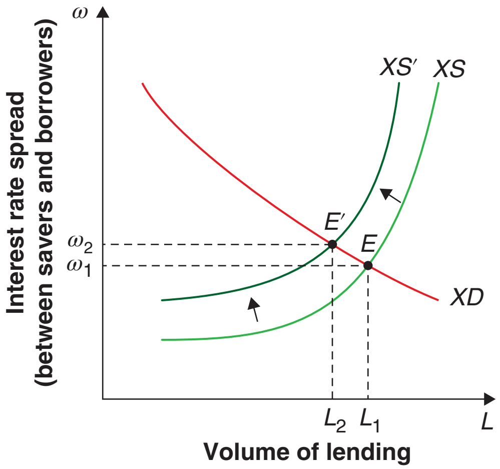
If there are financial intermediaries, there must be multiple interest rates – one that is paid to savers (i^s), and another that is charged from borrowers (i^b). Otherwise, of course, how would those intermediaries make any money? This market, depicting the supply of loans and the demand for loans, is shown in the upper panel of Figure 21.2. The IS curve below is drawn for a given level of spread.

As a result, the role of intermediation introduces a new channel for the amplification and propagation of economic shocks. For instance, suppose a high level of economic activity affects asset prices, and hence the net worth of financial intermediaries and borrowers. This will allow for additional borrowing at any level of spread (a shift of the XS curve to the right). This makes the IS curve flatter than what it would otherwise be: the same change in income would be associated with a smaller change in the interest rate paid to savers. This amplifies the effects on output of any shift in the LM/MP curves.

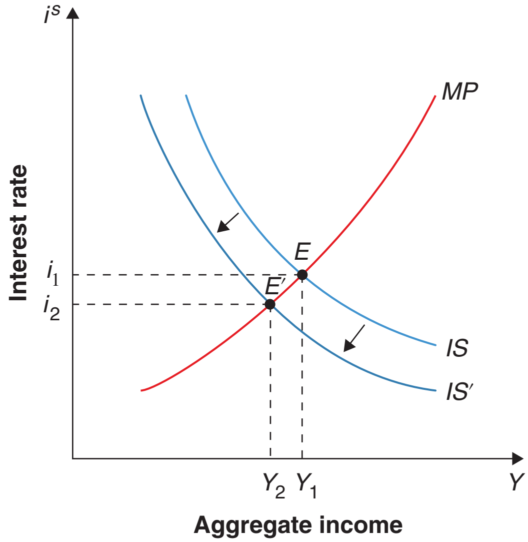
Even more interestingly, this lets us consider the effects of direct shocks to intermediation – beyond the amplification of other shocks. An upward shift of the XS curve (less credit available for any level of spread) means a downward shift to the IS curve – a larger equilibrium spread translated into less interest being paid to savers. This shock, illustrated in Figure 21.2, leads (in the absence of monetary policy compensating for the negative shock) to an output contraction with falling interest rates. Anything that impairs the capital of financial intermediaries (say, a collapse in the prices of mortgage-backed

Figure 21.2 Effects of a disruption of credit supply

A: A tightening of credit supply



B: Impact on the IS curve



securities they hold) or that tighten leverage constraints (say, they are required to post more collateral when raising funds because the market is suspicious of their solvency) will correspond to such an upward shift of the XS curve. If the IS curve is shifted far enough to the left, monetary policy may be constrained by the zero lower bound on interest rates. Does all of that sound familiar?

Needless to say, a simple IS-LM type of framework leaves all sorts of questions open in terms of the microfoundations behind the curves we've been fiddling around with. To that point we now turn.

21.2.2 | A model of quantitative easing

Now we focus on the role of the central bank balance and, more specifically, on the role of central bank reserves in the conduct of unconventional monetary policy. This emphasis has a practical motivation. As Figure 21.3 makes clear, the Federal Reserve (and other central banks) have issued reserves to purchase government bonds, private-sector bonds and other kinds of papers, dramatically enlarging the size of central bank balance sheets.

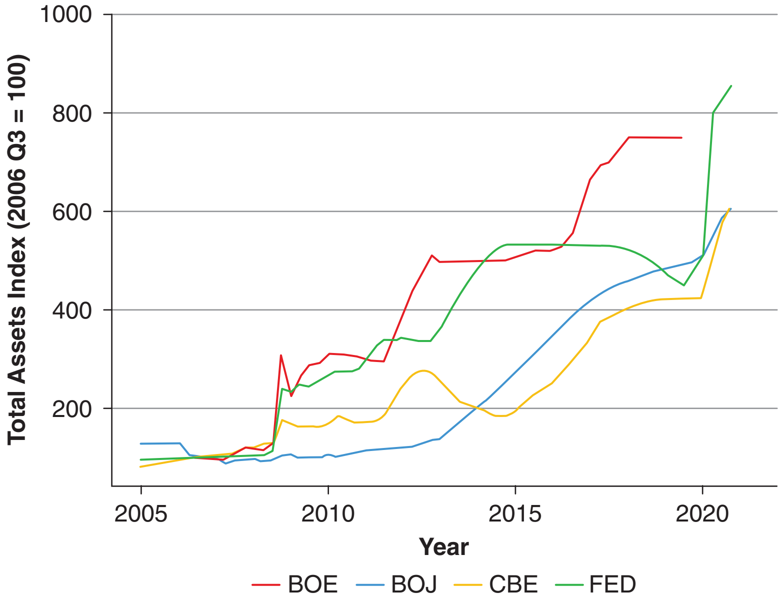
The assets in Figure 21.3 have been financed mostly with overnight interest paying voluntarily held deposits by financial institutions at the central bank. We call these deposits reserves for short.

As Reis (2016) emphasises, reserves have two unique features that justify this focus. First, the central bank is the monopoly issuer of reserves. As a monopoly issuer, it can choose the interest to pay on these reserves. Second, only banks can hold reserves. This implies that the aggregate amount of reserves in the overall banking system is determined by the central bank.

The liability side of a central bank balance sheet has two main components: currency (think of it as bank notes) and reserves. Together, currency and reserves add up to the monetary base. The central bank perfectly controls their sum, even if it does not control the breakdown between the two components of the monetary base.

These two properties of the central bank imply that the central bank, can in principle, choose both the quantity of the monetary base and the nominal interest rate paid on reserves. Whether it can also control the quantity of reserves, and do so independently of the interest rate that it pays, depends on the demand for reserves by banks[7].

Figure 21.3 Assets held by FED, ECB, BOE and BOJ



Before the 2008 financial crisis, central banks typically adjusted the volume of reserves to influence nominal interest rates in interbank markets. The zero lower bound made this policy infeasible during the crisis. Post-crisis, many central banks adopted a new process for monetary policy: they set the interest rate on reserves, and maintained a high level of reserves by paying an interest rate that is close to market rates (on bonds, say). In turn, changes in the reserve rate quickly feed into changes in interbank and other short rates.

Let D_t be the real value of a central bank-issued means of payment. You can think of it as central bank reserves. But following Diba and Loisel (2020) and Piazzesi et al. (2019), you can also think of it as a digital currency issued by the monetary authority and held directly by households[8]. In either case, the key feature of D_t is that it provides liquidity services: it enables parties to engage in buying, selling, and settling of balances. In what follows, we will refer to D_t using the acronym MP (means of payment, not be confused with our earlier use of MP for monetary policy), but do keep in mind both feasible interpretations. Later in this chapter we will show that the model developed here can also be extended (or reinterpreted, really) to study a more conventional situation in which only commercial banks have access to accounts at the central bank and households only hold deposits at commercial banks.

The simplest way to model demand for MP is to include it in the utility function of the representative household:

$$u_t = \left(\frac{\sigma}{\sigma - 1}\right) Z_t^{\left(\frac{\sigma-1}{\sigma}\right)}, \quad Z_t = C_t^{\alpha} D_t^{1-\alpha}, \tag{21.5}$$

where $\sigma > 0$ is the interemporal elasticity of substitution in consumption, and is a Cobb-Douglas weight with α that lies between 0 and 1. The representative household maximises the present discounted value of this utility flow subject to the following budget constraint:

$$\dot{D}_t + \dot{B}_t = Y_t + \left(i_t^b - \pi_t\right) B_t + \left(i_t^d - \pi_t\right) D_t - C_t, \tag{21.6}$$

where B_t is the real value of a nominal (currency-denominated) bond, issued either by the government or by the private sector, i_t^b is the nominal interest rate paid by the bond, and i_t^d is the nominal interest rate paid by the central bank to holders of D_t. (Income Y_t comprises household income and government transfers.) In accordance with our discussion above, the monetary authority controls this interest rate and the supply of MP [9].

Since we do not want to go into the supply side of the model in any detail here, we simply include a generic formulation of household income, which should include wage income but could have other components as well. Government transfers must be included because governments may wish to rebate to agents any seigniorage collected from currency holders.

Let total assets be $A_t = B_t + D_t$. Then we can write the budget constraint as

$$\dot{A}_t = Y_t + \left(i_t^b - \pi_t\right) A_t - \left(i_t^b - i_t^d\right) D_t - C_t. \tag{21.7}$$

In the household's optimisation problem, A_t is a state variable and D_t and C_t are the control variables. First order conditions are

$$\alpha Z_t^{\left(\frac{\sigma-1}{\sigma}\right)} = C_t \lambda_t \tag{21.8}$$

$$(1-\alpha) Z_t^{\left(\frac{\sigma-1}{\sigma}\right)} = \lambda_t D_t \left(i_t^b - i_t^d\right) \tag{21.9}$$

$$\dot{\lambda}_t = -\lambda_t \left(i_t^b - \pi_t - \rho\right), \tag{21.10}$$

where λ_t is the shadow value of household assets (the co-state variable in the optimisation problem). These conditions are standard for the Ramsey problem, augmented here by the presence of the MP. It follows from (21.8) and (21.9) in logs, denoted by small case letters, the demand function for MP is

$$d_t = c_t - \Delta_t, \tag{21.11}$$

where

$$\Delta_t = \log\left[\left(\frac{\alpha}{1-\alpha}\right)\left(i^b_t - i^d_t\right)\right]. \tag{21.12}$$

So, intuitively, demand for MP is proportional to consumption and decreasing in the opportunity cost $\left(i^b_t - i^d_t\right)$ of holding MP. Notice that this demand function does not involve satiation: as $i^b_t - i^d_t$ goes to zero, d_t does not remain bounded. From a technical point of view, it means that we cannot consider here a policy of $i^d_t = i^b_t$[10].

The appendix shows that in logs, the Euler equation is

$$\dot{c}_t = \sigma\left(i^b_t - \pi_t - \rho\right) + (1-\sigma)(1-\alpha)\dot{\Delta}_t. \tag{21.13}$$

Differentiating (21.11) with respect to time yields

$$\dot{c}_t - \dot{d}_t = \dot{\Delta}_t. \tag{21.14}$$

To close the model we need two more equations. One is the law of motion for real MP holdings, also in logs:

$$\dot{d}_t = \mu - \pi_t, \tag{21.15}$$

where μ is the rate of growth of the nominal stock of MP. Intuitively, the real stock rises with μ and falls with π. So μ and i^d_t are the two policy levers, with i^b_t endogenous (market-determined).

From (21.14) and (21.15) it follows that

$$\dot{c}_t = \mu - \pi_t + \dot{\Delta}_t. \tag{21.16}$$

This equation and the Euler equation (21.13) can be combined to yield

$$\dot{\Delta}_t = \frac{\sigma\left(i^b_t - \rho\right) - \mu + (1-\sigma)\pi_t}{\alpha + \sigma(1-\alpha)}. \tag{21.17}$$

Now, given the definition of Δ_t in (21.12),

$$i^b_t = \left(\alpha^{-1} - 1\right)e^{\Delta_t} + i^d_t, \tag{21.18}$$

which can trivially be included in (21.17)

$$\dot{\Delta}_t = \frac{\sigma\left[\left(\alpha^{-1} - 1\right)e^{\Delta_t} + i^d_t - \rho\right] - \mu + (1-\sigma)\pi_t}{\alpha + \sigma(1-\alpha)}. \tag{21.19}$$

Recall next that because the economy is closed all output is consumed, so $c_t = y_t$. If we again define $x_t \equiv y_t - \bar{y}$ as the output gap, the Euler equation becomes

$$\dot{x}_t = \sigma\left(i^b_t - \pi_t - r^n\right) + (1-\sigma)(1-\alpha)\dot{\Delta}_t, \tag{21.20}$$

where, as in previous sections, the natural rate of interest is $r^n \equiv \rho + \sigma^1 g$, and g is the exogenous rate of growth of the natural rate of output $\bar{y}$.

Next, with $c_t = y_t$ the MP demand function (21.11) becomes

$$d_t = y_t - \Delta_t, \tag{21.21}$$

which, in deviations from steady state, is

$$x_t = \left(d_t - \bar{d}\right) + \left(\Delta_t - \bar{\Delta}\right). \tag{21.22}$$

We close the model with the Phillips curve, using the same formulation as in this chapter and earlier:

$$\dot{\pi}_t = \rho\pi_t - \kappa x_t. \tag{21.23}$$

Replacing (21.22) in (21.23) we get

$$\dot{\pi}_t = \rho\pi_t - \kappa\left(d_t - \bar{d}\right) - \kappa\left(\Delta_t - \bar{\Delta}\right). \tag{21.24}$$

That completes the model, which can be reduced to a system of three differential equations in 3 unknowns, π_t, d_t and Δ_t, whose general solution is quite complex. But there is one case, that of log utility, which lends itself to a simple and purely graphical solution. On that case we focus next.

If $\sigma = 1$, then (21.19) simplifies to:

$$\dot{\Delta}_t = \left(\alpha^{-1} - 1\right)e^{\Delta_t} + i_t^d - \rho - \mu. \tag{21.25}$$

This is an unstable differential equation in Δ_t and exogenous parameters or policy variables. Thus, when there is a permanent shock, Δ_t jumps to the steady state. This equation does not depend on other endogenous variables $\left(x_t, d_t, \pi_t \text{ or } i_t^b\right)$, so it can be solved separately from the rest of the model. The evolution over time of Δ_t depends on itself and the policy parameters i_t^d and μ[11].

Now the Phillips curve and the law of motion for MP are a system of two differential equations in two unknowns, π_t and d_t, with $\left(\Delta_t - \bar{\Delta}\right)$ exogenously given. In matrix form the system is

$$\begin{bmatrix} \dot{\pi}_t \\ \dot{d}_t \end{bmatrix} = \Omega \begin{bmatrix} \pi_t \\ d_t \end{bmatrix} + \begin{bmatrix} \kappa\bar{d} - \kappa\left(\Delta_t - \bar{\Delta}\right) \\ \mu \end{bmatrix}, \tag{21.26}$$

where

$$\Omega = \begin{bmatrix} \rho & -\kappa \\ -1 & 0 \end{bmatrix}. \tag{21.27}$$

It is straightforward to see that $\text{Det}(\Omega) = -\kappa < 0$, and $\text{Tr}(\Omega) = \rho > 0$. It follows that one of the eigenvalues of Ω is positive and the other is negative. Since π_t is a jumpy variable and d_t is a sticky or state variable, we conclude that the 2×2 system is saddle-path stable, as seen in Figure 21.4.

Before considering the effects of shocks on the dynamics of this system, let us ask: why this model? What does it add to the standard NK formulation?

The first is realism. Since the Great Financial Crisis, many central banks have begun using the interest paid on reserves as an instrument of monetary policy. This policy alternative is not something one can study in conventional NK models.

Second, and more important, not only different interest rates, but the size and composition of the central bank's balance sheet now matter. Changes in the speed of MP creation and open market operations involving MP can affect both inflation and output. For a more general discussion of the role of the central bank's balance sheet, see Curdia and Woodford (2011).

Third, a technical but policy-relevant point: this model does not suffer from the problem of nonuniqueness of equilibrium that plagues NK models with an exogenous nominal interest rate, as we saw in Chapter 15. For further discussion, see Hall and Reis (2016) and Diba and Loisel (2020).

Figure 21.4 A model of central bank reserves

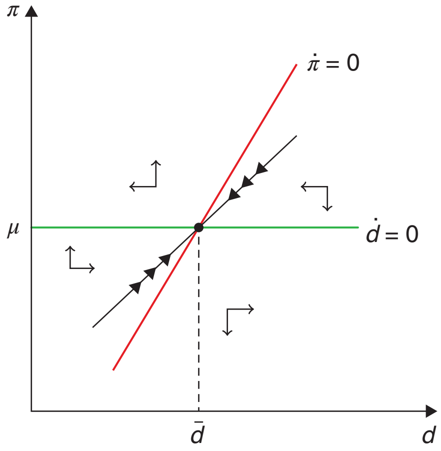
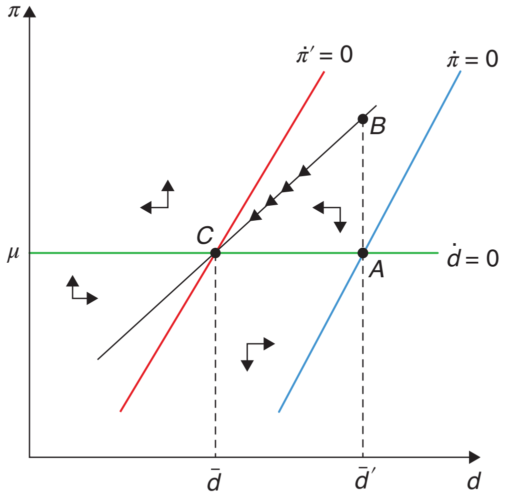


Figure 21.5 Reducing the rate on reserves



21.2.3 | Effects of monetary policy shocks

Consider first the effects of an unexpected and permanent reduction in i_t^d, one of the two policy tools the central bank has. Suppose that at time 0, i_t^d moves from i^d to $\underline{i}^d$, where $\underline{i}^d < i^d$. We show this in Figure 21.5.

Recall that in steady state the market rate of interest on bonds is pinned down by $i^b = \rho + \mu$. So, as i_t^d falls, $\bar{\Delta}$, the steady state gap between the two interest rates rises. We saw that in response to a permanent policy shock, Δ_t will immediately jump to its new (higher, in this case) steady state level. This means that we can look at the dynamics of π_t and d_t independently of Δ_t.

The other thing to notice is that as the steady state gap $\left(i^b - i^d\right)$ goes up, steady state demand for MP falls. In the phase diagram in Figure 21.5, this is reflected in the fact that the $\dot{\pi} = 0$ schedule moves to the left, and the new steady state is at point *C*. On impact, the system jumps up to point *B*, with inflation temporarily high. Thereafter, both inflation and real stocks of MP fall toward their new steady state levels.

What happens to consumption and output? The cut in i_t^d makes people want to hold less MP, but the stock of MP cannot fall immediately. What equilibrates the market for MP is a an upward jump in consumption (and output, given that prices are sticky). The temporary boom causes an increase in inflation above the rate μ of nominal MP growth, which over time erodes the real value of the stock of MP outstanding, until the system settles onto its new steady state.

In summary: the permanent cut in the interest rate paid on MP causes a temporary boom. Inflation rises and then gradually falls and so does output. All of this happens without modifying the pace of nominal MP growth. So, changes in the interest rate paid on central bank reserves (or on a digital means of payment) do serve as tool of monetary policy, with real effects.

Consider next the effects of an unexpected and permanent increase in μ, the other tool the central bank has at its disposal. Suppose that at time 0, policy moves from μ to $\bar{\mu}$, where $\bar{\mu} > \mu$. Recall again that in steady state the market rate of interest on bonds is pinned down by $i^b = \rho + \mu$. So, as μ rises and i^d remains constant, $\bar{\Delta}$, the steady state gap between the two interest rates will go up. But Δ_t will jump right away to $\bar{\Delta}$, so again we can look at the dynamics of the 2×2 system independently of Δ_t.

As the steady state gap $\left(i^b - i^d\right)$ rises, steady state demand for MP goes down. In the phase diagram in Figure 21.6, this is reflected in the fact that the $\dot{\pi} = 0$ schedule moves to the left. But now the $\dot{d} = 0$ schedule also shifts (upward), so that the new steady state is at point *F*. On impact, the system jumps up to point *E*, with inflation overshooting its new, higher, steady state level. Thereafter, both inflation and the real stock of MP fall toward their new steady state levels.

Figure 21.6 Increasing money growth

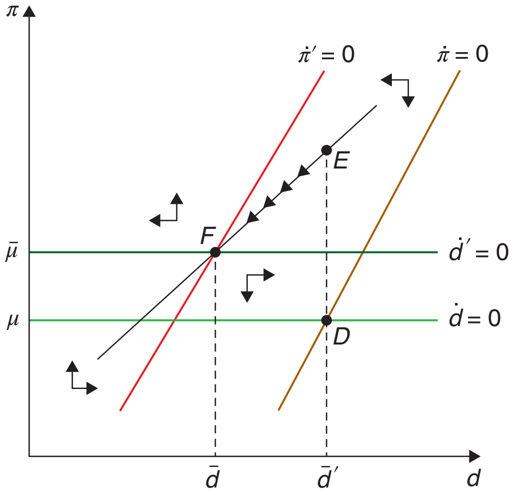
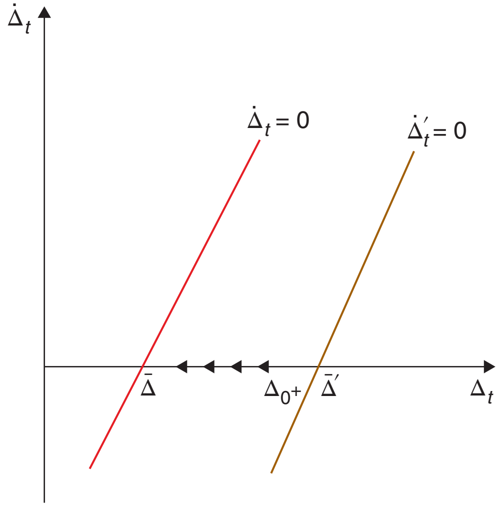


Figure 21.7 The dynamics of the interest rate spread



Note that the overshoot is necessary to erode the real value of MP, since in the new steady state agents will demand less of it. As in the previous case, inflation rises since consumption and output are temporarily above their steady state levels.

Finally, consider the effects of a temporary drop in i^d_t, the interest rate paid on MP. To fix ideas, consider the following unexpected shock occurring at time 0:

$$i^d_t = \begin{cases} \underline{i}^d < i^d & \text{for } \; 0 \le t < T \\ i^d & \text{for } \; t \ge T. \end{cases} \tag{21.28}$$

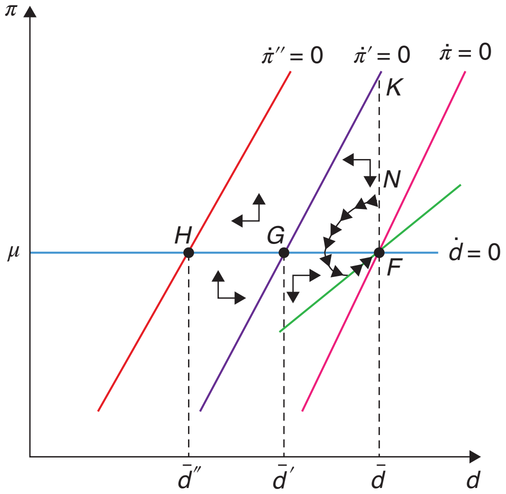
To sort out what happens it helps to begin by asking what is the trajectory of Δ_t. It rises on impact, but it does not go all the way up to $\bar{\Delta}'$, the level it would take on if the change were permanent. The differential Δ_t falls thereafter, so that it can jump at T when i^d_t goes back to its initial level, ensuring that Δ_t is back to its initial steady state level $\bar{\Delta}$ an instant after T (in contrast to the policy variable, i^b cannot jump).

Let Δ_{0^+} be the value of Δ_t once the unexpected shock happens at $t = 0$. It must be the case, by the arguments above, that $\bar{\Delta} < \Delta_0 + < \bar{\Delta}'$. You can see this evolution in the phase diagram in Figure 21.7, where we show the (linearised version of) the $\dot{\Delta}_t = 0$ schedule.

What are the implications for the dynamic behaviour of inflation and the real stock of MP? We can study that graphically in Figure 21.8 below. If the policy change were permanent, the $\dot{\pi} = 0$ schedule would have moved all the way to $\dot{\pi}'' = 0$, giving rise to a steady state at H. But the fact that $\Delta_0 - \bar{\Delta}' < 0$ offsets some of that leftward movement. So, the $\dot{\pi} = 0$ schedule moves to $\dot{\pi}' = 0$, creating a temporary (for an instant) steady state at G.

Ask what would happen if Δ_t were to remain at Δ_0 until time T. Inflation would jump up on impact. But it cannot go beyond point K, because if it did the system would diverge to the northwest afterwards. So, inflation would jump to a point like N. After the jump, the economy would begin to move following the arrows that correspond to the system with steady state at G.

Figure 21.8 A temporary decline in the rate on reserves



Of course, an instant after T, and because of the movement in Δ_t, the locus $\dot{\pi}' = 0$ begins to shift to the right. But this does not affect the qualitative nature of the adjustment path, because the system always lies to the right of the shifting $\dot{\pi}' = 0$ locus, and thus obeys the same laws of motion as it did an instant earlier. The evolution of inflation and real MP is guided by the need that, at T, the system must be on the saddle path leading to the initial steady state at point F.

You can see from the phase diagram that after the initial jump up, inflation falls between times 0 and T, and rises thereafter. The real value of MP drops initially due to the high inflation, but then gradually recovers as π_t falls below μ. One can show also that output goes through a boom between times 0 and T, takes a discrete drop at T when the interest rate i_t^d rises again, and recovers gradually until returning to its initial steady state level.

21.3 | Policy implications and extensions

21.3.1 | Quantitative easing

We emphasised above that in this model the monetary authority has access to two policy levers: an interest rate (i^d) and a quantity tool (μ) —or potentially, two interest rates, if the central bank chooses to engage in open market operations and use changes in quantities to target i^b. So we have gone beyond the realm of conventional policy, in which control of the single interest rate on bonds is the only alternative.[12]

We saw earlier that a dilemma arises when the nominal interest rate is against the zero lower bound. Can we use the model we have just built to study that conundrum? Is there a policy that can stabilise output and inflation when the lower bound binds? The answer is yes (subject to parameter values), and in what follows we explain how and why.

To fix ideas, let us go back to the situation studied earlier in this chapter, in which, because of lagging productivity growth, the natural rate of interest drops. Suppose initially $i^b = r^n > 0$, $i_t^d = 0$

and $\mu = \pi = 0$. Then the following shock hits

$$r_t^n = \begin{cases} \underline{r}^n < 0 & \text{for} \quad 0 \leq t < T \\ r^n > 0 & \text{for} \quad t \geq T. \end{cases} \tag{21.29}$$

So starting from r^n, at time 0 the natural rate of interest unexpectedly drops down to $\underline{r}^n < 0$ and it remains there until time T, when it returns to r^n and stays there forever.

Notice first that if $\sigma = 1$, during the duration of the shock the NKIS curve (21.20) becomes:

$$\dot{x}_t = (i_t^b - \pi_t - \underline{r}^n) \tag{21.30}$$

So $\dot{x}_t = \pi_t = 0$ would require $i_t^b = \underline{r}^n < 0$. But this is impossible if the zero lower bound is binding and hence i_t^b must be non-negative. Our first conclusion, therefore, is that for monetary policy to get around the zero lower bound problem we must focus on the case in which $\sigma \neq 1$. This is the case in which the utility function is not separable in consumption and liquidity (MP), so that that changes in the opportunity cost of holding liquidity have an impact on the time profile of consumption and aggregate demand.

If we go back to the case in which $\sigma \neq 1$, during the duration of shock the NKIS curve (21.20) becomes:

$$\dot{x}_t = \sigma(i_t^b - \pi_t - \underline{r}^n) + (1 - \sigma)(1 - \alpha)\dot{\Delta}_t \tag{21.31}$$

It follows from (21.31) that $\dot{x}_t = x_t = 0$ and $\pi_t = 0$ if and only if

$$\dot{\Delta}_t = \frac{\sigma \left[(\alpha^{-1} - 1)e^{\Delta_t} + i_t^d - \underline{r}^n\right]}{(\sigma - 1)(1 - \alpha)}. \tag{21.32}$$

where we have used $i_t^b = (\alpha^{-1} - 1)e^{\Delta_t} + i_t^d$. For simplicity, focus on the case $\sigma > 1$. In that case, the RHS of this equation is positive (recall $\underline{r}^n < 0$), so the interest gap Δ_t must rise gradually during the period of the shock.

At this point we have to take a stance on a difficult question: does the zero lower bound apply to i_t^d as well? If we interpret d_t narrowly, as reserves commercial banks hold at the central bank, the answer may be negative: it is not hard to think of liquidity or safety reasons why banks would want to hold reserves at the central bank even if they have to pay a cost to do so. But if we interpret i_t^d more broadly as a digital currency, then the answer could be yes, because if the nominal interest rate on reserves is negative, households could prefer to hold their liquidity under the mattress and look for substitutes as a means of payment. This is the standard "disintermediation" argument for the zero lower bound. To avoid wading into this controversy, in this section we assume $i_t^d \geq 0$.

Moreover, and to keep things very simple, we assume the central bank keeps i_t^d at its steady state level of zero throughout. In that case, the equation for the evolution of Δ_t (21.32) reduces to

$$\dot{\Delta}_t = \frac{\sigma \left[(\alpha^{-1} - 1)e^{\Delta_t} - \underline{r}^n\right]}{(\sigma - 1)(1 - \alpha)}. \tag{21.33}$$

Next, recall the liquidity demand function $d_t = c_t - \Delta_t$, which implies that if consumption is to be constant during the period of the shock, then $\dot{d}_t = -\dot{\Delta}_t$ That is to say, the interest gap can be rising only if the (real) stock of MP is falling. But since we are also requiring zero inflation during that period, real MP decline is the same as nominal MP decline, implying $\mu_t = -\dot{\Delta}_t < 0$.

So now we know what the time profile of Δ_t and d_t must be between times 0 and T. What about the initial and terminal conditions? Suppose we require $i_T^b = r^n$, so that the interest rate on bonds will be exactly at its steady state level at time T. Since i_t^d is constant at zero and Δ_t must be falling, it follows that i_t^b must be rising during the length of the shock. So i_t^b must have jumped down at time 0, which in turn means d_t must have jumped up at the same time.

In summary: if $\sigma > 1$, a policy that keeps output at "full employment" and inflation at zero, in spite of the shock to the natural interest rate, involves: a) discretely increasing the nominal and real stock of MP at the time of the shock, causing the interest rate on bonds to fall on impact in response to the shock, in what resembles QE;[13] b) allowing the nominal and real stock of MP to fall gradually during the period of the shock, in what resembles the "unwinding" of QE; c) once the shock is over, ensuring policy variables return to (or remain at) their steady state settings: $\mu = 0$ and $i_t^d = 0$ for all $t \geq T$.[14]

The intuition for why this policy can keep the economy at full employment is as follows. With two goods (in this case, consumption and liquidity services) entering the utility function, what matters for the optimal intertemporal profile of expenditure is not simply the real interest rate in units of consumption, but in units of the bundle Z_t that includes both the consumption good and the real value of MP. Because the nominal rate on bonds cannot fall below zero, what brings the real "utility-based" interest rate down to the full employment level is the behaviour of the "relative price" Δ_t. When $\sigma > 1$, Δ_t has to rise to achieve the desired effect. If, on the contrary, we assumed $\sigma < 1$, then Δ_t would have to fall over time the period of the shock.[15]

In the case $\sigma > 1$, the gradual increase in Δ_t follows an initial drop in the same variable, caused by a discrete increase in the nominal and real stock of MP. This "quantitative easing", if feasible, manages to keep the economy at full employment and zero inflation in spite of the shock to the natural rate of interest and the existence of a zero lower bound for both nominal interest rates.

21.3.2 | Money and banking

An objection to the arguments so far in this chapter is that digital currencies do not yet exist, so households do not have accounts at the central bank. In today's world, the only users of central bank reserves are commercial banks. But most households do use bank deposits for transactions.

This does not mean that our previous analysis is useless. On the contrary, with relatively small modifications, it is straightforward to introduce a banking system into the model. Piazzesi et al. (2019) carry out the complete analysis. Here, we just sketch the main building blocks.

A simplified commercial bank balance sheet has deposits and bank equity on the liability side, and central bank reserves and other assets (loans to firms, government bonds) on the asset side. Banks are typically borrowing-constrained: they can issue deposits only if they have enough collateral - where central bank reserves and government bonds are good collateral.

So now d_t can stand for (the log of) the real value of deposits held in the representative commercial bank, and i_t^d is the interest rate paid on those deposits. Because deposits provide liquidity services, i_t^d can be smaller than the interest rate on bonds, i_t^b.

The central bank does not control i_t^b or i_t^d directly. But banks do keep reserves at the central bank, and this gives the monetary authority indirect control over market rates. Denote by i_t^h the interest rate paid on central bank reserves. It is straightforward to show (see Piazzesi et al. (2019) for details) that optimal behaviour by banks leads to

$$\left(i_t^b - i_t^d\right) = \ell \left(i_t^b - i_t^h\right), \tag{21.34}$$

where $\ell < 1$ if banks are borrowing-constrained and/or have monopoly power.[16] Whenever $\ell < 1$, $\left(i_t^b - i_t^h\right)(1-\ell) = i_t^d - i_t^h > 0$ so that the rate on deposits and on central bank reserves are linked, with the former always above the latter. The central bank can affect the rate on deposits by adjusting both the quantity of reserves and the interest rate paid on them. Demand for deposits, as in the previous subsection, depends on the opportunity cost of holding deposits:

$$d_t = c_t - \log\left[\left(\frac{\alpha}{1-\alpha}\right)\left(i_t^b - i_t^d\right)\right]. \tag{21.35}$$

Using the equation above we have

$$d_t = c_t - \log\left[\left(\frac{\alpha}{1-\alpha}\right)\ell\left(i_t^b - i_t^h\right)\right] = c_t - \log\left[\left(\frac{\alpha}{1-\alpha}\right)\ell\right] - \log\left(i_t^b - i_t^h\right). \tag{21.36}$$

With this expression in conjunction with the dynamic NKIS curve, the NKPC, and the corresponding policy rules, we have a macro model almost identical to that of the earlier sections, and which can be used to analyse the effects of exogenous shocks and policy changes.

Aside from realism, this extended version has one other advantage: shocks to financial conditions can now become another source of business cycle variation that needs to be counteracted by monetary (and perhaps fiscal) policy. The parameter ℓ, reflecting conditions in the financial markets, the quality of the collateral, the extent of competition, etc., enter as shifters in the expression for deposit demand. To fix ideas, consider what happens if we continue with the policy arrangement of the previous subsection, with $i_t^b = 0$ and the interest rate on reserves (now labelled i_t^h) exogenously given. Then, and since d_t is a sticky variable that cannot jump in response to shocks, an unexpected change in ℓ would imply a change in consumption, and, therefore, in aggregate demand and output. So, in the presence of shocks to financial market conditions, monetary policymakers have to consider whether and how they want to respond to such shocks.

21.3.3 | Credit easing

So far the focus of this chapter has been on unconventional policies that involve changing the quantity of reserves by having the central bank carry out open market operations involving safe assets like government bonds. But at the zero lower bound, and if the interest rate on reserves is brought down to the level of the interest rate on bonds (a case of liquidity satiation, not considered above), then from the point of view of the private sector (of a commercial bank, say), central bank reserves and short-term, liquid government bonds become identical: they are both i.o.u´s issued by the state (or the consolidated government, if you wish), paying the same rate of interest. So, operations that involve swapping one for the other cannot have any real effects.

That is why, in the face of financial markets frictions and distortions, over the last decade and particularly since the Great Financial Crisis, central banks have turned to issuing reserves to purchase other kinds of assets, from corporate bonds to loans on banks´ balance sheets, in effect lending directly to the private sector. As mentioned at the outset, these are usually labelled credit easing policies, in contrast to the "quantitative easing" policies that only involve conventional open market operations.

Credit easing can be incorporated into a simple model like the one we have been studying in this chapter, or also into more sophisticated models such as those of Curdia and Woodford (2011) and Piazzesi et al. (2019). There are many obvious reasons why such policies can have real effects: one is that they can get credit flowing again when the pipes of the financial system become clogged or frozen in a crisis.

A related reason is that in this context policy can not only address aggregate demand shortfalls, but also help alleviate supply constraints — if, for instance, lack of credit keeps firms from having the necessary working capital to operate at the optimal levels of output. This all begs the question of what policy rules ought to look like in such circumstances, a fascinating subject we cannot address here, but about which there is a growing literature — beginning with the 2009 lecture at LSE in which Ben Bernanke, then Fed Chair, explained the Fed's approach to fighting the crisis, which stressed credit easing policies (Bernanke (2009).

21.4 | Appendix

The FOC, (21.8)-(21.10) repeated here for convenience, are

$$\alpha Z_t^{\left(\frac{\sigma-1}{\sigma}\right)} = C_t \lambda_t \tag{21.37}$$

$$(1-\alpha) Z_t^{\left(\frac{\sigma-1}{\sigma}\right)} = \lambda_t D_t \left(i_t^b - i_t^d\right) \tag{21.38}$$

$$\dot{\lambda}_t = -\lambda_t \left(i_t^b - \pi_t - \rho\right), \tag{21.39}$$

where we have defined

$$C_t^{\alpha} D_t^{1-\alpha} \equiv Z_t. \tag{21.40}$$

Combining the first two, we have demand for MP:

$$D_t = \frac{C_t}{\left(\frac{\alpha}{1-\alpha}\right)\left(i_t^b - i_t^d\right)}, \tag{21.41}$$

which in logs is

$$d_t = c_t - \Delta_t, \tag{21.42}$$

where

$$\Delta_t = \log\left[\left(\frac{\alpha}{1-\alpha}\right)\left(i_t^b - i_t^d\right)\right]. \tag{21.43}$$

Next, differentiating (21.38) with respect to time and then combining with (21.40) yields

$$\left(\frac{\sigma-1}{\sigma}\right)\frac{\dot{Z}_t}{Z_t} = \frac{\dot{C}_t}{C_t} - \left(i_t^b - \pi_t - \rho\right). \tag{21.44}$$

Or, in logs

$$\left(\frac{\sigma-1}{\sigma}\right)\dot{z}_t = \dot{c}_t - \left(i_t^b - \pi_t - \rho\right). \tag{21.45}$$

Using demand for MP from (21.42) in the definition of Z_t (21.41) yields

$$Z_t = C_t^{\alpha} D_t^{1-\alpha} = C_t \left(\frac{\alpha}{1-\alpha}\right)^{-(1-\alpha)} \left(i_t^b - i_t^d\right)^{-(1-\alpha)}. \tag{21.46}$$

Or, in logs

$$z_t = c_t - (1-\alpha)\Delta_t. \tag{21.47}$$

Differentiating (21.48) with respect to time yields

$$\dot{z}_t = \dot{c}_t - (1-\alpha)\dot{\Delta}_t. \tag{21.48}$$

Replacing the expression for $\dot{z}_t$ from (21.46) in (21.49) we obtain the Euler equation (21.13) used in the text:

$$\dot{c}_t = \sigma\left(i_t^b - \pi_t - \rho\right) + (1-\sigma)(1-\alpha)\dot{\Delta}_t, \tag{21.49}$$

which can be also written, perhaps more intuitively, as

$$\dot{c}_t = \sigma\left[i_t^b - \pi_t - \left(\frac{\sigma-1}{\sigma}\right)(1-\alpha)\dot{\Delta}_t - \rho\right]. \tag{21.50}$$

This way of writing it emphasises that the relevant real interest rate now includes the term $\left(\frac{\sigma-1}{\sigma}\right)(1-\alpha)\dot{\Delta}_t$, which corrects for changes in the relative price of the two items that enter the consumption function.

Notes

[1] On monetary policy during the pandemic, see Woodford (2020).

[2] A good review of the discussion can be found in Rogoff (2017).

[3] A technical clarification: in Chapter 15 we claimed that, in the absence of an activist interest rule, the canonical 2− equation New Keynesian model does not have a unique equilibrium. So why have no multiplicity issues cropped up in the analysis here? Because, to draw the phase diagram the way we did we assumed the central bank would do whatever it takes to keep $\pi_t = x_t = 0$ starting at T (including, perhaps, the adoption of an activist rule starting at that time). That is enough to pin down uniquely the evolution of the system before T_j because it must be exactly at the origin $\left(\pi_t = x_t = 0\right)$ at T. See Werning (2011) for the formal details behind this argument.

[4] Recall from Chapter 14 that $\kappa \equiv \alpha^2\eta > 0$, and α^{-1} is the expected length of a price quotation in the Calvo (1983) model. So as prices become perfectly flexible, κ goes to infinity.

[5] For details, see the discussion by Gertler on the paper by Eggertsson and Woodford (2003).

[6] See Woodford (2010) from which this discussion is taken.

[7] In particular, on whether banks' demand for liquidity has been satiated or not. See the discussion in Reis (2016).

[8] We will see later that, under some simple extensions, D_t can also be thought of as deposits issued by commercial banks. But let us stick with the digital currency interpretation for the time being.

[9] You may be wondering where currency is in all of this. We have not modelled it explicitly, but we could as long as it is an imperfect substitute for MP (meaning they are both held in equilibrium even though they have different yields — zero in nominal terms in the case of currency).

[10] According to Reis (2016), this is more or less what the Federal Reserve has tried to do since the Great Financial Crisis of 2007-09, thereby satiating the demand for liquidity.

[11] This very helpful way of solving a model of this type is due to Calvo and Végh (1996).

[12] Notice, however, that all the analysis so far (and what follows as well) assumes $i^d < i^b$. That is, there is an opportunity cost of holding reserves (or MP, if you prefer) and therefore liquidity demand by banks (or households, again, if you prefer) is not satiated. The situation is different when the interest rate on reserves is the same as the interest rate on government bonds. Reserves are a liability issued by one branch of government — the central bank. Bonds or bills are a liability issued by another

branch of government — the Treasury. The issuer is the same, and therefore these securities ought to have the same (or very similar) risk characteristics. If they also pay the same interest rate, then they become perfect substitutes in the portfolios of private agents. An operation involving exchanging reserves for bonds, or vice-versa, would have no reason to deliver real effects. A Modigliani-Miller irrelevance result would kick. However, there may be some special circumstances (fiscal or financial crisis, for instance) in which this equivalence breaks down. See the discussion in Reis (2016).

[13] QE involves issuing reserves to purchase bonds, and that is exactly what is going on here.

[14] Notice this policy is not unique. There are other paths for MP and i_t^d that could keep output and inflation constant. We have just chosen a particularly simple one. Notice also that in the sequence we described, the interest gap Δ_t jumps down on impact and then rises gradually until it reaches its steady level $r^n > 0$ at time T, but this trajectory is feasible as long as the shock does not last too long (T is not too large) and the shock is not too deep ($\underline{r}^n$ is not too negative). The constraints come from the fact that on impact Δ_t drops but can never reach zero (because in that case demand for MP would become unbounded). In other words, the central bank is not free to pick any initial condition for Δ_t, in order to ensure that, given the speed with which it must rise, it will hit the right terminal condition at time T. Part of the problem comes from the fact that we have assumed that the inflation rate in the initial steady state is zero, so the initial nominal interest rate on bonds is equal to the natural rate of interest. But, in practice, most central banks target inflation at 2 percent per year, which gives Δ_t "more room" to drop, so that central bankers can freely engage in the kind of policy we have described. Moreover, in the aftermath of the 2007-09 global financial crisis there were suggestions to raise inflation targets higher, to give central banks even "more room" in case of trouble.

[15] Dornbusch (1983) was the first to make this point.

[16] By contrast, in the absence of financial frictions and with perfect competition, $\ell = 1$ and $i_t^d = i_t^h$, so that the interest rate on deposits is equal to the rate paid on central bank reserves.

References

Beckworth, D. (2019). Facts, fears, and functionality of NGDP level targeting: A guide to a popular framework for monetary policy. *Mercatus Research Paper*.

Bernanke, B. S. (2009). The crisis and the policy response. Stamp Lecture, London School of Economics, January, 13, 2009.

Blanchard, O. & Galí, J. (2007). Real wage rigidities and the New Keynesian model. *Journal of Money, Credit and Banking*, *39*, 35–65.

Brunnermeier, M., Eisenbach, T., & Sannikov, Y. (2013). Macroeconomics with financial frictions: A survey. *Advances in Economics and Econometrics*.

Calvo, G. A. (1978). On the time consistency of optimal policy in a monetary economy. *Econometrica*, 1411–1428.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, *12*(3), 383–398.

Calvo, G. A. & Végh, C. A. (1996). Disinflation and interest-bearing money. *The Economic Journal*, *106*(439), 1546–1563.

Curdia, V. & Woodford, M. (2011). The central-bank balance sheet as an instrument of monetarypolicy. *Journal of Monetary Economics*, *58*(1), 54–79.

Diba, B. & Loisel, O. (2020). Pegging the interest rate on bank reserves: A resolution of New Keynesian puzzles and paradoxes. *Journal of Monetary Economics*.

Dornbusch, R. (1983). Real interest rates, home goods, and optimal external borrowing. *Journal of Political Economy*, *91*(1), 141–153.

Eggertsson, G. B. & Woodford, M. (2003). Zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity*, *2003*(1), 139–233.

Eggertsson, G. B. & Woodford, M. (2004). Policy options in a liquidity trap. *American Economic Review*, *94*(2), 76–79.

Gertler, M. & Karadi, P. (2011). A model of unconventional monetary policy. *Journal of Monetary Economics*, *58*(1), 17–34.

Hall, R. E. & Reis, R. (2016). *Achieving price stability by manipulating the central bank's payment on reserves*. National Bureau of Economic Research.

Hicks, J. R. (1937). Mr. Keynes and the "classics"; A suggested interpretation. *Econometrica*, 147–159.

Krugman, P. (1998). Japan's trap. http://web.mit.edu/krugman/www/japtrap.html.

Kydland, F. E. & Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, *85*(3), 473–491.

Piazzesi, M., Rogers, C., & Schneider, M. (2019). *Money and banking in a New Keynesian model*. Working Paper, Stanford.

Reis, R. (2016). Funding quantitative easing to target inflation. LSE Research Online, http://eprints.lse.ac.uk/67883/.

Rogoff, K. (2017). Dealing with monetary paralysis at the zero bound. *Journal of Economic Perspectives*, *31*(3), 47–66.

Summers, L. H. (2018). Secular stagnation and macroeconomic policy. *IMF Economic Review*, *66*(2), 226–250.

Sumner, S. B. (2014). Nominal GDP targeting: A simple rule to improve Fed performance. *Cato Journal. 34*, 315.

Werning, I. (2011). *Managing a liquidity trap: Monetary and fiscal policy*. National Bureau of Economic Research.

Woodford, M. (2010). Financial intermediation and macroeconomic analysis. *Journal of Economic Perspectives*, *24*(4), 21–44.

Woodford, M. (2016). *Quantitative easing and financial stability*. National Bureau of Economic Research. Working Paper 22285 http://www.nber.org/papers/w22285.

Woodford, M. (2020). Post-pandemic monetary policy and the effective lower bound.

CHAPTER 22

New developments in monetary and fiscal policy

In Chapters 19 and 20 we laid out the basics of monetary theory. Chapter 19 explored the relation between prices and money while Chapter 20 focused on the historical debate on whether monetary policy should be conducted through rules or with discretion. In Chapter 21 we discussed new challenges to monetary policy, particularly those that became clear after the Great Financial Crisis: how does the ZLB constrain the operation of monetary policy, and what is the role and scope for quantitative easing?

In this chapter we address three points that are currently being discussed. None of these are settled as of today, but we hope the presentation here will help introduce the issues.

The first topic is Alvin Hansen's secular stagnation dilemma, brought back to life by Larry Summers a few years ago. In Hansen's (1939) own words:

> This is the essence of secular stagnation, sick recoveries which die in their infancy and depressions which feed on themselves and leave a hard and seemingly immovable form of unemployment.

While by 1939 the U.S. economy was in full recovery, the idea brewed after a decade-long recession. The idea was that depressed expectations may lead to increased savings that do not find a productive conduit, further depressing aggregate demand. This idea blends with the "savings glut" referred to in Ben Bernanke's 2005 speech pointing to the fact that the world had been awash with savings in recent years. While unemployment has been low in recent years (as opposed to the 30s), there is a sense that recoveries are slow, and policy tools are ineffective to address this. Thus, our first section in this chapter deals with models that try to formalise this pattern: excessive savings leading to depressed aggregate demand that policies cannot counteract if the interest rate has a lower bound.

The second section deals with the need to build a monetary theory in a world without money. In future years, private crypto currencies, floating and fixed, and electronic payments will make the money supply issued by central banks increasingly irrelevant. Picture a world in which payments are done through mechanisms such as QR codes, electronic transfers, while base money demand falls to zero. Imagine that the replacement of cash has gone so far that only 1 dollar of base money is left.[1] Will prices double if that one dollar becomes two? We deal with these issues by discussing the so-called "fiscal theory of the price level", a long term effort proposed by economists such as Michael Woodford, Chris Sims, and John Cochrane. The theory focuses on the budget constraint of the government. The

How to cite this book chapter:
Campante, F., Sturzenegger, F. and Velasco, A. 2021. *Advanced Macroeconomics: An Easy Guide.*
Ch. 22. 'New developments in monetary and fiscal policy', pp. 345–362. London: LSE Press.
DOI: https://doi.org/10.31389/lsepress.ame.v License: CC-BY-NC 4.0.

government issues debt and charges taxes, and the price level is the one that equalises the real value of debt with the present discounted value of taxes. Think of taxes as the money sucked by the government from the system. If the real value of debt is higher than this, there is an excess demand that pushes prices upwards until the equilibrium is restored. We revisit the discussion on interest rules within this framework.

While the quantitative easing policies discussed in the previous chapter initially raised eyebrows and had many skeptics and detractors who had forecasted increasing inflation and interest rates, none of those predictions bore out. Inflation rates have remained low, even when interest rates reached historical lows. In fact, low persistent interest rates have raised the issue of the possibility of an unbounded fiscal policy: if $r < g$ can the government raise debt without bound? Does that make sense? This is one of the most significant policy debates today, particularly after the debt buildup as a result of the Covid pandemic. Furthermore, if $r < g$, can assets be unbounded in their valuations or include bubbles? And if they do, what is their welfare effect? This, in turn, opens a new set of policy questions: should the monetary authorities fight bubbles? In practical terms, should the Fed worry about stock market prices? Our final section tackles this issue. It resembles somewhat the discussion on optimality that we found in the OLG section of this book. Bubbles, when they exist, improve welfare. However, bubble-driven fluctuations are not welfare-improving and monetary authorities should attempt to avoid them. This is one the hottest and most debated current topics in monetary policy.

22.1 | Secular stagnation

As we have been discussing all along, recent years have shown very low interest rates, so low that they make the zero lower bound constraint something we need to worry about. In the previous chapter we showed how monetary policy could respond to this challenge, here we provide an alternative representation that allows for financial constraints and productivity growth to play a role.

During recent years, inflation has surprised on the downside and economic recovery has been sluggish. This combination has been dubbed secular stagnation, a name taken from Hansen's 1939 depiction of the U.S. economy during the Great Depression. The story is simple: low interest rates due to abundant savings are associated with depressed demand. But this lack of demand generates lower inflation, pushing up the real interest rate and strengthening the contractionary effect. We follow Eggertsson et al. (2017) in modelling all these effects in a simple framework.

Their model is an overlapping generations framework (we will need overlapping generations if we want to produce a low interest rate). In their specification, every individual lives for three periods. In period one the individual has no income and needs to borrow in order to consume. However, it is subject to a collateral constraint D_t (this will open the door for financial effects in the model). The individual generates income in the middle period and no income in old age (this will produce the need for savings). In summary, the individual maximises

$$maxE_t \left\{ \log(C_t^y) + \frac{1}{1+\rho} \log(C_{t+1}^m) + \left(\frac{1}{1+\rho}\right)^2 \log(C_{t+2}^o) \right\}. \tag{22.1}$$

subject to

$$C_t^y = \frac{D_t}{(1+r_t)}, \tag{22.2}$$

$$C_{t+1}^m = Y_{t+1}^m - D_t + B_{t+1}^m, \tag{22.3}$$

$$C_{t+2}^o = -(1+r_{t+1})B_{t+1}^m. \tag{22.4}$$

To make things interesting, we will consider the case in which the financing constraint is binding in the first period. This implies that the only decision is how much to borrow in middle age. You should be able to do this optimisation and find out that the desired savings are

$$B_t^m = -\frac{(Y_t^m - D_{t-1})}{2+\rho}, \tag{22.5}$$

which is the supply of savings in the economy. Equilibrium in the bond market requires that borrowing of the young equals the savings of the middle-aged so that $N_t B_t^y = -N_{t-1} B_t^m$, and denoting as usual, n as the rate of population growth,

$$(1+n)B_t^y = \frac{(1+n)}{(1+r_t)} D_t = -B_t^m = \frac{(Y_t^m - D_{t-1})}{2+\rho}, \tag{22.6}$$

This equation readily solves the interest rate for the economy

$$1 + r_t = (2+\rho)(1+n)\frac{D_t}{(Y_t^m - D_{t-1})}. \tag{22.7}$$

Notice that the interest rate can be lower than the growth rate, and even negative. The fact that individuals are cash constrained in the first period will also impact the interest rate: a tighter constraint today (a smaller D_t) leads to a fall in the interest rate. Notice also that a lowering of productivity growth, if it tightens the financing constraint due to lower expected future income, lowers the interest rate, as does a lower rate of population growth. These low interest rates are not transitory but correspond to the steady state of the economy.

Let's introduce monetary policy in this model. The real interest rate now is

$$(1+r_t) = (1+i_t)\frac{P_t}{P_{t+1}}, \tag{22.8}$$

where the notation is self explanatory. The problem arises if we impose a zero lower bound for the nominal interest rate ($i_t \geq 0$). In a deflationary equilibrium this may not allow the desired real interest rate and will be a source of problems. In order to have output effects we will assume firms produce using labour with production function

$$Y_t = L_t^{\alpha}. \tag{22.9}$$

The critical assumption will be that a fraction γ of workers will not accept a reduction in their nominal wages, just as Keynes suggested back in the 1930s. This implies that the nominal wage of the economy will be

$$\tilde{W}_t = \gamma W_{t-1} + (1-\gamma) W_t^{flex}, \tag{22.10}$$

where W_t^{flex} indicates the wage that clears the labour market. In order to compute the aggregate supply curve we first look for the steady state wage when the nominal constraint is binding. Dividing (22.10) by P_t and replacing W_t^{flex} by the marginal product of labour at full employment, $\alpha \bar{L}^{\alpha-1}$ we can see that the steady state wage is

$$w = \frac{(1-\gamma)\alpha\bar{L}^{\alpha-1}}{(1-\frac{\gamma}{\Pi})}, \tag{22.11}$$

where Π is gross inflation. Then, noting that firms will equate the marginal product of labour to the wage, we make the wage in (22.11) equal to the marginal product of labour $\alpha L^{\alpha-1}$. After some simplifications we get the equation

$$\frac{\gamma}{\Pi} = 1 - (1-\gamma)\frac{Y}{Y^f}^{\frac{1-\alpha}{\alpha}}, \tag{22.12}$$

where Y^f represents full employment output. Notice that this operates as a Phillips curve, higher inflation is associated with higher output. The intuition is straightforward, as inflation increases the real wages falls and firms hire more labour. Given the rigidities of the model, this is a steady state relationship. If inflation is sufficiently high, the nominal wage rigidity does not bind and the aggregate supply becomes vertical, as drawn in Figure 22.1.

Aggregate demand follows directly from (22.7) combined with the Fisher relation (22.8) and a Taylor rule such as

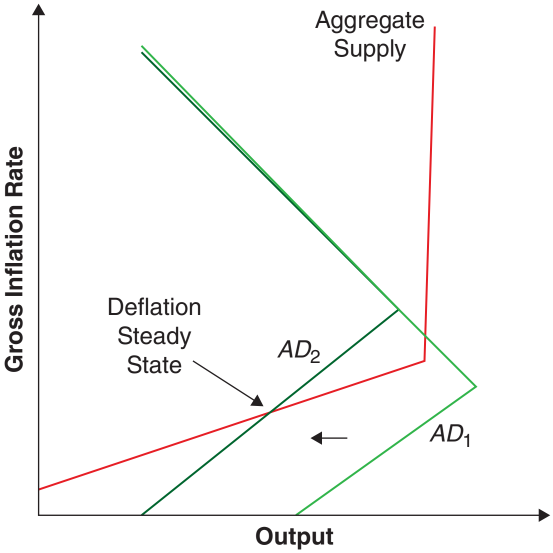
$$(1+i) = (1+i^*)\left(\frac{\Pi_t}{\Pi^*}\right)^{\phi_\Pi}. \tag{22.13}$$

Substituting both we obtain:

$$Y = D + (2+\rho)(1+n)D\frac{\Pi^{*\phi_\Pi}}{(1+i^*)}\frac{1}{\Pi^{\phi_\Pi - 1}}. \tag{22.14}$$

The upper portion of the AD curve in Figure (22.1), when inflation is high and the interest rate is unconstrained, depicts this relationship. As inflation increases, the central bank raises the nominal interest rate by more than one for one (since $\phi_\Pi > 1$), which, in turn, increases the real interest rate and reduces demand.

Figure 22.1 Shift in the AD-AS model



Now, in low inflation states where the interest rate hits its zero lower bound $i = 0$, this curve simplifies to

$$Y = D + (2 + \rho)(1 + n)D\Pi. \tag{22.15}$$

The key point of this equation is that it is upward sloping in Π. The reason is straightforward, as inflation decreases, the lower bound implies that the real rate increases because the interest rate cannot compensate. Therefore, in this range, a lower inflation means a higher real rate. For reasonable parameter values the AD curve is steeper than the AS curve as in Figure 22.1

With this framework, imagine a decrease in D. In (22.15) is is easy to see that this moves the aggregate demand to the left as shown in the graph. This pushes inflation and output downwards. The link between the financial crisis and a protracted stagnation process has been established.

22.2 | The fiscal theory of the price level

Let's imagine now a situation where people transact without money (in the current world of Venmo, electronic wallets, etc, this is not such a far-fetched assumption, and will become less and less far-fetched as time goes by). What pins down the price level?[2]

The fiscal theory of the price level, as its name suggests, focuses on the role of fiscal policy in determining the price level. For sure, it is the government that prints money and issues debt. It then mops up money through taxes. To build intuition, let's imagine the government issues debt that needs to be paid at the end of the period. Taxes will be used to that end and the fiscal result will be s_t. Fiscal theory postulates that

$$\frac{B_{t-1}}{P_t} = s_t. \tag{22.16}$$

The price level adjusts to equate the real value of debt to the expected surplus. Why is this the case? Imagine the price level is lower than the one that makes the real value of debt equal to the expected surplus. This means that at the end of the day the mopping up of money will be less than the value of debt. This implies that agents have an excess demand of resources to spend, which they use to push the price level up. If the price level makes the value of debt smaller than what will be mopped up, this implies that people are poorer and reduce spending. The price level comes down. In short, there is equilibrium when the real value of debt equals the expected surplus. Equation (22.16) is not the budget constraint of the government. The budget constraint is

$$B_{t-1} = s_t P_t + M_t. \tag{22.17}$$

The point is that there is no demand for M_t. This is what makes (22.16) an equilibrium condition. In the reasoning above we take s_t as exogenous, but we can also consider the case in which fiscal policy is endogenous in the sense that it adjusts in order to attain a particular P_t. When so we will say we have a passive fiscal police. We will come back to this shortly.

Of course this intuition can be extended to multiple periods. In an intertemporal setting the budget constraint becomes

$$M_{t-1} + B_{t-1} = P_t s_t + M_t + B_t E_t[\frac{1}{(1 + i_t)}], \tag{22.18}$$

but because people do not hold money in equilibrium, and using the fact that $(1+i)=(1+\rho)\frac{P_{t+1}}{P_t}$ we can write this as

$$B_{t-1}=P_ts_t+B_tE_t[\frac{1}{(1+i_t)}]=P_ts_t+B_tE_t[\frac{1}{(1+\rho)}\left(\frac{P_t}{P_{t+1}}\right)]. \tag{22.19}$$

It is easy to solve this equation by iterating forward (and assuming a transversality condition) to obtain

$$\frac{B_{t-1}}{P_t}=E_t\sum_{j=0}^{\infty}\frac{s_{t+j}}{(1+\rho)^j}. \tag{22.20}$$

This equation pins down the price level as the interplay of the real value of debt and the present discounted value of expected fiscal surpluses. Again, the right-hand side is the amount of dollars to be mopped up. The left-hand side is the real value of the assets in the private sector's hands. If the former is bigger than the future surpluses, there will be an aggregate demand effect that will push prices upward.

Imagine that the government issues debt but promises to increase the surpluses required to finance it. According to this view, this will have no effect on aggregate demand, or the price level. In fact, this explains why large fiscal deficits have had no effect on the price level. According to this view, to bring inflation up you need to issue debt and commit not to collect taxes! In short, you need to promise irresponsible behaviour. In Chris Sims' (2016) words,

> 'Fiscal expansion can replace ineffective monetary policy at the zero lower bound, but fiscal expansion is not the same thing as deficit finance. It requires deficits aimed at, and conditioned on, generating inflation. The deficits must be seen as financed by future inflation, not future taxes or spending cuts.'

Notice that, under this light, a large increase in assets in the balance sheet of the central bank, (such as the quantitative easing exercises of last chapter), does not affect the price level. The government (the Fed) issues debt in exchange of commercial assets. If both assets and liabilities pay similar rates this does not affect future surpluses, it just adds assets and liabilities for the same amount. This result is not dissimilar to what happens when a Central Bank accumulates (foreign currency) reserves in fixed exchange rate regimes. An increase in money liabilities accompanied by an increase in backing is understood not to be inflationary. In summary in this framework the large expansions that have come with quantitative easing do not necessarily generate inflation pressures.

Finally, the FTPL shows the interplay between fiscal and monetary policy. Imagine a central bank that increases the interest rate. If the fiscal surplus does not change this would decrease the surpluses and generate inflation. The reason why typically it is not assumed that an increase rate in the interest increases inflation is due to the fact that it is assumed that an increase in the interest rate will lead automatically to an increase in the fiscal surplus in order to pay the higher interest cost. But, according to this view, this response is not necessary. It is an assumption that we make.

22.2.1 | Interest rate policy in the FTPL

What does the FTPL have to say about the stability of interest rate policies? Let's analyse this within the context of the New Keynesian model we discussed in Chapter 15. As always, the behavioural equations include the New Keynesian IS curve (NKIS):

$$\log(Y_t)=E_t\left[\log(Y_{t+1})\right]-\sigma(i_t-E_t\pi_{t+1}), \tag{22.21}$$

and the New Keynesian Phillips Curve (NKPC):

$$\pi_t = \kappa y_t + \beta E_t \left[\pi_{t+1}\right]. \tag{22.22}$$

And an exogenous interest rule

$$i_t = \bar{i}. \tag{22.23}$$

Notice that we use the exogenous interest rate rule that delivered instability in the traditional NK framework. The key innovation now is an equation for the evolution of the debt to GDP level in the model. The dynamics of the log of the debt to GDP can be approximated by

$$d_{t+1} + s_{t+1} = d_t + (i_{t+1} - \pi_{t+1}) - g_{t+1}. \tag{22.24}$$

The debt to GDP ratio grows from its previous value with the primary deficit plus the real interest rate minus the growth rate.

The difference with the traditional NK model is that the NK model assumes that s_t adjusts to balance the budget to make debt sustainable, responding in a passive way, for example, to a change in the interest rate. This is the assumption that we lift here. Substituting the interest rate in the other equations we have a system of future variables on current variables with a coefficient matrix

$$\Omega = \begin{bmatrix} \frac{1}{\beta} & -\frac{\kappa}{\beta} & 0 \\ -\frac{\sigma}{\beta} & 1 + \frac{\kappa\sigma}{\beta} & 0 \\ -\frac{1}{\beta} & \frac{\kappa}{\beta} & 1 \end{bmatrix}. \tag{22.25}$$

The last column suggests that one of the eigenvalues is equal to one.[3] Of the other two, one can be shown to be smaller than one, and the other larger than one. Because the system has two jumpy variables (output and inflation) and one state variable (debt), the system is stable.

The above may have sounded like a bit of mathematical jargon. But here we are interested in the economics of the model. First of all, what does the above mean? And then, why does the model uniquely pin down the equilibrium here?

Remember that in the standard NK model we needed the Taylor principle, an interest rule that reacted with strength to deviations in inflation to pin down the equilibrium. The above suggests the the Taylor principle is not necessary in this context. Here the equilibrium is stable even with relatively stable interest rates. But why is this the case? The key difference here is that the interest rate change does not lead to any reaction in fiscal policy. Typically we ignore the implications of monetary policy on the government's budget constraint, but in doing so we are assuming (perhaps inadvertently) that fiscal policy *responds automatically* to monetary policy to keep the debt stable or sustainable. This may very well be the case, particularly in stable economies where fiscal stability is not at stake or questioned, but, at any rate, it is an assumption that we make.

What happens if we explicitly assume that there is no response of fiscal policy? Well, in that case the jump in the interest rate in an economy with sticky prices increases the real interest rate, but now there is only one path of inflation and output dynamics that insures the stability of the debt dynamics. The need to generate stability in the debt dynamics is what pins down the equilibria!

The increase in the real interest rate, though transitory, reduces the present value of surpluses (alternatively you can think of it as increasing the interest cost) leading to a higher level of inflation in equilibrium! The fact that higher interest rates imply a higher rate of inflation, is not necessarily contradictory with our previous findings, it just makes evident that in the previous case we were assuming that fiscal policy responded passively to insure debt stability.

The exercise helps emphasise that it is key to understand that relationship between fiscal and monetary policy. This relationship may become critical in the aftermath of the debt buildup as the result of Covid. If, imagine, interest rates increase at some point, what will be the response of fiscal policy? It is obvious that that channel cannot be ignored.

22.3 | Rational asset bubbles

Our final topic is the role of bubbles in asset prices and their implications for the economy, both in terms of efficiency and stability.

Perhaps the right place to start is a standard asset-pricing arbitrage equation:

$$r = \frac{\dot{q}_t}{q_t} + \frac{c}{q_t}, \tag{22.26}$$

where r is the real interest rate,[4] c the (constant) coupon payment, and q_t the price of the asset[5]. The equation states, as we have encountered multiple times, that the dividend yield plus the capital gain have to equal the opportunity cost of holding the asset.

That relationship can also be written as the differential equation

$$\dot{q}_t = rq_t - c. \tag{22.27}$$

You may want to review in the mathematical appendix to check out that the solution of this differential equation has the form

$$q_t = \int_0^\infty ce^{-rt}dt + Be^{rt} = \frac{c}{r} + Be^{rt}, \tag{22.28}$$

with B an arbitrary number (the fact that we use the letter B is not a coincidence). The solution has two terms. We call $\frac{c}{r}$ the fundamental value of the asset (naturally, the present discounted value of the coupon stream). The term Be^{rt} is the bubble term, which has no intrinsic value. We sometimes refer to this term as a "rational bubble".

Unless we arbitrarily impose a terminal condition (like requiring that q does not exceed some boundary after some given time period), then every value of q that satisfies the differential equation above is a candidate solution. The set of possible solutions is shown in Figure 22.2, which graphs the dynamics of the differential equation.

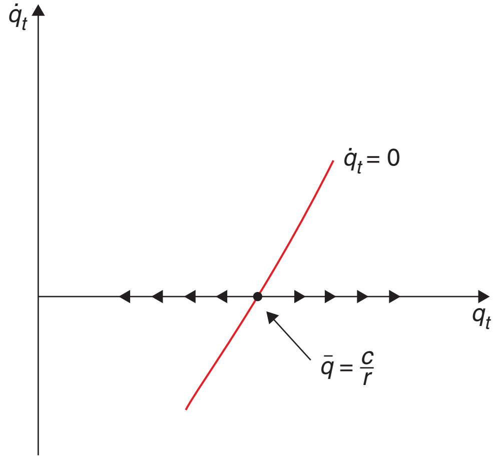
If q starts to the right of $\bar{q}$, it will go to infinity; if it starts to the left of $\bar{q}$, it will go to zero. All these paths satisfy the relevant arbitrage condition and the corresponding differential equation.

Before moving on, one important point to note is that the bubble term has a very tight structure: it grows at r. This is intuitive: the asset can price above its fundamental value only if agents expect this extra cost will also deliver the required rate of return ... forever. Thus, transitory increases in asset prices cannot be associated with bubbles.[6]

Can we rule out these bubbly paths? If $r > g$ – that is, the interest rate is larger than the growth rate – then the bubble eventually becomes so big that it is impossible for it to continue growing at the required rate (think of it becoming larger than the economy!). But if it cannot grow it cannot exist, and if it does not exist at a given moment in time by backward induction it cannot exist at anytime before. In dynamically efficient economies bubbles cannot exist.

That is the *unsustainable* element in bubbles; but there is also the *de-stabilising* element. Most borrowing contracts require collateral. Families use their homes as collateral; financial intermediaries

Figure 22.2 Solutions of the bubble paths



use stocks and bonds as collateral. When the prices of these assets become inflated, so does people's ability to borrow. So debt and leverage go up, which in turn may further stimulate the economy and causes asset prices to rise even more. The problem, of course, is that if and when asset prices come back to earth, it will find firms and businesses highly indebted, thus triggering a painful process of de-leveraging, recession, and bankruptcies.

So are bubbles all bad, then? Not necessarily. Imagine an overlapping generations ice cream economy in which people live for two days only. At any given time there are only two people alive: a 1-day-old person and a 2-day-old person. Each 1-day-old person gets an ice-cream cone in the evening. In this strange world, 1-day-olds do not like ice-cream, but 2-day-olds do. The problem is, anyone trying to store an ice cream cone overnight will only have a melted mess in her hands the next morning.

The 1-day-old person could kindly offer her cone to the 2 -day-old person, but the 2-day-old cannot offer anything in return, because they will be dead by morning. In this economy, therefore, one possible equilibrium is that all ice-cream melts and no one gets to eat any.

Unless, that is, one enterprising 2-day-old comes up with a new idea: give the 1-day-old a piece of paper (you can call it a dollar, or a bond if you prefer) in exchange for her cone, arguing that 24 hours later the then 2-day-old will be able to give that same piece of paper to a 1-day-old in exchange for a nice, fresh and un-melted ice cream cone! If today the 1-day-old agrees to the deal, and so do all successive 1-day-olds until kingdom come, then the problem will have been solved: all ice cream cones will be consumed by 2-day-olds who really appreciate their flavour, no ice cream will melt and welfare has improved.

But notice, the piece of paper that begins to change hands is like a bubble: it has no intrinsic value, and it is held only because of the expectation that others will be also willing to hold it in the future. In the simple story we have told the value of the bubble is constant: one piece of paper is worth one ice cream cone forever.[7] But in slightly more complicated settings (with population growth, for instance, or with changing productivity in the ice cream business), the value of the piece of paper could be rising over time, or it could even be falling! People might willingly exchange one cone today for 0.9 cones tomorrow simply because the alternative is so bad: eating no ice cream at all!

You may have found a resemblance between this ice-cream story and our discussion of dynamic efficiency when analysing the overlapping generations models in Chapter 8. When the return on capital is low (in this case it is -100% because if everyone saves the ice-cream cone overnight, then there is no ice-cream to consume tomorrow), finding a way to transfer income from the young to old is welfare enhancing.

In Samuelson's (1958) contribution that introduced the overlapping generations (OLG) model Samuelson already pointed out that bubble-like schemes (he called them "the social contrivance of money") could get economies out of that dynamic inefficiency and raise welfare[8].

But again, here is the rub: those schemes are inherently fragile. Each 1-day-old accepts the piece of paper if and only if they expect all successive 1-day-olds will do the same. If not, they have no incentive to do so. The conclusion is that, in a bubbly environment, even slight shifts in expectations can trigger big changes in behaviour and asset prices, what, in modern parlance, we call financial crises.

In another very celebrated paper, Samuelson's fellow Nobelist Jean Tirole (1985) took a version of the OLG model and analysed what kinds of bubbles could arise and under what conditions. The main conclusion is what we mentioned above: rational asset bubbles only occur in low interest rate OLG economies where the rate of interest is below the rate of growth of the economy. This is how Weil (2008) explains the intuition:

> 'This result should not be a surprise: I am willing to pay for an asset more than its fundamental value (the present discounted value of future dividends) only if I can sell it later to others. A rational asset bubble, like Samuelsonian money, is a hot potato that I only hold for a while-until I find someone to catch it.'

Tirole's paper gave rise to a veritable flood of work on bubbles, rational and otherwise. Recent surveys covering this new and fast-growing literature are Miao (2014) and Martin and Ventura (2018). One result from that literature is that there is nothing special in the 2 -generation Samuelson model that generates bubbles. Overlapping generations help, but other kinds of restrictions on trade (which is, essentially, what the OLG structure involves) can deliver very similar results.

In what follows we develop a simple example of rational bubbles using the Blanchard-Yaari perpetual youth OLG model that we studied in Chapter 8. In building this model we draw on recent work by Caballero et al. (2008), Farhi and Tirole (2012) and Galí (2020).

22.3.1 | The basic model

Consider an economy made up of overlapping generations of the Blanchard type with age-independent probability of death p (which by the law of large numbers is also the death rate in the population) and birth rate $n > p$. Together they mean population grows at the rate $n - p$. As in Blanchard (1985), assume there exist insurance companies that allow agents to insure against the risk of death (and, therefore, of leaving behind unwanted bequests)[9].

Let r_t be the interest rate and ρ the subjective rate of time discounting. Starting on the day of their birth, agents receive an endowment that cannot be capitalised and that declines at the rate $\gamma > \rho$. It is this assumption that income goes down as agents get older that will create a demand for a savings vehicle in order to move purchasing power from earlier to later stages of life.

The per capita (or, if you prefer, average) endowment received by agents who are alive at time t is (see Appendix 1 for details) a constant y.

Next let h_t stand for per capita (or average) household wealth. Appendix 2 shows that h_t evolves according to

$$\dot{h}_t = \left(r_t + p + \gamma\right) h_t - y. \tag{22.29}$$

To develop some intuition for this last equation, it helps to think of h_t as the present per capita value of the income flow y the household receives. What is the relevant rate at which the household should discount those future flows? The total discount rate is $(r_t + p + \gamma)$. In addition to the standard rate of interest r_t, the household must also discount the future by p, the instantaneous probability of death, and by γ, which is the rate at which an individual´s income falls over time.

The only way to save is to hold a bubble, whose per capita value is b_t. Arbitrage requires that per capita gains on the value of the bubble equal the interest rate:

$$\frac{\dot{b}_t}{b_t} = r_t - (n - p). \tag{22.30}$$

Recall that b_t is the per capita stock of the bubble, and $(n - p)$ is the rate of population growth. So this equation says that, by arbitrage, the percentage rate of growth of the (per capita) bubble must be equal to the (per capita) return on financial assets, which equals the interest rate net of population growth.

Finally, if utility is logarithmic, then (see Appendix 3) the per capita consumption function is

$$c_t = (p + \rho)\left(b_t + h_t\right), \tag{22.31}$$

so that consumption is a fixed portion of household wealth and financial wealth (the bubble). This condition mimics those we found for all Ramsey problems with log utility as we know from earlier chapters of this book. Because the economy is closed, all output must be consumed. Market-clearing requires

$$c_t = y = (p + \rho)\left(b_t + h_t\right). \tag{22.32}$$

Therefore, differentiating with respect to time we have

$$\dot{b}_t + \dot{h}_t = 0. \tag{22.33}$$

Replacing (22.29) and (22.30) in (22.33) yields

$$\left(r_t + p + \gamma\right) h_t + \left(r_t + p - n\right) b_t = y. \tag{22.34}$$

Combining this equation with the market-clearing condition (22.32) we get after some simplifications

$$r_t + \gamma - \rho = \frac{(p + \rho)(n + \gamma)b_t}{y}, \tag{22.35}$$

so that r_t is increasing in b_t: a larger bubble calls for a higher interest rate. Next we can use this equation to eliminate $r_t - n$ from the arbitrage equation (22.30) that governs the bubble:

$$\dot{b}_t = \frac{(p + \rho)(n + \gamma)b_t^2}{y} - (n - p + \gamma - \rho)b_t. \tag{22.36}$$

It follows that $\dot{b}_t = 0$ implies

$$[(p + \rho)(n + \gamma)b - (n - p + \gamma - \rho)y]b = 0. \tag{22.37}$$

This equation has solutions $b = 0$ and

$$b = \frac{(n - p + \gamma - \rho)y}{(p + \rho)(n + \gamma)} > 0. \tag{22.38}$$

Figure 22.3 describes the dynamic behaviour of the bubble. There are two steady states. The one with $b = 0$ is stable, while the one with with $b > 0$ is unstable. So bubbles can exist, but they are very fragile. Starting from the bubbly steady state, a minuscule shift in expectations is enough to cause the value of the bubble to start declining until it reaches zero.

Because this is a closed economy, per capita consumption must be equal to per capita income. But the same is not true of the consumption profiles of individual cohorts. It is easy to check that, in the bubbly steady state individual cohort consumption grows over time, while in the non-bubbly steady state, individual cohort consumption is flat. The bubble amounts to a savings vehicle that allows individuals to push consumption to later stages of life.

Put differently, in the steady state with no bubble, the interest rate is negative (you can see that the interest rate is $\rho - \gamma < 0$ by just assuming $b = 0$ in (22.35)). This is, trivially, smaller than the rate of population growth $(n - p) > 0$. So in that steady state the economy is dynamically inefficient. The existence of a bubble allows the economy to escape that inefficiency and settle on a golden rule steady state in which the interest rate is equal to the rate of population growth (to check this replace the level of b from (22.38) in (22.35) to find that $r = n - p$). But that equilibrium is fragile, as we have seen.

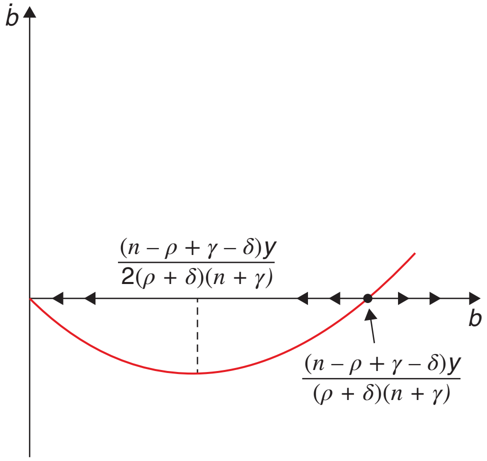
22.3.2 | Government debt as a bubble

Now define d_t as per capita public debt. The government budget constraint is

$$\dot{d}_t = \left(r_t - n + p\right) d_t + s, \tag{22.39}$$

where s is constant net spending per capita - the equivalent of the per capita primary deficit. Assume that all government spending is transfers to households (as opposed to government consumption), so

Figure 22.3 Bubbles in the Blanchard model



the law of motion of household wealth must be modified to read

$$\dot{h}_{\mathrm{t}} = \left(r_t + p + \gamma\right) h_t - y - s. \tag{22.40}$$

The consumption function is still (22.32), and $c_t = y$ which again implies that $\dot{h}_t + \dot{d}_t = 0$. We can then repeat the steps in the previous subsection replacing b_t by d_t to get an expression for the interest rate as a function of d_t:

$$\left(r_t + \gamma - \rho\right) y = (n + \gamma)(p + \rho)d_t, \tag{22.41}$$

which is virtually identical to (22.35). As before r_t is increasing in d_t: a larger stock of government debt requires a higher interest rate. Using this last expression in the government budget (22.39) constraint to eliminate $r_t - n$ we have

$$\dot{d}_t = \left(r_t - n + p\right) d_t + s = y^{-1}(n + \gamma)(p + \rho)d_t^2 - (n - p + \gamma - \rho)d_t + s. \tag{22.42}$$

Notice $\dot{d}_t = 0$ implies

$$(n + \gamma)(p + \rho)d^2 - (n - p + \gamma - \rho)yd + sy = 0. \tag{22.43}$$

This equation has two solutions, given by

$$\frac{d}{y} = \frac{(n - p + \gamma - \rho) \pm \sqrt{(n - p + \gamma - \rho)^2 - 4(n + \gamma)(p + \rho)\frac{s}{y}}}{2(n + \gamma)(p + \gamma)}. \tag{22.44}$$

Solutions exist as long as spending does not exceed a maximum allowable limit, given by

$$s = \frac{(n - p + \gamma - \rho)^2 y}{4(n + \gamma)(p + \rho)}. \tag{22.45}$$

If s is below this maximum, we have two equilibria, both with positive levels of debt. Figure (22.4) confirms that in this case the bubble is fragile: the steady state with the larger stock of debt is unstable. A slight shift in expectations will cause the value of the bubble to start declining.

It is straightforward to show that there are also two solutions for the real interest rate, given by

$$r = \frac{(n - p) - (\gamma - \rho) \pm \sqrt{((n - p) + (\gamma - \rho))^2 - 4(n + \gamma)(p + \rho)\frac{s}{y}}}{2}. \tag{22.46}$$

If $s = 0$, then

$$r = \frac{(n - p) - (\gamma - \rho) \pm ((n - p) + (\gamma - \rho))}{2}, \tag{22.47}$$

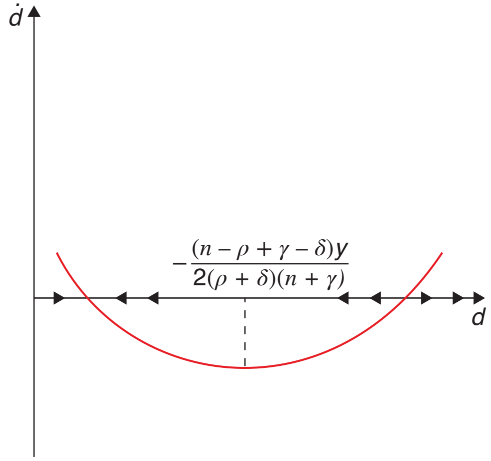
so $r = n - p$ or $r = -(\gamma - \rho)$, just as in the case of a pure bubble.

If $s > 0$, then in both equilibria $r < n - p$, so government debt turns out to be a bubble. In either steady state, the government budget constraint is

$$(n - p - r)d = s. \tag{22.48}$$

When the steady state interest rate is higher, so is the steady state value of the debt for a given primary deficit. One steady state is bubblier, with a larger valuation for public debt and a higher interest rate. But that steady state, as we saw graphically in the phase diagram above, is fragile. It only takes a shift in expectations to drive the economy out of that resting place and toward the alternative steady state with a lower valuation for public debt.

Figure 22.4 Government debt as a bubble



22.3.3 | Implications for fiscal, financial and monetary policy

As the discussion in the previous section should make clear, bubbles have huge implications for fiscal policy. We saw that in an economy with strong demand for liquidity, private sector agents may be willing to hold government debt even if that debt pays an interest rate that is lower than the rate of population growth, which is also the rate of growth of the economy.

This is good news for treasury officials and fiscal policymakers: the model suggests, they can run a primary deficit forever without ever having to raise taxes to retire the resulting debt. This is not just a theoretical *curiosum*. Today in most advanced economies, the real rate of interest is below the rate of economic growth (however paltry that rate of growth may be). This fact is motivating a deep rethinking about the limits of fiscal policy and the scope for a robust fiscal response not only to the Covid-19 pandemic, but also to the green infrastructure buildup that global warming would seem to require. Olivier Blanchard devoted his Presidential Lecture to the American Economic Association (2019) to argue that a situation in which $r < g$ for a prolonged period of time opens vast new possibilities for the conduct of fiscal policy[10].

But, at the same time, a bubbly world also bears bad news for those in charge of fiscal policy because, as we have seen above, bubbly equilibria are inherently fragile. Could it be that an advanced country issues a great deal of debt at very low interest rates and one day investors decide to dump it simply because of a self-fulfilling change in expectations? Hard to say, but it is not a possibility that can be entirely ignored. In fact, Blanchard (2019) acknowledges that arguments based on the possibility of multiple equilibria are "the most difficult to counter" when making the case for the increased fiscal space that low interest rates bring.

Bubbles also have vast implications for financial markets and financial regulation. The obvious concern, mentioned at the outset, is that asset bubbles typically end in tears, with overvaluation abruptly reversing itself and wrecking balance sheets. But here, also, the news is not all bad. Financial markets typically involve inefficient borrowing constraints that keep a subset of agents (especially small and medium enterprises) from undertaking positive net-present-value projects. Therefore, as Martin and Ventura (2018) emphasise, to the extent that bubbles pump up the value of collateral and relax borrowing constraints, they can promote efficiency and raise welfare as long as those bubbles do not burst [11].

Last but certainly not least, bubbles present difficult dilemmas for central banks and for monetary policy more generally. In the presence of sticky prices, if bubbles affect aggregate demand they also affect output and inflation, giving rise to bubble-driven business cycles. The implication is that standard monetary and interest rate rules need to be modified to take into account this new source of fluctuations. In some cases those modifications are relatively minor, but that is not always the case. Galí (2020) discusses the issues involved in greater detail than we can here.

22.4 | Appendix 1

Let $N_{t,\tau}$ be the size at time t of the cohort born at τ. The initial size of the cohort born at τ is nN_{τ}. In addition, the probability that someone born at τ is still alive at $t \geq \tau$ is $e^{-p(t-\tau)}$. It follows that

$$N_{t,\tau} = nN_{\tau}e^{-p(t-\tau)}. \tag{22.49}$$

Now, $N_t = N_{\tau}e^{(n-p)(t-\tau)}$, so

$$\frac{N_{t,\tau}}{N_t} = ne^{-p(t-\tau)}e^{-(n-p)(t-\tau)} = ne^{-n(t-\tau)}. \tag{22.50}$$

We conclude the relative size at time t of the cohort born at τ is $ne^{-n(t-\tau)}$ For any variable $x_{t,\tau}$ define the per capita (or average) x_t as

$$x_t = \int_{-\infty}^{t} x_{t,\tau}\left(\frac{N_{t,\tau}}{N_t}\right)d\tau, \tag{22.51}$$

$$x_t = \int_{-\infty}^{t} x_{t,\tau}ne^{-n(t-\tau)}d\tau. \tag{22.52}$$

For a person belonging to the cohort born on date τ, endowment income at time t is

$$y_{t,\tau} = \left(\frac{n+\gamma}{n}\right)ye^{-\gamma(t-\tau)}, \tag{22.53}$$

where y is a constant. Next define per capita (or average) endowment at time t as

$$y_t = \int_{-\infty}^{t} y_{t,\tau}ne^{-n(t-\tau)}d\tau, \tag{22.54}$$

$$y_t = y\left(\frac{n+\gamma}{n}\right)\int_{-\infty}^{t} e^{-\gamma(t-\tau)}ne^{-n(t-\tau)}d\tau, \tag{22.55}$$

$$y_t = y(n+\gamma)\int_{-\infty}^{t} e^{(n+\gamma)(t-t)}d\tau, \tag{22.56}$$

$$y_t = y. \tag{22.57}$$

22.5 | Appendix 2

The following derivation follows Blanchard (1985). Let the human wealth at time t of someone born on date τ be

$$h_{t,\tau} = \int_t^\infty y_{s,\tau} e^{-\int_t^s (r_v+p)dv} ds. \tag{22.58}$$

The income at time s of an individual is born on date τ is (this is the key declining income path assumption)

$$y_{s,\tau} = \left(\frac{n+\gamma}{n}\right) y e^{-\gamma(s-\tau)} = \left(\frac{n+\gamma}{n}\right) y e^{-\gamma(s-t)} e^{\gamma(t-t)}. \tag{22.59}$$

Therefore, the expression for $h_{t,\tau}$ can be written

$$h_{t,\tau} = e^{\gamma(\tau-t)} \int_t^\infty \left(\frac{n+\gamma}{n}\right) y e^{-\int_t^s (r_v+p+\gamma)dv} ds. \tag{22.60}$$

Next define per capita (or average) human wealth held by those still alive at t, given by

$$h_t = \int_{-\infty}^t h_{t,\tau} n e^{n(\tau-t)} d\tau. \tag{22.61}$$

Using the expression for $h_{t,\tau}$ the last equation can be written as

$$h_t = \int_{-\infty}^t e^{\gamma(\tau-t)} \left\{ \int_t^\infty \left(\frac{n+\gamma}{n}\right) y e^{-\int_t^s (r_v+p+\gamma)dv} ds \right\} n e^{n(\tau-t)} d\tau, \tag{22.62}$$

where the expression in curly brackets is the same for all agents. It follows that

$$h_t = \left\{ \int_t^\infty y e^{-\int_t^s (r_v+p+\gamma)dv} ds \right\} (n+\gamma) \int_{-\infty}^t e^{(n+\gamma)(\tau-t)} d\tau, \tag{22.63}$$

$$h_t = \int_t^\infty y e^{-\int_t^s (r_v+p+\gamma)dv} ds. \tag{22.64}$$

Finally, differentiating with respect to time t we arrive at

$$\dot{h}_t = \left(r_t + p + \gamma\right) h_t - y, \tag{22.65}$$

which is the equation of motion for human capital in the text.

22.6 | Appendix 3

One can show the individual Euler equation at time t for an agent born at date s is

$$\dot{c}_{t,\tau} = c_{t,\tau} \left(r_t - \rho\right). \tag{22.66}$$

The present-value budget constraint of this agent is

$$\int_t^\infty c_{s,\tau} e^{-\int_t^s (r_v+p)dv} ds = b_{t,\tau} + h_{t,\tau}. \tag{22.67}$$

Using the Euler equation here, we have

$$c_{t,\tau}\int_t^{\infty} e^{\int_t^s (r_v-\rho)dv} e^{-\int_t^s (r_v+p)dv} ds = b_{t,\tau} + h_{t,\tau}, \tag{22.68}$$

$$c_{t,\tau}\int_t^{\infty} e^{-(p+\rho)(s-t)} ds = b_{t,\tau} + h_{t,\tau}, \tag{22.69}$$

$$c_{t,\tau} = (p+\rho)\left(b_{t,\tau} + h_{t,\tau}\right). \tag{22.70}$$

Next derive the per capita consumption function, given by

$$c_t = \int_{-\infty}^{t} c_{t,\tau} n e^{-n(t-\tau)} d\tau. \tag{22.71}$$

Using (22.70) this becomes

$$c_t = (p+\rho)\int_{-\infty}^{t} \left(b_{t,\tau} + h_{t,\tau}\right) n e^{-n(t-\tau)} d\tau, \tag{22.72}$$

$$c_t = (p+\rho)\int_{-\infty}^{t} b_{t,\tau} n e^{-n(t-\tau)} d\tau + (p+\rho)\int_{-\infty}^{t} h_{t,\tau} n e^{-n(t-\tau)} d\tau, \tag{22.73}$$

$$c_t = (p+\rho)\left(b_t + h_t\right), \tag{22.74}$$

which is the per capita (or average) consumption function.

Notes

[1] Economists such as Ken Rogoff have been advocating the phasing out of cash, see Rogoff (2016). Recently, India implemented a drastic reduction of cash availability, which is analysed in Chodorow-Reich et al. (2020).

[2] This section follows mostly the work of John Cochrane as presented in his book Cochrane (2021), to which we refer for those interested in further exploration of this topic.

[3] If you find this statement confusing, remember that to find the characteristic equation you need to find the determinant of this matrix. Expanding by the last column means that the equation is $1-\lambda$ times the determinant of the upper left quadrant. This quickly indicates that 1 is one of the eigenvalues.

[4] Actually, the opportunity cost of an asset with similar risk characteristics.

[5] The assumption of a constant coupon is done for simplification but in no way necessary.

[6] Another interesting point is that if the bubble bursts, its value goes to zero: there is no gradual undoing of a bubble. Thus, the pattern of the unwinding of an asset price will tell you a lot about whether the previous surge was or was not a bubble or just a change in the perception on fundamentals.

[7] In the framework of (22.28) this would be a case in which the interest rate is zero.

[8] For a crystal-clear explanation of this see Weil (2008), on which this introduction draws.

[9] This is a slightly simplified version of the same model we presented in Chapter 8 and here follows Acemoglu (2009). Rather than having you go back to that model we have moved some computations to three appendices of this chapter for ease of reference.

[10] On the same topic, see also Reis (2020).

[11] Caballero and Krishnamurthy (2006) explore this tension in the context of bubbly capital flows to emerging markets.

References

Acemoglu, D. (2009). *Introduction to modern economic growth*. Princeton University Press.

Blanchard, O. (2019). Public debt and low interest rates. *American Economic Review*, *109*(4), 1197–1229.

Blanchard, O. J. (1985). Debt, deficits, and finite horizons. *Journal of Political Economy*, *93*(2), 223–247.

Caballero, R. J., Farhi, E., & Gourinchas, P.-O. (2008). An equilibrium model of "global imbalances" and low interest rates. *American Economic Review*, *98*(1), 358–93.

Caballero, R. J. & Krishnamurthy, A. (2006). Bubbles and capital flow volatility: Causes and risk management. *Journal of Monetary Economics*, *53*(1), 35–53.

Chodorow-Reich, G., Gopinath, G., Mishra, P., & Narayanan, A. (2020). Cash and the economy: Evidence from India's demonetization. *The Quarterly Journal of Economics*, *135*(1), 57–103.

Cochrane, J. H. (2021). The fiscal theory of the price level. *Manuscript. URL* https://www.johnhcochrane.com/research-all/the-fiscal-theory-of-the-price-level-1.

Farhi, E. & Tirole, J. (2012). Bubbly liquidity. *The Review of Economic Studies*, *79*(2), 678–706.

Galí, J. (2020). *Monetary policy and bubbles in a New Keynesian model with overlapping generations*. National Bureau of Economic Research.

Martin, A. & Ventura, J. (2018). The macroeconomics of rational bubbles: A user's guide. *Annual Review of Economics*, *10*, 505–539.

Miao, J. (2014). Introduction to economic theory of bubbles. *Journal of Mathematical Economics*, *53*, 130–136.

Reis, R. (2020). The constraint on public debt when r<g but g<m.

Rogoff, K. S. (2016). *The curse of cash*. Princeton University Press.

Samuelson, P. A. (1958). An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, *66*(6), 467–482.

Sims, C. A. (2016). Fiscal policy, monetary policy and central bank independence. https://bit.ly/3BmMJYE.

Tirole, J. (1985). Asset bubbles and overlapping generations. *Econometrica*, 1499–1528.

Weil, P. (2008). Overlapping generations: The first jubilee. *Journal of Economic Perspectives*, *22*(4), 115–34.