Optimal Monetary Policy with Learning by Doing

Chris Redl

ERSA working paper 490

January 2015
Abstract

I study the implications of learning by doing in production for optimal monetary policy using a basic New Keynesian model. Learning-by-doing is modeled as a stock of skills that accumulates based on past employment. The presence of this learning-by-doing externality breaks the 'divine coincidence' result, that by stabilising inflation the output gap will automatically be closed, for a variety of shocks that are important in explaining the business cycle. In this context, the policy maker must consider the impact on future productivity of any trade-off between output and inflation today. The appropriate inflation-output trade-off is between inflation today and the present value of deviations in the output gap. The approach to optimal monetary policy follows Woodford (2010) permitting a study of variations in key parameters and steady states which is uncommon in the literature that relies on a quadratic approximation to the utility function. Exploiting this variation I find that learning induces a small increase in the importance of the output gap under a cost-push shock for the (more realistic case) of a distorted steady state. The welfare costs of business cycles are shown to be significantly larger even under the optimal policy.

JEL classification: E52, J24, E31

Keywords: Monetary Policy, Labor Productivity, Inflation

Queen Mary, University of London. Email address: c.redl@qmul.ac.uk. I would like to thank Xavier Mateos-Planas, Giulio Fella, Winfried Koeniger, Renato Faccini, Kevin Sheedy, Jordi Gali, Martin Ellison and an anonymous referee for helpful comments and suggestions on this paper. Any errors are my own.


1 Introduction

What is the role of output in monetary policy? The Basic New Keynesian model (BNKM) considers a trade-off between inflation and the output gap. However, for a variety of shocks that are important in explaining movements in output at business cycle frequency, that is labour supply, technology and demand or preference shocks (Smets and Wouters (2003); Adolfson et al. (2007)); the BNKM does not indicate a policy trade-off between inflation and output. Rather optimal monetary policy is strict inflation targeting. This is because eliminating inflation entails that output behaves as it does in a flexible price world and absent any real rigidities or market imperfections the flexible price level of output is the welfare relevant level output used to construct the output gap. Thus zero inflation entails zero output gap. The ability to achieve zero inflation in turn depends on how marginal costs move with output and these shocks. Firms’ marginal costs are determined within-period by the output level relative to technology, the preferences of households to consume and supply labour. Having output adjust to offset these changes entails that marginal costs do not deviate from steady state. Since inflation is nothing but the present value of expected deviations in marginal costs in these models inflation is zero in every period. In a flexible price world output responds in exactly this way. It is as if by “divine coincidence” (Blanchard and Gali (2005)) that by only caring about inflation the policy maker can avoid welfare losses emanating from changes in the real economy. This contrasts sharply with how policy makers see central bank objectives especially since the financial crisis of 2007 (Dell’Ariccia et al. (2013)). As Blanchard and Gali (2005) note, in order to undo this result some non-trivial real imperfection is needed. Medium scale models used in policy analysis, e.g. (Smets and Wouters (2003)), avoid this problem by including real imperfections such as nominal wage rigidities (Erceg et al. (2000)) and/or consumption habits (Leith et al. (2012)) to augment the BNKM. Here I introduce an alternative imperfection, a learning-by-doing (LBD) externality into an otherwise standard New Keynesian model of Woodford (2010) which makes these shocks non-trivial for monetary policy.

LBD is modelled following Chang et al. (2002) whereby the productivity of workers depends on both an exogenous technology as well as an endogenous stock of skills of the workforce. These skills depend on the past levels of employment capturing the notion that workers skills may depreciate during periods of low employment. For tractability, these skills are a property of a representative household that makes a labour supply decision and thus skills do not vary across workers or firms.

The LBD externality breaks the divine coincidence through two channels. Firstly, a marginal cost channel whereby lower output and employment today entail lower skill levels in the future.
With lower skill levels, worker productivity falls raising marginal costs and inflation. If the policy maker wants to neutralise the effect of a shock on marginal costs today by engineering a drop in output that fully absorbs the impact then she must accept that future marginal costs will rise. Since inflation is nothing but expected discounted marginal costs this induces higher inflation today. In this way LBD induces a trade-off between inflation and output today and so breaks the divine coincidence 1. In the BNKM marginal costs depend only on current output implying that letting output fully absorb the effects of a shock in order to neutralise deviations in marginal costs does not have additional inflationary effects through higher future marginal costs. The second channel is the direct impact of skills on the utility of the household. Higher levels of skills mean that households need to work fewer hours to produce a unit of output. Lower output today means households have to work more in the future due to lower productivity to produce a given unit of output. The policy maker that reduces output to stabilise marginal costs today must consider that the household will have to work harder in the future creating an additional cost to reducing output in order to stabilise inflation. The optimal policy considers the impact of both these channels when deciding the appropriate inflation-output trade-off today.

I approach the monetary policy problem as a Ramsey problem as in Woodford (2010) with lump-sum taxation. This approach avoids the necessity of deriving a quadratic approximation to representative agents utility function which, although insightful, restricts the feasibility of studying parameter variation, for example the elasticity of intertemporal substitution in an open economy environment (Gali and Monacelli (2005); Wren-Lewis and Leith (2007)) and requires tedious derivations of the model equations to second order to study the behaviour of policy away from the efficient steady state2 (Benigno and Woodford (2005)). For this reason optimal policy in the case of the distorted steady state is understudied.

Using this framework I study 4 shocks in both the efficient and distorted steady state: three shocks where the divine coincidence holds in the Basic New Keynesian Model (BNKM), labour supply, technology and demand or preference shock; and a cost-push shock which creates an inflation-output trade-off even in the BNKM.

If the optimal allocation is the flexible price allocation then policy maker should target an inflation rate as close as possible to zero and the welfare relevant output gap will be constructed

---

1An important exception is a technology shock when $\sigma = 1$ as discussed in section 4.

2I found this approach infeasible in the context of LBD. In particular it was not possible to replace all linear terms from the objective function via second order approximations to the model equations due to the presence of the law of motion for the stock of endogenous skills. The efficient steady state is one where a production subsidy is assumed that offsets any reduction in steady state output due to imperfect competition or externalities (such as LBD). Analysis of the inefficient steady state below assumes a production subsidy where only the reduction in output due to imperfect competition is offset, see section 3.2.1 below.
as output less output under flexible prices. The introduction of the LBD externality means that the flexible price allocation is no longer optimal. I show that the LBD externality implies an additional forward looking component to policy decisions, considering the impact tomorrow of today’s choices on output (beyond inflation expectations). This is summarised in an inflation-output trade-off where the costs of inflation today are weighted against the present value of future output gaps. This entails a smoothing motive for the policy maker which depends on how households value current versus future consumption (the elasticity of intertemporal substitution).

I show that LBD introduces a greater role for stabilising the output gap under cost-push shocks but only in the distorted steady state. This is due to an interaction between time-consistent policy choices and the size of distortions in the steady state. A time-consistent policy maker must place a positive weight on the marginal cost and revenue conditions that firms face in this distorted steady state but the size of these distortions are significantly larger in the presence of LBD than without.

Diminishing returns to labour induce lower levels of steady state output. Learning amplifies this steady state effect by further reducing the productivity of workers when output is low. As shown in section 3.2.1, these effects are large relative to the case without learning. The larger is the gap between efficient and steady state output the greater the weight the policy maker must give to the impact of changes in output on marginal costs of firms operating in the inefficient steady state. These effects are not present in the efficient steady state since by assumption the appropriate production subsidy ensures firms produce the maximal level of steady state output. The welfare costs of shocks increase by 2-8 times the levels seen in the BNKM depending on the shock. This is due to additional costs from a non-trivial policy trade-off and, mechanically, from the endogenous propagation of shocks through LBD.

This paper is related to the literature on optimal monetary policy addressing the divine coincidence, studies of monetary policy decisions in the distorted steady state and attempts to merge endogenous growth theories with business cycles. Blanchard and Gali (2005) originally noted the problem of the “divine coincidence” and introduced rigid real wages as a means of creating a wedge between the flexible price level of output and the efficient level of output. A similar approach that applies the Calvo (1983) mechanism, developed by Erceg et al. (2000), to induce wage rigidity has become widespread in medium scale macro economic models such as Smets and Wouters (2003). Consumption habits are another addition to the core New Keynesian model popular in medium scale models that introduces a real imperfection capable of breaking the divine coincidence (Leith et al. (2012)). The present study is similar in spirit to Leith et al. (2012), however I study the implications of alternative dynamic externality on optimal policy.
The analysis of the optimal policy in New Keynesian models rarely includes a study of the case of the distorted steady state. Either these distortions are assumed to be small enough so that they would not materially alter policy makers response to shocks (Gali (2008); Woodford (2010)) or a production subsidy capable of supporting the efficient level of steady state output is assumed. Studies that have attempted to analyse monetary policy with steady state distortions have yielded important insights: The classic inflationary-bias result of Barro and Gordon (1983) is such a study and DePaoli (2009), exploring parameter variation in an open economy New Keynesian model with a distorted steady state, has found the conditions under which exchange rate stability should be part of the goals of monetary policy. Similarly Benigno and Benigno (2008) have characterised under which shocks and steady states there are gains from cooperation in monetary policy between countries. In addition to the appeal that production subsidies are relatively rare in practice there is also the additional concern that a real imperfection can lead to large steady state distortions requiring implausibly high subsidies. Here I will contrasts the results from both the efficient and distorted steady state highlighting why and when these matter.

Learning-by-doing externalities have been studied extensively in the growth literature (e.g. Grossman and Helpman (1993)) however the cyclical implications are less well understood. The studies that have are largely positive exercises in matching empirical regularities in the business cycle data. Chang et al. (2002) developed the reduced form LBD mechanism employed here where they showed that this mechanism delivers improved persistence in a real business cycle model based on Bayesian estimates of this learning mechanism using U.S. PSID data for 1971-1992. Tsuruga (2007) employed a similar mechanism to capture the hump-shaped response of output to a monetary policy shock. LBD has been fruitfully employed in an open economy setting. Johri and Lahiri (2008) have shown that learning at the firm level can help produce the persistence of real exchange rates found in the data. Benigno and Fornaro (2012) have shown that learning by doing generated by importing intermediate goods can help to explain the significant reserve accumulation seen in developing Asian countries in the last decade. The latter authors also analyse the normative implications of their learning mechanism showing that such reserve accumulation is optimal from the prespective of the accumulating country. In addition there is evidence that skills depreciation is an important cost of low levels of employment (Altug and Miller (1998); Sparber (2011); Hansen and Imrohoroglu (2009)). The reduced form LBD

---

3For context these are 10% in the BNKM when the elasticity of substitution between varieties of goods is 11 rising to 29% with moderate learning and 46% when learning is strong (these qualitative descriptions of the strength of learning are made concrete in Section 4 below). Clearly the imperfection of LBD introduces makes the supposition of production subsidies of this scale more implausible and helps to motivate a check of the results in the case where the subsidy is zero.
mechanism adopted here provides a tractable framework to study whether these costs alter the inflation-output trade-off of the policy maker.

The remainder of the paper is organised as follows. Section 2 presents a New Keynesian model with learning-by-doing and highlights the channels through which learning affects firms costs and households utility. Section 3 discussed the Ramsey optimal policy problem and presents the implications of learning for the steady state output and the welfare-relevant output gap. Section 4 discusses the quantitative results of the optimal monetary policy. Section 5 presents the welfare costs under the optimal policy. Section 6 presents results on the optimal inflation-output trade-off when the policy maker follows a Taylor rule. Section 7 concludes.
2 Model

The model presented here follows the canonical New Keynesian model of Woodford (2010) where learning-by-doing in production in the spirit of Chang et al. (2002) is introduced.

2.1 Households

The economy is cashless (Woodford (2003)) and populated by identical infinitely-lived households who choose their consumption, labour supply and holdings of nominal bonds to solve:

$$\max_{\{C_t,H_t,B_t\}_{t=t_0}} U_t = E_0 \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ u(C_t; \xi^C_t) - v(H_t; \xi^H_t) \right\}$$

$$= E_0 \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ \frac{C_t^{1-\frac{1}{\sigma}} \left( \xi^C_t \right)^{\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \frac{H_t^{1+\frac{1}{\psi}} \left( \xi^H_t \right)^{\frac{1}{\psi}}}{1+\frac{1}{\psi}} \right\}$$

s.t. $P_t C_t + Q_{t,t+1} B_t \leq B_{t-1} + W_t H_t + \Upsilon_t + T_t$

$C_t$ is an index of aggregate consumption, $H_t$ is hours of labour supplied, $W_t$ is the nominal wage, $T_t$ are net government transfers, $\Upsilon_t$ are profits from firms; and $\xi^C_t$ and $\xi^H_t$ are shocks to preferences for consumption and labour supply respectively. Households have access to complete asset markets where they can trade one-period bonds, $B_t$, at a price $Q_{t,t+1}$. Aggregate consumption, $C_t$, and the price level, $P_t$, are defined with the Dixit-Stiglitz aggregators over individual consumption good varieties, $C_t(i)$, and their prices, $P_t(i)$:

$$C_t = \left[ \int_0^1 C_t(i)^{\frac{\epsilon-1}{\epsilon}} \, di \right]^{\frac{1}{\epsilon-1}}$$

$$P_t = \left[ \int_0^1 P_t(i)^{1-\epsilon} \, di \right]^{\frac{1}{1-\epsilon}}$$

(2)

Where $\epsilon$ is the elasticity of substitution between varieties of goods. The solution to the households problem, (1), entails the following intra-temporal labour supply condition and bond price:
\[
\frac{W_t}{P_t} = \frac{v_h}{u_c} = \left( \frac{C_t}{\xi_t} \right)^{\frac{1}{2}} \left( \frac{H_t}{\xi_t} \right)^{\frac{1}{2}}
\]

\[
Q_{t+1} = \beta \frac{u_C(C_{t+1}, \xi_{t+1})}{u_C(C_t, \xi_t)} \frac{P_t}{P_{t+1}} = \beta \left( \frac{Y_t \xi_t}{Y_{t+1} \xi_{t+1}} \right)^{\frac{1}{2}} \frac{P_t}{P_{t+1}}
\]

Where the last equality follows from imposing the market clearing which requires that \( Y_t = C_t \).

### 2.2 Firms

#### 2.2.1 Production

There is a continuum of monopolistically competitive firms where each variety of good, indexed by \( i \in [0, 1] \), is supplied by a single producer. The \( i^{th} \) firms buys labour hours, \( N(i) \), from households on a competitive labour market. The productivity of workers depends on the aggregate level of exogenous technology, \( A_t \), as well as the level of endogenous aggregate worker skills, \( X_t \), that alters the effective unit of labour supplied by households. Firms face a diminishing returns to production, governed by \( \alpha \), in employing this effective unit of labour:

\[
Y_t(i) = A_t(X_tN_t(i))^{1-\alpha}
\]

The aggregate stock of workers’ skills \( X_t \) evolves depending on past levels of employment, a form of learning-by-doing, as in Chang et al. (2002)\(^4\):

\[
X_t = X_{t-1}^{\phi}N_{t-1}^{\mu}
\]

I will follow Chang et al. (2002) in that this learning is external to firms, that is, they do not internalise the effects of employment on productivity of their workers\(^5\). In the current context this can be motivated by the fact that worker productivity depends on the economy-wide level of skills and each producer’s employment decision, \( N_t(i) \), contributes only infinitesimally to the aggregate stock of skills, \( X_t \).

\(^4\)This nests the BNKM analysed in Woodford (2010) when \( \mu \to 0 \)

\(^5\)If a single producer or a few large producers supplied output and internalised the impact of employment levels today on their future marginal costs then the labour demand would no longer be static as here, but would instead include a dynamic term similar to the second term on the RHS of the planners first order condition (31) given below in section 3.1. This is not considered since it would mean the externality poses no policy problem for the monetary policy maker to counteract.
Each producer faces a downward sloping demand curve for their variety of goods based on the Dixit-Stiglitz preferences described above:

\[ Y_t(i) = \left( \frac{P_t(i)}{P_t} \right)^{-\epsilon} Y_t \quad (7) \]

\( Y_t \) is the aggregate demand for the consumption basket \( C_t \) defined in (2). Producers are subject to Calvo (1983) price rigidities which implies that \( P_t(i) \) need not equal the aggregate price level \( P_t \) as only a subset of firms are able to reset prices each period giving rise to a measure of cross-sectional price dispersion:

\[ \Delta_t \equiv \frac{1}{\hat{\epsilon}} \int_0^1 \left( \frac{P_t(i)}{P_t} \right)^{-\epsilon(1+\eta)} \, di \geq 1 \quad (8) \]

Where \( \eta = \frac{\alpha}{1-\alpha} \). Using (5) and (7), this now allows us to relate the level of skills to the past levels of output and the dispersion of prices:

\[ N_t(i) = \left( \frac{Y_t(i)}{A_t} \right)^{1+\eta} \frac{1}{X_t} \quad (9) \]

\[ N_t \equiv \int_0^1 N_t(i) \, di = \left( \frac{1}{A_t} \right)^{1+\eta} \frac{1}{X_t} \int_0^1 Y_t(i)^{(1+\eta)} \, di = \left( \frac{Y_t}{A_t} \right)^{1+\eta} \frac{\Delta_t}{X_t} \quad (10) \]

(10) indicates that for a given supply of output, \( Y_t \), improved technology or skills requires fewer workers whereas greater price dispersion acts to reduce labour productivity. Why? Consumers want to consume an equal amount of each variety of goods produced by the different firms (since Dixit-Stiglitz preferences entail that each these differentiated goods have an equal weight in the consumption basket). Thus the demand for aggregate output falls when this output exhibits a greater price dispersion across varieties of goods. Even if technology and skills are unchanged, dispersed output levels will aggregate up to a lower \( Y_t^6 \).

Thus (lagged) price dispersion in affecting the level of aggregate output, affects the demand for labour and thus employment altering the evolution of skills:

\[ X_t = X_{t-1}^{\phi_x-\mu} \left( \frac{Y_{t-1}}{A_{t-1}} \right)^{\mu(1+\eta)} \Delta_{t-1}^{\mu} \quad (11) \]

\(^6\)This is not simply a feature of Calvo pricing. Rotemberg pricing implies the same so long as the price level today is not identical to the price level last period, that is, it will matter in the presence of shocks. There is no effect of price dispersion or price adjustment costs on labour productivity in the steady state.
Here we can see that higher output raises skill levels but so does price dispersion. This is due to the requirement for more labour to produce a given level of output when prices are dispersed.

2.2.2 Price setting

Producers are subject to Calvo (1983) price rigidities whereby they face a fixed probability, $\omega$, of being able to reset their price each period. Thus each firm takes into account that the price chosen today, $t$, has a probability of survival of $\omega^{T-t}$, after $T$ periods have passed. Thus a firm able to reset their price at time $t$ will solve the following problem:

$$\max_{\{P_t(i)\}_{t=t_0}} E_t \sum_{T=t}^{\infty} \omega^{T-t} Q_{t,T} \Pi(P_t(i), P_T, Y_T, X_T; \xi_T)$$ (12)

$\xi_T$ refers to the entire collection of shocks that affect firms pricing decision, $\xi_T' = [A_T \; \xi_T^C \; \xi_T^H \; \mu_P^T]$. $\mu_P^T$ refers to a shock to firms desired steady state mark-up, $\epsilon_{-1}$ and $A_T$ refers to the level of exogenous technology. $Q_{t,T}$ is the value placed on nominal profits returned to the household $T$ periods hence (see equation (4)). For the $i^{th}$ firm nominal profits in period $T$ are simply nominal revenues less costs:

$$(1 - \tau)P_t(i) \left( \frac{P_t(i)}{P_T} \right)^{-\epsilon} Y_T - W_T N_T(i)$$ (13)

Nominal revenue is $(1 - \tau)P_t(i) Y_t(i)$ where I have used (7) and $\tau$ is a production tax or subsidy levied by the government. In order to see how (13) can be written as a function of $P_t(i), P_T, Y_T, X_T, \xi_T$ only, as in (12), I make use of (3) and (10) to decompose the nominal cost term $W_T N_T(i)$. Applying the intratemporal optimality condition of households, (3), nominal wages must equal the marginal rate of substitution:

$$W_T = P_T \left( \frac{C_T}{\xi_T^C} \right)^{\frac{1}{\sigma}} \left( \frac{N_T}{\xi_T^H} \right)^{\frac{1}{\psi}}$$ (14)

Market clearing in this closed economy requires that $Y_t = C_t$ and $\int_0^1 N_t(i) di = N_t = H_t$. Using the latter and (3) and (10), the nominal wage becomes:

$$W_T = P_T \left( \frac{Y_T}{\xi_T^C} \right)^{\frac{1}{\sigma}} \left( \frac{1}{\xi_T^H} \right)^{\frac{1}{\psi}} \left( \frac{Y_T}{A_T} \right)^{\frac{1 + \eta}{\psi}} \left( \frac{\Delta_T}{X_T} \right)^{\frac{1}{\psi}}$$ (15)
The required number of employees, $N_T(i)$:

$$N_T(i) = \left( \frac{Y_T(i)}{A_T} \right)^{1+\eta} \frac{1}{X_T} = \left( \frac{Y_T}{A_T} \right)^{1+\eta} \frac{1}{X_T} \left( \frac{P_t(i)}{P_T} \right)^{-\epsilon(1+\eta)} \quad (16)$$

Thus $Q_{t,T}(P_t(i), P_T, Y_T, X_T; \xi_T)$ in (12) can be written as a function of $P_t(i), P_T, Y_T, X_T, \xi_T$ only. The first order conditions for profit maximisation is:

$$E_t \sum_{T=t}^\infty \omega^{T-t} Q_{t,T} \Pi_t(P_t(i), P_T; Y_T, \xi_T) = 0 \quad (17)$$

All firms able to reset their price will make the same choice (as they are identical) thus $P_t(i) = P_t^*$. Given the assumptions made above a convenient closed form relationship characterising aggregate supply in the economy can be derived (see appendix):

$$\left( \frac{P_t^*}{P_t} \right)^{\frac{1}{1+\eta}} \quad (18)$$

$F_t$ captures the expected future nominal (marginal) revenue and $K_t$ captures expected future nominal (marginal) costs. These functions are the key forward looking variables in the model that lead to the New Keynesian Phillips curve. They are defined by:

$$F_t = E_t \sum_{T=t}^\infty (\omega \beta)^{T-t} f(Y_T; \xi_T^C) \left( \frac{P_T}{P_t} \right)^{\epsilon-1} \quad (19)$$

$$K_t = E_t \sum_{T=t}^\infty (\omega \beta)^{T-t} k(Y_T, X_T, \Delta_T; \xi_T) \left( \frac{P_T}{P_t} \right)^{\epsilon(1+\eta)} \quad (20)$$

$$f(Y_T; \xi_T) = (1 - \tau)Y_T^{1-\frac{1}{\sigma}} \left( \xi_T^C \right)^{\frac{1}{\sigma}} \quad (21)$$

$$k(Y_T, X_T, \Delta_T; \xi_T) = \mu_t^P (1 + \eta) \left( \frac{Y_T}{A_T} \right)^{1+\chi} \left( \frac{1}{X_T} \right)^{1+\psi} \left( \frac{\Delta_T}{\xi_T^H} \right)^{\frac{1}{\psi}} \quad (22)$$

Exogenous variations in $\mu_t^P$ will be studied as cost-push shocks below and $\chi \equiv (1 + \frac{1}{\psi})(1 + \eta) - 1$. From (22) we can see that a higher skills will induce lower marginal costs for firms, that is higher levels of worker skills ($X_T$) raises their marginal product which intern requires firms to hire fewer workers at the given wage, reducing marginal costs. This formulation is very convienient in that (19) and (20) can be written recursively, where $\Pi_t = P_t/P_{t-1}$:
\[ F_t = f(Y_t; \xi_t) + \omega \beta E_t \Pi_{t+1}^{-1} F_{t+1} \]  

(23)

\[ K_t = k(Y_t, X_t, \Delta_t; \xi_t) + \omega \beta E_t \Pi_{t+1}^{(1+\eta)} K_{t+1} \]  

(24)

The Calvo scheme entails that the price index evolves according to:

\[ P_t^{1-\epsilon} = (1 - \omega) (P^*_t)^{1-\epsilon} + \omega P_{t-1}^{1-\epsilon} \]  

(25)

Which can be used in conjunction with (18) to yield an equation governing the behaviour of inflation each period, analogous to an aggregate supply relation:

\[ \frac{1 - \omega \Pi_t^{-1}}{1 - \omega} = \left( \frac{F_t}{K_t} \right)^{\frac{\epsilon}{1+\epsilon}} \]  

(26)

As noted by Woodford (2010) this description is equivalent to the New Keynesian Phillips Curve (NKPC) if one log-linearises (19),(20) and (26):

\[ \pi_t = \kappa \left[ \hat{Y}_t - \left( \frac{1 + \psi^{-1}}{\chi + \sigma^{-1}} \right) \hat{X}_t + \left( \frac{\psi^{-1}}{\chi + \sigma^{-1}} \right) \hat{\Delta}_t + u'_\xi \hat{\xi}_t \right] + \beta E_t \pi_{t+1} + O(||\xi^2||) \]  

(27)

\[ \kappa = \frac{(1 - \omega)(1 - \omega \beta)(\chi + \sigma^{-1})}{\omega(1 + \epsilon \eta)}; \quad u'_\xi = (\chi + \frac{1}{\sigma})^{-1} \begin{bmatrix} -1 - \sigma^{-1} - \psi^{-1} \end{bmatrix} \]

\[ \xi'_t = \begin{bmatrix} A_t & \xi'_C & \xi'_H & \mu'_P \end{bmatrix} \]

Where for each variable \( \hat{Z}_t \equiv \ln Z_t - \ln \bar{Z} \) and \( \bar{Z}_t = Z_t - \bar{Z} \) and \( \pi_t \equiv \bar{\Pi}_t \). Since the log-linearised version of the law of motion for price dispersion, see equation (28) in section 2.3, is the deterministic equation \( \hat{\Delta}_t = \omega \hat{\Delta}_{t-1} \) if \( \hat{\Delta}_{t-1} = O(||\xi^2||) \) then \( \hat{\Delta}_t = O(||\xi^2||) \forall t \). Thus terms in \( \hat{\Delta}_t \) can be ignored to a first order. If additionally \( \hat{X}_t = 0 \) then this is simply the standard NKPC. The channel from higher skills to lower marginal costs and thus inflation is evident from (27). A higher stock of skills reduces firms marginal costs and thus reduces inflation.

Households optimal holding of one period bonds, as described by (4), is linked to the choice of the short term nominal interest rate set by the policy maker via the no arbitrage relation on bonds, \( 1 + i_t = [E_t Q_t]^{-1} \). The combination of these conditions yields the New Keynesian IS curve (when log-linearised). The final component required to close the model is a statement of
how the interest rate will be chosen, this is implied by the Ramsey planners choice of allocations described below in section 3.2.

2.3 Price dispersion, skills and welfare

A law of motion linking cross-sectional price dispersion, $\Delta_t$, to aggregate inflation, $\Pi_t$, can be derived from (8) and (25):

$$\Delta_t = \omega \Delta_{t-1} \Pi_t^{(1+n)} + (1 - \omega) \left( \frac{1 - \omega \Pi_t^{(1+n)}}{1 - \omega} \right)$$  \hspace{1cm} (28)

This link is key to explaining the welfare implications of inflation in New Keynesian models. The impact of skills and price dispersion on welfare can now be seen by imposing market clearing and substitution of (10) into the utility function of the household:

$$U_{t_0} = E_0 \sum_{t=t_0}^{\infty} \beta^{t-t_0} \{ U(Y_t, X_t, \Delta_t; \xi_t) \}$$

$$U_{t_0} = E_0 \sum_{t=t_0}^{\infty} \beta^{t-t_0} \{ u(Y_t; \xi^C_t) - v(Y_t, X_t, \Delta_t; \xi_t) \}$$

$$U_{t_0} = E_0 \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ \frac{Y_t^{1-\frac{1}{\sigma}} \left( \xi^C_t \right)^{\frac{1}{\sigma}}} {1 - \frac{1}{\sigma}} - \left( \frac{\xi^H_t}{Y_t} \right)^{-\frac{1}{\psi}} \left( \frac{Y_t}{A_t} \right)^{(1+n)(1+\frac{1}{\psi})} \left( \frac{\Delta_t}{X_t} \right)^{(1+\frac{1}{\psi})} \right\} \hspace{1cm} (29)$$

From (29) we can see that the impact of both skills and price dispersion on period utility is equal but opposite in the sense that $U_{X_t, X_t} = -U_{\Delta_t, \Delta_t}$. The intuition behind this result is that greater price dispersion implies greater output dispersion which leads to a composition of output that is less valued by households (see discussion in section 2.2.1). To produce a given unit output, more labour hours are required the higher is price dispersion. This effect reduces utility through the disutility of work. Higher levels of productivity, either exogenous ($A_t$) or endogenous ($X_t$), directly increase the output of workers requiring fewer hours of work to produce a given unit of output. Thus the higher is endogenous productivity (say from high level of activity in the past), the lower is the disutility created from producing a unit of output. The value of this channel in reducing the negative impact of inflation on household welfare is evaluated by solving the Ramsey problem for this economy.
A competitive equilibrium in this economy is a sequence of allocations and prices such that markets clear and household’s utility (1) and firms profits (12) are maximised. This is summarised by \{F_t, K_t, \Delta_t, \Pi_t, Y_t, X_t\}_{t=0}^\infty satisfying the households optimality conditions (3) and (4); the law of motion for skills (11); the definition of the forward looking measures of marginal costs and revenue for firms (23) and (24); the firms first order condition summarised in aggregate supply realtion (26); the law of motion for price dispersion (28) together with a description for the exogenous stochastic processes \( \xi_t \).

3 Optimal monetary policies with learning by doing

3.1 Social planners problem

The social planner maximises the households utility subject to the resource constraints captured in equation (10) and the law of motion for skills (11). However the social planner will never choose to have any price dispersion in this economy since this requires the household to work harder\(^7\). Thus the social planner solves:

\[
\max_{\{Y_t\}_{t=t_0}} U_{t_0} = E_0 \sum_{t=t_0}^\infty \beta^{t-t_0} \left\{ u(Y_t; \xi_C^t) - v(Y_t, X_t; \xi_t) \right\}
\]

s.t. \( X_t = X_{t-1}^{\phi_x - \mu} \left( \frac{Y_{t-1}}{A_{t-1}} \right)^{\mu(1+\eta)} \)

In the standard New Keynesian model the optimal rule requires that the marginal benefit of an additional unit of output is just compensated by the additional disutility of producing that output: \( u_{Y,t} = v_{Y,t} \). In the presence of learning this static rule becomes dynamic with the additional benefit today of working being a lower disutility of work tomorrow:

\[
u_Y(Y_t; \xi_C^t) + \beta E_t \left\{ -v_X(Y_{t+1}, X_{t+1}; \xi_{t+1}) \frac{\partial X_{t+1}(Y_t, X_t, \Delta_t; \xi_t)}{\partial Y_t} \right\} = v_Y(Y_t, X_t; \xi_t)
\]

\(^7\)In the BNKM there are only loses from higher price dispersion. However in this context there may be offsetting gains due to the positive effect of higher skills and thus a lower disutility of labour. In fact, however, this effect is not large enough to compensate for the costs, in terms of disutility of labour, that it attracts. Higher price dispersion can induce higher skills tomorrow (see equation (11)). However, to a first order approximation, any increase in \( \hat{\Delta}_t \) will induce an increase of \( \mu \) on \( \hat{X}_{t+1} \). To a first order, the law of motion for \( \Delta_t \), around the zero inflation steadystate, is \( \Delta_t = \omega \Delta_{t-1} \). Thus this will induce an increase in \( \Delta_{t+1} \) of \( \omega \). Since \( \omega > \mu \) for all reasonable parametrisations, the social planner will never find it optimal to create price dispersion to increase productivity as this is always overwhelmed by the decrease in productivity induced by more price dispersion.
Where \( v_X < 0 \). This optimal rule is used to characterise the efficient level of output, \( Y_t^e \). This level of output will not, in general, be feasible when the policy maker must make her choices subject to competitive equilibrium. Optimal choices subject to competitive equilibrium are Ramsey Policies.

### 3.2 Ramsey policies

The Ramsey policy is a choice of \( \{F_t, K_t, \Delta_t, \Pi_t, Y_t\} \) for all \( t \geq t_0 \) to maximise (29) while satisfying (23),(24), the forward looking relations capturing marginal costs and revenues for firms; (26), the aggregate supply relation relating these forwarding looking variables to inflation; the law of motion for skills (11); and (28), which links inflation to the welfare relevant measure of cross-sectional price dispersion; given \( \Delta_{t_0-1} \) and \( X_{t_0-1} \). \( X_t \) is not an explicit choice variable for the policy maker since the choice of \( Y_{t-1} \) and \( \Delta_{t-1} \) take into account their influence on \( X_t \). This facilitates comparison with the literature where the policy maker is thought of as choosing \( \{\Pi_t, Y_t\} \) only. As currently described this problem is time inconsistent. Due to the forward looking conditions (23)-(24) policy makers know that the choices made today will have an effect on expectations formed in the previous period, since firms will form expectations based on these policies, and this constrains the choices they can make. However this constraint is not binding when \( t = t_0 \) leading to time inconsistent policy choices between the initial point \( t_0 \) and periods thereafter \( t > t_0 \). I adopt the solution proposed by Woodford (2010) which he calls optimal policy from a “timeless perspective”. The policy maker undertakes precommitments, at \( t_0 - 1 \), to certain values of the forward looking variables in the next period, \( K_{t_0} \) and \( F_{t_0} \), that that are consistent with the optimal choices in future periods. These precommitments are captured in the values of the lagrange multipliers (see problem statement below) \( \phi_{1,t_0-1} \) and \( \phi_{2,t_0-1} \) which govern whether the forward looking constraints bind; or in other words, they control the temptation to

---

8Contrasting (31) with (39) illustrates why this level of output may not be feasible: the planner would need to take account of the forward looking behaviour of firms captured in the Lagrange multipliers associated with firms’ forward looking constraints to the Ramsey problem.

9The New Keynesian IS curve will not enter the Ramsey problem as the nominal interest rate does not appear in the utility function of the household nor in any of the constraints to the problem. Thus output can be thought of as chosen directly by the policy maker and the nominal interest rate (the instrument of monetary policy) that this requires backed out of this IS relation.

10As shown in section 2.2.2 on price setting, combining the two forward looking equations for \( F_t \) and \( K_t \) along with the aggregate supply curve (26) and log-linearising produces the standard New Keynesian Phillips Curve (27). In a standard linear quadratic approach to optimal monetary policy, e.g. Woodford (2003); Gali (2008), this would allow the problem to be stated in terms of just 2 variables aggregate inflation, \( \Pi_t \), and aggregate output, \( Y \). Since here I do not pursue the linear quadratic approach I use a full set of equilibrium conditions which includes variables \( \{F_t, K_t, \Delta_t, \Pi_t, Y_t\} \). Where LBD absent this would be identical to the linear quadratic approach typically used, as shown in Woodford (2010).
raise Πt₀ above a level consistent with the forward looking constraints. The precommitment values for Kₜ₀ and Fₜ₀ are the steady state values of Fₜ and Kₜ, denoted by K and F, for the Ramsey problem below. The choices of K and F are a function of steady state output and have been constrained to be consistent with future choices, that is choices made under the same constraints the Ramsey planner faces in future periods (see steady state solution in Section 3.2.1 and Appendix II for details). Thus, the Ramsey plan from a “timeless perspective” requires that the Ramsey planner treat forward looking behaviour in a way that is consistent with the initial conditions (or steady state) of the Ramsey problem.

The Ramsey problem outlined above can be described by the Lagrangian:

$$\max_{\{F_t, K_t, \Delta_t, \Pi_t, Y_t\}_{t=0}^{\infty}} L_t = E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} L(Y_t, X_t, F_t, K_t, \Delta_t; \theta_t, \phi_t, \xi_t)$$

Where

$$L(Y_t, X_t, F_t, K_t, \Pi_t, \Delta_t; \theta_t, \phi_t, \xi_t) = u(Y_t; \xi_t^c) - v(Y_t, X_t, \Delta_t; \xi_t) +$$

$$\theta_t \left[ \Delta_t - \omega \Delta_t \Pi_t^{(1+\eta)} - (1 - \omega) \left( \frac{1-\omega \Pi_t^{1-1}}{1-\omega} \right)^{\frac{1+\eta}{1-\eta}} \right]$$

$$+ \phi_{1,t} \left[ F_t - f(Y_t; \xi_t) \right] - \omega \phi_{1,t-1} \left[ \Pi_t^{1-1} F_t \right] + \phi_{2,t} \left[ K_t - k(Y_t, X_t, \Delta_t; \xi_t) \right] + \omega \phi_{2,t-1} \left[ \Pi_t^{(1+\eta)} K_t \right]$$

$$+ \phi_{3,t} \left[ \frac{1-\omega \Pi_t^{1-1}}{1-\omega} - \left( \frac{F_t}{K_t} \right)^{(1+\eta)} \right]$$

Where $\phi_t' = [\phi_{1,t} \phi_{2,t} \phi_{3,t}].$ The multipliers $\phi_{1,t-1}, \phi_{2,t-1}$ will capture the precommitments i.e. they will be the values consistent with the steady state solution of the model under the same constraints the Ramsey planner faces for $t > t_0.$ I consider the local dynamics near the zero inflation steady state of the model. The complete first-order conditions (FOCs) are described in the appendix. Here I will focus on the optimal rule for output and price dispersion since these rules involve additional terms due to LBD. The optimal rule for the the choice of output is:

$$u_Y(Y_t; \xi_t^c) \left( 1 - \phi_{1,t}(1 - \tau)(1 - \sigma^{-1}) \right) +$$

$$\beta E_t \left\{ -v_X(Y_{t+1}, X_{t+1}, \Delta_{t+1}; \xi_{t+1}) - \phi_{2,t+1}k_X(Y_{t+1}, X_{t+1}, \Delta_{t+1}; \xi_{t+1}) \right\} \frac{\partial X_{t+1}(Y_t, X_t, \Delta_t; \xi_t)}{\partial Y_t} =$$
\[
(1 + \phi_{2,t} \mu_t P (1 - \chi)(1 + \psi^{-1})\Delta_t^{-1}) v_Y(Y_t, X_t, \Delta_t; \xi_t)
\]

This optimal rule is analogous to that of the social planner, (31), however when choosing \(Y_t\) in addition to considering the reduction in labour disutility tomorrow, \(-v_{X,t+1}\), the Ramsey planner also must consider the reduction in the marginal costs of firms tomorrow, \(-k_{X,t+1}\) as well as additional terms linked to the marignal revenue, \((1 - \tau)(1 - \sigma^{-1})\), and marginal cost, \(\mu_t P (1 - \chi)(1 + \psi^{-1})\Delta_t^{-1}\), of producing output in a decentralised economy which are weighted by the degree to which firms pricing decisions, captured in \(F_t\) and \(K_t\), are a binding constraint on the planner’s choice of \(Y_t\). These weights are \(\phi_{2,t}\) and \(\phi_{1,t}\).

The optimal choice for \(\Delta_t\) is:

\[
th_t + \beta \{-v_X(Y_{t+1}, X_{t+1}, \Delta_{t+1}; \xi_{t+1}) - \phi_{2,t+1} k_X(Y_t, X_t, \Delta_t; \xi_t)\} \frac{\partial X_{t+1}(Y_t, X_t, \Delta_t; \xi_t)}{\partial \Delta_t} = \]

\[
+v_\Delta(Y_t, X_t, \Delta_t; \xi_t) + \phi_{2,t} \Delta(Y_t, X_t, \Delta_t; \xi_t) + \theta_{t+1} \omega \Pi_{t+1}^{(1+\eta)}
\]

(34)

A similar pattern is seen as with output; when deciding how much price dispersion to tolerate the Ramsey planner considers the benefit of increased price dispersion on future skills resulting in a benefit to having greater price dispersion today. However the increase in skills due to greater price dispersion (from a higher labour input requirement for firms) is in general smaller than the increase in skills from greater output since:

\[
D_t \equiv \frac{\partial X_{t+1}/\partial Y_t}{\partial X_{t+1}/\partial \Delta_t} = (1 + \eta) \frac{\Delta_t}{Y_t} \geq 1
\]

This is true since \(D_t \to 1\) when \(Y_t \to 1 + \eta\) and \(\Delta_t \to 1\). However it can be seen from (26) that for output to rise to such a high level above it’s steady state value\(^{11}\), inflation would have to rise which would push \(\Delta_t\) above 1 (28). Thus the key channel through which skills operates is via it’s role on increasing the productivity of households through higher levels of past output.

3.2.1 Steadystate

In order to study the optimal response to shocks I linearise these conditions around the zero inflation steady state consistent with the optimality onditions of the Ramsey problem\(^ {12}\). This

\(^{11}\)For context, steady state output, \(\bar{Y} = 1 + \eta\) would require a production subsidy larger than the output of the entire economy (a subsidy of approximately 158% given the baseline parametrisation).

\(^{12}\)Thus this is not the optimal steady state of the social planner, but only the optimal steady state from the
is the steady state associated with the above 5 FOCs and the 4 constraints. The steady state is characterised by \( \{ \bar{F}, \bar{K}, \bar{\Delta}, \bar{\Pi}, \bar{Y}, \bar{X}, \bar{\theta}, \bar{\phi} \} \) that solves these 9 equations when \( \xi_t = \bar{\xi} \). The zero inflation steady state has \( \bar{\Delta} = \bar{\Pi} = 1 \). This immediately implies that \( \bar{K} = \bar{F} \) from the aggregate supply relation (26). From this result we can find the relationship between steady state output and the production subsidy \( \tau \) as well as the steadystate stock of skills, using (11):

\[
f(\bar{Y}) = k(\bar{Y}, \bar{X}, \bar{\Delta}) \iff \frac{v_h(\bar{Y}, \bar{X}, \bar{\Delta})}{u_c(\bar{Y})} = \frac{1 - \bar{\tau}}{\epsilon} \iff \bar{Y} = \left( \frac{1 - \bar{\tau}}{\epsilon} \right)^{\frac{1}{\gamma(1 - \phi_x + \mu)/(1 + \chi)}} \tag{35}
\]

\[
\bar{X} = \bar{Y} \frac{\mu(1 + \eta)}{1 - \phi_x + \mu} \tag{36}
\]

This result illustrates the role of the production subsidy in determining whether this steady state is the (constrained) efficient or inefficient one. From this we can see that the production subsidy will offset any distortion due to monopolistic competition if \( \tau = 1 - \left( \frac{\epsilon}{\epsilon - 1} \right) \). If this is the case and we further assume that \( \alpha = 0 \), that is, constant returns to scale then \( \bar{Y} = 1 \) regardless of the learning parameters \( \mu \) and \( \phi_x \). However when diminishing returns to scale are present \( (\alpha > 0) \) or if the production subsidy is too small to neutralise the steadystate effects of monopolistic competition then these learning parameters can affect \( \bar{Y} \). To illustrate, suppose the production subsidy neutralises monopolistic competition and \( \alpha = \frac{1}{3} \). The impact of stronger learning (higher \( \mu \)) in this environment is shown in Figure 1 by \( Y^{ce} \).

Thus the effect of stronger learning is to reduce \( Y^{ce} \) even when \( \tau = 1 - \frac{\epsilon}{\epsilon - 1} \). In competitive equilibrium firms don’t internalise the impact on future productivity of lower output today this has the effect when combined with dimishing returns to labour of amplifying the decline in steady state output that these diminishing returns induce\(^\text{13}\). This can be constrasted with the level of efficient output consistent with the social planners solution: equation (31) in steady state requires the following production subsidy for \( \bar{Y}^e = Y^{ce} \):

\[
\tau^e = 1 - \left( \frac{\epsilon}{\epsilon - 1} \right) \left( \frac{1}{1 - \beta \mu} \right) \tag{37}
\]

When this subsidy is in place and the determinanition of output thus includes a forwarding looking element we see that the presence of learning (\( \mu \)) acts to undo the contractionary impact of diminishing returns if learning is strong enough. The trade-off between the strength of

---

\(^{13}\)Without dimishing returns an increase in \( \mu \) induces higher steadystate output
Figure 1: Steady state output and learning

![Graph showing steady state output and learning](image)

Figure plots the value of $\bar{Y}$ for 2 different values of $\tau$. The efficient case of $\tau = \tau^e$ which rises with $\mu$ and the usual subsidy that offsets only monopolistic competition $\tau = 1 - \left(\frac{\epsilon}{\epsilon - 1}\right)$. The values for other parameters are as assumed in Table 1.

Diminishing returns and learning leading to higher output is described in Figure 2. The graph shows the level of $\mu$ whereby an increase in $\mu$ induces higher output. For low levels of $\alpha$ output rises for all increases in $\mu$, however as diminishing returns becomes stronger a higher level of $\mu$ is needed before learning induces an increase in output.

Figure 2: Parameter combinations where LBD dominates diminishing returns to labour

![Graph showing parameter combinations](image)

This combination of $(\mu, \alpha)$ are those where increases in the strength of learning from past employment will raise the efficient level steady state output (shaded region). Said differently, any further increases in $\mu$ will increase $\bar{Y}^e$. For the unshaded region increases in $\mu$ lead to a decline in $\bar{Y}^e$.

Note that if an efficient subsidy is not in place then the size of the distortions in output
in competitive equilibrium relative to the efficient equilibrium are very large with LBD and relatively small for the standard model. To illustrate suppose \( \tau = 0 \). In the standard model this will induce a distortion of output of approximately 4.5% - which can be seen from (35) when the learning parameters are set to 0. However with \( \mu = 0.15 \) and \( \phi_x = 0.8 \) this gap becomes 13% and 20% if \( \mu = 0.25 \).

Using the FOCs it is also possible to derive the values of the lagrange multipliers in (32). The following results will be useful in the section to follow:

\[
\tilde{\phi}_2 = -\tilde{\phi}_1 = \frac{U_Y + \beta U_X \frac{\partial X}{\partial Y}}{k_Y - f_Y + \beta k_X \frac{\partial X}{\partial Y}} \tag{38}
\]

These are the steadystate levels that the Ramsey planner commits to at \( t = t_0 - 1 \) which bind her for future periods. The precommitments of the planner, which constrain her to behave in a consistent way at \( t_0 \) and all other periods, entail that the shadow value of the forward looking constraints (\( \phi_1 \) and \( \phi_2 \)) are relevant to the steady state level of output. Intuitively, both steady state output and the path of output when it evolves according to the choice of a planner acting from a “timeless perspective” reduce the influence of time in a particular way. The connection between the level of the production subsidy, \( \tau \), and the level of these lagrange multipliers on the forward looking constraints (\( F_t \) and \( K_t \)) can be seen by recognising that the numerator in (38) is in fact the steadystate version of equation (31) which states \( U_Y + \beta U_X \frac{\partial X}{\partial Y} = 0 \). Thus when the optimal subsidy is in place the value of these multipliers is zero. When \( \tau \leq \tau^e \) then \( \tilde{\phi}_2 \geq 0 \) and the precommitment made by the policy maker bind for \( t \geq t_0 \). How the presence of a steady state distortion influences the policy makers target level of output and thus response to shocks is discussed next.

### 3.3 The target level of output and the output gap

#### 3.3.1 Basic New Keynesian Model

What is the appropriate level of output for the monetary authority to target? Consider the case without learning. The target level of output ought to be the level of output that is consistent with price stability (the steady state assumption) as well as the constraints on the monetary authority in achieving this. Following Benigno and Woodford (2005) this can be seen to be the first order condition for the Ramsey problem, (33) with the omission of the terms depending on \( X_t \), when prices are flexible:
\[ U_Y(Y_t^*, 1; \xi_t) = \bar{\phi}_1 f_Y(Y_t^*; \xi_t) + \bar{\phi}_2 k_Y(Y_t^*, 1; \xi_t) \quad (39) \]

The lagrange multipliers governing the behaviour of (23) and (24) take on their steady state values as under flexible prices since firms pricing decisions no longer have any dynamic considerations implying \( f(Y_t; \xi_t) = k(Y_t, 1; \xi_t) \forall t \). This equation states the target output is a function of exogenous shocks only\(^{14}\). Intuitively, since the only friction present in the model (nominal rigidities) is removed when prices are flexible, the only driver of the target level of output are shocks hitting the economy. In the case with an additional frictions (learning) discussed below this will not be the case as the target level of output will depend on the optimal evolution of workers’ skills. This allows us to easily relate the efficient and natural (flexible price) levels of output to this target level based on the steady-state production subsidy.

The efficient level of output maximises utility subject only to technology and exogenous shocks hitting the economy, restating (31):

\[ U_Y(Y_t^e, 1, \xi_t) = 0 \quad (40) \]

The natural level of output must be consistent with firms (static) price setting decision when prices are flexible. The aggregate supply relation, (26), entails that \( F_t = K_t \forall t \) which in turn requires:

\[ f(Y_t^n; \xi_t) = k(Y_t^n, 1; \xi_t) \forall t. \quad (41) \]

\[ (1 - \tau)u_Y(Y_t^n; \xi_t) = \left( \frac{\epsilon}{\epsilon - 1} \right) v_Y(Y_t^n, 1; \xi_t) \quad (42) \]

Now the links between \( Y_t^e, Y_t^n \) and \( Y_t^* \) can be clarified. If \( \tau = 1 - \frac{\epsilon}{\epsilon - 1} \) then (42) is just the statement \( U_Y(Y_t^n, 1, \xi_t) = 0 \) which means \( Y_t^e = Y_t^n \), that is, the efficient and natural (or flexible price) level of output are the same since there is no additional friction to drive a wedge between them e.g. a learning externality (see the next section). This then also implies \( \bar{\phi}_1 = -\bar{\phi}_2 = \frac{U_Y}{k_y - f_Y} = 0 \) and thus \( Y_t^e = Y_t^n = Y_t^* \) from (39).

\(^{14}\)A log-linearised version of the equation around the zero inflation steady state reveals that:

\[
Y_t^* = \frac{1}{(U_{yy} - \phi_1 f_{yy} + \phi_2 k_{yy}) Y} \left( \bar{\phi}_1 f_{y\xi} + \bar{\phi}_2 k_{y\xi} - \bar{U}_{y\xi} \right) \xi_t
\]
3.3.2 Learning-by-doing

The target level of output is given by (33) when prices are fully flexible:

\[ U_Y(Y_t^e, X_t^e, 1; \xi_t) + \beta E_t \left\{ U_X(Y_{t+1}^e, X_{t+1}^e, 1; \xi_{t+1}) - \phi_2 k_X(Y_{t+1}^e, X_{t+1}^e, 1; \xi_{t+1}) \right\} \frac{\partial X_{t+1}^e}{\partial Y_t^e} = 0 \]

\[ \phi_1 f_y(Y_t^e; \xi_t) + \phi_2 k_y(Y_t^e, X_t^e, 1; \xi_t) \]

(43)

Comparison with (39) shows that targetting output is no longer a function of \( \xi_t \) only but now includes dynamic consideration of \( (X_{t+1}^e, Y_{t+1}^e; \xi_t) \). Moreover the monetary authority cannot know the target level of output unless they know the target level of skills, \( X_t^e \). Of course, \( X_t^e \) is nothing but a summary of \( \{Y_{t+1}^e\}_{T=t_0}^{t-1} \). As such the target level of skills is simply given by

\[ X_t^e = (X_{t-1}^e)^{\phi_x - \mu} \left( \frac{Y_{t-1}^e}{A_{t-1}} \right)^{\mu(1+\eta)} \]

The efficient level of output would see a social planner choosing output according to the rule:

\[ U_Y(Y_t^e, X_t^e, 1; \xi_t) + \beta E_t \left\{ U_X(Y_{t+1}^e, X_{t+1}^e, 1; \xi_{t+1}) \right\} \frac{\partial X_{t+1}^e}{\partial Y_t^e} = 0 \]

where skills are \( X_t^e = (X_{t-1}^e)^{\phi_x - \mu} \left( \frac{Y_{t-1}^e}{A_{t-1}} \right)^{\mu(1+\eta)} \). We can see that the result of the previous section continues to hold: if \( \phi_1 = \phi_2 = 0 \), due to the appropriate subsidy, then \( Y_t^e = Y_t^e \). The natural level of output must be consistent with the analogous version of 42:

\[ f(Y_t^n; \xi_t) = k(Y_t^n, X_t^n, 1; \xi_t) \forall \tau. \]

(44)

\[ (1 - \tau) u_Y(Y_t^n; \xi_t) = \left( \frac{\epsilon}{\epsilon - 1} \right) v_Y(Y_t^n, X_t^n, 1; \xi_t) \]

(45)

Comparison between \( Y_t^e \) and \( Y_t^n \) shows that even if \( \tau = \tau^e = 1 - \left( \frac{\epsilon}{\epsilon - 1} \left( \frac{1}{1 - \beta \mu} \right) \right) \Rightarrow \phi_1 = \phi_2 = 0 \), (45) is not a restatement of (43) as in the BNKM. This is because firms do not take into account the dynamic effects of their hiring decision today on costs tomorrow. To define an output gap that is zero in the zero inflation steady state regardless of the production subsidy, i.e. even in the case of the distorted steady state that is studied below, I define the output gap as \( y_t^g = Y_t - Y_t^e \) and the skills gap as \( x_t^g = X_t - X_t^e \).

To more clearly see the implications of the forward looking nature of optimal policy I log-linearise (33) and (43) to find the following inflation-output gap trade-off\(^{15}\) (derivation in the

\(^{15}\)I have here assumed for simplicity that \( \hat{\Delta}_{t-1} = O(||\xi^2||) \Rightarrow \hat{\Delta}_t = O(||\xi^2||) \forall \tau \) as in the discussion of the
\[ \zeta_{\pi} \pi_t = \lambda E_t \sum_{j=0}^{\infty} \beta^j \left( \frac{k_X \frac{\partial X}{\partial Y}}{f_Y - k_Y} \right)^j \left\{ \bar{Y} (\Omega_Y \Delta y^g_{t+j} + \beta \gamma_Y \Delta y^g_{t+j+1}) + \bar{X} (\Omega_X \Delta x^g_{t+j} + \beta \gamma_X \Delta x^g_{t+j+1}) \right\} \]

(46)

\[ \lambda = \frac{1}{(f_Y - k_Y)} < 0 \]

Where \( \zeta_{\pi} < 0, \Omega_Y < 0, \Omega_X > 0, \gamma_Y < 0 \& \gamma_X > 0 \) (under the baseline calibration in Table 3) are functions of the model parameters and the steady state level of output (details are in the appendix). Note that \( \left( \frac{k_X \frac{\partial X}{\partial Y}}{f_Y - k_Y} \right) > 0 \). Without LBD this rule would be that found by Woodford (2010) p.60:

\[ \zeta_{\pi} \pi_t = \Upsilon_Y \Delta y^g_t \]

(47)

\[ \Upsilon_Y = \lambda \bar{Y} \left( U_{YY} - \tilde{\rho}_1 (f_{YY} - k_{YY}) \right) > 0 \]

In the BNKM the optimal inflation-output gap trade-off\(^\text{16}\) is between inflation and the growth in the contemporaneous output gap. We can now compare these to rules to highlight how LBD changes the policy makers behaviour. The presence of skills creates a link between todays decisions and tomorrows state of the world making the inflation-output trade-off one where the entire present value of output gap changes is considered when thinking about what is the right level of inflation. The forward looking nature of this rule introduces an output gap smoothing motive not present in the BNKM. The effective discount rate \( \beta \left( \frac{k_X}{f_Y - k_Y} \right) \) depends on how strong is the influence of skills on marginal costs, \( k_X \), relative to the gap between the marginal revenue, \( f_Y \), and marginal cost, \( k_Y \), of producing output. How much does policy change with this forward looking rule? This is discussed in the following 2 sections.

\(^{16}\) Note that both (46) and (47) capture a trade-off between \( \pi_t \) and \( \Delta y^g_t \) since \( \frac{\Upsilon_Y}{\zeta_{\pi}} < 0 \) for the former and \( \frac{\Sigma_Y}{\Sigma_{\pi}} < 0 \). Thus the coefficients have opposite signs.
3.4 The case for price stability

When target output moves in proportion with the flexible-price level of output then the goal of monetary policy is maximal price stability (i.e. to attempt to replicate the flexible price equilibrium response to shocks). Woodford (2010); Benigno and Woodford (2005) have shown that this is true even for the case of the distorted steady state, \( \tau < \tau^e = 1 - \frac{\epsilon}{\epsilon - 1} \) in the BNKM. Here I show that this result does not hold under LBD.

The response of \( Y^n_t \) to shocks is governed by (44) which can be rewritten as:

\[
(1 - \tau)u_Y(Y^n_t; \xi_t) = \left( \frac{\epsilon}{\epsilon - 1} \right) v_Y(Y^n_t, X^n_t, 1; \xi_t)
\]

and in log-linear form is:

\[
\hat{Y}^n_t - \frac{1 + \psi^{-1}}{\sigma^{-1} + \chi} \hat{X}^n_t = -u'_\xi \xi_t \quad (48)
\]

Where \( u_\xi \) is as defined in (27). The response of \( Y^*_t \) is governed by (43) which can be rewritten:

\[
\frac{1 + \tilde{\phi}_2(1 - \tau)(1 - \sigma^{-1})}{1 + \tilde{\phi}_2(1 + \chi) \left( \frac{\epsilon}{\epsilon - 1} \right)} u_Y(Y^*_t; \xi_t) - \beta E_t v_X(Y^*_t+1, X^*_t+1, 1; \xi_{t+1}) \frac{\partial X^*_t+1}{\partial Y^*_t} = v_Y(Y^*_t, X^*_t, 1; \xi_t) \quad (49)
\]

Using the steadystate value of \( \tilde{\phi}_2 \) and under the assumption that LBD is zero, these equations show that \( Y^*_t = Y^n_t \) and the result of price stability holds as in Woodford (2010). However learning entails an additional term focusing on the implications for the disutility of labour tomorrow. The presence of \( \xi_{t+1}, Y^*_t+1, X^*_t+1 \) will create a smoothing motive for target output that is absent in the behaviour of \( Y^n_t \). The role of LBD depends importantly on the elasticity of intertemporal substitution, \( \sigma \), as expected when a smoothing motive is present. If \( \sigma > 1 \) then \( \hat{Y}^*_t < \hat{Y}^n_t \) whereas for \( \sigma < 1 \) then \( \hat{Y}^*_t > \hat{Y}^n_t \). However the quantitative impact is relatively small (Figure 3). The reasons for this behaviour are discussed in the following section.
4 The optimal response to shocks

The optimal response of the Ramsey planner is studied by log-linearising the 5 FOCs and 4 constraints around the steady state described in section 3.2.1. A perturbation approach is pursued as described in Schmitt-Grohe and Uribe (2004). The Ramsey solution is studied under shocks to $\xi_t' = \begin{bmatrix} A_t & \xi_t^C & \xi_t^H & \mu_t^P \end{bmatrix}$. The model is calibrated to a quarterly frequency where shocks are temporary but persistent AR(1) processes (see calibration in Table 3). The simulations compare three models, the Basic New Keynesian model (i.e. the model with $\mu = 0$), a model with moderate LBD ($\mu = 0.15$) and a model with strong LBD ($\mu = 0.25$); in two steadystates, the distorted steady state without any production subsidy and the Pareto efficient steady state supported by an efficient production subsidy.

Broadly the findings are as follows. The LBD mechanism undoes the 'divine coincidence' (Blanchard and Gali (2005)) that closing the output gap is entailed by complete price stability. This is because the path of output has stronger implications for marginal costs and thus inflation. This implies that labour supply, preference/demand and technology shocks are non-trivial matters for monetary policy. However the strength of this effect is determined by the divergence between the target level of output, $Y_t^*$, and the flexible price level of output, $Y_t^\pi$. Since this divergence is relatively small (as seen in the preceding section) the resultant inflation and output gap deviations are small also. In addition there is a motive to reduce the fall in the output gap in the face of cost-push shocks however this effect is only present in the distorted steady state. This due to a drop in $Y_t^*$ that doesn’t occur in the efficient steady state combined with an unchanged
The results for a technology shock, $A_t$, indicate a departure from optimal policy in the standard model: so long as $\sigma \neq 1$ there is an inflation-output trade-off (see figure 4). The ‘divine coincidence’ that there is no such trade-off in the standard model is broken by a real imperfection (as suggested by the authors noting this coincidence Blanchard and Gali (2005)): a learning-by-doing externality. Without such a real imperfection monetary policy is trivial in response to a variety of shocks (Woodford (2010)): labour supply, demand/preference shocks and technology shocks. The policy maker simply ensures that the nominal interest rate matches the path of the real interest rate consistent with the flexible price equilibrium (the target real interest rate) and the economy will replicate that flexible price equilibrium with an output gap and inflation rate of zero. How the policy maker responds to this dynamic externality depends on how future consumption is valued relative to consumption today, parametrised by $\sigma$. When $\sigma > 1$ households place more value on output growth. The planner achieves this by returning output to steady state more quickly which entails returning skills to steady state more quickly. This has the cost of inducing deflation via the marginal cost channel as skills recover more quickly; but this cost is more than compensated for by the faster output growth. The reverse holds when $\sigma < 1$. When $\sigma = 1$ the planner chooses to have output match the path of technology. This entails that the current level of employment is sufficient to produce this level of output as
A cost-push shock creates inflationary pressure as firms raise their desired markups (see Figure 5). This forces the planner to face a trade-off between stabilising the output gap and inflation even in the Basic New Keynesian model. The optimal response is price level targetting with initial inflation and a small subsequent deflation. The presence of positive inflation permits a smaller negative output gap than if the policy maker cared only about inflation. As is well known, to make this gap smaller the policy maker must tolerate higher inflation Gali (2008).

How does LBD affect these results? Near the efficient steady state the optimal response is broadly similar with LBD: price-level targetting is achieved by engineering a hump-shaped drop in output (see figure 5). There are additional costs to this choice of output in the case of LBD whereby output falls further as skills depreciate. However these costs are not large enough to make accepting larger swings in inflation worthwhile. An interesting pattern emerges when we consider the response of the policy maker who operates near the distorted steady state. When $\tau < \tau^e$ the behaviour of the target level of output changes to take account of the precommitments made in the steady state by the policy maker operating from a ’timeless perspective’ (see equation 39). These precommitments are captured in $\bar{\phi}_1$ and $\bar{\phi}_2$. The larger the steady state gap between output and its efficient level the larger these weights become. These weights bind the policy
maker to make decisions taking into account the impact of changes in output on the marginal costs and revenues of firms that operate near $\bar{Y}$. Marginal revenue is proportional to $f_Y(\bar{Y})$ and marginal costs are described by $k_Y(\bar{Y}, \bar{X}, \bar{I})$ and $k_X(\bar{Y}, \bar{X}, \bar{I})$. $f_Y(\bar{Y})$ is the same in both the BNKM and LBD models. The differences in results are driven by the behaviour of marginal costs. When $\bar{Y} = 1$ marginal costs are the same in the standard and LBD model since $\bar{X} = 1$, as can be seen from the steady state result (36). When $\bar{Y} > 1$ marginal costs are lower in the LBD model due to the beneficial effects of higher skills on productivity however, $\bar{Y} < 1$ induces low levels of skills pushing marginal costs above those in the standard model (see Figure 6). Due to this endogenous productivity channel steady state output is much lower when $\tau < \tau^e$ with LBD than without (see figure 1). For these monopolistically competitive firms a drop in output lowers marginal costs and raises marginal revenue. Thus the planner aims to accommodate this drop in output more in the distorted steady state. For this reason $Y^*_t$ falls with the cost-push shock near a distorted steady state but is unchanged near the efficient steady state. The drop in $Y^*_t$ combined with a similar path for $Y_t$ as in the efficient steady state case implies a smaller negative output gap. This mechanism is large with LBD and inconsequential without it.

Shocks to preferences over consumption and hours worked do not induce any inflation-output trade-off in the BNKM. As with a technology shock this 'divine coincidence' result does not
hold with LBD. Moreover the nature of the shocks makes offsetting the influence of skills more challenging than the case of technology shocks (where output matching movements in exogenous technology result in no change in employment). Here the planner wishes to make output fall in response to a negative shock to consumption or labour supply (see Figures 9 & 10) since output is less valued by households. In the case of the preference shock to consumption the results are similar to those of a technology shock (with $\sigma < 1$) with positive inflation and a larger output decline. The policy maker now faces a negative output gap to achieve price level targetting.

The labour supply shock is slightly different in that deflation is experienced on impact with subsequent inflation (the reverse of the consumption shock experience). Why? The Ramsey plan involves engineering a small positive skills gap as skills value in reducing the disutility of labour is now higher\(^\text{17}\). The same effect applies to the value of skills in controlling firms marginal costs as hiring labour has become more expensive. The positive skills gap for the first few quarters after the shock creates expectations that marginal costs will be lower than they otherwise would be without this positive gap. At the time the shocks this drop in expected future marginal costs is felt in as a mild deflation (depending on the strength of skills). The majority of the negative output and skills gap occurs only after inflation has largely recovered from the shock and is near steady state thus delaying the impact of the skills-marginal cost channel.

5 Welfare

Welfare comparisons are based on steady state output changes required to make the household indifferent between experiencing the shock and enjoying that level of output: a value for $\zeta$ satisfying:

$$\sum_{t=t_0}^{\infty} \beta^{t-t_0} U(\bar{Y}(1 - \zeta), \bar{X}(\zeta), 1; \bar{\xi}) = E_0 \sum_{t=t_0}^{\infty} \beta^{t-t_0} U(Y_t, X_t, \Delta_t; \xi_t)$$

(50)

Where by $\bar{X}(\zeta)$ I emphasise that the decline in $\bar{Y}$ required for this equality to hold will also require a change in $\bar{X}$. The results are presented in Table 1 & Table 2.

\(^{17}\text{From equation (29) we can see that } \frac{\partial^2 U(Y_t, X_t, \Delta_t; \xi_t)}{\partial X_t \partial \xi_t} > 0\)
Table 1: Welfare with $\tau = \tau^e$

(a) Loss in terms of % of steady state output ($\zeta$)

<table>
<thead>
<tr>
<th>Model / Shocks</th>
<th>Cost-push</th>
<th>Technology</th>
<th>Preference</th>
<th>Labour Supply</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Standard NK model</td>
<td>-0.0077</td>
<td>-0.0070</td>
<td>-0.0100</td>
<td>0.2325</td>
</tr>
<tr>
<td>(2) LBD $\mu = 0.15$</td>
<td>-0.0140</td>
<td>-0.0301</td>
<td>-0.0228</td>
<td>0.7288</td>
</tr>
<tr>
<td>(3) LBD $\mu = 0.25$</td>
<td>-0.0184</td>
<td>-0.0558</td>
<td>-0.0342</td>
<td>1.2387</td>
</tr>
</tbody>
</table>

(b) Relative to Standard NK model

<table>
<thead>
<tr>
<th>Model / Shocks</th>
<th>Cost-push</th>
<th>Technology</th>
<th>Preference</th>
<th>Labour Supply</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) LBD $\mu = 0.15$</td>
<td>1.8147</td>
<td>4.2820</td>
<td>2.2839</td>
<td>3.1352</td>
</tr>
<tr>
<td>(3) LBD $\mu = 0.25$</td>
<td>2.3877</td>
<td>7.9280</td>
<td>3.4279</td>
<td>5.3284</td>
</tr>
</tbody>
</table>

Table 2: Welfare with $\tau = 0$

(a) Loss in terms of % of steady state output ($\zeta$)

<table>
<thead>
<tr>
<th>Model / Shocks</th>
<th>Cost-push</th>
<th>Technology</th>
<th>Preference</th>
<th>Labour Supply</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Standard NK model</td>
<td>-0.0078</td>
<td>-0.0059</td>
<td>-0.0085</td>
<td>0.2345</td>
</tr>
<tr>
<td>(2) LBD $\mu = 0.15$</td>
<td>-0.0147</td>
<td>-0.0248</td>
<td>-0.0162</td>
<td>0.7508</td>
</tr>
<tr>
<td>(3) LBD $\mu = 0.25$</td>
<td>-0.0198</td>
<td>-0.0457</td>
<td>-0.0219</td>
<td>1.3016</td>
</tr>
</tbody>
</table>

(b) Relative to Standard NK model

<table>
<thead>
<tr>
<th>Model / Shocks</th>
<th>Cost-push</th>
<th>Technology</th>
<th>Preference</th>
<th>Labour Supply</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) LBD $\mu = 0.15$</td>
<td>1.8724</td>
<td>4.1777</td>
<td>1.8952</td>
<td>3.2020</td>
</tr>
<tr>
<td>(3) LBD $\mu = 0.25$</td>
<td>2.5249</td>
<td>7.6916</td>
<td>2.5709</td>
<td>5.5507</td>
</tr>
</tbody>
</table>

Absolute losses are in percent of steady state output. The baseline calibration in Table 3 is used. As may be expected the welfare costs of introducing an additional imperfection to the NK model raises the costs of business cycles. However it is less surprising that even under the optimal Ramsey policy these costs rise significantly. This can be ascribed to the fact that the divine coincidence no longer holds meaning that additional costs need to be incurred either in terms of higher inflation or output volatility. The increase in the welfare cost is most notable for technology shocks. This is because the technology shock leads to the largest initial drop in output which combined with the endogenous propagation introduced by skills leads to a long departure of inflation from zero.
6 Optimal Simple Rules

Describing monetary policy in terms of Taylor rules is useful for (at least) two reasons: first, it directly and simply explains the trade-off between output and inflation in terms of interest rate policy and, second, it provides a comparison to a large body of empirical work estimating these Taylor Rules for central banks (Taylor (1993) and Clarida et al. (1997)).

To find the optimal weights for inflation and output in the interest rule for the policy maker I solve the competitive equilibrium described in the last paragraph of section 2 where monetary policy is described by the Taylor Rule $i_t = \gamma_y(y_t / \bar{y}) + \gamma_\pi \pi_t$ (following Leith et al. (2012)). This means adding the Taylor rule to the system comprised of the households optimality conditions (3) and (4), the law of motion for skills (11), the definition of the forward looking measures of marginal costs and revenue for firms (23) and (24), the firms first order condition summarised in aggregate supply realtion (26), the law of motion for price dispersion (28) and the description for the exogenous stochastic processes $\xi_t$. I assume that an optimal produciton subsidy is in place, thus the model is solved around the efficient steady state. The optimal weights $\gamma_y$ and $\gamma_\pi$ are found from a minimising the welfare loss described in (50) by solving the model at each point on a grid of 300 values for $\gamma_y \in [0, 10]$ and $\gamma_\pi \in (1, 2]$\footnote{Repeated simulations with $\gamma_\pi \in (1, 10]$ showed that the optimal weight on inflation is always near 1 when $\mu > 0.11$ and for $\mu < 0.11$, $\gamma_y/\gamma_\pi \to 0$. Since the latter result is preserved with a smaller grid I use it to gain better accuracy with fewer simulations over the range of interest for $\mu$.}. This is done for each value of the parameter $\mu$ governing the strength of the feedback from employment to skills ranging from a low of 0.1 to a high of 0.4. Welfare losses are measured as an average all shocks, i.e. technology, preference, labour supply and cost-push. Thus these weights are ones that are optimal from the perspective of a policy maker concerned equally with each of these shocks.

For low levels of the learning parameter, $\mu \leq 0.11$, $\gamma_y/\gamma_\pi \to 0$ as in the BNKM. For increases in the learning parameter in the range $\mu \in [0.11, 0.27]$, it is optimal for the policy maker to put a higher (but small) weight on output variations ($\gamma_y/\gamma_\pi=2\%$). High dependence of skills on past hours worked, $\mu \in [0.28, 0.33]$ substantially raise the value of output fluctuations to the policy maker where the weight on output is half that of inflation. For very high levels of learning, $\mu > 0.34$, output variations matter almost as much as movements in inflation with $\gamma_y/\gamma_\pi=87\%$.

This exercise suggests that for learning by doing to have a quantitatively significant impact on the operating policy of a central bank that uses a Taylor Rule, learning needs to be much stronger than that measured by Chang et al. (2002) for the US economy. However it may well be that this learning effect is stronger in other economies and is plausibly increasing in relevance as exogenous technical progress leads to higher levels of depreciation in worker skills.
Optimal weights $\gamma_y$ and $\gamma_\pi$ are found from a minimising the welfare loss described in (50) by searching across a grid of 300 values for $\gamma_y \in [0, 10]$ and $\gamma_\pi \in (1, 2]$. For each value these weights the competitive equilibrium described in the last paragraph of section 2 is solved where monetary policy is described by the Taylor Rule $i_t = \gamma_y \left( \frac{y_t}{\bar{y}} \right) + \gamma_\pi \pi_t$. This is done for values of the learning feedback parameter from employment to skills, $\mu \in [0.1, 0.4]$. This requires 9300 simulations.

7 Conclusion

This paper studies the implications for monetary policy of introducing learning-by-doing in production into an otherwise standard Basic New Keynesian model. The time-consistent Ramsey policies are studied in the neighbourhood of the distorted and efficient steady state. The presence of learning introduces two new channels through which output matters for the policy maker. Firstly a marginal cost channel whereby changes in output today lead to proportionate changes in worker productivity tomorrow. Secondly, the presence of learning directly affects the disutility of labour creating an incentive for the policy maker to avoid raising this disutility by letting skills depreciate. The presence of this LBD externality breaks the 'divine coincidence' result, that by stabilising inflation the output gap will automatically be closed, for a variety of
shocks that are considered important in explaining business cycle. I find that skills induce a small increase in the importance of the output gap under a cost-push shock but only for the (more realistic case) of a distorted steady state. The reason for this is due to an interaction between time-consistent policy choices and significant steady state distortions to output due to the presence of LBD on productivity. The welfare costs of business cycles are shown to be significantly larger when learning effects are strong even under the optimal policy.

The approach to optimal monetary policy pursued here, following Woodford (2010), allows for convenient study of different steady states by avoiding the need to derive a purely quadratic approximation to the representative households utility. This approach may be fruitful in studying time-consistent policy choices where steady state distortions can be large, due to significant real imperfections. Korinek (2010) and Benigno and Fornaro (2012) have shown that the results of the growth literature where imported technology drives learning externalities can have large welfare effects in an open economy setting. Neither of these studies have nominal rigidities and thus do not study the implications for monetary policy in the context of an open economy. Building on these results drawing on the framework applied here to study monetary policy for the small open economy may be a fruitful avenue for further research.
References


URL http://ideas.repec.org/a/bla/restud/v65y1998i1p45-85.html

URL http://ideas.repec.org/h/eee/macchp/1-11.html

URL http://www.nber.org/papers/w0807


URL http://cep.lse.ac.uk/pubs/download/dp1161.pdf

URL http://ideas.repec.org/a/tpr/jeurec/v3y2005i6p1185-1236.html

URL http://www.nber.org/papers/w11806

URL http://www.jstor.org/stable/3083260

URL http://www.nber.org/papers/w6254

URL http://ideas.repec.org/p/imf/imfspn/2010-03.html


URL http://www.nber.org/papers/w17880

URL http://ideas.repec.org/a/eee/moneco/v46y2000i2p281-313.html


URL http://ideas.repec.org/a/bla/restud/v72y2005i3p707-734.html

URL http://ideas.repec.org/b/mtp/titles/0262570971.html

URL http://ideas.repec.org/a/eee/jetheo/v144y2009i6p2293-2309.html
URL http://ideas.repec.org/a/eee/inecon/v76y2008i2p223-236.html

URL http://ideas.repec.org/p/nbr/nberwo/15826.html


URL http://ideas.repec.org/a/red/issued/09-154.html


URL http://ideas.repec.org/a/tpr/jeurec/v1y2003i5p1123-1175.html


URL http://ideas.repec.org/a/eee/eecrev/v51y2007i5p1107-1125.html


URL http://www.nber.org/papers/w8573
8 Appendix I: Additional Figures & Tables

8.1 Figures

All simulations presented here have used the baseline calibration presented in Table 3.

8.1.1 Optimal Response to shocks with efficient subsidy

The efficient subsidy for the the standard model is $\tau^e = 1 - \epsilon$ and $\tau^e = 1 - \left(\frac{\epsilon}{\epsilon-1}\right)\left(\frac{1}{1-\beta\mu}\right)$ for the model with learning by doing. These values are assumed in simulations below.

Figure 8: Cost-push shock with $\tau = \tau^e$
Figure 9: Technology shock with $\tau = \tau^e$
Figure 10: Labour Supply shock with $\tau = \tau^e$
Figure 11: Preference shock with $\tau = \tau^e$
8.1.2 Optimal Response to shocks with $\tau = 0$

Only the response to a cost-push shock is significantly altered thus the IRFs for the remaining shocks are not reported here.

Figure 12: Cost-push shock with $\tau = 0$
## 8.2 Tables

Table 3: Baseline Calibration

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.99</td>
<td>Consistent with 4% annual interest rate</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.33</td>
<td>Consistent with a labour share of $\frac{2}{3}$</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.8</td>
<td>Attanasio (1999). Unless stated otherwise.</td>
</tr>
<tr>
<td>$\psi$</td>
<td>1</td>
<td>Dyrda et al. (2012)</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>11</td>
<td>Leith et al. (2012)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.75</td>
<td>Average price duration of 4 quarters, Klenow and Malin (2010)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.15</td>
<td>Consistent with range in Chang et al. (2002). Unless stated otherwise.</td>
</tr>
<tr>
<td>$\phi_x$</td>
<td>0.79</td>
<td>Chang et al. (2002)</td>
</tr>
<tr>
<td>$\sigma_{\mu}^2$</td>
<td>0.0016</td>
<td>Smets and Wouters (2003)</td>
</tr>
<tr>
<td>$\sigma_A^2$</td>
<td>0.0071</td>
<td>Gali and Monacelli (2005)</td>
</tr>
<tr>
<td>$\sigma_H^2$</td>
<td>0.0166</td>
<td>Smets and Wouters (2003)</td>
</tr>
<tr>
<td>$\sigma_C^2$</td>
<td>0.0028</td>
<td>Smets and Wouters (2003)</td>
</tr>
<tr>
<td>$\rho_{\mu}$</td>
<td>0.8</td>
<td>Leith et al. (2012)</td>
</tr>
<tr>
<td>$\rho_A$</td>
<td>0.91</td>
<td>Adolfson et al. (2007)</td>
</tr>
<tr>
<td>$\rho_H$</td>
<td>0.93</td>
<td>Smets and Wouters (2003)</td>
</tr>
<tr>
<td>$\rho_C$</td>
<td>0.88</td>
<td>Smets and Wouters (2003)</td>
</tr>
</tbody>
</table>

Shocks are assumed to be uncorrelated and follow an AR(1) process with persistence $\rho$ and standard deviation $\sigma^e$. The model is calibrated to quarterly frequency.
9 Appendix II: Derivations

9.1 Derivation of Aggregate Supply relation (26)

The firm able to reset their price at time $t$ will solve the following problem:

$$
\max_{\{P_t(i)\}_{t=t_0}} \mathbb{E}_t \sum_{T=t}^{\infty} \omega^{T-t} Q_{t,T} \Pi(P_t(i), P_T, Y_T, X_T; \xi_T)
$$

(51)

$\xi_T$ refers to the entire collection of shocks that affect firms pricing decision,

$$
\xi_T = \begin{bmatrix} A_T & \xi_C^T & \xi_H^T & \mu_P^T \end{bmatrix}. \\
\mu_P^T \text{ refers to a shock to firms desired steady state mark-up.}
$$

$Q_{t,T}$ is the value placed on nominal profits returned to the household $T$ periods hence:

$$
Q_{t,T} = \beta^{T-t} \frac{u_C(C_T, \xi_C^T)}{u_C(C_t, \xi_C^T)} P_t \left( \frac{Y_T}{Y_t} \xi_C^T \right)^{\frac{1}{\sigma}}
$$

(52)

For the $i^{th}$ firm nominal profits in period $T$ are simply nominal revenues less costs:

$$
(1 - \tau) P_t(i) \left( \frac{P_t(i)}{P_T} \right)^{-\epsilon} Y_T - W_T N_T(i)
$$

(53)

Substituting $W_T N_T(i)$ using (15), (16) and (4), the firms problem is:

$$
\max_{\{P_t(i)\}_{t=t_0}} \mathbb{E}_t \sum_{T=t}^{\infty} \omega^{T-t} \left( \frac{Y_t}{\xi_t^C} \right)^{\frac{1}{\sigma}} \{ (1 - \tau) P_t(i)^{1-\epsilon} \left( \frac{P_t(i)}{P_T(i)} \right) Y_T^{(1-\sigma)} \left( \xi_C^i \right)^{\frac{1}{\sigma}} \\
- P_t \left( \frac{P_t(i)}{P_T} \right)^{\sigma(1+\eta)} \left( \frac{Y_T}{A_T} \right)^{(1+\chi)} \frac{\Delta_{TL}^T}{X_T^{(1+\frac{1}{\psi})}} \frac{1}{\left( \xi_H^T \right)^{\frac{1}{\psi}}} \}
$$

The first order conditions for profit maximisation is:

$$
\mathbb{E}_t \sum_{T=t}^{\infty} \omega^{T-t} \{ (1 - \tau) \left( \frac{P_t(i)}{P_T} \right)^{\sigma(1+\eta)} \left( \frac{P_t(i)}{P_T(i)} \right) Y_T^{(1-\sigma)} \left( \xi_C^i \right)^{\frac{1}{\sigma}} \\
- \left( \frac{\epsilon}{\epsilon - 1} \right) (1 + \eta) \left( \frac{P_t(i)}{P_T} \right)^{-\sigma(1+\eta)} \left( \frac{P_t(i)}{P_T} \right) Y_T^{(1+\chi)} \left( \xi_C^i \right)^{\frac{1}{\sigma}} \frac{\Delta_{TL}^T}{X_T^{(1+\frac{1}{\psi})}} \frac{1}{\left( \xi_H^T \right)^{\frac{1}{\psi}}} \}
$$

Which can be rearranged as, after imposing $P_t(i) = P^*_t$: 

44
\[
E_t \sum_{T=t}^{\infty} (\omega \beta)^{T-t} \left\{ (1 - \tau) \left( P_t^* \right)^{-\epsilon} \left( \frac{P_t}{P_T} \right) Y_T^{(1-\frac{1}{\sigma})} \left( \xi_T^{\frac{1}{\sigma}} \right) \right\} = 1 \tag{54}
\]

Multiplying by \( \left( \frac{P_t^*}{P_t} \right)^{(1+\epsilon)} \) and regarrranging we have:

\[
\left( \frac{P_t^*}{P_t} \right)^{(1+\epsilon)} = \frac{E_t \sum_{T=t}^{\infty} (\omega \beta)^{T-t} \left\{ (1 - \tau) Y_T^{(1-\frac{1}{\sigma})} \left( \xi_T^{\frac{1}{\sigma}} \right) \right\} \left( \frac{P_t}{P_T} \right)^{\epsilon-1}}{E_t \sum_{T=t}^{\infty} (\omega \beta)^{T-t} \left\{ (1 + \eta) \left( \frac{Y_T}{A_T} \right)^{(1+\chi)} \Delta_T^{\frac{1}{\psi}} (\xi_T^{\frac{1}{\psi}}) \left( \xi_T^{\frac{1}{\psi}} \right) \right\} \left( \frac{P_t}{P_T} \right)^{(1+\epsilon)}} \tag{55}
\]

Which rearranged gives equation (26):

\[
\left( \frac{P_t^*}{P_t} \right) = \left( \frac{F_t}{K_t} \right)^{\frac{1}{1+\epsilon}} \tag{56}
\]

### 9.2 Linearised New Keynesian Phillips Curve (27)

The linearisation is around the zero inflation steady state described by the Ramsey solution (fully described in 8.3 of this appendix). The law of motion for the price dispersion (equation 28), the forward looking relations \( F_t, K_t \) (equations 23, 24) and the aggregate supply relation (equation 26); can be linearised as:

\[
\hat{\Delta}_t = \omega \hat{\Delta}_{t-1} \tag{57}
\]

\[
F_t = (1 - \omega \beta) \left[ f_y \hat{Y}_t + f_k^2 \hat{\xi}_t \right] + \omega \beta E_t \left[ (\epsilon - 1) \hat{P}_{t+1} + \hat{F}_{t+1} \right] \tag{58}
\]

\[
K_t = (1 - \omega \beta) \left[ k_y \hat{Y}_t + k_X \hat{X}_t + k_\Delta \hat{\Delta}_t + k_\xi \hat{\xi}_t \right] + \omega \beta E_t \left[ \epsilon (1 + \eta) \hat{P}_{t+1} + \hat{K}_{t+1} \right] \tag{59}
\]

\[
\hat{P}_t = \frac{1 - \omega}{\omega} \frac{1}{1 - \epsilon \eta} (\hat{K}_t - \hat{F}_t) \tag{60}
\]

From (57) it can be seen that if \( \hat{\Delta}_{t_0-1} = O(||\xi^2||) \) then \( \hat{\Delta}_t = O(||\xi^2||) \forall t \). Differencing (59) from (58) and substituting out \( (\hat{K}_t - \hat{F}_t) \) using (60) we have (27) in the main text.

### 9.3 Solution to Ramsey problem in section 3.2

The Ramsey problem outlined above can be described by the Lagrangian:
\[ \max_{\{F_t, K_t, \Delta_t, \Pi_t, Y_t\}} \mathcal{L}_{t_0} = E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} L(Y_t, X_t, F_t, K_t, \Pi_t, \Delta_t; \theta_t, \phi_t, \xi_t) \]  

(61)

Where

\[ L(Y_t, X_t, F_t, K_t, \Pi_t, \Delta_t; \theta_t, \phi_t, \xi_t) = U(Y_t, X_t, \Delta_t; \xi_t) + \phi_{1,t} \left[ F_t - f(Y_t; \xi_t) - \omega \beta E_t \Pi_{t+1}^{(1+\eta)} F_{t+1} \right] + \phi_{2,t} \left[ K_t - k(Y_t, X_t, \Delta_t; \xi_t) - \omega \beta E_t \Pi_{t+1}^{(1+\eta)} K_{t+1} \right] + \theta_t \left[ \Delta_t - \omega \Delta_{t-1} \Pi_{t}^{(1+\eta)} - (1 - \omega) \left( \frac{1 - \omega \Pi_{t}^{(1+\eta)} - 1}{1 - \omega} \right)^{\frac{(1+\eta)}{\eta+1}} \right] + \phi_{3,t} \left[ \frac{1 - \omega \Pi_{t}^{(1+\eta)} - 1}{1 - \omega} - \left( \frac{F_t}{K_t} \right)^{\frac{1}{1+\eta}} \right] \]

Or equivalently

\[ L(Y_t, X_t, F_t, K_t, \Pi_t, \Delta_t; \theta_t, \phi_t, \xi_t) = U(Y_t, X_t, \Delta_t; \xi_t) + \theta_t \left[ \Delta_t - \omega \Delta_{t-1} \Pi_{t}^{(1+\eta)} - (1 - \omega) \left( \frac{1 - \omega \Pi_{t}^{(1+\eta)} - 1}{1 - \omega} \right)^{\frac{(1+\eta)}{\eta+1}} \right] + \phi_{1,t} \left[ F_t - f(Y_t; \xi_t) - \omega \phi_{1,t-1} \left[ \Pi_{t-1}^{(1+\eta)} F_t \right] + \phi_{2,t} \left[ K_t - k(Y_t, X_t, \Delta_t; \xi_t) \right] + \omega \phi_{2,t-1} \left[ \Pi_{t}^{(1+\eta)} K_t \right] \right] + \phi_{3,t} \left[ \frac{1 - \omega \Pi_{t}^{(1+\eta)} - 1}{1 - \omega} - \left( \frac{F_t}{K_t} \right)^{\frac{1}{1+\eta}} \right] \]

Where the multipliers \( \phi_{1,t-1}, \phi_{2,t-1} \) will capture the precommitments made at \( t_0 \) i.e. they will be the values consistent with the steady state solution of the model (which holds for \( t_0 - 1 \)) under the same constraints the Ramsey planner faces for \( t > t_0 \).

The first-order conditions (FOCs) of the above problem when the Ramsey planner chooses \( \{F_t, K_t, \Delta_t, \Pi_t, Y_t\} \) are:

\[ \frac{\partial \mathcal{L}_{t_0}}{\partial Y_t} = 0 : \quad U_Y(Y_t, X_t, \Delta_t; \xi_t) + \beta \{ U_X(Y_{t+1}, X_{t+1}, \Delta_{t+1}; \xi_{t+1}) - \phi_{2,t+1} k_X(Y_t, X_t, \Delta_t; \xi_t) \} \frac{\partial X_{t+1}}{\partial Y_t} = \]

\[ + \phi_{1,t} f_y(Y_t; \xi_t) + \phi_{2,t} k_y(Y_t, X_t, \Delta_t; \xi_t) \]  

(62)

\[ \frac{\partial \mathcal{L}_{t_0}}{\partial \Delta_t} = 0 : \quad U_\Delta(Y_t, X_t, \Delta_t; \xi_t) + \theta_t + \beta \{ U_X(Y_{t+1}, X_{t+1}, \Delta_{t+1}; \xi_{t+1}) - \phi_{2,t+1} k_X(Y_t, X_t, \Delta_t; \xi_t) \} \frac{\partial X_{t+1}}{\partial \Delta_t} = \]

\[ + \phi_{2,t} k_\Delta(Y_t, X_t, \Delta_t; \xi_t) + \theta_{t+1} \omega \Pi_{t+1}^{(1+\eta)} \]  

(63)

\[ \frac{\partial \mathcal{L}_{t_0}}{\partial \Pi_t} = 0 : \quad \phi_{3,t} \frac{\omega (1+\eta)}{1-\omega} P(\Pi_t) \left( \frac{1+\eta}{1-\eta} \right) \Pi_t^{-2} K_t + \phi_{1,t-1} \omega ( \epsilon - 1 ) \Pi_t^{-2} \Phi_t + \phi_{2,t-1} \omega \epsilon ( 1+\eta ) \Pi_t^{(1+\eta)-1} K_t \]

46
\[ +\theta_t \left[ \omega e(1 + \eta)\Delta_{t-1}\Pi_t^{(1+\eta)^{-1}} - \omega e(1 + \eta)\Delta_{t-1}P(\Pi_t)^{(1+\eta)^{-1}}\Pi_t^{\epsilon-2} \right] = 0 \quad (64) \]

\[ \frac{\partial L_{t0}}{\partial K_t} = 0 : \phi_{3,t}P(\Pi_t)^{(1+\eta)^{-1}} + \phi_{2,t} - \phi_{2,t-1}\omega\Pi_t^{\epsilon(1+\eta)} = 0 \quad (65) \]

\[ \frac{\partial L_{t0}}{\partial F_t} = 0 : -\phi_{3,t} + \phi_{1,t} - \phi_{1,t-1}\omega\Pi_t^{-1} = 0 \quad (66) \]

In order to study the optimal response to shocks I linearise these conditions around the optimal steady state\(^{19}\). This is the steady state associated with the above 5 FOCs and the 4 constraints. The optimal steady state is characterised by \( \{\bar{F}, \bar{K}, \bar{\Delta}, \bar{\Pi}, \bar{Y}, \bar{X}, \bar{\theta}, \bar{\phi} \} \) that solves these 9 equations when \( \xi_t = \bar{\xi} \). This steady state includes the following conditions:

\[ f(\bar{Y}) = k(\bar{Y}, \bar{X}, \bar{\Delta}) \iff \frac{\nu_h(\bar{Y}, \bar{X}, \bar{\Delta})}{\nu_c(\bar{Y})} = \frac{1 - \bar{\tau}}{\epsilon^{-1}} \iff \bar{Y} = \left( \frac{1 - \bar{\tau}}{\epsilon^{-1}(1 + \eta)} \right)^{1 - \phi_{x} + \mu} \quad (67) \]

\[ \bar{X} = \bar{Y}^{\mu(1+\eta)} \quad (68) \]

\[ \bar{\Delta} = \bar{\Pi} = 1 \quad (69) \]

\[ \bar{K} = \bar{F} = (1 - \omega\beta)^{-1}f(\bar{Y}) \quad (70) \]

\[ \bar{\phi}_2 = -\bar{\phi}_1 = \frac{U_Y + \beta U_X \frac{\partial X}{\partial Y}}{k_Y - f_Y + \beta \frac{\partial X}{\partial Y} k_X} \quad (71) \]

\[ \bar{\phi}_3 = (1 - \omega)\bar{\phi}_1 \quad (72) \]

\(^{19}\) Thus this is not the optimal steady state of the social planner, but only the optimal steady state from the perspective of a planner that must take actions subject to competitive equilibrium.
\[
\hat{\theta} = \frac{U_{\Delta} + \beta U_X \frac{\partial X}{\partial X}}{1 - \omega} + \left( k_{\Delta} + \beta k_X \frac{\partial X}{\partial X} \right) \left( U_Y + \beta U_X \frac{\partial X}{\partial Y} \right) \frac{1}{1 - \omega \left( k_Y - f_Y + \beta k_X \frac{\partial X}{\partial Y} \right)}
\]

(73)

These are the steadystate levels that the Ramsey planner commits to at \( t = t_0 \) which bind her for future periods. The values of \( \Psi_F(K_{t0}, F_{t0}) \) and \( \Psi_K(K_{t0}, F_{t0}) \) are:

\[
\Psi_F(K_{t0}, F_{t0}) = \left[ \frac{1}{\omega} - \left( \frac{1 - \omega}{\omega} \right) \left( \frac{\bar{F}}{K} \right) \right] \hat{F} = \bar{F}
\]

(74)

\[
\Psi_K(K_{t0}, F_{t0}) = \left[ \frac{1}{\omega} - \left( \frac{1 - \omega}{\omega} \right) \left( \frac{\bar{F}}{K} \right) \right] \hat{K} = \bar{K}
\]

(75)

Finally, note the substitution made in (23) and (24):

\[
\Pi_t = \left[ \frac{1}{\omega} - \left( \frac{1 - \omega}{\omega} \right) \left( \frac{F_t}{K_t} \right) \right]^{\frac{1}{1+\eta}}
\]

(76)

### 9.4 Derivation of \( \pi - g^t \) trade-off of Section 3.3.2

The FOC for output in the Ramsey problem (32) is:

\[
\frac{\partial \zeta_{t0}}{\partial Y_{t0}} = 0: \quad U_Y(Y_t, X_t, \Delta_t; \xi_t) + \beta E_t \left\{ U_X(Y_{t+1}, X_{t+1}, \Delta_{t+1}; \xi_{t+1}) - \phi_{t, t+1} k_X(Y_t, X_t, \Delta_t; \xi_{t+1}) \right\} \frac{\partial X_{t+1}}{\partial Y_{t0}} = \]

\[
+ \phi_{t, t} f_Y(Y_t; \xi_t) + \phi_{t, t} k_y(Y_t, X_t, \Delta_t; \xi_t)
\]

(77)

Linearised around the steady state described in Section 8.3 of this appendix this becomes:

\[
\ddot{Y}_{t} \Omega_{Y} \ddot{Y}_{t} + \ddot{X}_{t} \Omega_{X} \ddot{X}_{t} + \Omega_{\Delta} \ddot{\Delta}_{t} + \Omega_{\xi} \ddot{\xi}_{t}
\]

\[
- \beta E_t \left\{ \gamma_{Y} \dot{Y}_{t+1} + \gamma_{X} \dot{X}_{t+1} + \gamma_{\Delta} \dot{\Delta}_{t+1} + \gamma_{\xi} \dot{\xi}_{t+1} \right\}
\]

\[
(k_Y - f_Y) \ddot{\phi}_{t, t} - \beta E_t k_X \ddot{\phi}_{t, t+1} = 0
\]

(78)

Where for each variable \( \ddot{Z}_t \equiv \ln Z_t - \ln \bar{Z} \) and \( \ddot{Z}_t = Z_t - \bar{Z} \). The equation defining the target level of output is:

\[
U_Y(Y_t^*, X_t^*, 1; \xi_t) + \beta E_t \left\{ U_X(Y_{t+1}^*, X_{t+1}^*, 1; \xi_{t+1}) - \ddot{\phi}_{t, t} k_X(Y_{t+1}^*, X_{t+1}^*, 1; \xi_{t+1}) \right\} \frac{\partial X_{t+1}}{\partial Y_{t}^*} = \]

48
\[ \hat{\phi}_1 f_y(Y_t^*; \xi_t) + \hat{\phi}_2 k_y(Y_t^*; X_t^*, 1; \xi_t) \]

Linearising this gives:

\[ U_{t_0} = E_0 \sum_{t=t_0}^{\infty} \beta^{t-t_0} \{ U(Y_t, X_t, \Delta_t; \xi_t) \} \]

\[ \bar{Y}_t \hat{\Omega} \hat{Y}_t^* + \bar{X}_t \hat{\Omega} \hat{X}_t^* + \Omega' \hat{\xi}_t + \beta E_t \{ \bar{Y}_t \gamma_Y \hat{Y}_{t+1} + \bar{X}_t \gamma_X \hat{X}_{t+1} + \gamma' \hat{\xi}_{t+1} \} = 0 \]  

(80)

The coefficients in both (78) and (80) are defined by:

\[ \Omega_Y = U_{YY} + \beta \frac{\partial^2 X}{\partial Y^2} (U_X - \tilde{\phi}_2 k_X) - \tilde{\phi}_1 (f_{YY} - k_{YY}) < 0; \quad \Omega_X = U_{YX} + \beta \frac{\partial^2 X}{\partial Y \partial X} (U_X - \tilde{\phi}_2 k_X) + \tilde{\phi}_1 k_{YY} > 0 \]

\[ \Omega_\xi = U'_{\xi} + \beta \frac{\partial^2 X}{\partial Y \partial \xi} (U_X - \tilde{\phi}_2 k_X) - \tilde{\phi}_1 (f'_{Y \xi} - k'_{Y \xi}) \]

\[ \gamma_Y = \frac{\partial X}{\partial Y} (U_{XY} - \tilde{\phi}_2 k_{XY}) > 0; \quad \gamma_X = \frac{\partial X}{\partial Y} (U_{XX} - \tilde{\phi}_2 k_{XX}) < 0; \quad \gamma_\xi = \frac{\partial X}{\partial Y} (U'_{\xi} - \tilde{\phi}_2 k'_{X \xi}) \]

The signs for these coefficients assume the baseline calibration of the model given in Table 3. (78) and (80) can be combined into a statement in terms of the output gap \( y_t^g = \bar{Y}_t - \hat{Y}_t^* \) and the skills gap as \( x_t^g = \hat{X}_t - \hat{X}_t^* \):

\[ \bar{Y}_t \hat{\Omega} y_t^g + \bar{X}_t \hat{\Omega} x_t^g + \beta E_t \left\{ \bar{Y}_t \gamma_Y y_{t+1}^g + \bar{X}_t \gamma_X x_{t+1}^g - k_X \frac{\partial X}{\partial Y} \hat{\phi}_{2,t+1} \right\} \]

\[ = (f_Y - k_Y) \hat{\phi}_{1,t} \]  

(81)

Where for simplicity I have again used the assumption that \( \hat{\Delta}_{t_0-1} = \mathcal{O}(||\xi^2||) \Rightarrow \hat{\Delta}_t = \mathcal{O}(||\xi^2||) \forall t \) and ignored price dispersion. For ease of notation call \( A_t \equiv \bar{Y}_t \gamma_Y y_{t+1}^g + \bar{X}_t \gamma_X x_{t+1}^g \) and \( B_{t+1} \equiv \bar{Y}_t \gamma_Y y_{t+1}^g + \bar{X}_t \gamma_X x_{t+1}^g \). Thus (81) can be written as:

\[ (f_Y - k_Y) \hat{\phi}_{1,t} = \beta \left( k_X \frac{\partial X}{\partial Y} \right) E_t \hat{\phi}_{1,t+1} + A_t + \beta E_t B_{t+1} \]  

(82)

Where I have used the fact that \( \hat{\phi}_{1,t} = -\hat{\phi}_{2,t} \). Iterating on (82) and assuming no bubble solutions we have:

\[ \hat{\phi}_{1,t} = \frac{1}{(f_Y - k_Y)} E \sum_{j=0}^{\infty} \beta^j \left( \frac{k_X \partial X}{f_Y - k_Y} \right)^j \{ A_{t+j} + \beta E_t B_{t+j+1} \} \]  

(83)

\[ \text{This follows from a linearisation of (64), (65) and (66) as proved by Woodford (2010), page 58.} \]
Note that \( \left( \frac{k_X \frac{\partial X}{\partial Y}}{f_Y - k_Y} \right) > 0 \). Linearising the conditions (64), (65) and (66) yields the following relationship between inflation and the multiplier \( \tilde{\phi}_{1,t} \):

\[
\zeta \pi_t = \Delta \tilde{\phi}_{1,t}
\]  

(84)

\[
\zeta = -\frac{\bar{\theta}}{K} \epsilon (1 + \chi) - \frac{\omega}{1 - \omega} (1 + \epsilon \eta)
\]

Taking first differences of (83) to replace the term \( \Delta \tilde{\phi}_{1,t} \) in (84) we have:

\[
\zeta \pi_t = \frac{1}{(f_Y - k_Y)} E_t \sum_{j=0}^{\infty} \beta^j \left( \frac{k_X \frac{\partial X}{\partial Y}}{f_Y - k_Y} \right)^j \{ \Delta A_{t+j} + \beta E_t \Delta B_{t+j+1} \}
\]  

(85)

Which can be rearranged using the definitions of \( A_t \) and \( B_{t+1} \) as equation (46) in the main text:

\[
\zeta \pi_t = \lambda E_t \sum_{j=0}^{\infty} \beta^j \left( \frac{k_X \frac{\partial X}{\partial Y}}{f_Y - k_Y} \right)^j \{ \bar{Y} (\Omega_Y \Delta y_{t+j}^g + \beta \gamma_Y \Delta y_{t+j+1}^g) + \bar{X} (\Omega_X \Delta x_{t+j}^g + \beta \gamma_X \Delta x_{t+j+1}^g) \}
\]  

(86)

\[
\lambda = \frac{1}{(f_Y - k_Y)} < 0
\]