The Welfare Losses of Adaptive Learning Dynamics

Atanas Christev and Sergey Slobodyan
Heriott-Watt University and CERGE-EI

February 16, 2016

Abstract

In this paper we study optimal monetary policy with adaptive learning. We establish a connection between the higher speed of convergence in the mean learning dynamics and the lower welfare losses associated with learning algorithms characterized by the existence of invariant distribution of agents' beliefs in equilibrium. We show this result using a purely-forward looking New Keynesian model with an expectations-based monetary policy rule, and relate it to richer versions of the model in the literature on optimal policy. The algorithm of computing the welfare losses can be generalized to derive the welfare criterion in a closed form solution for different values of the parameter space governing the policy choices.

First draft April 2015. Preliminary please don't quote.

1 Introduction

Monetary policy design influences equilibrium outcomes. It is well established that active monetary policy in models of adaptive learning is helpful in avoiding multiple equilibria, see Bullard and Mitra (2002) and Woodford (2003), and Woodford (2001) for a discussion of the Taylor principle. This is important as the economy does not coordinate on the "wrong" set of self-confirming beliefs which may generate detrimental outcomes to welfare. In these models agents' beliefs place an additional constraint on monetary policy choices, as the equilibria have to be stable under adaptive learning. Equally important, as pointed our in a recent survey by Eusepi and Preston (2015), is to assess how monetary policy performs in the transition to rational expectation (or the invariant distribution of beliefs).

In this paper we use a standard New Keynesian model to show that the speed of convergence of the mean dynamics under constant gain learning, commonly adopted in the literature, is essential to guarantee minimal welfare losses. An updating algorithm with a constant gain is well suited to the problems of tracking slowly changing parameters of the economy, see Benveniste Metiver and Priouret (1990). Private sector agents who do not know their environment exactly and thus need to resort to learning to understand the model of the economy might suspect the environment to be non-stationary as well, which justifies reliance on constant gain algorithms. Their initial beliefs need not be in the vicinity of the rational expectations ones. The speed of convergence of the beliefs to their rational expectations values, or the mean dynamics speed of convergence, turns out to be an essential feature to gleaning the model' welfare properties. The convergence speed, and thus the welfare characteristics of the model, could be manipulated by the policymaker choosing parameters of a Taylor-type monetary policy rule.

This paper considers optimal monetary policy and welfare in dynamic general equilibrium with constant gain learning in order to study the robustness of activist monetary policy advice in this type of settings. Our work follows in the tradition of Orphanides and Williams (2007) and Orphanides and Williams (2008), and is closely related to the original contributions of Gaspar et al. (2006) and Gaspar et al. (2011) and Molnar and Santoro (2014). We show that under some conditions, faster mean dynamics speed of convergence leads to lower welfare losses associated with the invariant distribution of agents' beliefs under adaptive learning.

In dynamic models with Taylor-type rules, considered in the paper, results depend critically on the expectations of the private sector regarding future monetary policy actions (or the future path of policy). Hence the problem of monetary policy essentially amounts to facilitating or shaping expectations to rule out private sector beliefs that may send the economy toward an equilibrium that is suboptimal in terms of welfare. This is the main contribution of the paper: under what conditions in the learning model are welfare losses minimal? How can policy be designed to ensure this outcome or engender this self-fulfilling process? Can policy help navigate evolving (rational) beliefs towards the inflation target without creating excessive fluctuations in private sector expectations with adverse welfare effects?

The rest of the paper is organized as follows. In the Section 2, we describe a simplified New Keynesian model with adaptively learning agents and compute invariant distribution of the agents' beliefs. In Section 3, we compute the welfare, associated with a particular monetary policy. In Section 4, we compare our results for a special case of *iid* mark-up shock to those in the literature, and discuss the connection between the welfare criterion that we use to the one commonly used in the literature. Section 5 describes the connection between welfare criterion and the mean dynamics convergence speed, and Section 6 concludes by demonstrating how our set-up could be extended to more complicated settings.

2 The Model

Consider the following NK model:

$$\begin{aligned} x_t &= \hat{E}_t x_{t+1} - \varphi(i_t - \hat{E}_t \pi_{t+1}), \\ \pi_t &= \beta \hat{E}_t \pi_{t+1} + \kappa x_t + u_t. \end{aligned} \tag{1}$$

In general, expectations could be different from mathematical expectations, which is reflected in the notation $\hat{E}$. We assume that cost-push shock u_t is an AR(1) process with coefficient of correlation ρ and *iid* innovation ϵ_t.

Using a simple, expectations based, linear monetary policy rule, which does not include any lagged variables, we could write the model as

$$\begin{bmatrix} \pi_t \\ x_t \end{bmatrix} = Q + B\hat{E}_t \begin{bmatrix} \pi_{t+1} \\ x_{t+1} \end{bmatrix} + Cu_t.$$

Here Q is a 2×1 vector which equals zero, B a 2×2 matrix, and C a 2×1 vector.

Suppose that the agents are using (constant gain) recursive least squares (RLS) to learn parameters of the model. They are uncertain of the long-term means of inflation π_t and output gap x_t which under rational expectations (RE) are both zero, and thus use a learning rule with an intercept.

In the RE equilibrium, both inflation π_t and output gap x_t are proportional to the contemporaneous cost-push shock. We assume that the agents don't have access to the current value u_t and could observe only u_{t-1}, and perceive the law of motion of the state variables as

$$\left[\begin{array}{c} \pi_t \\ x_t \end{array} \right] = A + K u_{t-1}, \tag{2}$$

where both A and K are 2×1 vectors. The equation (2) represents the agents' Perceived Law of Motion, or PLM.

Given this PLM, the agents' expectations are given by

$$\hat{E}_t \left[\begin{array}{c} \pi_{t+1} \\ x_{t+1} \end{array} \right] = A + \rho K u_{t-1},$$

and therefore

$$\left[\begin{array}{c} \pi_t \\ x_t \end{array} \right] = Q + B\left(A + \rho K u_{t-1}\right) + C\left(\rho u_{t-1} + \epsilon_t\right) = Q + BA + \rho\left(BK + C\right) u_{t-1} + C\epsilon_t,$$

which represents the Actual Law of Motion, or ALM. Therefore, the T-map is given by

$$\left[\begin{array}{c} A \\ K \end{array} \right] \rightarrow \left[\begin{array}{c} Q + BA \\ \rho BK + \rho C \end{array} \right],$$

and the E-stability ODE by

$$\begin{aligned} \dot{A} &= Q + (B - I) A, \\ \dot{K} &= \rho C + (\rho B - I) K. \end{aligned} \tag{3}$$

If all roots of B are inside unit circle, the above ODE is globally stable, and thus the unique REE of the model is stable as well.

More compactly, write the agents' beliefs as

$$\Psi^T = \left[A \;\; \vdots \;\; K \right].$$

The PLM is then

$$Y_t = \begin{bmatrix} \pi_t \\ x_t \end{bmatrix} = \begin{bmatrix} A & \vdots & K \end{bmatrix} \cdot \begin{bmatrix} 1 \\ u_{t-1} \end{bmatrix} = \Psi^T \cdot U_t + C\epsilon_t,$$

and the ALM could be written as

$$Y_t = \begin{bmatrix} Q + BA & \vdots & \rho C + \rho BK \end{bmatrix} \cdot U_t + C\epsilon_t. \tag{4}$$

2.1 An One-Dimensional Model

In this paper, following Ferrero (2007), we assume an expectations-based policy rule with a coefficient on $\hat{E}_t x_{t+1}$ such that the model, effectively, becomes one-dimensional, because the output gap can be expressed as a function of inflation expectations:

$$\begin{aligned} \pi_t &= \beta \hat{E}_t \pi_{t+1} + \kappa x_t + u_t, && \text{(5a)} \\ x_t &= \hat{E}_t x_{t+1} - \varphi(i_t - \hat{E}_t \pi_{t+1}), && \text{(5b)} \\ i_t &= \gamma_x \hat{E}_t x_{t+1} + \gamma_\pi \hat{E}_t \pi_{t+1} + \gamma_u u_t, && \text{(5c)} \\ \gamma_x &= 1/\varphi, && \text{(5d)} \end{aligned}$$

then,

$$\begin{aligned} \pi_t &= (\beta + \kappa\varphi(1 - \gamma_\pi)) \hat{E}_t \pi_{t+1} + (1 - \kappa\varphi\gamma_u) u_t, && \text{(6a)} \\ x_t &= \varphi(1 - \gamma_\pi) \hat{E}_t \pi_{t+1} + \varphi\gamma_u u_t. && \text{(6b)} \end{aligned}$$

This simplification allows us to consider only the first equation above in order to compute the agents' beliefs. As the coefficients $Q = 0$, $B = \beta + \kappa\varphi(1 - \gamma_\pi)$, and $C = 1 - \kappa\varphi\gamma_u$ are now scalars, the agents' beliefs, A and K, are scalars as well.

The agents update their beliefs in a constant gain RLS learning step,

$$\begin{aligned} \Psi_{t+1}^T &= \Psi_t^T + \gamma R_t^{-1} U_t \left(Y_t - \Psi_t^T \cdot U_t\right)^T \\ &= \Psi_t^T + \gamma R_t^{-1} U_t \left(\begin{bmatrix} Q + (B - I) A & \vdots & \rho C + (\rho B - I) K \end{bmatrix} \cdot U_t + C\epsilon_t \right)^T \\ &= \Psi_t^T + \gamma \left\{ \underbrace{R_t^{-1} \cdot U_t U_t^T \cdot \begin{bmatrix} Q + (\beta + \kappa\varphi(1 - \gamma_\pi) - 1) A \\ \rho C + [\rho(\beta + \kappa\varphi(1 - \gamma_\pi)) - 1] K \end{bmatrix} + R_t^{-1} U_t \epsilon_t C}_{\mathcal{H}} \right\} \end{aligned} \tag{7}$$

where γ is the (small) constant gain, and $\mathcal{H}$ denotes the right-hand-side of the agents' updating rule written in the form of a stochastic recursive algorithm (SRA). In general, (7) is a matrix equation; it is usually studied by vectorization so that $\theta = vec(\Psi)$ is a *vector* of beliefs. In the example above, Ψ is a vector, and thus $\theta = \Psi$.[1]

2.2 The Invariant Distribution of Beliefs

(Evans and Honkapohja, 2001, Theorem 7.9), following Benveniste et al. (1990), study the SRAs similar to (7) by deriving a linear approximating diffusion in the form

$$d\tilde{\theta}_t = D_\theta h\left(\overline{\theta}\right) \cdot \tilde{\theta}_t dt + \mathcal{R}^{1/2}\left(\overline{\theta}\right) \cdot dW_t, \tag{8}$$

where $\tilde{\theta}_n$ are centered and scaled beliefs, equal to $\gamma^{-1/2} \cdot \left(\theta_n - \bar{\theta}\right)$, and $\bar{\theta}$ is the stationary point of the deterministic part of (7) which corresponds to the RE equilibrium of the model[2]. $D_\theta h\left(\overline{\theta}\right)$ is the Jacobian of the approximating ODE, related to the deterministic part of $\mathcal{H}$, and $\mathcal{R}$ a matrix that depends on covariances among components of $\mathcal{H}$, see below. Using this linear approximation, they show that the beliefs θ converge to an invariant distribution

$$\theta \sim N\left(\overline{\theta}, \gamma\mathcal{C}\right), \tag{9}$$

[1]In what follows, we ignore the approximating diffusion for elements of the matrix of second moments R, as well as the corresponding elements of the matrices $D_\theta h\left(\bar{\theta}\right)$ and $\mathcal{R}$. This is because the model is similar to the one evaluated by Evans and Honkapohja (2001), in Section 14.4, with variance-covariance matrix of the invariant distribution being block-diagonal in agents' beliefs about coefficients θ and beliefs about elements of the second moments matrix R.

An alternative is to assume that the agents are using generalized stochastic gradient (GSG) rule and know the exact form of the equilibrium matrix of the second moments R, in which case the SRA for elements of R is not present.

Moreover, as demonstrated in Kolyuzhnov, Bogomolova, and Slobodyan (2014), even when the variance-covariance matrix of invariant distribution is not block-diagonal, ignoring this effect might lead to vanishingly small numerical errors.

[2]The formula (7) is a consequence of Theorem 2, Chapter 3, Part I, and the corresponding Theorem 15, Chapter 4, Part II, of Benveniste et al. (1990). Benveniste et al. (1990) use for $\tilde{\theta}_n$ a slightly different definition, $\gamma^{-1/2} \cdot (\theta_n - \theta(t_n))$, where $\theta(t_n)$ is the solution path of the approximating ODE with some initial point θ_0. Taking $\bar{\theta}$ as an initial point for the ODE solution path, we get $\theta(t_n) \equiv \bar{\theta}$, and the formula (7).

where

$$\mathcal{C} = \int_0^\infty e^{sD_\theta h(\overline{\theta})} \mathcal{R} e^{s\left[D_\theta h(\overline{\theta})\right]^T} ds.$$

This integral could be computed as a solution of a Lyapunov equation

$$D_\theta h\left(\overline{\theta}\right) \cdot \mathcal{C} + \mathcal{C} \cdot \left[D_\theta h\left(\overline{\theta}\right)\right]^T + \mathcal{R} = 0. \tag{10}$$

The necessary and sufficient condition for the invariant distribution (9) to exist is that all real parts of eigenvalues of $D_\theta h\left(\overline{\theta}\right)$ were negative, *i.e.*, just the E-stability condition (plus, other technical conditions, assumptions A.2, for example, of Chapter 2, Part I of Benveniste et al. (1990), section 2.2.3.1).

The Jacobian matrix $D_\theta h\left(\overline{\theta}\right)$ of the approximating ODE (3), evaluated at the REE $\overline{\theta}$ which is a stationary point of the (3),[3] is given by a 2×2 matrix

$$D_\theta h\left(\overline{\theta}\right) = \begin{bmatrix} \beta + \kappa\varphi\left(1 - \gamma_\pi\right) - 1 & 0 \\ 0 & \rho\left(\beta + \kappa\varphi\left(1 - \gamma_\pi\right)\right) - 1 \end{bmatrix} = \begin{bmatrix} B - 1 & 0 \\ 0 & \rho B - 1 \end{bmatrix}.$$

At the REE, $A = 0$ and $\rho C + \left[\rho\left(\beta + \kappa\varphi\left(1 - \gamma_\pi\right)\right) - 1\right] K = 0$. Thus, the first term of (7) equals zero identically, and we could write the rest as

$$\mathcal{H} = R^{-1} U_t \epsilon_t C = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_u^{-2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ u_{t-1} \end{bmatrix} \cdot C\epsilon_t = \begin{bmatrix} C \\ C\sigma_u^{-2} u_{t-1} \end{bmatrix} \epsilon_t.$$

The matrix $\mathcal{R}$ in the expression for approximating diffusion is calculated as

$$\mathcal{R}^{ij}\left(\overline{\theta}\right) = \sum_{k=-\infty}^{\infty} cov\left(\mathcal{H}^i\left(\overline{\theta}, Y_k^{\overline{\theta}}\right), \mathcal{H}^j\left(\overline{\theta}, Y_0^{\overline{\theta}}\right)\right).$$

In our example, $\mathcal{R}$ is evaluated to

$$\mathcal{R} = C^2 \begin{bmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_u^{-4} \sigma_u^2 \sigma_\epsilon^2 \end{bmatrix} = C^2 \begin{bmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_u^{-2} \sigma_\epsilon^2 \end{bmatrix},$$

[3]The updating equations for Ψ^T were written in a matrix form. We, hovewer, could always stack the elements of Ψ^T differently:

$$\theta = vec\left(\Psi^T\right) = \begin{bmatrix} A \\ K \end{bmatrix}.$$

In this simple example, where both A and K are scalars, the difference between θ and Ψ is not important. It will become relevant for the general policy rule case, where both A and K are vectors.

and the expression $e^{sD_\theta h(\bar\theta)}\mathcal{R}e^{s\left[D_\theta h(\bar\theta)\right]^T}$ equals to

$$\begin{bmatrix} \exp\left[(B-1)\,s\right] & 0 \\ 0 & \exp\left[(\rho B-1)\,s\right] \end{bmatrix}\cdot C^2\begin{bmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_u^{-2}\sigma_\epsilon^2 \end{bmatrix}\cdot\begin{bmatrix} \exp\left[(B-1)\,s\right] & 0 \\ 0 & \exp\left[(\rho B-1)\,s\right] \end{bmatrix}$$
$$= C^2\begin{bmatrix} \exp\left[2\,(B-1)\,s\right]\sigma_\epsilon^2 & 0 \\ 0 & \exp\left[2\,(\rho B-1)\,s\right]\cdot\sigma_u^{-2}\sigma_\epsilon^2 \end{bmatrix},$$

which after integrating over s from 0 to ∞ gives

$$\mathcal{C} = \frac{C^2\sigma_\epsilon^2}{2}\begin{bmatrix} \frac{1}{1-B} & 0 \\ 0 & \frac{\sigma_u^{-2}}{1-\rho B} \end{bmatrix}. \tag{11}$$

The mean dynamics convergence speed in this example equals $1-B$. This value is decreasing in B while both non-zero terms of the matrix $\mathcal{C}$ are increasing in B. Therefore, an *increase* in convergence speed leads to unambiguous *decrease* in the spread of the invariant distribution of agents' beliefs around its RE-consistent values. As we show in the next sections, making the distribution of the agents' beliefs tighter leads to lower welfare losses evaluated over the invariant distribution of beliefs. This channel operates in addition to the usual inflation-output gap trade-off which is present under RE.

3 Welfare Under Learning

3.1 Welfare Losses

Given a particular parameter draw from the invariant distribution of beliefs, the evolution of the state variable vector $Y_t = (\pi_t, x_t)^T$ is given as some VAR(1) process, which allows us to compute unconditional variances of π and x for this value of beliefs. In a typical monetary policy model, including ours, the welfare-theoretic loss function takes the form:

$$\mathbb{E}_0\left\{\sum_{t=0}^{\infty}\beta^t L_t\right\} \tag{12}$$

with a period welfare loss function defined as $L_t = \pi_t^2 + \lambda x_t^2$ for a suitably chosen non-negative weight λ, or, more generally, as a quadratic form over the state variable vector Y_t. The distribution of Y_t is a function of the agents' beliefs, while the beliefs are distributed as a multivariate normal variable.

We define the objective of the monetary policy to be an unconditional expectation of the period loss function over the invariant distribution of the agents' beliefs, or

$$W^{ID} = E^{ID}\left[\pi_t^2\right] + \lambda E^{ID}\left[x_t^2\right]. \tag{13}$$

We use ID to mean the invariant distribution here.

In the adaptive learning literature, in order to justify usage of approximating ODEs such as (3) or approximating diffusions such as (8), it is assumed that the evolution of the state vector Y happens on a much faster scale than the evolution of the beliefs. Therefore, to compute W^{ID}, we, first, evaluate the expected value of a quadratic form of Y for given beliefs (A and K in our example), and then compute the expectation of this object over the multivariate normal distribution of A and K. In general, it might be hard to expect a closed-form solution in cases more complex than presented above. We then need to rely on numerical results.

The welfare criterion W^{ID} is related to the literature on monetary policy, where objectives (12) and $W = E\left[\pi_t^2\right] + \lambda E\left[x_t^2\right]$ are taken to be interchangeable at the REE, Eusepi et al. (2012). We take this correspondence a step further and compute W^{ID} which is an expectation of W over (9). We compare our results to some recent studies discussing optimal monetary policy under learning using a standard welfare criterion (12), such as Molnar and Santoro (2014).

3.2 The welfare loss computation

Let's consider the system (6). Under learning, the inflation process is given by the ALM

$$\pi_t = BA + \rho\left(C + BK\right)u_{t-1} + C\epsilon_t, \tag{14}$$

with $B = \beta + \kappa\varphi\left(1 - \gamma_\pi\right)$ and $C = 1 - \kappa\varphi\gamma_u$. The inflation beliefs are derived from the PLM,

$$\begin{aligned} \pi_t &= A + Ku_{t-1} + \epsilon_t, \\ \hat{E}_t\pi_{t+1} &= A + \rho K u_{t-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} x_t &= \varphi\left(1-\varphi_\pi\right)\hat{E}_t\pi_{t+1} - \varphi\gamma_u u_t = \frac{B-\beta}{\kappa}\left(A+\rho K u_{t-1}\right) + \frac{C-1}{\kappa}u_t = \\ &= \frac{B-\beta}{\kappa}A + \rho\left(\frac{B-\beta}{\kappa}K + \frac{C-1}{\kappa}\right)u_{t-1} + \frac{C-1}{\kappa}\epsilon_t. \end{aligned}$$

Computation of W^{ID}, detailed in the Appendix, produces the following result:

Result 1 The value of the welfare criterion W^{ID} is given by

$$\begin{aligned} W^{ID} &= C^2\sigma_\varepsilon^2\left(1+\frac{\rho^2}{\left(1-\rho^2\right)\left(1-B\rho\right)^2}\right) \\ &+ C^2\sigma_\varepsilon^2\cdot\gamma\frac{B^2}{2}\left(\frac{1}{1-B}+\frac{\rho^2}{1-B\rho}\right) \\ &+ \lambda\times\left\{\frac{\sigma_\varepsilon^2}{1-\rho^2}\left[\left(\frac{C-1}{\kappa}\right)^2 + 2C\frac{B-\beta}{\kappa}\frac{C-1}{\kappa}\frac{\rho^3}{1-B\rho} + \left(\frac{B-\beta}{\kappa}\right)^2\frac{C^2\rho^4}{\left(1-B\rho\right)^2}\right]\right\} \\ &+ \lambda\times C^2\sigma_\varepsilon^2\cdot\gamma\frac{\left(\frac{B-\beta}{\kappa}\right)^2}{2}\left(\frac{1}{1-B}+\frac{\rho^2}{1-B\rho}\right). \end{aligned} \tag{15}$$

The first two lines in the above expression correspond to $E\left[\pi_t^2\right]$ taken over the invariant distribution, while the last two lines reflect the contribution of $\lambda E\left[x_t^2\right]$ to the welfare. The second and the fourth lines are proportional to the value of the constant gain γ and thus reflect the spread of the invariant distribution of beliefs, while the first and the third lines are equivalent to the result one would get by computing a weighted average of inflation and output gap variances under rational expectations.

Before we proceed to numerical derivation of the optimal monetary policy using the above expression, note that it simplifies significantly when the cost-push shock is *iid*, $\rho = 0$. In this case, W^{ID} reduces to

$$W^{ID} = \left(C^2 + \lambda\left(\frac{C-1}{\kappa}\right)^2\right)\sigma_\varepsilon^2 + \gamma C^2\sigma_\varepsilon^2\frac{B^2+\lambda\left(\frac{B-\beta}{\kappa}\right)^2}{2\left(1-B\right)}. \tag{16}$$

In this case in the limit of $\gamma \to 0$ (Rational Expectations), W^{ID} depends only on C (and thus the Taylor rule coefficient γ_u) but not B (Taylor rule coefficient γ_π). Therefore, in this limit the optimal value of γ_π is not defined. However, for any $\gamma > 0$, there is a non-trivial trade-off with respect to B and thus γ_π which delivers optimal value of γ_π that is independent of γ.

4 Comparisons with the literature

4.1 Molnar & Santoro

Optimal policy under learning has been studied before in the model we use, in particular by Molnar and Santoro (2014). They used exactly the same model as we do but utilized expression (12) for the purposes of deriving the optimal monetary policy. Their approach is limited to the case $\rho = 0$, as for non-*iid* mark-up shocks their solution algorithm becomes intractable. Optimal monetary policy derived by them, equation (6c), takes the same form as (5c-d). The only potential difference arises because in that model, the agents use so-called "steady-state learning", with only a constant entering the set of regressors used in their forecasting functions. In our example, they use a constant and past cost-push shock u_{t-1}, see (2). However, it turns out that for $\rho = 0$, the expression for W^{ID}, (16), is the same, whether the agents' PLM includes u_{t-1} or not.[4] Therefore, the model set-up described by (5) is identical to that of Molnar and Santoro (2014).

Before comparing our results, we have to note the following. As mentioned in the Section 3.2, for $\rho = 0$ our approach cannot obtain optimal coefficient γ_π in the RE limit of $\gamma \to 0$, while Molnar and Santoro (2014) are able to compute a well-defined coefficient δ_π^{AU} for all values of γ: as $\gamma \to 0$, δ_π^{AU} converges to the value obtained for optimal monetary policy under discretion by Evans and Honkapohja (2003), $\delta_\pi^{AU} = 1 + \frac{1}{\varphi}\frac{\beta\kappa}{\kappa^2+\lambda}$. Note, however, that when $\rho = 0$, the model (1) implies that the expected values of both inflation and the output gap under RE are identically zero; thus, the value δ_π^{AU} could be changed to an arbitrary number, and this will still deliver a monetary policy rule which prescribes the optimality condition under discretion, $\frac{\kappa}{\lambda}\pi_t + x_t = 0$. Therefore, there is an essential indeterminacy in the coefficients of the expectations-based Taylor rule under both RE *and* under adaptive learning as $\gamma \to 0$. The only reason a unique coefficient δ_π^{AU} is obtained by Molnar and Santoro (2014) for $\gamma \to 0$ is that their method of solution is similar to that of Evans and Honkapohja (2003), who combine equations of the model with the optimality condition under discretion. This method picks a unique set of Taylor rule coefficients out of a continuum that delivers the

[4] In general, this result does not have to be true: when the agents include into their PLM a variable that does not affect RE allocation, learning-related fluctuations in the beliefs about this variable will introduce additional variability in forecasts. This extra variability could translate into additional welfare costs.

same optimal allocation, described by $\frac{\kappa}{\lambda}\pi_t + x_t = 0$.

Our solution method is different: we explicitly compute the welfare criterion and then optimize with respect to γ_π and γ_u. In situations where the welfare is independent of γ_π, we do not have any additional mechanism for selecting among the resulting continuum. Note that this indeterminacy is not observed for the coefficient of the Taylor rule γ_u, and that our results are close to those of Molnar and Santoro (2014).

This essential indeterminacy is broken down when $\rho \neq 0$ and the expectations become a function of the current cost-push shock. However, in this case Molnar and Santoro (2014) are unable to derive any results, and thus we cannot compare them to our numbers.

On the other hand, as the constant gain γ increases, the spread of the invariant distribution grows, and the contribution to total welfare derived from this source[5] starts to dominate. This explains why for larger γ Molnar and Santoro (2014) results start converging to ours.

Table 1. Optimal values of γ_π/γ_u for model (1) when $\rho = 0$. First line: M&S, **Second line**: our results.

γ	λ	0.05	0.10	0.15	0.20	0.25
0.01		1.56/0.34	1.32/0.18	1.22/0.12	1.17/0.09	1.14/0.07
		2.03/0.34	**1.74/0.18**	**1.60/0.13**	**1.52/0.10**	**1.46/0.08**
0.05		1.84/0.36	1.55/0.20	1.42/0.13	1.34/0.10	1.29/0.08
		2.03/0.37	**1.74/0.20**	**1.60/0.14**	**1.52/0.12**	**1.46/0.10**
0.10		1.92/0.39	1.62/0.22	1.49/0.16	1.41/0.12	1.35/0.10
		2.03/0.40	**1.74/0.23**	**1.60/0.17**	**1.52/0.13**	**1.46/0.11**
0.20		1.95/0.45	1.67/0.27	1.53/0.20	1.45/0.16	1.40/0.14
		2.03/0.46	**1.74/0.28**	**1.60/0.21**	**1.52/0.17**	**1.46/0.15**
0.50		1.95/0.63	1.68/0.42	1.55/0.34	1.48/0.28	1.42/0.25
		2.03/0.62	**1.74/0.42**	**1.60/0.34**	**1.52/0.29**	**1.46/0.25**

The comparison above shows that despite utilizing different approaches to defining optimal monetary policy under adaptive learning, our paper achieves the results close to those of Molnar and Santoro (2014), especially regarding the coefficient of the Taylor rule reacting to the current value of the cost-push shock, γ_u. The optimal γ_π derived by our method is independent of the constant gain γ. Note that even in the limit of large γ, our approach calls for a more aggressive monetary policy as the Taylor rule inflation feedback

[5]The term proportional to $\frac{1}{1-B}\left[B^2 + \lambda\left(\frac{B-\beta}{\kappa}\right)^2\right]$.

coefficient γ_π is always larger.

This larger value of optimal γ_π is due to the beliefs spread. As mentioned already, the mean dynamics convergence speed in this model equals $1-B$ and thus is increasing in γ_π. More aggressive monetary policy leads to higher speed and smaller spread of the invariant distribution. This effect is represented by the term $1/(1-B)$ in (16). The term $B^2+\lambda\left(\frac{B-\beta}{\kappa}\right)^2$ in the same expression gives the basic trade-off between inflation and output gap variances. Presence of the spread-related term $1/(1-B)$ leads to minimum loss achieved at lower B, and thus higher γ_π and the mean dynamics convergence speed.

To further contrast our findings, we report results for $\rho = 0.5$.

Table 2. Optimal values of γ_π/γ_u for model (1) when $\rho = 0.5$

γ	λ 0.05	0.10	0.15	0.20	0.25
0.01	1.99/0.34	1.55/0.18	1.40/0.12	1.33/0.08	1.28/0.06
0.05	1.96/0.37	1.60/0.19	1.46/0.12	1.39/0.08	1.34/0.06
0.10	1.94/0.40	1.62/0.21	1.49/0.13	1.42/0.10	1.37/0.07
0.20	1.93/0.45	1.63/0.25	1.51/0.17	1.44/0.13	1.39/0.10
0.50	1.92/0.60	1.65/0.37	1.53/0.28	1.46/0.22	1.41/0.19

Comparing this to results reported in Table 1, it is obvious that persistent cost-push shock leads to less aggressive monetary policy (lower γ_π) being optimal. Optimal response to the cost-push shock (γ_u) also appears to be accentuated. As we mentioned previously, our approach appears dominant in the sense of being more analytically tractable, as we are capable of deriving the welfare criterion in a closed form (15) for arbitrary value of ρ. Our approach could be extended to richer model settings. The next subsection demonstrates how to derive the optimal monetary policy in a model utilized by Gaspar et al. (2011). Even more complex models could be handled easily without running into the "curse of dimensionality" problem inherent in the standard approach in this previous studies.

4.2 Gaspar, Smets & Vestin

Gaspar et al. (2011) (GSV in what follows) studied a question of optimal monetary policy in a model with intrinsic inflation persistence and a cost-push shock. As in Molnar and Santoro (2014), the agents in GSV are learning a simple PLM, consisting only of the past inflation. Similarly, the cost-push shock is *iid*. Given that the learning is not of a "steady-state learning" nature any longer, GSV do take into account of the second moment of the regressors present in the agents' forecasting rule (only inflation in this case).

In order to illustrate our approach, we study a problem similar to that of GSV. The law of motion for inflation is given by

$$\pi_t - \gamma\pi_{t-1} = \beta\left(E^*_{t-1}\pi_{t+1} - \gamma\pi_t\right) + \kappa x_t + u_t. \tag{17}$$

In contast to GSV, we take u_t to be an AR(1) process with coefficient of autocorrelation ρ and innovation ϵ_t. The agents' subjective expectations are formed at time $t-1$, not t as in GSV. We also assume that the Central Bank has direct control over the output gap x_t and uses it as a policy instrument, and looks for an optimal policy rule within the following class:

$$x_t = \delta_{\widehat{\pi}} E^*_{t-1}\pi_{t+1} + \delta_\pi \pi_{t-1} + \delta_u u_{t-1}.$$

This policy rule is similar to the one studied by GSV, who state that exact (nonlinear) optimal monetary policy rule is a function of u_t, π_{t-1}, agents' beliefs, and the second moment R_t, cf. their equation (32). Note, however, that in further discussion the dependence on R_t is partially ignored. In particular, they assume that the derivative of the value function with respect to R_t is zero. The loss function is given by

$$W^{ID} = E^{ID}\left[\partial\pi_t^2\right] + \lambda E^{ID}\left[x_t^2\right],$$

where $\partial\pi_t = \pi_t - \gamma\pi_{t-1}$, see GSV.

Plugging the assumed policy rule into (17), we obtain the following reduced form model

$$\pi_t = \underbrace{\frac{\beta+\kappa\delta_{\widehat{\pi}}}{1+\beta\gamma}}_{\alpha_1} E^*_{t-1}\pi_{t+1} + \underbrace{\frac{\gamma+\kappa\delta_\pi}{1+\beta\gamma}}_{\alpha_2}\pi_{t-1} + \underbrace{\frac{\rho+\kappa\delta_u}{1+\beta\gamma}}_{\alpha_3} u_{t-1} + \underbrace{\frac{1}{1+\beta\gamma}}_{\alpha_4}\epsilon_t.$$

Finally, writing the agents' PLM as

$$\pi_t = a\pi_{t-1} + bu_{t-1}, \tag{18}$$

we get the Actual Law of Motion (ALM) in the form

$$\pi_t = \left(\alpha_1 a^2 + \alpha_2\right)\pi_{t-1} + \left(\alpha_1 b\left(a+\rho\right) + \alpha_3\right)u_{t-1} + \alpha_4\epsilon_t. \tag{19}$$

Denoting the equilibrium of the model by $\left(\overline{a}, \overline{b}\right)^T$, we can see that the E-stability conditions are given as

$$\begin{aligned} 2\alpha_1 \overline{a} &< 1, \\ \alpha_1 \left(\overline{a} + \rho\right) &< 1. \end{aligned}$$

Out of the two solutions for $\overline{a}$, we select the one that is real and non-explosive, $|\overline{a}| < 1$.

We then proceed with the computations as in the Section 2.2 above. A significant difference is that the matrix $\mathcal{R}$ is not diagonal due to non-zero correlation between π_{t-1} and u_{t-1}. As a result, the variance-covariance matrix of the invariant distribution, $\mathcal{C}$, cannot be computed explicitly as in (11), and has to be obtained as a solution to the Lyapunov equation (10).

Final complication is that due to (19), variance of π_t and its covariance with u_{t-1} are now complex (and not just polynomial) functions of the beliefs a and b, making exact analytical computation of $E^{ID}\left[\pi_t^2\right]$ and $E^{ID}\left[x_t^2\right]$ impossible. There are two ways to approach the computation. One is to generate 1000 draws from the computed invariant distribution of beliefs for every grid point in the $(\delta_{\widehat{\pi}}, \delta_{\pi}, \delta_u)$ space, compute $E\left[\pi_t^2\right]$ and $E\left[x_t^2\right]$ for every draw using (19), and then average over the draws. Another is to use the delta method, which requires computing 2^{nd} and 4^{th} order derivatives of variances of pseudo-inflation difference $\pi_t - \gamma\pi_{t-1}$ and output gap x_t as a function of beliefs.[6] For the purposes of the comparison below, we use the delta method.

As in the previous section, we compare our results to GSV using $\rho = 0$ and the same limited PLM as that paper.[8] As our approach is not suffering from the curse of dimensionality, it could easily be extended to $\rho > 0$ or additional lags in the model. We can incorporate learning about the transmission mechanism where the Central Bank is using interest rate as a policy instrument, exactly as was done in the example considered in Molnar and Santoro (2014).

[6] We need to compute expectation of $\sigma^2_{\partial\pi}$ and σ^2_x over the invariant distribution of beliefs which are assumed to be normally distributed. Spelling out Taylor expansion of $\sigma^2_{\partial\pi}$ and σ^2_x, and noting that expectations of 1^{st} and 3^{rd} order terms are zero, we only compute the second and the fourth derivatives.

When $\rho = 0$, the belief b could be discarded, in which case $\sigma^2_{\partial\pi}$ and σ^2_x as well as the relevant derivatives could be written in a closed form[7].

For $\rho \neq 0$, we compute the derivatives, which are now matrix objects, numerically.

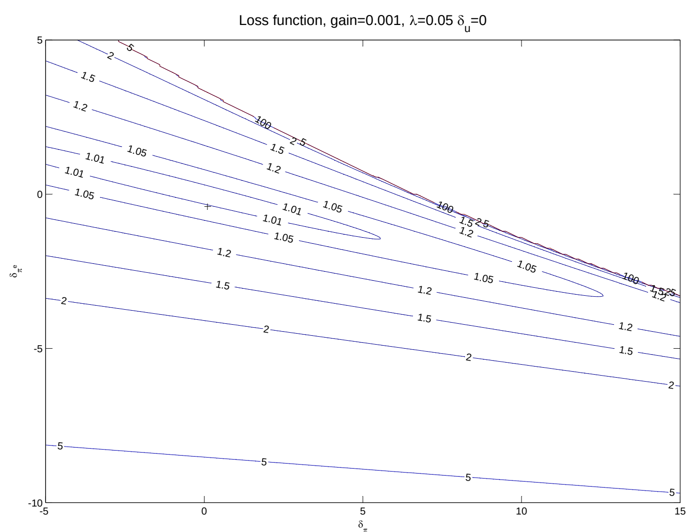
[8] In that paper, $\pi_t = a\pi_{t-1}$ was used as the agents' PLM, which is justifiable given assumption that u_t is an iid process.

We plot optimal choices of the policy parameters δ_π, δ_{π^e} and δ_u for different set of values of the gain and λ. We observe that for a low gain value of the learning algorithm, the loss function surface is flatter than it is for 0.02 gain parameter for example. Adaptive learning with a higher constant gain appears to tilt the loss function minimum significantly deeper.

Table 3. Optimal Monetary Policy for different values of the gain and λ

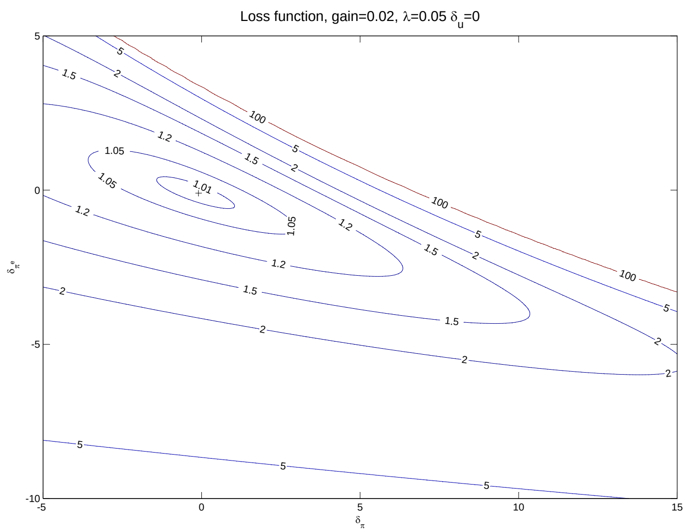
gain	λ	0.05	0.10	0.20	0.25
0.001		0.10/-0.40/-0.15	0.05/-0.20/-0.10	0.00/ 0.00/-0.05	0.00/ 0.00/-0.05
0.02		-0.10/-0.10/-0.20	-0.05/-0.10/-0.10	-0.05/ 0.00/-0.05	0.00/-0.10/-0.05
0.1		-1.30/ 0.90/-0.25	-0.80/ 0.60/-0.10	-0.50/ 0.40/-0.05	-0.40/ 0.30/-0.05

Figure 1: Contour Plot of the Loss Function



Note:

Figure 2: Contour Plot of the Loss Function



Note:

5 Welfare Losses and the Speed of Convergence

The policy parameters — γ_π and γ_u in the model of section 3, and $\delta_{\widehat{\pi}}$, δ_π, and δ_u in that of section 4 —in general have two effects on the welfare. The first one is a familiar trade-off between inflation and output gap volatility. In the simple NK model, more aggressive monetary policy (higher γ_π) leads to lower inflation variance but (in presence of cost-push shocks) higher output gap variance. The policymaker's weight on output gap stabilization in the objective function, λ, then determines the optimal values of policy parameters. This is a familiar story from optimal monetary policy literature.

The second effects is related to the spread of the agents' beliefs. Suppose that the equation (10) is not a matrix but a scalar. Then mean dynamics convergence speed equals $D_\theta h\left(\overline{\theta}\right)$. A lower speed means that $D_\theta h\left(\overline{\theta}\right)$, a negative number, becomes smaller in absolute value. Trivially, for any given $\mathcal{R}$, the variance $\mathcal{C}$ increases, which means that the invariant distribution of beliefs is more spread out for any value of the constant gain. In a model with zero steady-state inflation and output gaps non-zero average inflation and output gap are costly from the welfare perspective; thus, more spread out beliefs lead to higher average welfare losses. In the model of section 4.2, the effect is more subtle but similar: as the beliefs become more spread out, inflation occasionally becomes extremely persistent, leading to very high inflation and output gap variance and thus higher losses.

The argument in the matrix case is similar. A lower mean dynamics convergence speed, in general, will imply existence of at least one direction along which the variance-covariance matrix $\mathcal{C}$ is large, which leads to higher values of both inflation and output gap variances.

The interplay between these two effects then invariably leads towards selecting policy parameters that imply tighter beliefs distribution and thus a higher speed of convergence than the one which would obtain in a model without learning. This argument is in general valid for any objective function, (12) or (13). Under (12), a lower mean dynamics speed of convergence will lead, other things equal, to a larger response of the beliefs to a given forecasting error. This sensitivity is translated into higher losses following a forecasting error, which forces the model to be biased towards policy rules that are less sensitive to such errors — the ones with higher mean dynamics convergence speed. The fact that for large constant gain values both our approach and Molnar and Santoro (2014) are producing similar optimal policy clearly shows that the above argument is indeed valid for either (12) or (13).

6 Further Extensions and Conclusion

There is no obvious limit to the class of dynamic models which could be tackled using the approach proposed in this paper. Most models in the literature allow at least numerical computation of the REE, $D_\theta h\left(\overline{\theta}\right)$, and $\mathcal{R}$, and thus of the variance-covariance matrix of beliefs. $\mathcal{C}$. The computations are not more demanding than solving Lyapunov matrix equations. The evaluation of the welfare criterion (13) can then be performed either using the delta method, or by direct simulation and averaging as in section 4.2. The resulting optimal monetary policy thus can be obtained in a wide variety of empirically-relevant models, which would resist both Molnar and Santoro (2014) approach due to non-linearity, and Gaspar et al., 2011 direct attack due to the "curse of dimensionality".

Currently, we study only one-period-ahead expectations (so-called “Euler equation learning”. We plan to generalize our results to the anticipated utility framework in future work.

References

Benveniste, A., M. Métivier, and P. Priouret (1990). *Adaptive Algorithms and Stochastic Approximations*. Berlin Heidelberg New York: Springer Verlag.

Bullard, J. and K. Mitra (2002). Learning about monetary policy rules. *Journal of Monetary Economics 49*, 1105–1129.

Eusepi, S., M. Giannoni, and B. Preston (2012). Long-term debt pricing and monetary policy transmission under imperfect knowledge. *Federal Reserve Bank of New York Staff Reports 547*.

Eusepi, S. and B. Preston (2015). The science of monetary policy: An imperfect knowledge perspective.

Evans, G. W. and S. Honkapohja (2001). *Learning and Expectations in Macroeconomics*. Pricenton, Oxford: Princeton University Press.

Evans, G. W. and S. Honkapohja (2003). Adaptive learning and monetary policy design. *Journal of Money, Credit and Banking 35*, 1045–1072.

Ferrero, G. (2007). Monetary policy, learning and the speed of convergence. *Journal of Economic Dynamics and Control 31*, 3006–3041.

Gaspar, V., F. Smets, and D. Vestin (2006). Adaptive learning, persistence and optimal monetary policy. *Journal of the European Economic Association 4*, 376–385.

Gaspar, V., F. Smets, and D. Vestin (2011). Inflation expectations, adaptive learning and optimal monetary policy. In B. M. Friedman and M. Woodford (Eds.), *Handbook of Monetary Economics*, Volume 3B, pp. 1055–1095. Elsevier.

Molnar, K. and S. Santoro (2014). Optimal monetary policy when agents are learning. *European Economic Review 66*, 39–62.

Orphanides, A. and J. C. Williams (2007). Robust monetary policy with imperfect information. *Journal of Monetary Economics 54*, 1406–1435.

Orphanides, A. and J. C. Williams (2008). Learning, expectations formation, and the pitfalls of optimal control monetary policy. *Journal of Monetary Economics 55*, S80–S96.

Woodford, M. (2001). The taylor rule and optimal monetary policy. *American Economic Review 91*, 232–237.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy.* Pricenton, Oxford: Princeton University Press.

A Computing the Loss Function

Let's re-write (14) as

$$\pi_t = \begin{pmatrix} BA & \rho(C+BK) \end{pmatrix} \begin{bmatrix} 1 \\ u_{t-1} \end{bmatrix} + \epsilon_t = \left(\begin{pmatrix} 0 & \rho C \end{pmatrix} + B\tilde{\theta}^T \right) \begin{bmatrix} 1 \\ u_{t-1} \end{bmatrix} + C\epsilon_t.$$

Here the vector $\tilde{\theta}$ is two-dimensional, $\tilde{\theta}^T = \begin{pmatrix} A & \rho K \end{pmatrix}$, with the stationary point $\bar{\theta}^T = \begin{pmatrix} 0 & \rho^2 C/(1-\rho B) \end{pmatrix}$. Writing the expression for $E\left[\pi_t^2\right]$, we get

$$E\left[\pi_t^2\right] = C^2\sigma_\varepsilon^2 + E\left\{ \left(\begin{pmatrix} 0 & \rho C \end{pmatrix} + B\tilde{\theta}^T \right) \cdot \begin{bmatrix} 1 \\ u_{t-1} \end{bmatrix} \cdot (1 \; u_{t-1}) \left[B\tilde{\theta} + \begin{matrix} 0 \\ \rho C \end{matrix} \right] \right\}.$$

In the adaptive learning methodology, it is assumed that the evolution of the state variables (u, π, and x) is much faster than the evolution of beliefs which could be taken fixed for purposes of computing expectations with respect to the state variables. Therefore, if we talk about welfare averaged over invariant distribution of beliefs, we need to, first, compute the expected values of π_t^2 and x_t^2 as a function of the beliefs θ, and then average out the welfare over the invariant distribution of θ.

Computing $E\left[\pi_t^2\right]$ for given θ, we get

$$E\left[\pi_t^2\right] = C^2\sigma_\epsilon^2 + E\left\{ \left(\begin{pmatrix} 0 & \rho C \end{pmatrix} + B\tilde{\theta}^T \right) \cdot \Sigma \cdot \left(B\tilde{\theta} + \begin{bmatrix} 0 \\ \rho C \end{bmatrix} \right) \right\},$$

where

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}.$$

Next, we take expectations over the beliefs $\tilde{\theta}$. We get

$$E\left[\pi_t^2\right] = C^2\sigma_\epsilon^2 + (0,\rho C)\cdot\Sigma\cdot\begin{bmatrix}0\\ \rho C\end{bmatrix} + 2B\cdot E\left[\tilde{\theta}^T\right]\cdot\Sigma\cdot\begin{bmatrix}0\\ \rho C\end{bmatrix} \tag{20a}$$

$$+B^2\cdot E\left[\tilde{\theta}^T\cdot\Sigma\cdot\tilde{\theta}\right]. \tag{20b}$$

The second term in (20) equals $\rho^2C^2\sigma_u^2$, while the third is given by

$$2B\cdot\begin{pmatrix}0 & \rho^2C/(1-\rho B)\end{pmatrix}\cdot\begin{bmatrix}0\\ \rho\sigma_u^2\end{bmatrix} = 2B\sigma_u^2\frac{C^2\rho^3}{1-\rho B}.$$

The sum of the first three terms in (20), therefore, equals

$$C^2\sigma_\epsilon^2 + C^2\rho^2\sigma_u^2 + 2B\sigma_u^2\frac{C^2\rho^3}{1-\rho B} = C^2\sigma_\varepsilon^2\left(1+\frac{\rho^2}{1-\rho^2}\frac{1+\rho B}{1-\rho B}\right). \tag{21}$$

Computing (20), we get

$$B^2\cdot E\left[\tilde{\theta}^T\cdot\Sigma\cdot\tilde{\theta}\right] = B^2\cdot trace\left(\Sigma\gamma\mathcal{C}\right) + B^2\cdot E\left[\tilde{\theta}\right]^T\cdot\Sigma\cdot E\left[\tilde{\theta}\right] =$$

$$B^2\cdot trace\left(\Sigma\cdot\gamma C^2\frac{\sigma_\epsilon^2}{2}\begin{bmatrix}\frac{1}{1-B} & 0\\ 0 & \frac{\sigma_u^{-2}}{1-\rho B}\end{bmatrix}\right) + B^2\cdot\begin{pmatrix}0 & \rho^2C/(1-\rho B)\end{pmatrix}\cdot\Sigma\cdot$$

$$= \gamma C^2\frac{\sigma_\epsilon^2}{2}B^2\cdot\left(\frac{1}{1-B}+\frac{\rho^2}{1-B\rho}\right) + \frac{\rho^4C^2B^2}{(1-\rho B)^2}\sigma_u^2.$$

Finally, adding up the last term in the line above with (21), we get

$$E\left[\pi_t^2\right] = C^2\sigma_\epsilon^2\cdot\left(1+\frac{\rho^2}{(1-\rho^2)(1-\rho B)^2}\right) + \gamma C^2\frac{\sigma_\epsilon^2}{2}B^2\cdot\left(\frac{1}{1-B}+\frac{\rho^2}{1-B\rho}\right).$$

Proceeding in a similar fashion, we write

$$E\left[x_t^2\right] = \left(\frac{C-1}{\kappa}\right)^2\sigma_\epsilon^2 + E\left\{\left(\begin{pmatrix}0 & \rho\frac{C-1}{\kappa}\end{pmatrix} + \frac{B-\beta}{\kappa}\tilde{\theta}^T\right)\cdot\Sigma\cdot\left(\frac{B-\beta}{\kappa}\tilde{\theta} + \begin{bmatrix}0\\ \rho\frac{C-1}{\kappa}\end{bmatrix}\right)\right\} =$$

$$\left(\frac{C-1}{\kappa}\right)^2\sigma_\epsilon^2 + \left(0,\rho\frac{C-1}{\kappa}\right)\cdot\Sigma\cdot\begin{bmatrix}0\\ \rho\frac{C-1}{\kappa}\end{bmatrix} + 2\frac{B-\beta}{\kappa}E\left[\tilde{\theta}^T\right]\cdot\Sigma\cdot\begin{bmatrix}0\\ \rho\frac{C-1}{\kappa}\end{bmatrix} +$$

$$+\left(\frac{B-\beta}{\kappa}\right)^2\cdot E\left[\tilde{\theta}^T\cdot\Sigma\cdot\tilde{\theta}\right].$$

Using the expression for $E\left[\tilde{\theta}^T\cdot\Sigma\cdot\tilde{\theta}\right]$ from (22), multiplying by λ, and summing with $E\left[\pi_t^2\right]$ derived above, gives the formula for W^{ID} presented in the paper, equation (15).