The Response of Prices to Technology and Monetary Policy Shocks under Rational Inattention

Luigi Paciello*

Northwestern University

November 2007

(Job Market Paper)

Abstract

The speed of inflation adjustment to aggregate technology shocks is substantially larger than to monetary policy shocks. Prices adjust very quickly to technology shocks, while they only respond sluggishly to monetary policy shocks. This evidence is hard to reconcile with existing models of stickiness in prices. I show that the difference in the speed of price adjustment to the two types of shocks arises naturally in a model where price setting firms optimally decide what to pay attention to, subject to a constraint on information flows. In my model, firms pay more attention to technology shocks than to monetary policy shocks when the former affects profits more than the latter. Furthermore, strategic complementarities in price setting generate complementarities in the optimal allocation of attention. Therefore, each firm has an incentive to acquire more information on the variables that the other firms are, on average, more informed about. These complementarities induce a powerful amplification mechanism of the difference in the speed with which prices respond to technology shocks and to monetary policy shocks.

1 Introduction

I present a model that is consistent with the empirical evidence that prices respond much more quickly to technology shocks than to monetary policy shocks. I show that this response pattern arises naturally in a framework based on imperfect information with an endogenous choice of information structure similar to Sims (2003). In my model, the only obstacle that

[1] * I am particularly grateful to Martin Eichenbaum, Giorgio Primiceri and Mirko Wiederholt for continous comments, support and suggestions. I have also benefited from very useful discussions with Larry Christiano, Lars Hansen, and Eva Nagypal and seminar partecipants at Northwestern Univeristy. Any errors are my own. E-mail address: l-paciello@northwestern.edu. Tel.: +1-847-287-6520.

firms have when changing their prices is that they might not be well informed about the realizations of the shocks of the economy. The ability of a firm to adjust its price quickly to a particular shock depends on how well informed the firm is about the realization of that shock. The more attention a firm chooses to pay to a given shock, the more informed the firm is about the realizations of that shock. Similar to Sims (2003), I assume there is a limit on the total attention the firm can pay to the different shocks impacting on the economy. Therefore, if the firm allocates more attention to technology shocks, it must allocate less attention to monetary policy shocks. In my model, the firm will optimally choose to allocate more attention to those particular shocks that most reduce profits when prices are not adjusted properly. Since technology shocks affect profits more than monetary policy shocks, the firm will allocate more attention to technology shocks than to monetary policy shocks.

Other things being equal, this effect helps to rationalize the observed differential speed with which prices respond to technology shocks and to monetary policy shocks. However, this effect alone is not large enough to quantitatively account for the differential response. Fortunately, complementarities in price setting generate complementarities in firms' decision about which information to acquire[1]. These complementarities induce firms to acquire and process more information on the same variables that other firms are more informed about. The reallocation of attention in favor of technology shocks, and away from monetary policy shocks, generates a large amplification in the difference with which prices respond to technology shocks and to monetary policy shocks.

I choose the parameters governing firms' information processing capabilities such that the loss each firm faces from not being perfectly informed is a very small fraction of profits. The degree of strategic complementarity in price setting in my model is similar to the degree of strategic complementarity in price setting generally adopted in the large literature investigating the implications of price stickiness for the dynamics of macroeconomic variables[2]. As it turns out, under my assumptions, firms respond to technology shocks roughly as they would under complete information. In contrast, firms respond much more slowly to monetary policy shocks than they would under complete information.

There is a large empirical literature investigating how macroeconomic variables respond to monetary policy shocks. In this literature, there is substantial consensus that inflation

[1] Hellwig and Veldkamp (2007) theoretically study the role of strategic complementarities in information choices.

[2] See Woodford (2003) for a review.

responds slowly to monetary policy shocks[3]. A more recent literature investigates the effects of technology shocks using structural vector autoregression (SVAR) models. Papers in this literature consistently find that prices respond in general very quickly to technology shocks[4]. Paciello (2007) studies the differential speed in the adjustment of prices to technology and monetary policy shocks in the context of SVAR models using a variety of alternative identification schemes, sub-samples, and data from different countries. I show that the basic findings of the SVAR literature with respect to the difference in the speed with which prices respond to technology shocks and monetary policy shocks are very robust. The same patterns that hold for the United States also hold for Canada, France, Japan, and the United Kingdom. I argue that the SVAR results for the United States reflect a negative, and statistically significant, correlation between quarterly aggregate total factor productivity growth and different measures of aggregate inflation.

The different speed with which prices respond to technology shocks and to monetary policy shocks is not easy to reconcile with existing models of price stickiness. For instance, Smets and Wouters (2003, 2007) estimate a large-scale dynamic stochastic general equilibrium model with many nominal and real frictions, using U.S. and European data. In their paper, sticky prices are modeled using Calvo style time-dependent contracts. Smets and Wouters (2003, 2007) find that the response of prices to technology shocks is very similar to the response of prices to monetary policy shocks, in terms of speed of adjustment and persistence. In a related literature, other authors model nominal frictions as arising from the presence of menu costs[5]. These costs generate state-dependent pricing. In these models, firms can adjust prices any time they wish by paying a menu cost. To the best of my knowledge, the impact of menu costs has not been analyzed in an environment where there are both aggregate technology shocks and monetary policy shocks. In general, the frequency of response of prices to technology shocks will be large if these shocks are large. Once firms have paid the menu cost, they can adjust prices to all realized shocks. Therefore, if firms adjust prices very frequently to aggregate technology shocks, they will most likely adjust prices frequently to monetary policy shocks. Menu costs models would then have a difficult time in accounting for the different speed with which prices respond to technology and monetary policy shocks.

[3]See for example Christiano, Eichenbaum and Evans (1999).

[4]See for example Shapiro and Watson (1998) or Altig, Christiano, Eichenbaum and Linde (2005).

[5]See for example Gertler and Leahy (2006), Golosov and Lucas (2006), Midrigan (2006), Nakamura and Steinsson (2007).

The model I propose is related to Woodford (2002). Woodford (2002) uses an incomplete information model to explain the sluggish response of prices to nominal shocks. He argues that such a framework could potentially deliver a differential response of prices to aggregate supply shocks relative to nominal demand shocks, if firms were relatively more informed about the former than they were about the latter. However, he leaves open the question of why firms should choose to be relatively more informed about some shocks. Sims (2003) and Mackóviak and Wiederholt (2007) study the endogenous optimal choice of the information structure. In particular, Mackóviak and Wiederholt (2007) focus on the differential response of prices to aggregate nominal shocks versus idiosyncratic shocks in a framework with limited information-processing capabilities, and with an exogenous process for nominal spending. Firms use signals to set prices, but in their paper the signals are endogenous. Firms decide to be relatively more informed about idiosyncratic shocks because the latter have a larger impact on the profit-maximizing price. Furthermore, when firms pay limited attention to aggregate conditions, there is a lower incentive for other firms to pay attention to aggregate conditions. The model I propose differs from Maćkoviak and Wiederholt (2007) in at least two dimensions. The first difference is that I introduce two types of aggregate shocks. This assumption has important consequences, as it allows me to not only provide an explanation for the differential speed of adjustment of prices to such shocks, but it generates a large difference in the allocation of attention by price setters across shocks through complementarities in price setting. The second difference is that I embed the attention allocation problem in a more standard general equilibrium framework that captures the roles of different actors in influencing the differential responses of prices, with particular emphasis on the central bank.

The rest of the paper is organized as follows: in section 2, I introduce the main empirical facts that motivate the paper. In section 3, I describe the model. In section 4, I derive the solution of the attention allocation problem in the special case when the model is calibrated to be static. In section 5, I outline the numerical procedure to solve the model, choose the parameters, and comment on the results. Section 6 contains an analysis of the roles of complementarities, monetary policy, and information structure in the differential response of prices. Section 7 concludes.

2 Facts

Paciello (2007) investigates in details the responses of aggregate prices to monetary policy and technology shocks using SVAR models. Here, I report the results from the benchmark estimation procedure for the U.S. economy. I use a SVAR methodology to document the responses of aggregate prices to total factor productivity (TFP) shocks and monetary policy shocks. To this aim, I consider the following reduced form VAR:

$$Y_t = \Pi(L)Y_{t-1} + u_t,$$

where Y contains all the variables of interest, and $\Pi(L)$ is a lag operator of order p. The covariance matrix of the vector of reduced-form residuals, u_t, is Σ. The variables I include in the benchmark specification are the growth rate in labor productivity, the Federal Reserve Funds rate (FFR), the GDP deflator inflation, commodity inflation, the logarithm of per-capita hours worked, the logarithms of the ratios of consumption and investment to output, the logarithm of money velocity and the logarithm of labor productivity adjusted real hourly wages[6]. In this specification, the Federal Reserve Fund rate is the monetary policy instrument, although results hold for other choices of instruments too. The sample period is 1959:2 - 2007:2[7] and, based on the Akaike criterion, I choose the number of lags to be four, even if results are robust to different choices. Identification in the structural VAR literature amounts to providing enough restrictions to recover the decomposition of the estimated matrix of variance covariance of the reduced form VAR:

$$\Sigma = A_0 A_0'.$$

From this relationship and imposed restrictions, there is a unique mapping from u_t to the vector of orthogonal structural shocks, ϵ_t, such that $u_t = A_0\epsilon_t$. Once this map is defined, it

[6] This specification is similar to the one used by Altig, Christiano, Eichenbaum and Linde (2005), Francis and Ramey (2005). Results would be unchanged to more parsimonious specifications.

[7] The following variables were obtained from DRI Basic Economic Database. Nominal gross output is measured by GDPD, real gross output by GDPQ. Nominal investment is GCD (household durables consumption) plus GPI (gross private domestic investment). Nominal consumption is measured by GCN (nondurables) plus GCS (services) plus GCE (government consumption). Per capita hours worked are measured by LBMNU (Nonfarm business hours) divided by P16 (US population above 16). Real wages per capita are measured by LBCPU (nominal hourly non-farm business compensation) divided by the price index and P16. The price index is GDP/GDPQ. Commodity price index is an index over commodities available from DRI.

I obtained the Federal Funds rate (FedFunds) and M2 (M2MSL) from FRED. Monthly data were converted into quarterly frequency averaging over the quarter.

is possible to estimate the series of structural shocks and the responses of the variables into the system to such shocks. Since I am interested in two structural shocks, I only need to give conditions to define the mapping from u_t to the neutral and monetary policy technology shocks. I identify the column of A_0 relative to the neutral technology shocks through long run restrictions as in Gali (1999), using a property of standard neoclassical models, where the only type of shock having an impact on labor productivity in the long run is a permanent technology shock. The column of A_0 relative to the monetary policy shocks is identified as in Christiano et al (2003), relying on the assumption that the Federal Reserve set the monetary policy instrument after some other variables have been realized. This means that there is a subset of variables in Y, the ones in the Federal Reserve' feedback rule, to which the monetary policy shock is orthogonal. I therefore assume that all variables in the VAR enter the feedback rule except for the velocity of money.

The results presented in $Figure$ 1 show that a positive TFP shock has a sudden impact on the GDP deflator, with inflation dropping contemporaneously to the shock and then quickly converging to zero. In particular, a one basis point increase in TFP reduces prices on impact by approximately 0.35 basis points. The two standard deviations error bands confirm that this result is significant at a 5 percent significance level. On the inverse, the GDP deflator responds very slowly to a FFR shock, with the peak of the response taking place approximately twelve quarters after the shock. In particular, following a negative one basis point shock to the FFR, we have to wait approximately six quarters before inflation is positive and statistically different from zero. But even then, the magnitude of inflation is no larger than 0.08 basis points. Table 1 contains the variance decomposition of the forecast error for inflation in terms of fractions of total variance. The first result is that the TFP shock accounts for most of the variance of the forecast error of inflation for the first 10 quarters. The second is that on the inverse the monetary policy shock explains a marginal proportion. Hence technology shocks are a much more important determinant of the volatility of inflation than monetary policy shocks.

These findings support the thesis that prices respond much more quickly to a technology shock than they do to a monetary policy shock. Most of the adjustment to the former takes place along with the shock, while most of the response to the latter materializes several quarters after the shock. The differential speed of adjustment in prices is very large, suggesting that the two shocks generate very different incentives for firms to adjust their prices accordingly.

3 The model economy

I introduce a dynamic general equilibrium model with three types of actors: households, firms and central bank. Since I am interested in firms' price-setting behavior, I assume that these have limited information processing capabilities of the type suggested by Sims (2003). For tractability, I assume that households and central bank have complete information[8]. Households choose consumption, bond holdings, investments in physical capital, amount of working hours and capital services to supply to firms. The central bank sets nominal rates following a Taylor type rule. There is a constant return to scale production function common to all producers, which use labor, capital and intermediate inputs as factors of production. The only two exogenous shocks are an aggregate neutral technology shock and a monetary policy shock.

3.1 Households

The household side of the economy is modeled along the same lines as that of Smets and Wouters (2007). Households have complete information. They maximize expected discounted utility given by:

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\ln \left(C_t - bC_{t-1} \right) - \frac{\psi_0}{1+\psi_l} \; L_t^{1+\psi_l} \right), \tag{1}$$

where $\beta \in (0,1)$ is the discount factor, C_t is the households aggregate consumption, L_t denotes the household supply of labor, b is the coefficient defining the degree of habit persistence in preferences, ψ_0 and ψ_l determine respectively the level and the convexity of the disutility of labor. A complete set of Arrow-Debreu contingent securities, $V_{t+1}(\omega)$, is traded in the economy. The household budget constraint and the technology to accumulate capital at period t can be written as:

$$\begin{aligned} & P_t C_t + \frac{B_t}{R_t} + P_t I_t + \int_{\Omega} g_t(\omega) V_{t+1}(\omega) \, d\omega \\ = \; & B_{t-1} + P_t W_t L_t + P_t \left(r_t^k u_t - \Psi(u_t) \right) K_t + V_t + P_t \Pi_t, \end{aligned} \tag{2}$$

$$K_{t+1} = (1-\delta) K_t + \left(1 - S\left(\frac{I_t}{I_{t-1}} \right) \right) I_t, \tag{3}$$

[8] Adams (2007) uses similar assumptions of complete information on households and central banks to study the optimal monetary policy in an economy where firms have limited information processing capabilities.

where K_t is the stock of physical capital at the beginning of period t, u_t is the capital utilization rate so that $u_t K_t$ is the total service of capital at time t, I_t is the level of investments, R_t is the gross nominal interest rate on the risk free bonds B_t, W_t and r_t^k are respectively the real wage and the rental rate of capital in period t, Π_t is the dividend received from full ownership in the firms, P_t is the price of the unique final good of the economy and $g_t(\omega)$ is the set of prices of state contingent securities. The function $S\left(\frac{I_t}{I_{t-1}}\right)$ represents the installation (disinstallation) costs associated with accumulating (decumulating) stock of capital, and similarly to Altig, Christiano, Eichenbaum and Linde (2005) satisfies $S(1) = S'(1) = 0$, and $S''(1) > 0$. This captures the idea that installation costs are smaller for smoother growth rates in investments[9]. The cost of capital utilization is captured by the function $\Psi(u_t)$. As in Smets and Wouters (2007), I assume $u_t = 1$ and $\Psi(u_t) = 0$ on the non-stochastic balanced growth path.

Knowing the history up to time t, the household chooses the quantities $\{C_t, B_t, I_t, K_{t+1}, L_t, u_t\}$ and the optimal holdings of state contingent securities, $V_{t+1}(\omega)$, so to maximize the expected discounted utility in (1) subject to $(2) - (3)$.

The composite final good, Y_t, is a Dixit-Stiglitz aggregator over the set of differentiated goods indexed by z,

$$Y_t = \left(\int_0^1 Y_t(z)^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}, \tag{4}$$

where θ is the elasticity of substitution across different varieties. I assume that Y_t is aggregated by the household and can be used indifferently for consumption, investments or production as an intermediate input.

3.2 Monetary Policy

The monetary policy authority sets short term nominal interest rates, R_t, following a Taylor type rule described by:

$$\frac{R_t}{\bar{R}} = \left(\frac{R_{t-1}}{\bar{R}}\right)^{\rho_r} \left(\frac{1+\pi_t}{1+\bar{\pi}}\right)^{\phi_\pi} \left(\frac{1+y_t}{1+\bar{y}}\right)^{\phi_y} e^{\varepsilon_t^r}, \tag{5}$$

[9] Although capital adjustment costs do not play any role in the differential response of prices to the two aggregate shocks, they turn out to be important in order to have a drop in nominal rates, R_t, following a negative shock to ε^r.

where $\varepsilon_t^r \sim N(0, \sigma_r^2)$ is the *iid* shock to the policy rule, $\bar{R}$ and $\bar{\pi}$ are the non-stochastic steady state values of nominal interest rates and inflation, π_t is the inflation rate at t, y_t is the growth rate in real value added output[10] at time t, and $\bar{y}$ is the non-stochastic steady state value of output growth. Orphanides (2003b) has shown that a rule specified in terms of output growth is at least as well representative of the actual monetary policy in the United States as a rule specified in terms of output levels[11]. The reliance of information regarding growth rates, as opposed to natural-rate gaps, is also consistent with verbal descriptions of policy considerations and is easy to communicate, since output growth rates are usually used to describe the state of the economy. Orphanides and Williams (2003, 2006) also show that the rule expressed in terms of growth rates in output is to be preferred to the rule expressed in terms of levels of output, when the state of the economy, and, in particular, potential output are unknown. In such a case, a rule specified in differences reduces the volatility of inflation and output induced by errors in the perception of the output gap. Related to this argument is the fact that I am assuming the central bank has complete information, which means it perfectly observes current output growth and inflation. If I were to model the rule depending on the levels, I should have scaled the potential output level by the state of technology in order to have a stationary output gap, as in my model there is a non-stationary stochastic component of the technology process. In that case, assuming complete information on the side of the central bank would have implied that the central bank perfectly knows the current state of technology, which is arguable as sustained by Orphanides and Williams (2003, 2006). In contrast, the specification in terms of output growth requires the central bank to only observe current inflation and output growth, and to know their steady state values, which is equivalent to estimate a time trend. I believe that it is realistic to assume a central bank has enough information processing capabilities to implement such a rule.

3.3 Modeling the limited information capability

Here I introduce the tools used in this paper to model the limited information capability of firms. I need to define a measure to quantify the reduction in uncertainty coming from

[10] Real value added output is the sum of real aggregate consumption and investment, $C_t + I_t$.

[11] For example, similar to Justiniano and Primiceri (2005):

$$\frac{R_t}{\bar{R}} = \left(\frac{R_{t-1}}{\bar{R}}\right)^{\rho_r} \left(\frac{1+\pi_t}{1+\bar{\pi}}\right)^{\phi_\pi} \left(\frac{Y_t}{A_t}\right)^{\phi_y} e^{\varepsilon_t^r}$$

information processing. I build on the seminal work of Sims (2003) and use the concept of entropy to measure uncertainty in economic models. The larger is the entropy of a random variable, the larger is the uncertainty about its realizations. The entropy H of a stationary multivariate normally distributed random variable, $x^T = (x_1, x_2, ...x_T)$, equals:

$$H(x^T) = \frac{1}{2}\log_2\left[(2\pi e)^T \left|\Omega_{x^T}\right|\right],$$

where $|\Omega_{x^T}|$ is the determinant of the variance-covariance matrix of x^T. Therefore, a normal random variable has an entropy that depends only on the second moments of the distribution. Close to the definition of entropy is the definition of conditional entropy of $x^T = (x_1, x_2, ...x_T)$ given $s^T = (s_1, s_2, ...s_T)$:

$$H(x^T \mid s^T) = \frac{1}{2}\log_2\left[(2\pi e)^T \left|\Omega_{x^T|s^T}\right|\right],$$

where x^T and s^T must have a joint multivariate normal distribution, and where $\left|\Omega_{x^T|s^T}\right|$ is the determinant of the conditional-covariance matrix of x^T given s^T. I then define the reduction in uncertainty about a vector of multivariate normally distributed random variables x^T, from observing a vector of multivariate normally distributed random variables s^T, as the difference between the entropy of x^T and the conditional entropy of x^T given s^T :

$$\mathtt{I}(x^T; s^T) = H(x^T) - H(x^T \mid s^T).$$

This measure is called mutual information. I can then define the information flow between two stochastic processes as the average per period amount of information that one process contains about another process. If x^T and s^T are the first T realizations of the processes $\{x_t\}$ and $\{s_t\}$, then the information flow can be defined as:

$$\bar{\mathtt{I}}\left(\{x_t\}; \{s_t\}\right) = \lim_{T\to\infty}\frac{1}{T}\mathtt{I}(x^T; s^T). \tag{6}$$

In this paper, restricting information processing capabilities means restricting the average information processed by an agent per period. The information flow defined in (6) is the measure used for it. In the case of stationary multivariate normally distributed random variables the information flow reduces to:

$$\bar{\mathtt{I}}\left(\{x_t\}; \{s_t\}\right) = \lim_{T\to\infty}\frac{1}{T}\log_2\left(\frac{\left|\Omega_{x^T}\right|}{\left|\Omega_{x^T|s^T}\right|}\right),$$

and it is independent of the realizations of the signal process. When the process $\{s_t\}$ is completely uninformative about the realizations of the process $\{x_t\}$, as for example if $\{s_t\}$

is a constant, the conditional variance-covariance matrix is identical to the unconditional one, and the implied information flow is zero. When the process $\{s_t\}$ is perfectly revealing about the realizations of $\{x_t\}$, there is no more uncertainty about the latter, and $\left|\Omega_{x^T|s^T}\right|$ is zero, implying an infinite information flow. A process $\{s_t\}$ that is not fully revealing, but contains some information about the realization of $\{x_t\}$ will imply a finite and strictly positive information flow.

3.4 Firms

There is a continuum of Dixit-Stiglitz monopolistically competitive firms of mass one, and indexed by z. Each firm specializes in the production of a differentiated product. Like Basu (1995) and Nakamura and Steinsson (2007)[12], I assume that all products serve both as final output in consumption and investments, and as intermediate inputs into the production process of other products. Incorporating intermediate inputs into the production function increases the degree of strategic complementarity in price setting. Being that prices of intermediate inputs are directly linked to the aggregate price, the rigidity of prices to shocks is therefore amplified and transmitted to firms through rigidity of intermediate inputs prices. In this structure, there is no first product that is made without the use of other products[13]. Each firm z uses an index of intermediate inputs, $X_t(z)$, for production, which is, for simplicity, assembled by the household as in (4). The production function of firm z is then:

$$Y_t\left(z\right) = A_t\left(K_t\left(z\right)^{\alpha} L_t\left(z\right)^{1-\alpha}\right)^{1-\mu} X_t(z)^{\mu},$$

where $Y_t\left(z\right)$ is the gross output of firm z, A_t is the aggregate productivity variable common to all firms, which follows an exogenous stochastic process defined by:

$$\ln\frac{A_{t+1}}{A_t} = \gamma_a + \rho_a \ln\frac{A_t}{A_{t-1}} + \varepsilon^a_{t+1}, \tag{7}$$

[12] Basu (1995) and Nakamura and Steinsson (2007) apply this structure to a menu costs type model, obtaining a high degree of strategic complementarity in price setting. Furthermore, Nakamura and Steinsson (2007) show that this type of complementarities is well suited to explain the high rigidity of aggregate prices to demand shock, and the high frequency of price changes due to idiosyncratic productivity shocks.

[13] As sustained by Basu (1995) this is well representative of the U.S. economy: "Input-output studies certainly do not support the chain of production view, where goods move in only one direction down the stages of processing. Even the most detailed input-output tables show surprisingly few zeros. In its discussion of the 1977 input-output table, the BEA (1984 p. 50) notes that the table "shows heavy interdependence among industries. Seventy-six of the 85 industries shown in the table required inputs of at least 40 commodities, and 52 industries required inputs of at least 50 commodities."

where ε_{t+1}^{a} is normally distributed, $\varepsilon_{t+1}^{a} \sim N\left(0, \sigma_a^2\right)$, and is *iid* over time. $K_t\left(z\right)$ is the amount of capital services rent from households, and $L_t(z)$ is the labor input hired from the households by firm z. Total demand for good z, $Y_t\left(z\right)$, is:

$$Y_t\left(z\right) = Y_t \left(\frac{P_t\left(z\right)}{P_t}\right)^{-\theta},$$

where aggregate demand, Y_t, is:

$$Y_t = C_t + I_t + \int_0^1 X_t\left(z\right) dz + \Psi\left(u_t\right) K_t.$$

Each firm has three decisions to take at each period t. The firm has to choose the optimal price, $P_t\left(z\right)$, at which it is willing to sell any quantity demanded, and the optimal mix of inputs, both in terms of ratio of capital to labor, $k_t\left(z\right) \equiv \frac{K_t(z)}{L_t(z)}$, and in terms of ratio of intermediate inputs to the other factors of production, $x_t\left(z\right) \equiv \frac{X_t(z)}{K_t(z)^{\alpha} L_t(z)^{1-\alpha}}$. I assume there are three separate decision makers at each firm, one responsible for the choice of the selling price, one responsible for the optimal capital-to-labor ratio and one responsible for the intermediate-inputs ratio[14]. For tractability, the firm is not choosing the optimal basket of intermediate inputs, $X_t\left(z\right)$, which is assembled by the household[15] as in (4). Formally the problem of the price setter in each period t, at the firm z, is choosing $P_t\left(z\right)$ so to solve:

$$\max_{P_t(z)} E\left[\sum_{\tau=t}^{\infty} \Lambda_{\tau} \pi\left(P_{\tau}\left(z\right), k_{\tau}\left(z\right), x_{\tau}\left(z\right), \upsilon_{\tau}\right) \mid s_{zp}^{t}\right] \tag{8}$$

where Λ_{τ} is the discount factor[16] between period t and $t+\tau$, and $s_{zp}^{t} = \left\{s_{zp,1}, s_{zp,2}, ..., s_{zp,t}\right\}$ denotes the realization of the signal process up to time t for the price setter at firm z. Finally, $\upsilon_t = \left(Y_t, P_t, A_t, W_t, r_t^k\right)$ is the vector of realizations of the aggregate variables outside the control of firm z. The optimization problems of the other two decision makers are similar and therefore reported in appendix A. Up to this point, the decision problem at firm z is quite standard. Each agent makes an optimal decision conditional on its information set. If the information set contained all the realizations of current and past variables in the economy,

[14] This assumption is similar to the one used by Mankiw and Reis (2006). They assume that at each firm there is a price setting agent with incomplete information and an input decision maker with complete information. One difference is that I allow for incomplete information for each decision maker at each firm, but do not allow incomplete information on the household side.

[15] An equivalent assumption would be that there is a separate decision maker at each firm that assembles the basket of intermediate inputs in complete information.

[16] $\Lambda_{t+j} = \beta^j \frac{C_t - bC_{t-1}}{C_{t+j} - bC_{t+j-1}}$.

we would be in the conventional case considered in the literature on monopolistically competitive firms applied to macroeconomic models: firms would price with constant markups to nominal marginal costs, and the optimal input choice would be defined by the relative price of production factors. In such a case, it would make no difference whether there are three separate decision makers or only one, as choices are made on the basis of the same information set. In this paper, the information sets are endogenous. The optimal signal process $\{s_{zp,t}\}$ is chosen by the price setter in period zero and satisfies a constraint on the average flow of information,

$$\bar{\mathtt{I}}\left(\left\{P^{\dagger}_{a,t}(z), P^{\dagger}_{r,t}(z)\right\}; \{s_{zp,t}\}\right) \leq \kappa_p \tag{9}$$

where $\left\{P^{\dagger}_{a,t}(z), P^{\dagger}_{r,t}(z)\right\}$ is the vector of stochastic processes for the complete information optimal responses to the two aggregate shocks. The sum of these two processes delivers the optimal complete information price level, $P^{\dagger}_{t}(z) = P^{\dagger}_{a,t}(z) + P^{\dagger}_{r,t}(z)$. Therefore $P^{\dagger}_{t}(z)$ is the price level the price-setter at firm z would choose if she had complete information, or equivalently if $\kappa_p \to +\infty$. In addition to choosing the price level at any period t, in period zero the price setter at firm z solves the following problem:

$$\max_{\{s_{zp,t}\}\in S} E\left[\sum_{t=0}^{\infty} \Lambda_t \pi\left(P^{*}_{t}(z), k_t(z), x_t(z), \upsilon_t\right)\right] \tag{10}$$

subject to (9), where $P^{*}_{t}(z)$ solves (8) at each period t. The attention allocation problems for the other two decision makers are similar and reported in appendix A.

The three decision makers at each firm are indexed by $j = p$, k, x, indicating respectively the price setter, the decision maker for the capital-to-labor ratio and the decision maker for the intermediate-inputs ratio. Each decision maker is endowed with information processing resources that allow her to process on average κ_j bits of information per period[17]. The allocation of κ_j across separate decision makers is optimal, in the sense that the marginal value of additional information across the three agents at each firm is identical, and $\kappa = \sum_j \kappa_j$ is the total of information-processing resources at each firm; κ is chosen so that the overall marginal value of information at the firm level is very small[18], implying a relatively small

[17] In information theory, the flow of information is measured in bits. One bit is the flow of information necessary to completely reduce uncertainty about the realization of a discrete random variable with two equally likely outcomes. See Cover and Thomas (1991) for more details.

[18] In principle it would be an easy exercise to set up a cost function, or a market for information processing capabilities. But given there is no microeconomic empirical evidence on such a structure, it is equivalent to calibrate directly the equilibrium value of κ.

friction: firms would invest very few resources to acquire more information processing capabilities at the equilibrium. Intuitively, my model is equivalent to an organization structure where there are three separate managers at each firm, a marketing manager in charge of the price choice, a production manager in charge of the optimal mix of capital and labor, and a purchasing manager, responsible for the optimal level of intermediate inputs relative to the other factors of production. On top of the three managers, there is a CEO that allocates optimally the firm total information processing resources, κ, across the three managers in period zero. Each decision maker uses its information processing capability to acquire and process information on those variables that most matter for its choice. Although each decision maker maximizes the same profit function, the optimal choice of the variable she is in control of, depends potentially on different factors. For example the decision maker in charge of the price level has to process information on the impact its choice has on the relative demand of the firm. On the inverse, the two decision makers for the capital labor and intermediate-inputs ratios do not need direct information on demand, as they minimize the cost of production for any level of demand. I believe that the decision process at the firm level is a complex activity that involves many individuals, each of them in charge of a piece of the decision process[19]. Therefore distributing the decision powers across several individuals seems more realistic.

3.5 Restrictions on the set of signals for the benchmark model

I assume that signals cannot contain information about future realizations of shocks, ε_t^a and ε_t^r. This removes any forecasting power over shocks that have not yet been realized. This assumption is not controversial as long as exogenous shocks are assumed to be independent over time, and this is the case for this paper. Second, I restrict the signals to follow stationary Gaussian processes:

$$\{s_{zj,t}, \varepsilon_t^a, \varepsilon_t^r\} \text{ is a stationary Gaussian process.} \tag{11}$$

This assumption allows having a closed form expression for the information flow and facilitates the computation of a solution for the optimal signal structure, as it reduces to the

[19] Zbaracki, Levy and Bergen (2007) study the decision process for a price cut at a large manufacturing firm. They report that, although the reasons behind the price cut are understood and supported by all agents, the decision about how to do that is a very complex activity. Different individuals, in the same firm, use different economic models to make optimal choices, each consistent with their own objectives, but also each competing with the others.

choice of variance-covariance matrices[20]. I assume that firms acquire and process information about the two types of shocks separately. This means that the signal firm z receives at time t is a vector that can be partitioned into two subvectors, one containing information about $\{\varepsilon_t^a\}$ and one containing information about $\{\varepsilon_t^r\}$:

$$\{s_{zja,t}, \varepsilon_t^a\} \text{ and } \{s_{zjr,t}, \varepsilon_t^r\} \text{ are independent.} \tag{12}$$

This assumption is probably extreme, as in reality the two processing activities may have some overlapping, and hence there might be some learning about one shock by processing information about the other. I will relax this assumption later in the paper and show that not only results do still hold, but they are actually reinforced. Finally, I assume that all the noise in the signal is idiosyncratic, conveying the idea that all the information is available but the limited information processing capability generates idiosyncratic errors in the processing of available information.

3.6 Markets clearing conditions and resource constraint

In equilibrium, the markets for labor, capital and intermediate goods clear in each period t: *i)* $L_t = \int_0^1 L_t(z)\,dz$, *ii)* $u_t K_t = \int_0^1 K_t(z)\,dz$, *iii)* $X_t = \int_0^1 X_t(z)\,dz$. Also, the bonds and state contingent securities markets clear at each period t and state ω : $B_t = 0$, $V_t(\omega) = 0$. Finally, the resource constraint is satisfied in any period t :

$$Y_t = C_t + I_t + X_t + \Psi(u_t) K_t. \tag{13}$$

4 The solution to the static version of the model

In this section I solve a static version of the model introduced in section 3. This will provide useful insights and intuitions into how the attention allocation determines the differential speed of adjustment of prices to the two aggregate shocks, and how complementarities and monetary policy affect the responses of prices to these shocks. I impose α

[20] When the objective function is quadratic, this assumption is not binding, because Gaussian signals turn out to be optimal. See Maćkowiak and Wiederholt (2007) for a proof. I will obtain a quadratic objective with a second order Taylor expansion. Then the normality assumption is not very restrictive as long as such approximation is not a bad one.

equal to zero, so that labor and intermediate inputs are the only inputs in production: $Y_t(z) = A_t L_t(z)^{1-\mu} X_t(z)^{\mu}$. I also assume that the decision maker choosing the optimal intermediate-inputs ratio, x_t, has complete information. Therefore in this economy only the price setter faces an attention allocation problem. I impose that A_t is *iid*, hence $\ln A_t = \varepsilon_t^a$. I also assume no habit persistence in the utility function, $b = 0$. Finally, I restrict the monetary policy rule to be static, assuming $\rho_r = 0$, and to take the form:

$$\frac{R_t}{\bar{R}} = \left(\frac{P_t}{\bar{P}}\right)^{\phi_\pi} \left(\frac{C_t}{\bar{C}}\right)^{\phi_y} e^{\varepsilon_t^r},$$

where C_t is aggregate demand and coincides with real value added output, and the rule targets the deviation of the price level from steady state. I solve the model through a log-linearization around the non-stochastic steady state. The solution procedure for the attention allocation problem has two steps. In the first step I formulate a guess for aggregate prices and I solve for the dynamics of the model implied by the guess. In the second step I solve the attention allocation problem of the price setter, aggregate prices over firms, and then solve for the guess. The log-deviation of aggregate prices from steady state at time t is a linear function of the realizations of two *iid* shocks at time t, which are the sole state variables[21]:

$$\hat{P}_t = \eta_r \varepsilon_t^r + \eta_a \varepsilon_t^a.$$

The optimal price of firm z under complete information in log-deviations from steady state is given by:

$$\hat{P}_t^{\dagger}(z) = \hat{P}_t + \xi \hat{C}_t - (1 + \psi_l)\, \varepsilon_t^a, \tag{14}$$

where $\xi = (1 + \psi_l)(1 - \mu)$ is the degree of strategic complementarities in price setting, as defined in Woodford (2003)[22], and $\hat{C}_t$ is the log deviation of real demand from steady state[23]. A larger share of intermediate inputs in total costs, μ, implies a larger degree of strategic complementarity in price setting. Given the guess for aggregate prices and the solution for $\hat{C}_t$ in terms of the two fundamental shocks, I obtain a linear equation that links the log deviations of complete information optimal price, $\hat{P}_t^{\dagger}(z)$, to the two fundamental shocks:

$$\hat{P}_t^{\dagger}(z) = \left(\left(1 - \xi\frac{1+\phi_\pi}{1+\phi_y}\right)\eta_r + \vartheta_r\right)\varepsilon_t^r + \left(\left(1 - \xi\frac{1+\phi_\pi}{1+\phi_y}\right)\eta_a + \vartheta_a\right)\varepsilon_t^a, \tag{15}$$

where $\vartheta_r = -\frac{\xi}{1+\phi_y}$, and $\vartheta_a = \frac{\xi}{1-\mu}$. The shock ε^i, $i = a, r$, has an impact on the complete information price directly through parameter ϑ_i, and indirectly through the feedback from

[21] All variable with a hat are intended in log-deviations from the steady state.

[22] A larger ξ means a lower degree of strategic complementarity in price setting.

[23] See appendix C for details on these derivations.

aggregates prices. The magnitude of the latter is determined by the degree of strategic complementarities in prices, and the monetary policy rule. A larger degree of strategic complementarities, a lower ξ, implies everything else equal a larger feedback from aggregate prices. This is intuitive as more complementarities in price setting imply that the action of each price setter is influenced more by the average action of the other price setters.

In order to solve for the attention allocation problem in (10), I take a log-quadratic approximation of the sum of the discounted expected profits in (10), expressed in terms of log deviations from steady state. The optimal allocation of attention problem reduces to[24]:

$$\begin{aligned}
& \min_{\{s_{zp,t}\}\in S} \omega_1 E\left(\hat{P}_t^*(z) - \hat{P}_t^\dagger(z)\right)^2 \\
& s.t. \\
i) \quad : \quad & \hat{P}_t^\dagger(z) = \left(\left(1-\xi\frac{1+\phi_\pi}{1+\phi_y}\right)\eta_r + \vartheta_r\right)\varepsilon_t^r + \left(\left(1-\xi\frac{1+\phi_\pi}{1+\phi_y}\right)\eta_a + \vartheta_a\right)\varepsilon_t^a, \\
ii) \quad : \quad & \hat{P}_t^*(z) = E\left[\hat{P}_t^\dagger(z) \mid s_{zp}^t\right], \\
iii) \quad : \quad & \bar{\mathrm{I}}\left(\{\varepsilon_t^a, \varepsilon_t^r\}; \{s_{zp,t}\}\right) \leq \kappa_p.
\end{aligned} \tag{16}$$

In solving for the optimal signal process, the price setter minimizes the mean square error in price setting. Since the objective is quadratic, the optimal price choice in any period t, $\hat{P}_t^*(z)$, will be the projection of $\hat{P}_t^\dagger(z)$ on the realizations of the signal process up to time t. Under the restrictions on S in $(11)-(12)$, the signals take the form of true value plus noise,

$$s_{azp,t} = \varepsilon_t^a + \upsilon_a u_{zt}^a \tag{17}$$

$$s_{rzp,t} = \varepsilon_t^r + \upsilon_r u_{zt}^r \tag{18}$$

where u_{zt}^a and u_{zt}^r are *iid* normally distributed with zero mean and unitary variance. After some algebra, the attention allocation problem in (16) reduces to[25]:

$$\begin{aligned}
& \min_{\{\upsilon_a\geq 0, \upsilon_r\geq 0\}} \omega_1 \left[\frac{\left(\tilde{\xi}\eta_a + \vartheta_a\right)^2 \sigma_a^2}{1+\frac{\sigma_a^2}{\upsilon_a^2}} + \frac{\left(\tilde{\xi}\eta_r + \vartheta_r\right)^2 \sigma_r^2}{1+\frac{\sigma_r^2}{\upsilon_r^2}}\right] \\
& s.t. \\
i \quad : \quad & \left(1+\frac{\sigma_a^2}{\upsilon_a^2}\right)\left(1+\frac{\sigma_r^2}{\upsilon_r^2}\right) \leq 2^{2\kappa}
\end{aligned} \tag{19}$$

where I have defined for simplicity the variable $\tilde{\xi} \equiv 1 - \xi\frac{1+\phi_\pi}{1+\phi_y}$, which represents the degree of feedback from aggregate prices to individual firm complete-information optimal prices, and

[24] See Appendix A for details.

[25] See Appendix C for more details.

depends on the degree of complementarities and the monetary policy. The problem in (19) has a very intuitive interpretation. Firm z chooses the precision of each signal, υ_i, facing the constraint that the product of the two signal-to-noise ratios cannot exceed an upper bound coming from limited information processing capabilities. In the case of an interior solution, the optimal signal-to-noise ratio for each fundamental shock is given by:

$$1 + \frac{\sigma_a^2}{\upsilon_a^{*2}} = 2^{\kappa} \frac{\tilde{\xi}\eta_a + \vartheta_a}{\tilde{\xi}\eta_r + \vartheta_r} \frac{\sigma_a}{\sigma_r}, \tag{20}$$

$$1 + \frac{\sigma_r^2}{\upsilon_r^{*2}} = 2^{\kappa} \frac{\tilde{\xi}\eta_r + \vartheta_r}{\tilde{\xi}\eta_a + \vartheta_a} \frac{\sigma_r}{\sigma_a}, \tag{21}$$

A larger signal-to-noise ratio for a shock means being relatively more informed about that shock. The signal-to-noise ratios will be larger, the larger the upper bound on information flow, κ, is: the larger the information processing capability at each firm, then the smaller the firm's error as it processes any variable.

I use (20) – (21) and the fact that:

$$\hat{P}_t = \int_0^1 E\left[\hat{P}_t^{\dagger}(z) \mid s_{zp}^t\right] dz,$$

to solve for the fixed point, (η_a, η_r), and to determine the response of the aggregate price level to the fundamental shocks at an interior solution. The fixed point at an interior solution[26] is:

$$\eta_a = \vartheta_a \frac{\left(1 - \tilde{\xi} + \tilde{\xi} 2^{-2\kappa}\right) - 2^{-\kappa}\frac{1}{\gamma}}{\left(1 - \tilde{\xi}\right)^2 - 2^{-2\kappa}\tilde{\xi}^2}, \tag{22}$$

$$\eta_r = \vartheta_r \frac{\left(1 - \tilde{\xi} + \tilde{\xi} 2^{-2\kappa}\right) - 2^{-\kappa}\gamma}{\left(1 - \tilde{\xi}\right)^2 - 2^{-2\kappa}\tilde{\xi}^2}. \tag{23}$$

where

$$\gamma \equiv \frac{\vartheta_a}{\vartheta_r} \frac{\sigma_a}{\sigma_r}, \tag{24}$$

[26] The conditions for an interior solution are:

$$\begin{cases} \tilde{\xi} \leq \frac{1 - \frac{1}{\gamma} 2^{-\kappa}}{1 - 2^{-2\kappa}} & \text{if} \quad \gamma \leq 1 \\ \tilde{\xi} \leq \frac{1 - \gamma 2^{-\kappa}}{1 - 2^{-2\kappa}} & \text{if} \quad \gamma > 1 \end{cases}$$

The corner solutions are derived in appendix C.

is the parameter defining the relative impact of a shock on the loss function. The response of prices to the two shocks is proportional to the direct impact each shock has on the complete information optimal choice, represented by ϑ_i, $i = a, r$. A larger γ means that, everything else being equal, there is a larger impact of the technology shock on the objective function, and hence it is more costly to be uninformed about that shock. The larger κ, the more responsive aggregate prices are to the two shocks. As κ converges to infinity, the price responses converge to the complete information counterparts, $\frac{\vartheta_i}{1-\xi}$.

4.1 Complementarities and trade-off in attention allocation: the amplification mechanism

I derive an expression that links the relative precision of signals at an interior solution, $\dfrac{1+\frac{\sigma_a^2}{\upsilon_a^{*2}}}{1+\frac{\sigma_r^2}{\upsilon_r^{*2}}}$, to the coefficient γ, and to another coefficient, χ that I refer to as the *attention multiplier*:

$$\frac{1+\frac{\sigma_a^2}{\upsilon_a^{*2}}}{1+\frac{\sigma_r^2}{\upsilon_r^{*2}}} = \gamma^2\chi^2, \tag{25}$$

$$\chi \equiv \frac{1-\left(1-\xi\frac{1+\phi_\pi}{1+\phi_y}\right)\left(1+\frac{1}{\gamma}2^{-\kappa}\right)}{1-\left(1-\xi\frac{1+\phi_\pi}{1+\phi_y}\right)\left(1+\gamma 2^{-\kappa}\right)}. \tag{26}$$

For $\gamma > 1$, there is an *initial* incentive at the firm level to process more information on technology shocks because either they are more volatile, σ_a is larger than σ_r, or they have a larger impact on the complete information profit-maximizing price, ϑ_a is larger than ϑ_r. The *attention multiplier*, χ, will amplify or reduce the incentive to process more information on the technology shocks depending on the degree of strategic complementarity in price setting, ξ, and on the monetary policy rule. If the degree of strategic complementarity in price setting is large enough, or monetary policy is not too much more aggressive on inflation than it is on output, then there will be an amplification of the allocation of attention in favor of the shock that would already receive more attention, given the *initial* incentive implied by the value of γ. A larger degree of strategic complementarity in price setting, a smaller ξ, implies a larger feedback from aggregate prices to the firm level complete information optimal price, $\hat{P}_t^{\dagger}(z)$. This causes a larger difference in the allocation of attention as price-setters at each firm reallocate resources from one shock to the other, eventually making aggregate prices respond even more to technology shocks and even less to monetary policy shocks, and triggering new reallocations until the fixed point is reached. Therefore, through the positive feedback from aggregate prices, each price-setter has an incentive to allocate more resources to acquire information on the same type of shocks that other firms acquire more information on. This mechanism can potentially cause a large diversion of attention towards the technology shocks. In fact, the *attention multiplier*, χ, has no upper bound:

$$\lim_{\xi \to \iota^-} \chi = +\infty, \quad \forall \gamma > 1$$

where $\iota = \frac{\gamma 2^{-\kappa}}{1+\gamma 2^{-\kappa}} \frac{1+\phi_y}{1+\phi_\pi}$. This result is particularly important, as it implies that no matter how small the *initial* incentives to allocate more attention to the technology shock are, hence how close γ is to 1, it is always possible to have a large difference in the allocation of attention across the two shocks, by choosing a high enough degree of strategic complementarity in price setting. This is appealing as it implies that such a framework can naturally generate a very different response of aggregate prices to the two aggregate shocks, despite that in principle the impact of such shocks on the variability of the profit-maximizing price is very similar under complete information. This means that it can achieve a large difference in the responsiveness of prices to shocks when standard models of price stickiness cannot. For example, consider a case where γ is equal to 2 , and ϕ_y is equal to ϕ_π. If κ is equal to 1, then ι is 0.5. This means that a degree of strategic complementarity, $1 - \xi$, close to 0.5 would imply a multiplier, χ, close to infinity. If κ is equal to 3, then ι is 0.2, and then, for a degree of strategic complementarity close to 0.8, the *attention multiplier* would be close to infinity. These levels of strategic complementarities are not unreasonable if compared to those typically assumed in the literature on sticky prices[27].

The degree of strategic complementarity in price setting and the upper bound on the information processing capabilities are not the only determinants of the attention multiplier. The monetary policy has a central role too. In fact, a monetary policy authority more aggressive on prices, or less aggressive on output, reduces the differential allocation of attention, and the differential speed in price adjustment to the two shocks. For a given increase in prices, a more aggressive policy on prices, a larger ϕ_π, causes real rates to be larger and current real demand, C_t, to be smaller. Then, everything else being equal, a smaller change in C_t causes a smaller change in the complete information profit-maximizing price, $\hat{P}_t^\dagger(z)$, in (14). Therefore, the variability of $\hat{P}_t^\dagger(z)$ is reduced in response to each shock. However, this

[27] Woodford (2003) suggests a degree of strategic complementarity in price setting, $1-\xi$, between 0.85 and 0.9.

also reduces the difference in the variability of $\hat{P}_t^{\dagger}(z)$ due to the two shocks, which then feedbacks into the allocation of attention inducing a smaller difference in the attention allocation across the two shocks. Therefore, a more aggressive monetary policy on prices reduces the feedback from aggregate prices to firms level complete information profit-maximizing prices, $\hat{P}_t^{\dagger}(z)$, inducing lower complementarities in the allocation of attention. A similar argument holds for a less aggressive monetary policy on output.

In this section, I set ϕ_π equal to 1.5 and ϕ_y equal to 0.5. I assume that $\frac{\sigma_a}{\sigma_r}$ is equal to 1, and that μ is equal to 0.75. I also impose ψ_l equal to 1. This parameterization implies a value of γ equal to 6. The implied degree of strategic complementarity in price setting, $1-\xi$, is 0.5. At this value the feedback from aggregate prices to firm level complete information optimal prices, $\tilde{\xi}$ is positive at 0.17. In $Figure$ 2, I plot the price responses to the two shocks under rational inattention as a fraction of the response under perfect information, and expressed as a function of κ. The closer the fraction is to 1, the closer the price responses under rational inattention are to the ones under complete information. With low values of κ the firm will pay attention only to the technology shocks, ε_t^a, not responding at all to the monetary policy shocks, ε_t^r. As κ increases, the response to the technology shocks converges quickly to the complete information one, while the one to the monetary policy shocks has a much slower convergence. In $Figure$ 3, I plot the value of the *attention multiplier*, χ, as a function of κ. For small enough values of κ there is a corner solution in attention allocation. As κ increases the *attention multiplier* converges, as expected, to 1, but remains substantially large for intermediate values. In $Figure$ 4, I plot the attention multiplier χ, as a function of ξ, setting κ equal to 3. For low values of ξ, and therefore for large degrees of strategic complementarities in prices, the *attention multiplier* gets particularly large, pushing towards a corner solution where all the attention is allocated to the technology shocks. In $Figures$ 5 and 6, I plot the relative responses of prices to shocks to ε_t^a and ε_t^r as a function of both ξ and κ. A larger κ increases the relative responses of prices to both shocks, while a larger ξ reduces strategic complementarities in prices, and everything else being equal, increases price responses to both shocks. It has to be said that a value of γ equal to 6 is already a very large incentive to allocate more information processing resources to the technology shocks. This reduces the need for a particularly high degree of strategic complementarities in prices to generate a large difference in the allocation of attention across the two aggregate shocks. We will see, however, that this will not be the case in the "full-blown" dynamic model parameterized in section 5.

5 The numerical solution to the model

In this section I solve the dynamic model introduced in section 3 with numerical methods. In subsection 5.1, I describe the numerical routine, in subsection 5.2, I choose the parameters of the model and in subsection 5.,3 I comment the results of the attention allocation problems and the implied dynamics of aggregate prices.

5.1 The solution routine

I apply a two-step solution procedure[28]. In the first step I formulate a guess for the aggregate price, $\hat{P}_t$, a guess for the aggregate capital-to-labor ratio, $\hat{k}_t$, and a guess for the aggregate intermediate-inputs ratio, $\hat{x}_t$, all in log-deviations from the non-stochastic balanced growth path, and solve for the dynamics of the model economy.

In the second step, I solve for the optimal allocation of attention of each decision maker, given the processes for the endogenous variables of the model economy obtained in the first step. In order to solve each agent's attention allocation problem, I take a log-quadratic expansion of the sum of the discounted expected profits around the non-stochastic balanced growth path[29]. In order to save on space, I express the attention allocation problems of the three decision makers in terms of the variable $\hat{\zeta}_{j,t}(z)$, which I define in the following way:

$$\hat{\zeta}_{j,t}(z) \equiv \begin{cases} \hat{P}_t(z), & j = p \\ \hat{k}_t(z), & j = k \\ \hat{x}_t(z), & j = x \end{cases} .$$

The attention allocation problem for the decision maker choosing $\hat{\zeta}_{j,t}(z)$ at firm z, can be

[28] See Appendix B for more details.

[29] As discussed by Sims (2006) p. 161, and Maćkoviak and Wiederholt (2007) pp. 35-37, solving the attention allocation problem through a second order Taylor expansion of the objective function allows for a good approximation of the solution, as long as departures from complete information are not significant. At the value of κ considered in this paper, the marginal value of additional information is low at the firm level, implying potentially small departures from the solution obtained. See Appendix A for more details.

then expressed as:

$$\min_{\{s_{zj,t}\}\in S} \omega_j E\left(\hat{\zeta}^{*}_{j,t}(z)-\hat{\zeta}^{\dagger}_{j,t}(z)\right)^2 \tag{27}$$
$$s.t.$$
$$i) \ : \ \hat{\zeta}^{*}_{j,t}(z)=E\left[\hat{\zeta}^{\dagger}_{j,t}(z)\mid s^{t}_{zj}\right], \tag{28}$$
$$ii) \ : \ \bar{\mathrm{I}}\left(\left\{\zeta^{\dagger}_{aj,t}(z),\zeta^{\dagger}_{rj,t}(z)\right\};\{s_{zj,t}\}\right)\leq\kappa_j, \tag{29}$$
$$iii) \ : \ \hat{\zeta}^{\dagger}_{j,t}(z)=\zeta^{\dagger}_{aj,t}(z)+\zeta^{\dagger}_{rj,t}(z) \tag{30}$$

where $\omega_j > 0$, $\hat{\zeta}^{\dagger}_{j,t}(z)$ is the log-deviation from the non-stochastic balanced growth path of the optimal choice of $\zeta^{\dagger}_{j,t}(z)$ in the case of a perfectly informed decision maker j, and $\hat{\zeta}^{*}_{j,t}(z)$ is the projection of $\hat{\zeta}^{\dagger}_{j,t}(z)$ on the realization of signals for decision maker j, up to time t, and at firm z. The processes for $\left\{\zeta^{\dagger}_{aj,t}(z),\zeta^{\dagger}_{rj,t}(z)\right\}$ are obtained from the first step. I can then solve the attention allocation problems in $(27)-(30)$, obtaining the implied processes for aggregate prices, capital-to-labor ratio and intermediate-inputs ratio:

$$\begin{aligned}\hat{P}^{*}_{t} &= \int_0^1 \hat{P}^{*}_{t}(z)\,dz,\\ \hat{k}^{*}_{t} &= \int_0^1 \hat{k}^{*}_{t}(z)\,dz,\\ \hat{x}^{*}_{t} &= \int_0^1 \hat{x}^{*}_{t}(z)\,dz.\end{aligned}$$

I then update the guess and start again from the first step, iterating until convergence.

5.2 Calibration

I set the discount factor β equal to 0.99. The depreciation rate is equal to 0.025. The elasticity of value added output with respect to capital, α, is assumed to be 0.36, a value roughly consistent with observed income shares. I set the habit parameter b equal to 0.7, and the inverse of the Frisch's elasticity, ψ_l, equal to 1, similar to Altig, Christiano, Eichenbaum, and Linde (2005). I choose ψ_0 so that on the non-stochastic balanced growth path households supply an amount of labor equal to one. The dynamics of capital adjustment costs around the non-stochastic balanced growth path are shaped by the second derivative of the capital adjustment cost function evaluated at steady state, $S''(1)$. I set the capital adjustment cost parameter, $S''(1)$, equal to 5. This is larger than the value estimated by Altig, Christiano, Eichenbaum, and Linde (2005), but it is slightly smaller than the one obtained by Smets and Wouters (2007). The elasticity of the cost of capital utilization, $\sigma_\psi = \frac{\psi''(1)}{\psi'(1)}$, is set to 0.5, which is similar to the value estimated by Burnside and Eichenbaum (1996). I choose the elasticity of substitution across goods, θ, and the share of intermediate inputs in total costs, μ, following Nakamura and Steinsson (2007)[30]. Therefore, I set θ equal to 4, and μ equal to 0.75. From input-output tables relative to the U.S. economy, Nakamura and Steinsson (2007) estimate that the weighted average of the share of intermediate inputs in revenues is approximately 56 percent. Then, given the average markup implied by θ, the steady state share of intermediate inputs in total costs of production is 0.75.

The parameters in the Taylor rule, ρ_r, ϕ_π and ϕ_y, are obtained by estimating the rule[31] on the U.S. data from 1959:2 to 2007:2. I estimate the Taylor rule through an efficient GMM estimator similar to Clarida, Gali and Gertler (2000). The instruments set includes the four lags of r_t, π_t and y_t, and the four lags of inflation in commodity prices, of M2 growth and of the "spread" between the ten years and the three months U.S. treasury bonds[32]. Table 3 contains the results of the estimation with associated robust standard errors in parenthesis. Therefore, ρ_r, ϕ_π and ϕ_y are set equal to $0.96, 0.12$ and 0.2 respectively. The test of overidentifying restrictions rejects the null at one percent significance level. The autocorrelation coefficient, ρ_a, and the constant, γ_a, are chosen according to the estimates of an AR(1) process on an estimate of the U.S. quarterly growth rate in TFP[33], from 1959:2 to 2007:2. The estimated autoregressive coefficient cannot be statistically distinguished from zero, therefore I set $\rho_a = 0$. The standard deviations of the two shocks, σ_a and σ_r, are obtained respectively from the standard deviation of the U.S. quarterly growth rate in TFP, and from the standard deviation of the residual of the estimated Taylor rule, over the period

[30]"Berry et al. (1995) and Nevo (2001) find that markups vary a great deal across firms. The value of θ I choose implies a markup similar to the mean markup estimated by Berry et al. (1995) but slightly below the median markup found by Nevo (2001). Broda and Weinstein (2006) estimate elasticities of demand for a large array of disaggregated products using trade data. They report a median elasticity of demand below 3. Also, Burstein and Hellwig (2006) estimate an elasticity of demand near 5 using a menu cost model. Midrigan (2005) uses θ= 3 while Golosov and Lucas (2006) use θ= 7."

[31]The equation I estimate is:

$$r_t = c + \rho_r r_{t-1} + \phi_\pi \pi_t + \phi_y y_t + u_t^r,$$

where r_t is the Federal Fund rate, π_t is the log-difference in the GDP price deflator, and y_t is the deviation of the growth rate of output from a linear trend.

[32]Quarterly measures were computed averaging over months.

[33]Fernald (2007) estimates a quarterly series for the U.S. TFP growth rate trough a Solow residual accounting technique similar to Basu, Fernald, Kimball (2004).

1959:2-2007:2. The standard deviation of the U.S. quarterly growth rate in TFP is about 4 times the standard deviation of the residual from the Taylor rule[34]. In mapping the estimated standard deviation of the TFP growth rate to the standard deviation of the technology shock in the model, I have to adjust for the fact that the TFP growth rate has been estimated according to a model with a value added production function with no intermediate inputs[35]. Therefore, I need to scale the standard deviation of the estimated TFP growth rate by $1-\mu$. Since μ has been set equal to 0.75, the ratio of standard deviations of shocks in the model, $\frac{\sigma_a}{\sigma_r}$, is set equal to 1. Finally the total information processing capabilities at the firm level, κ, is chosen so that in equilibrium the loss each firm faces from not being completely informed is a relatively small fraction of profits. Hence I choose κ equal to 4.

5.3 Results

In $Figure$ 7, I plot the responses of inflation and output to a one basis point shock to ε^a and ε^r in the model under complete information, $\kappa \to +\infty$. Not surprisingly, almost all of the adjustment in prices to ε^a takes place in two quarters, while all of the adjustment in prices following the shock to ε^r takes place in the period of impact of the shock. Under complete information, in fact, a one basis point positive shock to ε^a reduces prices by about 12 basis points on impact, and about 11 basis points after two quarters. A one basis point negative shock to ε^r increases prices by approximately 8.5 basis points along with the shock. Since the relative standard deviation of ε^a and ε^r is set equal to 1, and given that the impact of a technology shock on the complete information aggregate price level is larger than the impact of an equally sized monetary policy shock, there is an *initial* incentive for the firm to pay more attention to technology shocks than to monetary policy shocks, but such an incentive is relatively small. Under complete information, in fact, the long-run impact of a one basis point shock to ε^a on prices is about 30 percent larger than the long-run impact of a one basis point shock to ε^r. Intuitively this *initial* incentive is the dynamic counterpart of the variable γ I derived in the static version of the model. Therefore, if this model has to generate a large differential in the response of prices to the two shocks, it must come from the *attention multiplier*. Given that the monetary policy authority is substantially more aggressive on output than it is on inflation, and that the share of intermediate inputs in total costs, μ, is

[34] I obtain similar results if I use the standard deviations of the estimated TFP and monetary policy shocks from the VAR.

[35] See Appendix D for details.

0.75, the feedback from aggregate prices to firm level complete information profit-maximizing price, $\hat{P}_t^{\dagger}(z)$, is substantial, inducing a large *attention multiplier.* In $Figure$ 8, I plot the impulse responses of output and inflation in the model with limited information processing capabilities, with κ equal to 4. Prices adjust quickly to the ε^a shock, with almost all of the adjustment taking place in the first two quarters. In contrast, prices adjust very sluggishly to the ε^r shock, inducing a large real effect of the monetary policy shock. The response of output to the ε^r shock is very persistent and takes many quarters to converge to zero. The optimal allocation of κ across the different decision makers is such that 50 percent of κ is allocated to price decision maker, 33 percent is allocated to the intermediate-inputs ratio decision maker and the remaining to the capital-to-labor ratio decision maker. The price decision maker allocates almost all of its information processing resources to the technology shocks. The other two decision makers allocate similar resources to the technology and monetary policy shocks, as the capital-to-labor ratio and the intermediate-inputs ratio have similar impacts across the two shocks on the variability of profits. At equilibrium the marginal value of additional information processing resources at the firm level is small. Each firm faces a loss that is in the order of 1/1000 of its discounted sum of non-stochastic balanced growth path profits, where the loss is computed relative to the case the firm had complete information, $\kappa \to \infty$, and everything else being equal.

6 Complementarities, monetary policy and signals structure

In this section I investigate the roles of strategic complementarity in price setting, monetary policy and the role of restrictions on the signals space for the results obtained above. Lastly, I discuss potential extensions and shortcomings.

6.1 The role of complementarities

I reduce the share of intermediate inputs in total steady state costs, μ, from 0.75 to 0.5. A value of μ equal to 0.75 implied in section 4 a value of ξ equal to 0.5. With μ dropping to 0.5, ξ increases to 1, and the degree of strategic complementarity in price setting is substantially reduced. To have an idea of how low complementarities are in this model, it helps to consider the fact that Woodford (2003) recommends a value of ξ between 0.15 and 0.1 in models of sticky prices.

With a smaller degree of strategic complementarity in price setting there are two effects that reduce the difference in the allocation of attention across the two shocks for the price setter. The first is a direct effect that goes through the reallocation of attention at the price setter level: smaller complementarities in price setting induce smaller complementarities in the allocation of attention across price setters, and, everything else being equal, reduces the differential in the allocation of attention across shocks for each price setter. The second effect relates to the reallocation of information processing resources, κ, at the firm level: smaller complementarities in price setting induce a larger variability of aggregate prices in response to aggregate shocks, and, everything else being equal, increase the incentive to allocate more resources to process information about prices than to process information about the capital-to-labor ratio and the intermediate-inputs ratio. At the equilibrium, 75 percent of total information processing resources, κ, is allocated to the price setter, 16 percent is allocated to the intermediate-inputs ratio decision maker and the remaining to the capital-to-labor ratio decision maker. The price setter allocates 57 percent of its attention to technology. Therefore the differential in attention allocation across the two shocks is substantially smaller for this decision maker relative to the case with larger complementairity in price setting. Capital-to-labor and intermediate-inputs ratios will not be very responsive to the monetary policy shocks, pushing the respective decision makers to allocate almost all of their attention to the technology shocks. In $Figure$ 9, I plot the responses of inflation to a one basis point shock to ε^a and ε^r, in the model with and without limited information processing capabilities. Now, the adjustment of prices to ε^r takes place in two quarters, and therefore, the real effect of the monetary policy shocks are small and inflation is not very persistent. Most of the adjustment in prices to ε^a takes place in two quarters, similar to the benchmark calibration. Since $\frac{\sigma_a}{\sigma_r}$ is set equal to 2, the *initial* impact of a one basis point shock to ε^a on profits is about 30 percent larger than the *initial* impact of a one basis point shock to ε^r, and therefore the price setter's *initial* incentive to allocate more attention to technology is similar to the benchmark calibration[36]. However, with low complementarities in price setting, the *initial* incentive does not get amplified and the difference in the allocation of attention across the two shocks will be small.

[36]By the same argument in section 5.2, lowering μ to 0.5, causes the relative standard deviation of shocks, $\frac{\sigma_a}{\sigma_r}$, to be set equal to 2.

6.2 The role of monetary policy

In this paragraph I modify the parameterization of the monetary policy rule defined in (5). First, I decrease ϕ_y from 0.2 to 0.1, making the monetary authority less aggressive on output growth. A less aggressive monetary policy on output growth affects the speed of adjustment of prices to the two aggregate shocks mainly through three channels. The first two channels have to do with the allocation of attention at the price-setter level, the third channel is related to the allocation of attention decision at the firm level. A less aggressive monetary policy on output growth, reduces the variability of prices following technology shocks, therefore reducing the *initial* incentive to allocate attention to these shocks. The drop in ϕ_y is so large that the price setter has an *initial* incentive to allocate more attention to monetary policy shocks, being prices relatively more volatile to those shocks than to technology shocks under complete information. The second impact of the lower ϕ_y is on the price setter's allocation of attention that takes place through a lower *attention multiplier:* a less aggressive monetary policy on output growth reduces the feedback from aggregate prices to firms level complete information profit-maximizing prices, inducing lower complementarities in the allocation of attention. Finally there is a reallocation of information processing resources at the firm level: more resource devoted to process information on the variability of prices to monetary policy shocks cause prices to be more responsive to such shocks, and, as a consequence, the capital-to-labor ratio and the intermediate-inputs ratio to have a smaller variability following monetary policy shocks, as the economy dynamics following these shocks are closer to the complete information counterparts. It follows that the price decision becomes relatively more important than the other two decisions. Therefore more information processing capabilities are allocated to the price setter. In fact, 76 percent of κ goes to the price setter, 17 percent goes to the intermediate-inputs decision maker, and the residual goes to the capital-to-labor ratio decision maker. The price setter allocates 58 percent of her information processing capabilities to monetary policy shocks, and 42 percent to technology shocks. Unlike the benchmark specification, there is no substantial amplification in the differential allocation of attention in favor of the monetary policy shocks. In $Figure$ 10, I plot the responses of inflation to a one basis point shock to ε^a and ε^r, in the model with and without limited information processing capabilities. The speed of adjustment of prices to the two aggregate shocks is similar, with prices adjusting slightly more quickly to ε^r than to ε^a.

In the second modification to the parameterization of the Taylor rule, I increase ϕ_π from 0.12 to 0.3, holding the other parameters at the values of the benchmark calibration. This induces a more aggressive monetary policy on inflation, which reduces, everything else being equal, the variability of prices to any shock. The impact on the allocation of attention is similar to the impact caused by a decrease in the degree of strategic complementarity in price setting. The first direct effect goes through the reallocation of attention at the price setter level: a more aggressive monetary policy on inflation reduces the feedback from aggregate prices to firms level complete information profit-maximizing prices, inducing lower complementarities in the allocation of attention, and therefore more attention devoted to monetary policy shocks relative to the benchmark parameterization. The second, indirect, effect relates to the reallocation of information processing resources at the firm level: when price setters allocate more resources to process information on the variability of prices following monetary policy shocks, aggregate prices become more responsive to such shocks, and, as a consequence, the capital-to-labor ratio and the intermediate-inputs ratio become less responsive to monetary policy shocks. It follows that the price setter is allocated more resources relative to the benchmark parameterization. In contrast to the case in which I changed ϕ_y, a change in ϕ_π has no substantial impact on the *initial* incentive for the price setter to allocate more attention to technology shocks. As in the benchmark parameterization, there is an *initial* incentive to allocate more resources to process information about technology shocks, but here there is no large amplification of this incentive through the *attention multiplier.* At the solution, the price setter allocates 56 percent of her information processing capabilities to the technology shocks. Similarly to the case of a less aggressive monetary policy on output growth, the capital-to-labor and intermediate-inputs ratios will be not very responsive to the monetary policy shocks, and hence the respective decision makers will allocate almost all of their attention to technology shocks. In equilibrium, 64 percent of κ is allocated to the price setter. In $Figure$ 11, aggregate prices have similar speeds of adjustments to the two aggregate shocks.

The ability of the model to generate the differential response in prices depends, then, on the fact that the monetary policy, estimated over the sample period 1959:2-2007:2, has been relative more aggressive on output growth than it has been on inflation. When I estimate the same policy rule on the sub-sample[37] from 1979:3, to 2007:2, I obtain that the monetary

[37] This sub-sample includes period from the Volcker's presidency to 2007.

policy has been relatively more aggressive on inflation than it has been on output growth[38]. In such a case the model would imply a much smaller amplification of the differential in attention allocation. However, this does not mean necessarily the model would not capture the differential response on prices in that sub-sample: the estimated standard deviation of the TFP shock is, in fact, about 8.5 times the estimated standard deviation of the monetary policy shock over that sub-sample. The differential response of prices would, then, be driven more by the relative standard deviations of the shocks and less by the *attention multiplier*.

6.3 The role of the signal structure

So far I have assumed that attending to technology and monetary policy shocks are separate activities. This means that each decision maker is always able to distinguish between the two types of shocks. In this section I investigate what happens in the static version of the model of section 4, when I remove the independency assumption in (12). Specifically, suppose that the price-setter at firm z can choose signals of the form:

$$s_{zp,t} = \begin{cases} \hat{C}_t + \nu_c u^c_{z,t} \\ \hat{P}_t + \nu_p u^p_{z,t} \\ \hat{R}_t + \nu_r u^r_{z,t} \\ \hat{L}_t + \nu_l u^l_{z,t} \end{cases}, \tag{31}$$

where u^j_t is assumed to be *iid* and normally distributed with zero mean and unitary variance. In contrast to the signals in $(17) - (18)$, the signals in (31) have the property that each signal contains information about both technology and monetary policy shocks. This signal structure conveys the idea that each decision maker processes information coming from signals based on realizations of variables that are actually available in the real world.

The price setter at firm z solves the attention problem in (19), by choosing the precision of each signal, ν_j, $j = c, p, r, l$, in (31), and subject to the information flow constraint:

$$\kappa_p \geq \bar{\mathrm{I}}\left(\left\{\hat{P}^{\dagger}_t(z)\right\}; \{s_{zp,t}\}\right). \tag{32}$$

[38] The estimated parameters on the sub-sample 1979:3-2007:2 are:

$$\rho_r = 0.86$$
$$\phi_\pi = 0.37$$
$$\phi_y = 0.14$$

In *Figure* 12, I plot the price responses to the two shocks under rational inattention, and relative to perfect information as a function of κ. For large values of information processing capabilities, κ, firms make very small mistakes in setting their prices relative to the complete information optimal choices. In this economy, the complete information profit-maximizing price for each firm, $\hat{P}_t^{\dagger}(z)$, coincides with the aggregate price, since, except that for the realization of the signals, firms are identical. Therefore, for large κ, the error coming from limited information processing capabilities is relatively small and the aggregate price is a very good statistic for the optimal firm level price. Firms will basically acquire and process information almost only on aggregate prices. As κ decreases, the errors firms make in setting prices increase, and the aggregate price becomes a less valuable statistic for the optimal price. Hence firms will increase the precision of the other signals relative to the one for the aggregate price. In particular, they will process relatively more information about aggregate demand, which appears directly in the equation for complete information profit-maximizing price in (14). There is, however, a characteristic of aggregate demand that makes it very different from aggregate prices: while the covariance of aggregate prices with $\hat{P}_t^{\dagger}(z)$ is positive independently of the type of shock, the sign of the covariance of aggregate demand with $\hat{P}_t^{\dagger}(z)$ depends on the type of shock. In particular, demand is negatively correlated with $\hat{P}_t^{\dagger}(z)$, conditioning on the technology shock, while it is positively correlated with $\hat{P}_t^{\dagger}(z)$, conditioning on the monetary policy shock. Since a firm faces, on average, a larger loss when it is uninformed about the technology shocks than when it is uninformed about the monetary policy shocks, it decides to respond with a decrease in prices to an increase in demand. Therefore, for low values of k, aggregate prices are less informative and hence receive a relatively lower weight while more attention is devoted to aggregate demand, causing prices to respond with the *wrong* sign to the monetary policy shocks: prices raise after a positive shock to ε_t^r.

6.4 Final considerations and extensions

The household side of the economy has been modeled without any friction. Rationally inattentive households would, most likely, not reduce the ability of the model I propose in accounting for the difference in the speed of adjustment of prices to the two aggregate shocks. In the case in which households are monopolistic suppliers of factors, limited information processing capabilities on them would introduce inertia in the response of factor prices to shocks. This would have an asymmetric impact on the response of prices to the two aggregate shocks. After a monetary policy shock, less responsive factor prices mean less responsive nominal marginal costs, and, therefore, less responsive aggregate prices when everything else is equal. After a technology shock, less responsive factor prices would imply more responsive nominal marginal costs. This is because if factor prices react less, real marginal costs would move more, as they are affected directly by technology. In fact, the latter reduces the impact on real marginal costs of technology as it moves in the opposite direction. Therefore, rationally inattentive households could, in principle, also amplify the difference in the allocation of attention of the price setter across technology and monetary policy shocks. Understanding how big this effect is, is left for future research.

There is a role for information sharing at the firm level. If one decision maker was taking all the decisions at the firm level, therefore pooling information processing capabilities, the outcome would be at least as good as the one obtained by having three separate decision makers. Although it might not be completely realistic having only one agent taking all the decisions at the firm level, it would, most likely, not change substantially the results obtained in this paper in terms of differential speed of price adjustment to the two aggregate shocks. The reason is that the price decision is much more important for the firm that the other two decisions. In fact, the weight in the component of the loss function due to errors in pricing is an order of magnitude larger than the weights on the components of the loss function due to errors in input choices. However, there might also be other effects going on and further research is eventually needed on this dimension.

7 Conclusions

I have shown that a model in which price setters have limited information processing capabilities provides a natural explanation for the difference in the speed of adjustment of prices to neutral technology shocks and monetary policy shocks. Price setters allocate more attention to technology shocks because it is relatively more costly to be uninformed about those shocks than about monetary policy shocks. Therefore aggregate prices respond quicker to technology shocks than to monetary policy shocks. The result is driven by the large difference in the allocation of attention by the price setter across the two shocks. The large difference in the allocation of attention is, in great proportion, generated by the interaction of complementarities in price setting with limited information processing capabilities. Complementarities in price setting induce complementarities in the optimal allocation of attention: any price setter at each firm has an incentive to acquire more information on the same variables other firms are, on average, more informed about. Since there is an upper bound on the average flow of information processed by each decision maker, when more attention is paid to technology shocks, less attention is necessarily paid to monetary policy shocks.

The monetary policy authority plays a major role in the determination of the differential speed of price adjustment to the two aggregate shocks. A more aggressive monetary policy on inflation, or a less aggressive policy on output, reduces the multiplicative effect on the difference in the speed of price adjustment to technology and monetary policy shocks coming from the allocation of attention.

8 References

1. Adam, Klaus. 2007. "Optimal Monetary Policy with Imperfect Common Knowledge." Journal of Monetary Economics. Vol. 54(2), 276-301, 2007.

2. Altig, David, Lawrence J., Martin Eichenbaum, and Jesper Linde. 2005."Firm-Specific Capital, Nominal Rigidities and the Business Cycle", National Bureau of Economic Research working paper 11034.

3. Ball, Lawrence and David Romer. 1990. "Real Rigidities and the Non-Neutrality of Money." Review of Economic Studies, 57, 183-203.

4. Basu, Susanto. 1995. "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare." American Economic Review, vol. 85(3), pages 512-31, June.

5. Basu, Susanto, Fernald John and Kimball Miles Spencer. 2004. "Are Technology Improvements Contractionary?" American Economic Review, vol. 96(5), pages 1418-1448, December.

6. Berry, Steven, Levinsohn, James and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium," Econometrica, vol. 63(4), pages 841-90, July.

7. Bils, Mark and Peter Klenow. 2004. "Some Evidence on the Importance of Sticky Prices." Journal of Political Economy, 112, 947-985.

8. Broda, Christian and David E. Weinstein. 2006. "Globalization and the Gains from Variety," The Quarterly Journal of Economics, MIT Press, vol. 121(2), pages 541-585, May.

9. Burnside,Craig and Martin Eichenbaum. 1996. "Factor-Hoarding and the Propagation of Business-Cycle Shocks." The American Economic Review, Vol. 86, No. 5. (Dec., 1996), pp. 1154-1174.

10. Burstein, Ariel and Christian Hellwig. 2007. "Prices and Market Shares in a Menu Cost Model". mimeo UCLA.

11. Christiano, Lawrence J., Martin Eichenbaum, and Charles Evans. 1999. "Monetary Policy Shocks: What Have We Learned and to What End?" In "Handbook of Macroeconomics", edited by John B. Taylor and Michael Woodford, Elsevier, New York.

12. Christiano, Lawrence J., Martin Eichenbaum, and Robert Vigfusson. 2003. "The Response of Hours to a Technology Shock: Evidence Based on Direct Measures of Technology". National Bureau of Economic Research working paper 10254.

13. Christiano, Lawrence J., Martin Eichenbaum, and Charles Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy". Journal of Political Economy, 2005, vol. 113, no. 1.

14. Clarida, Richard, Jordi Galí and Mark Gertler. 2000. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory". The Quarterly Journal of Economics, vol. 115(1), pages 147-180, February.

15. Cover, Thomas M., and Joy A. Thomas (1991). "Elements of Information Theory." John Wiley and Sons, New York.

16. Fernald, John. 2007. "A Quarterly, Utilization-Corrected Series on Total Factor Productivity". Federal Reserve Bank of San Francisco Discussion Paper.

17. Fisher, D.M. Jonathan. 2006. "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks". Journal of Political Economy, Volume 114.

18. Galí, Jordi, and Mark Gertler. 1999. "Inflation Dynamics: A structural Econometric Analysis." Journal of Monetary Economics, 44 (October): 195-222.

19. Gali, Jordi. 1999. "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" American Economic Review. 89 (March): 249-271.

20. Gertler, Mark and John Leahy. 2006. "A Phillips Curve with an Ss Foundation," NBER Working Papers 11971.

21. Golosov, Mikhail and Robert Lucas. 2006. "Menu Costs and Phillips Curves." Journal of Political Economy, vol. 115(2), pages 171-199, April.

22. Hellwig, Christian and Laura Veldkamp. 2007. "Knowing What Others Know: Coordination Motives in Information Acquisition". Discussion paper, New York University.

23. Justiniano, Alejandro and Giorgio Primiceri. 2005. "The Time Varying Volatility of Macroeconomic Fluctuations". National Bureau of Economic Research working paper 12022.

24. Leeper, Eric M., Christopher A. Sims, and Tao Zha. 1996. "What Does Monetary Policy Do?" Brookings Papers on Economic Activity, 1996:2, 1-63.

25. Lucas, Robert E. Jr. 1972. "Expectations and the Neutrality of Money". Journal of Economic Theory, 4, 103-124.

26. Lucas, Robert E. Jr. 1973. "Some International Evidence on Output-Inflation Trade-Offs." American Economic Review, 63, 326-334.

27. Mankiw, N. Gregory, and Ricardo Reis. 2006. "Pervasive Stickiness". The American Economic Review, Volume 96, Number 2, May 2006 , pp. 164-169(6).

28. Maćkowiak, Bartosz, and Mirko Wiederholt. 2007. "Optimal Sticky Prices under Rational Inattention". CEPR discussion paper 6243.

29. Midrigan, Virgiliu. 2006. "Menu Costs, Multi-Product Firms, and Aggregate Fluctuations." Discussion paper, Ohio State University.

30. Mondria, Jordi. 2006. "Financial Contagion and Attention Allocation". Discussion paper, Princeton University.

31. Moscarini, Giuseppe. 2004. "Limited Information Capacity as a Source of Inertia". Journal of Economic Dynamics and Control, 28, 2003-2035.

32. Nakamura, Emi, and Jón Steinsson. 2007. "Monetary Non-Neutrality in a Multi-Sector Menu Cost Model". Harvard University.

33. Nakamura, Emi and Jón Steinsson. 2007b. "Five Facts About Prices: A Reevaluation of Menu Cost Models". Discussion paper, Harvard University.

34. Nevo, Aviv. 2001."New Products, Quality Changes and Welfare Measures Computed From Estimated Demand Systems." Review of Economics and Statistics, 2003, v85(2,May), 266-275.

35. Orphanides, Athanasios. 2003a. "Monetary Policy Evaluation with Noisy Information". Journal of Monetary Economics, 50, 605-631.

36. Orphanides, Athanasios. 2003b. "Historical monetary policy analysis and the Taylor rule." Journal of Monetary Economics, vol. 50(5), pages 983-1022, July.

37. Orphanides,Athanasios and John C. Williams, 2003. "Imperfect Knowledge, Inflation Expectations, and Monetary Policy," NBER Working Papers 9884.

38. Orphanides,Athanasios and John C. Williams, 2006. "Inflation targeting under imperfect knowledge," Working Paper Series 2006-14, Federal Reserve Bank of San Francisco.

39. Paciello, Luigi. 2007. "The Response of Prices to Technology and Monetary Policy Shocks: An Empirical Investigation." Northwestern University Discussion paper.

40. Phelps, Edmund S. 1970. "Introduction: The New Microeconomics in Employment and Inflation Theory". In Microeconomic Foundations of Employment and Inflation Theory, edited by Edmund S. Phelps et al., Norton, New York.

41. Reis, Ricardo. 2006. "Inattentive Producers." Review of Economic Studies. 73, 793-821.

42. Sims, A. Christopher. 1999. "Stickiness." Carnegie-Rochester Conference Series on Public Policy, 49, 317-356.

43. Sims, A. Christopher."Implications of Rational Inattention". Journal of Monetary Economics, Volume 50, Number 3, April 2003 , pp. 665-690(26).

44. Sims, A. Christopher. 2006. "Rational Inattention: Beyond the Linear Quadratic Case." American Economic Review Papers and Proceedings, 96, 158-163.

45. Smets, Frank and Ralf Wouters. 2003. "An estimated stochastic Dynamic General Equilibrium Model of the Euro Area. "Journal of European Economic Association, 1, 1123-1175.

46. Smets, Frank and Ralf Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach". American Economic Review, vol. 97(3), pages 586-606, June.

47. Uhlig, Harald. 2005. "What are the effects of monetary policy on output? Results from an agnostic identification procedure." Journal of Monetary Economics, Volume 52, Issue 2, March 2005, Pages 381-419.

48. Williams, John and Athnasios Orphanides. 2006. "Inflation targeting under imperfect knowledge." Finance and Economics Discussion Series 2006-20, Board of Governors of the Federal Reserve System.

49. Woodford, Michael. 2002. "Imperfect Common Knowledge and the Effects of Monetary Policy". In "Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps", edited by Philippe Aghion et al., Princeton University Press, Princeton and Oxford.

50. Woodford, Michael. 2003. "Interest and Prices. Foundations of a Theory of Monetary Policy." Princeton University Press, Princeton and Oxford.

51. Zbaracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta and Mark Bergen. 2004. "Managerial and Customer Costs of Price Adjustments: Direct Evidence from Industrial Markets." Review of Economics and Statistics, 86, 514-533.

52. Zbaracki, Mark J., Daniel Levy, Shantanu Dutta and Mark Bergen. 2007. "The Anatomy of a Price Cut: Discovering Organizational Sources of the Costs of Price Adjustment." Emory University Economics Working Paper.

9 Appendix A: The solution to the rational inattention problem

In this appendix I state the optimization problem for each of the three choices the firm has to make and take a second-order Taylor expansion around the non-stochastic balanced growth path.

The price setter problem is choosing $P_t^*(z)$ so to maximize (8) and $\{s_{zp,t}\}$ so to maximize (10) subject to (9). Similarly, the decision maker for the capital-to-labor ratio choses $k_t^*(z)$ in each period t to solves:

$$\max_{k_t(z)} E\left[\sum_{\tau=t}^{\infty} \Lambda_\tau \pi\left(P_\tau(z), k_\tau(z), x_\tau(z), \upsilon_\tau\right) \mid s_{kp}^t\right], \tag{33}$$

and chooses the optimal signal process $\{s_{zk,t}\}$ so to solve:

$$\max_{\{s_{zk,t}\}\in S} E\left[\sum_{t=0}^{\infty} \Lambda_t \pi\left(P_t(z), k_t^*(z), x_t(z), \upsilon_t\right)\right] \tag{34}$$

$$s.t. \tag{35}$$

$$\kappa_k \geq \bar{\mathrm{I}}\left(\left\{k_{a,t}^\dagger(z), k_{r,t}^\dagger(z)\right\}; \{s_{zk,t}\}\right) \tag{36}$$

where $\left\{k_{a,t}^\dagger(z), k_{r,t}^\dagger(z)\right\}$ is the vector of stochastic processes for the complete information optimal responses to the two aggregate shocks. Lastly, the decision maker for the intermediate-inputs ratio sets $x_t^*(z)$ in each period t to solve:

$$\max_{x_t(z)} E\left[\sum_{\tau=t}^{\infty} \Lambda_\tau \pi\left(P_\tau(z), k_\tau(z), x_\tau(z), \upsilon_\tau\right) \mid s_{xp}^t\right]. \tag{37}$$

and chooses the optimal signal process $\{s_{zx,t}\}$ in period zero, so to solve:

$$\max_{\{s_{zx,t}\}\in S} E\left[\sum_{t=0}^{\infty} \Lambda_t \pi\left(P_t(z), k_t(z), x_t^*(z), \upsilon_t\right)\right] \tag{38}$$

$$s.t. \tag{39}$$

$$\kappa_x \geq \bar{\mathrm{I}}\left(\left\{x_{a,t}^\dagger(z), x_{r,t}^\dagger(z)\right\}; \{s_{zx,t}\}\right) \tag{40}$$

where $\left\{x_{a,t}^\dagger(z), x_{r,t}^\dagger(z)\right\}$ is the vector of stochastic processes for the complete information optimal responses to the two aggregate shocks.

Non-stationary variables are scaled in the following way: $c_t \equiv \frac{C_t}{A_{t-1}^{(1-\alpha)(1-\mu)}}$, $y_t \equiv \frac{Y_t}{A_{t-1}^{(1-\alpha)(1-\mu)}}$, $w_t \equiv \frac{W_t}{A_{t-1}^{(1-\alpha)(1-\mu)}}$, $i_t \equiv \frac{I_t}{A_{t-1}^{(1-\alpha)(1-\mu)}}$, $H_t \equiv \frac{K_t}{A_{t-1}^{(1-\alpha)(1-\mu)}}$, $\tilde{H}_t \equiv \frac{K_t}{A_{t-1}^{(1-\alpha)(1-\mu)}}$, $k_t^s \equiv \frac{k_t}{A_{t-1}^{(1-\alpha)(1-\mu)}}$, $\lambda_t \equiv \Lambda_t A_{t-1}^{(1-\alpha)(1-\mu)}$, $x_t^s \equiv \frac{x_t}{A_{t-1}^{(1-\alpha)^2(1-\mu)}}$, $a_t \equiv \frac{A_t}{A_{t-1}}$. The other (stationary) endogenous variables are

R_t, r_t^k, π_t, μ_t, u_t where $\mu_t\lambda_t$ is the Lagrangian multiplier on the capital accumulation equation, and hence μ_t is the real price of installed capital. I define a variable with a hat to be the log deviation from its steady state, hence $\hat{\mathrm{x}}_t = \ln(x_t) - \ln(\bar{x})$. Define the real profit function of firm z at time t as:

$$\pi\left(P_t(z), k_t(z), x_t(z), Y_t, P_t, A_t, W_t, r_t\right) \equiv \left(\frac{P_t(z)}{P_t} - \frac{Q\left(k_t(z), x_t(z), W_t, r_t\right)}{A_t}\right)\left(\frac{P_t(z)}{P_t}\right)^{-\theta} Y_t, \tag{41}$$

where Y_t is aggregate demand, $k_t(z) \equiv \frac{K_t(z)}{L_t(z)}$ is the capital-to-labor ratio at firm z, and $x_t(z) \equiv \frac{X_t(z)}{K_t(z)^{\alpha} L_t(z)^{1-\alpha}}$ is the ratio of intermediate inputs to the composite input derived from capital and labor. $Q\left(k_t(z), x_t(z), W_t, r_t\right)$ is given by:

$$Q\left(k_t(z), x_t(z), W_t, r_t\right) \equiv \left(r_t^k k_t(z)^{1-\alpha} x_t(z)^{-\mu} + W_t k_t(z)^{-\alpha} x_t(z)^{-\mu} + x_t(z)^{1-\mu}\right). \tag{42}$$

I then construct the discounted profit function, multiplying the profit function by the (scaled) discount factor, λ_t:

$$\tilde{\pi}\left(\frac{P_t(z)}{P_t}, k_t^s(z), x_t^s(z), Y_t, A_t, W_t, r_t, \lambda_t\right) \equiv \lambda_t \frac{\pi\left(P_t(z), k_t(z), x_t(z), Y_t, P_t, A_t, W_t, r_t\right)}{A_{t-1}^{(1-\alpha)(1-\mu)}}$$

where $q_t = \frac{Q(k_t(z), x_t(z), W_t, r_t)}{A_{t-1}^{(1-\alpha)(1-\mu)}}$. Hence I express the profit function $\tilde{\pi}$ as a function of variables in log-deviations from the balanced growth path:

$$\breve{\pi}\left(\hat{P}_t(z) - \hat{P}_t, \hat{k}_t^s(z), \hat{x}_t^s(z), \hat{y}_t, \hat{a}_t, \hat{w}_t, \hat{r}_t, \hat{\lambda}_t\right) \equiv \tilde{\pi}\left(e^{\hat{P}_t(z)-\hat{P}_t}, \bar{k}e^{\hat{k}_t^s}, \bar{x}e^{\hat{x}_t^s}, \bar{y}e^{\hat{y}_t}, e^{\hat{a}_t}, \bar{w}e^{\hat{w}_t}, \bar{r}e^{\hat{r}_t}, \bar{\lambda}e^{\hat{\lambda}_t}\right).$$

I then take a second order Taylor expansion of $\breve{\pi}$ around the non-stochastic balanced growth path:

$$\begin{aligned}
& \breve{\pi}\left(\hat{P}_t(z) - \hat{P}_t, \hat{k}_t^s(z), \hat{x}_t^s(z), \hat{y}_t, \hat{a}_t, \hat{w}_t, \hat{r}_t, \hat{\lambda}_t\right) \\
\approx\; & \breve{\pi}(0,0,0,0,0,0,0,0) + \breve{\pi}_1\left(\hat{P}_t(z) - \hat{P}_t\right) + \breve{\pi}_2\hat{k}_t^s(z) + \breve{\pi}_3\hat{x}_t^s(z) + \breve{\pi}_4\hat{y}_t + \breve{\pi}_5\hat{a}_t + \breve{\pi}_6\hat{w}_t + \breve{\pi}_7\hat{r}_t + \breve{\pi}_8\hat{\lambda}_t + \\
& + \frac{\breve{\pi}_{11}}{2}\left(\hat{P}_t(z) - \hat{P}_t\right)^2 + \frac{\breve{\pi}_{22}}{2}\hat{k}_t^s(z)^2 + \frac{\breve{\pi}_{33}}{2}\hat{x}_t^s(z)^2 + \frac{\breve{\pi}_{44}}{2}\hat{y}_t^2 + \frac{\breve{\pi}_{55}}{2}\hat{a}_t^2 + \frac{\breve{\pi}_{66}}{2}\hat{w}_t^2 + \frac{\breve{\pi}_{77}}{2}\hat{r}_t^2 + \frac{\breve{\pi}_{88}}{2}\hat{\lambda}_t^2 + \\
& + \breve{\pi}_{12}\left(\hat{P}_t(z) - \hat{P}_t\right)\hat{k}_t^s(z) + \breve{\pi}_{13}\left(\hat{P}_t(z) - \hat{P}_t\right)\hat{x}_t^s(z) + \breve{\pi}_{14}\left(\hat{P}_t(z) - \hat{P}_t\right)\hat{y}_t + \breve{\pi}_{15}\left(\hat{P}_t(z) - \hat{P}_t\right)\hat{a}_t \\
& + \breve{\pi}_{16}\left(\hat{P}_t(z) - \hat{P}_t\right)\hat{w}_t + \breve{\pi}_{17}\left(\hat{P}_t(z) - \hat{P}_t\right)\hat{r}_t + \breve{\pi}_{18}\left(\hat{P}_t(z) - \hat{P}_t\right)\hat{\lambda}_t + \\
& + \breve{\pi}_{23}\hat{k}_t^s(z)\hat{x}_t^s(z) + \breve{\pi}_{24}\hat{k}_t^s(z)\hat{y}_t + \breve{\pi}_{25}\hat{k}_t^s(z)\hat{a}_t + \breve{\pi}_{26}\hat{k}_t^s(z)\hat{w}_t + \breve{\pi}_{27}\hat{k}_t^s(z)\hat{r}_t + \breve{\pi}_{28}\hat{k}_t^s(z)\hat{\lambda}_t + \\
& + \breve{\pi}_{34}\hat{x}_t^s(z)\hat{y}_t + \breve{\pi}_{35}\hat{x}_t^s(z)\hat{a}_t + \breve{\pi}_{36}\hat{x}_t^s(z)\hat{w}_t + \breve{\pi}_{37}\hat{x}_t^s(z)\hat{r}_t + \breve{\pi}_{38}\hat{x}_t^s(z)\hat{\lambda}_t + \\
& + \breve{\pi}_{45}\hat{y}_t\hat{a}_t + \breve{\pi}_{46}\hat{y}_t\hat{w}_t + \breve{\pi}_{47}\hat{y}_t\hat{r}_t + \breve{\pi}_{48}\hat{y}_t\hat{\lambda}_t + \breve{\pi}_{56}\hat{a}_t\hat{w}_t + \breve{\pi}_{57}\hat{a}_t\hat{r}_t + \breve{\pi}_{58}\hat{a}_t\hat{\lambda}_t + \\
& + \breve{\pi}_{67}\hat{w}_t\hat{r}_t + \breve{\pi}_{68}\hat{w}_t\hat{\lambda}_t + \breve{\pi}_{78}\hat{r}_t\hat{\lambda}_t.
\end{aligned}$$

In order to simplify further the problem, I subtract from the second order Taylor expansion for profits under incomplete information, the equivalent expression when the decision maker has complete information, everything else being equal. For example in the case of the price setter:

$$\breve{\pi}\left(\hat{P}_t^*(z)-\hat{P}_t,\hat{k}_t(z),\hat{x}_t(z),\hat{y}_t,\hat{a}_t,\hat{w}_t,\hat{r}_t,\hat{\lambda}_t\right)-\breve{\pi}\left(\hat{P}_t^\dagger(z)-\hat{P}_t,\hat{k}_t(z),\hat{x}_t(z),\hat{y}_t,\hat{a}_t,\hat{w}_t,\hat{r}_t,\hat{\lambda}_t\right).$$

The latter does not influence the attention allocation problem. Maximizing the discounted sum of profits relative to the optimal signal structure is equivalent to maximize it in deviations from the value under complete information, as the latter is independent of the signal choice. Therefore the objective for the attention allocation problem of the price setter is approximated by:

$$\begin{aligned}&E\sum_{t=0}^{\infty}\left[\breve{\pi}\left(\hat{P}_t^*(z)-\hat{P}_t,\hat{k}_t^s(z),\hat{x}_t^s(z),\hat{y}_t,\hat{a}_t,\hat{w}_t,\hat{r}_t,\hat{\lambda}_t\right)-\breve{\pi}\left(\hat{P}_t^\dagger(z)-\hat{P}_t,\hat{k}_t^s(z),\hat{x}_t^s(z),\hat{y}_t,\hat{a}_t,\hat{w}_t,\hat{r}_t,\hat{\lambda}_t\right)\right]\\ \approx\;&-\omega_1E\left(\hat{P}_t^*(z)-\hat{P}_t^\dagger(z)\right)^2\end{aligned}$$

where $\omega_1=\frac{\bar{Y}(\theta-1)}{2(1-\beta)}$, and where I have used the results from the second order Taylor expansion and the fact that:

$$i)\ :\ E\breve{\pi}_1=0, \tag{43}$$

$$ii)\ :\ E\left(\begin{array}{c}\breve{\pi}_{11}\left(\hat{P}_t^\dagger(z)-\hat{P}_t\right)+\breve{\pi}_{12}\hat{k}_t^s(z)+\breve{\pi}_{13}\hat{x}_t^s(z)+\\ \breve{\pi}_{14}\hat{y}_t+\breve{\pi}_{15}\hat{a}_t+\breve{\pi}_{16}\hat{w}_t+\breve{\pi}_{17}\hat{r}_t+\breve{\pi}_{18}\hat{\lambda}_t\end{array}\right)=0. \tag{44}$$

From (44), and computing the values of $(\breve{\pi}_{11},\breve{\pi}_{12},\breve{\pi}_{13},\breve{\pi}_{14},\breve{\pi}_{15},\breve{\pi}_{16},\breve{\pi}_{17},\breve{\pi}_{18})$ from (41), it is possible to show that the complete information optimal log-price coincides with the log-deviation of aggregate nominal marginal costs:

$$\begin{aligned}\hat{P}_t^\dagger(z)&=\hat{P}_t+\hat{q}_t-\hat{a}_t\\&=\hat{P}_t+(1-\mu)\left(\alpha\hat{r}_t^k+(1-\alpha)\hat{w}_t\right)-\hat{a}_t.\end{aligned}$$

The attention allocation problem for the price setter reduces to:

$$\min_{\{s_{zp,t}\}\in S}\omega_1E\left(\hat{P}_t^*(z)-\hat{P}_t^\dagger(z)\right)^2 \tag{45}$$

$$\begin{aligned}&s.t.\\ i)\ &:\ \hat{P}_t^*(z)=E\left[\hat{P}_t^\dagger(z)\mid s_{zp}^t\right],\\ ii)\ &:\ \kappa_p\geq\bar{\mathrm{I}}\left(\left\{\hat{P}_{at}^\dagger(z),\hat{P}_{rt}^\dagger(z)\right\};\{s_{zp,t}\}\right)\end{aligned}$$

where the optimal choice of prices, $\hat{P}_t^*(z)$, is a projection of the complete information price over the signal realization up to time t. This is due to the fact that the objective function is quadratic. It can be proven that the objective function is well defined in the sense that at the optimal solution it is finite.

Similarly, it can be shown that the objective in the attention allocation problem for the capital labor ratio choice can be approximated with a second order Taylor expansion by:

$$-\omega_2 E\left(\hat{k}_t^*(z) - \hat{k}_t^\dagger(z)\right)^2$$

where $\omega_2 = \frac{\bar{Y}(\theta-1)}{2(1-\beta))}\frac{\alpha(1-\alpha)}{\theta}$. The complete information optimal choice for the capital labor ratio depends on the relative ratio of real wages to rental rates:

$$k_t^\dagger(z) = \frac{\alpha}{1-\alpha}\frac{W_t}{r_t^k},$$

then by log-linearizing the above expression:

$$\hat{k}_t^\dagger(z) = \hat{W}_t - \hat{r}_t^k,$$

Then the attention allocation problem is:

$$\begin{aligned} & \min_{\{s_{zk,t}\}\in S} \omega_2 E\left(\hat{k}_t^*(z) - \hat{k}_t^\dagger(z)\right)^2 \\ & s.t. \\ i) \quad : \quad & \hat{k}_t^*(z) = E\left[\hat{k}_t^\dagger(z) \mid s_{zk}^t\right], \\ ii) \quad : \quad & \kappa_k \geq \mathrm{I}\left(\left\{\hat{k}_{at}^\dagger(z), \hat{k}_{rt}^\dagger(z)\right\}; \{s_{zk,t}\}\right). \end{aligned} \tag{46}$$

Finally, the objective in the attention allocation problem for the intermediate-inputs ratio to the other factors, can be approximated with a second order Taylor expansion by:

$$-\omega_3 E\left(\hat{x}_t^*(z) - \hat{x}_t^\dagger(z)\right)^2$$

where $\omega_3 = \frac{\bar{Y}(\theta-1)}{2(1-\beta)}\frac{\mu(1-\mu)}{\theta}$. The optimal intermediate-inputs ratio under complete information is given by:

$$x_t^\dagger(z) = \frac{\mu}{1-\mu}\frac{\left(r_t^k\right)^\alpha W_t^{1-\alpha}}{\alpha^\alpha(1-\alpha)^{1-\alpha}},$$

and then by log-linearizing the expression above I obtain :

$$\hat{x}_t^\dagger(z) = \alpha\hat{r}_t^k + (1-\alpha)\hat{W}_t.$$

Then the attention allocation problem for the production department is:

$$
\begin{aligned}
& \min_{\{s_{zx,t}\}\in S} \omega_3 E\left(\hat{x}_t^*(z)-\hat{x}_t^\dagger(z)\right)^2 \\
& s.t. \\
i) \quad : \quad & \hat{x}_t^*(z)=E\left[\hat{x}_t^\dagger(z)\mid s_{zx}^t\right], \\
ii) \quad : \quad & \kappa_x \geq \bar{\mathrm{I}}\left(\left\{\hat{x}_{at}^\dagger(z),\hat{x}_{rt}^\dagger(z)\right\};\{s_{zx,t}\}\right).
\end{aligned}
\tag{47}
$$

Notice that in the information flows for all three problems, I have replaced the levels with the logs. Given that this is a monotonic transformation the value of the information flow will be unchanged, and hence the level of uncertainty of the two processes is the same.

9.1 Appendix B: the solution routine

This is a two step procedure. In the first step the endogenous variables are scaled, and the first order conditions are log-linearized around the non-stochastic balanced growth path for given guesses for $\{g_{p,t}\}$, $\{g_{k,t}\}$ and $\{g_{x,t}\}$, obtaining a linear state space representation. In the second step the rational inattention problems are solved and the guesses are verified and updated.

9.1.1 Step 1

There are fourteen endogenous variables $\left(c_t,y_t,w_t,i_t,H_t,\tilde{H}_t,k_t^s,\lambda_t,\hat{x}_t^s,R_t,r_t^k,\pi_t,\mu_t,u_t\right)$. After log-linearizing the scaled model around its balanced growth path, I obtain the fourteen equations that define the equilibrium. Three of these equations are directly liked to the rational inattention problems. The equations directly involving the rational inattention problems are the one defining aggregate prices, the one relative to the aggregate capital-to-labor ratio, and the one relative to the intermediate-inputs ratio. These problems are solved in step 2, but in step 1 I give a guess for aggregate prices, the ratio of aggregate capital to labor, and the ratio of intermediate-inputs to capital and labor. Formally, the three guesses take the form of the sum of stationary $MA\left(T\right)$ processes[39]:

$$g_{p,t} \equiv g^a_{p,t} + g^r_{p,t} = \sum_{l=0}^{T-1} \gamma^a_{p,l}\varepsilon^a_{t-l} + \sum_{l=0}^{T-1} \gamma^r_{p,l}\varepsilon^r_{t-l}, \tag{48}$$

$$g_{k,t} \equiv g^a_{k,t} + g^r_{k,t} = \sum_{l=0}^{T-1} \gamma^a_{k,l}\varepsilon^a_{t-l} + \sum_{l=0}^{T-1} \gamma^r_{k,l}\varepsilon^r_{t-l}, \tag{49}$$

$$g_{x,t} \equiv g^a_{x,t} + g^r_{x,t} = \sum_{l=0}^{T-1} \gamma^a_{x,l}\varepsilon^a_{t-l} + \sum_{l=0}^{T-1} \gamma^r_{x,l}\varepsilon^r_{t-l}. \tag{50}$$

The equilibrium condition that defines the aggregate price is:

$$\hat{P}_t = \int_0^1 E\left[\hat{P}^\dagger_t\left(z\right) \mid s^t_{zp}\right] dz,$$

which I express as:

$$\hat{P}_t = \hat{P}^\dagger_t + g_{p,t},$$

which implies that the equilibrium price can be expressed as the price that would prevail under complete information for the decision maker, plus a process that depends on the realizations of the only two exogenous variables in the model. The condition above can then be manipulated to obtain an expression that does not depend on price level. This is important as the price level is not stationary. Therefore:

$$\begin{aligned} \hat{P}_t &= \hat{P}^\dagger_t\left(z\right) + g_{p,t} \\ \hat{P}_t &= \hat{P}_t + \left(1-\mu\right)\left(\alpha r^k_t + \left(1-\alpha\right)\hat{w}_t\right) - \hat{a}_t + g_{p,t} \\ 0 &= \left(1-\mu\right)\left(\alpha r^k_t + \left(1-\alpha\right)\hat{w}_t\right) - \hat{a}_t + g_{p,t} \end{aligned} \tag{51}$$

where the second equation derives from the definition of $\hat{P}^\dagger_t\left(z\right)$. In a similar way I obtain a condition for aggregate capital-to-labor ratio,

$$\hat{k}_t = \int_0^1 E\left[\hat{k}^\dagger_t\left(z\right) \mid s^t_{zk}\right] dz,$$

and express it as:

$$\begin{aligned} \hat{k}_t &= \hat{k}^\dagger_t\left(z\right) + g_{k,t} \\ &= \hat{W}_t - \hat{r}^k_t + g_{k,t}, \\ \hat{k}^s_t &= \hat{w}_t - \hat{r}^k_t + g_{k,t}. \end{aligned} \tag{52}$$

[39] Stationarity comes from the fact that these differences converge to zero after a one time shock.

Finally, the condition defining the aggregate intermediate-inputs ratio is given by:

$$\hat{x}_t = \int_0^1 E\left[\hat{x}_t^{\dagger}(z) \mid s_{zx}^t\right] dz,$$

and it is expressed as:

$$\begin{aligned}
\hat{x}_t &= \hat{x}_t^{\dagger}(z) + g_{x,t} \\
&= (1-\alpha)\hat{W}_t + \alpha \hat{r}_t^k + g_{x,t}. \\
\hat{x}_t^s &= (1-\alpha)\hat{w}_t + \alpha \hat{r}_t^k + g_{x,t}
\end{aligned} \tag{53}$$

I then formulate the guess for the $6T$ parameters, $\left\{\gamma_{p,l}^a,\ \gamma_{p,l}^r,\ \gamma_{k,l}^a,\ \gamma_{k,l}^r,\ \gamma_{x,l}^a,\ \gamma_{x,l}^r\right\}_{l=0}^{T-1}$, for a large T. Once this is done, I can solve the model, represented by the sixteen equations, three of which are $(51), (49), (53)$,and obtain a state space representation. In particular I can obtain the responses of $\hat{P}_t^{\dagger}(z)$, $\hat{k}_t^{\dagger}(z)$ and $\hat{x}_t^{\dagger}(z)$ to the two shocks, which is all I need to solve the rational inattention problems:

$$\hat{P}_t^{\dagger}(z) \equiv \hat{P}_t^{\dagger a}(z) + \hat{P}_t^{\dagger r}(z) = \sum_{l=0}^{T-1} \eta_{p,l}^a \varepsilon_{t-l}^a + \sum_{l=0}^{T-1} \eta_{p,l}^r \varepsilon_{t-l}^r, \tag{54}$$

$$\hat{k}_t^{\dagger}(z) \equiv \hat{k}_t^{\dagger a}(z) + \hat{k}_t^{\dagger r}(z) = \sum_{l=0}^{T-1} \eta_{k,l}^a \varepsilon_{t-l}^a + \sum_{l=0}^{T-1} \eta_{k,l}^r \varepsilon_{t-l}^r, \tag{55}$$

$$\hat{x}_t^{\dagger}(z) \equiv \hat{x}_t^{\dagger a}(z) + \hat{x}_t^{\dagger r}(z) = \sum_{l=0}^{T-1} \eta_{x,l}^a \varepsilon_{t-l}^a + \sum_{l=0}^{T-1} \eta_{x,l}^r \varepsilon_{t-l}^r \tag{56}$$

9.1.2 Step 2

In the second step I solve for the attention allocations problems. In order to save on space, I express the attention allocation problems of the three decision makers in terms of the variable $\hat{\zeta}_{j,t}(z)$, which I define in the following way:

$$\hat{\zeta}_{j,t}(z) \equiv \begin{cases} \hat{P}_t(z), & j = p \\ \hat{k}_t(z), & j = k \\ \hat{x}_t(z), & j = x \end{cases}.$$

Given the assumption of independent signals in (12), I can express the attention problem for the choice of signals for decision maker j as:

$$\begin{aligned}
&\min_{(\{s_{zja,t}\},\{s_{zjr,t}\})\in S} \omega_j E\left(\hat{\zeta}^{\dagger}_{ja,t}(z) - \hat{\zeta}^{*}_{ja,t}(z)\right)^2 + \omega_j E\left(\hat{\zeta}^{\dagger}_{jr,t}(z) - \hat{\zeta}^{*}_{jr,t}(z)\right)^2 \qquad (57)\\
&s.t.\\
i) \quad &: \quad \hat{\zeta}^{*}_{ja,t}(z) = E\left[\hat{\zeta}^{\dagger}_{ja,t}(z) \mid s^t_{zja}\right],\\
ii) \quad &: \quad \hat{\zeta}^{*}_{jr,t}(z) = E\left[\hat{\zeta}^{\dagger}_{jr,t}(z) \mid s^t_{zjr}\right],\\
iii) \quad &: \quad \kappa_j \geq \mathtt{I}\left(\left\{\hat{\zeta}^{\dagger}_{ja,t}(z)\right\}; \{s_{zja,t}\}\right) + \mathtt{I}\left(\hat{\zeta}^{\dagger}_{jr,t}(z); \{s_{zjr,t}\}\right),
\end{aligned}$$

I can then solve separately six attention allocation problems, two for each of the three decision makers, as the objective functions are separable, and the information flow constraints are additive. One can show that the objective function in (57) is finite at a solution. One can also show that, in this framework, $\mathtt{I}\left(\left\{\hat{\zeta}^{\dagger}_{ja,t}(z)\right\}; \{s_{zja,t}\}\right) = \mathtt{I}\left(\left\{\Delta\hat{\zeta}^{\dagger}_{ja,t}(z)\right\}; \{s_{zja,t}\}\right)$. Also, Maćkoviak and Wiederholt (2007) show that the attention allocation problem can be solved directly in terms of conditional expectations. I report only the solution procedure to attention allocation to shock $\{\varepsilon^a_t\}$ for choice j. The procedure for $\{\varepsilon^r_t\}$ is identical. Therefore, consider the optimal attention allocation for $\hat{\zeta}_{ja,t}(z)$ relative to the $\{\varepsilon^a_t\}$ process. The signal and the optimal price process $\left(\left\{\hat{\zeta}^{\dagger}_{ja,t}(z)\right\}; \{s_{zpa,t}\}\right)$ are normally distributed and the variable to process information about is univariate. I can then express the attention allocation problems as:

$$\begin{aligned}
&\min_{(\{a_{j,l}\},\{b_{j,l}\})} \omega_j \lim_{T\to\infty}\left[\sigma_a^2 \sum_{l=0}^{T-1}\left(\tilde{a}_{j,l} - \eta^a_{j,l}\right)^2 + \sum_{l=0}^{T-1}\tilde{b}^2_{j,l}\right]\\
&s.t.\\
i) \quad &: \quad \kappa_{ja} \geq \mathtt{I}\left(\left\{\Delta\hat{\zeta}^{\dagger}_{ja,t}(z)\right\}; \left\{\Delta\hat{\zeta}^{*}_{ja,t}(z)\right\}\right),\\
ii) \quad &: \quad \Delta\hat{\zeta}^{*}_{ja,t}(z) = E\left[\Delta\hat{\zeta}^{\dagger}_{ja,t}(z) \mid s^t_{zja}\right] \equiv \lim_{T\to\infty}\left[\sum_{l=0}^{T-1} a_{j,l}\varepsilon^a_{t-l} + \sum_{l=0}^{T-1} b_{j,l}u^a_{t-l}\right],\\
ii) \quad &: \quad u_{t-l} \;\; iid \; \sim N(0,1),\\
iv) \quad &: \quad \tilde{a}_{j,l} = \sum_{i=0}^{l} a_{j,l}; \; \tilde{b}_{j,l} = \sum_{i=0}^{l} b_{j,l},
\end{aligned}$$

where the information flow is defined by:

$$\mathtt{I}\left(\left\{\Delta\hat{\zeta}^{\dagger}_{ja,t}(z)\right\}; \left\{\Delta\hat{\zeta}^{*}_{ja,t}(z)\right\}\right) = \lim_{T\to\infty}\frac{1}{T}\left(0.5\log_2\left|(2\pi e)^T\Omega_{\Delta\hat{\zeta}^{\dagger}_{ja,t}}\right| - 0.5\log_2\left|(2\pi e)^T\Omega_{\Delta\hat{\zeta}^{\dagger}_{ja,t}\mid\Delta\hat{\zeta}^{*}_{ja,t}}\right|\right),$$

where $\Omega_{\Delta\hat{\zeta}^{\dagger}_{ja,t}}$ is the variance-covariance matrix of $\left\{\Delta\hat{\zeta}^{\dagger}_{ja,t}\right\}$ and $\Omega_{\Delta\hat{\zeta}^{\dagger}_{ja,t}|\Delta\hat{\zeta}^{*}_{ja,t}}$ is the variance-covariance matrix of $\left\{\Delta\hat{\zeta}^{\dagger}_{ja,t}\right\}$ conditional on $\left\{\Delta\hat{\zeta}^{*}_{ja,t}\right\}$. Both are stationary objects. The implied process for $\Delta\hat{\zeta}_{t,ja}$ is then

$$\Delta\hat{\zeta}_{ja,t} = \lim_{T\to\infty}\sum_{l=0}^{T-1} a^{*}_{j,l}\varepsilon^{a}_{t-l}.$$

In order to solve the model, I fix a large T and solve the model for that T. I can then update the guess in the following way:

$$\gamma^{a'}_{j,l} = (1-\varkappa)\,\gamma^{a}_{j,l} + \varkappa\left(\tilde{a}^{*}_{j,l} - \eta^{a}_{j,l}\right),$$

where $\varkappa$ is a constant chosen small enough to ensure convergence. Notice that at a solution,

$$\tilde{a}^{*}_{j,l} - \eta^{a}_{j,l} = \gamma^{a}_{j,l}.$$

In the same way I solve for the process $\{\varepsilon^{r}_{t}\}$ and update for the corresponding guesses. Once I have all the new guesses, I start again from step 1 and iterate until convergence.

9.2 Appendix C: Solution to the static model

The log-linearized equations defining the solution to the static version in section 4 of the model are:

$$\begin{aligned}
\hat{Y}_t &= (1-\mu)\,\hat{C}_t + \mu\hat{X}_t, && (58)\\
\hat{C}_t &= -\hat{R}_t - \hat{P}_t, && (59)\\
\hat{R}_t &= \phi_\pi\hat{P}_t + \phi_y\hat{C}_t + \varepsilon^{r}_{t}, && (60)\\
\hat{Y}_t &= \mu\hat{X}_t + (1-\mu)\,\hat{L}_t + \varepsilon^{a}_{t}, && (61)\\
\hat{W}_t &= \psi_l\hat{L}_t + \hat{C}_t, && (62)\\
\hat{X}_t &= \hat{L}_t + \hat{W}_t, && (63)\\
\hat{P}^{\dagger}_t(z) &= \hat{P}_t + (1-\mu)\,\hat{W}_t - \varepsilon^{a}_{t}. && (64)
\end{aligned}$$

By substituting (58) into (61) I get an expression for hours worked, $\hat{L}_t = \hat{C}_t - \frac{\varepsilon^{a}_{t}}{1-\mu}$, which I then substitute into (62), to get an expression for real wages as a function of demand and technology, $\hat{W}_t = (1+\psi_l)\,\hat{C}_t - \psi_l\frac{\varepsilon^{a}_{t}}{1-\mu}$. Substituting the latter into (64) gives:

$$\hat{P}^{\dagger}_t(z) = \hat{P}_t + \xi\hat{C}_t - (1+\psi_l)\,\varepsilon^{a}_{t}.$$

where $\xi = (1+\psi_l)(1-\mu)$. Using equations (59) and (60), I obtain an expression for $\hat{C}_t$ as a function of $\hat{P}_t$ and ε_t^r, $\hat{C}_t = -\frac{1+\phi_\pi}{1+\phi_y}\hat{P}_t - \frac{\varepsilon_t^r}{1+\phi_y}$. Finally, using the last result and the guess for aggregate prices, $\hat{P}_t = \eta_r\varepsilon_t^r + \eta_a\varepsilon_t^a$, I obtain an expression for complete information optimal prices as a function of the shocks only:

$$\begin{aligned}
\hat{P}_t^\dagger(z) &= \left(\left(1-\xi\frac{1+\phi_\pi}{1+\phi_y}\right)\eta_r + \vartheta_r\right)\varepsilon_t^r + \left(\left(1-\xi\frac{1+\phi_\pi}{1+\phi_y}\right)\eta_a + \vartheta_a\right)\varepsilon_t^a, \\
\vartheta_a &\equiv -(1+\psi_l), \\
\vartheta_r &\equiv -(1+\psi_l)(1-\mu)\frac{1}{1+\phi_y}.
\end{aligned}$$

I can then solve the price setter attention problem expressing it as:

$$\begin{aligned}
&\min_{\{s_{zp,t}\}\in S} \omega_1 E\left(\hat{P}_t^*(z) - \hat{P}_t^\dagger\right)^2 \\
&s.t. \\
i) \quad &: \quad \hat{P}_t^\dagger(z) = \left(\left(1-\xi\frac{1+\phi_\pi}{1+\phi_y}\right)\eta_r + \vartheta_r\right)\varepsilon_t^r + \left(\left(1-\xi\frac{1+\phi_\pi}{1+\phi_y}\right)\eta_a + \vartheta_a\right)\varepsilon_t^a, \\
ii) \quad &: \quad \hat{P}_t^*(z) = E\left[\hat{P}_t^\dagger(z) \mid s_{zpt}\right], \\
iii) \quad &: \quad s_{zprt} = (\xi\eta_r + \vartheta_r)\varepsilon_t^r + u_{zt}^r \\
iv) \quad &: \quad s_{zpat} = (\xi\eta_a + \vartheta_a)\varepsilon_t^a + u_{zt}^a \\
v) \quad &: \quad u_{zt}^r \perp u_{zt}^a,\ u_{zt}^a \,\tilde{}\, N(0,\upsilon_a^2),\ u_{zt}^r \,\tilde{}\, N(0,\upsilon_r^2) \\
vi) \quad &: \quad \frac{1}{2}\log_2\left(1+\frac{\sigma_a^2}{\upsilon_a^2}\right) + \frac{1}{2}\log_2\left(1+\frac{\sigma_r^2}{\upsilon_r^2}\right) \leq \kappa_p
\end{aligned}$$

Using the constraints $(i)-(v)$, and solving for the unconditional expectation, the objective the problem becomes:

$$\begin{aligned}
&\min_{\left\{\frac{\sigma_a^2}{\upsilon_a^2}\geq 0, \frac{\sigma_r^2}{\upsilon_r^2}\geq 0\right\}} \omega_1\left[\frac{\left(\tilde{\xi}\eta_a + \vartheta_a\right)^2\sigma_a^2}{1+\frac{\sigma_a^2}{\upsilon_a^2}} + \frac{\left(\tilde{\xi}\eta_r + \vartheta_r\right)^2\sigma_r^2}{1+\frac{\sigma_r^2}{\upsilon_r^2}}\right] \\
&s.t. \\
i \quad &: \quad \left(1+\frac{\sigma_a^2}{\upsilon_a^2}\right)\left(1+\frac{\sigma_r^2}{\upsilon_r^2}\right) \leq 2^{2\kappa}.
\end{aligned}$$

where $\tilde{\xi} = 1-\xi\frac{1+\phi_\pi}{1+\phi_y}$. The interior solution to the problem above is:

$$1+\frac{\sigma_a^2}{\upsilon_a^{*2}} = 2^\kappa\frac{\tilde{\xi}\eta_a + \vartheta_a}{\tilde{\xi}\eta_r + \vartheta_r}\frac{\sigma_a}{\sigma_r}, \tag{65}$$

$$1+\frac{\sigma_r^2}{\upsilon_r^{*2}} = 2^\kappa\frac{\tilde{\xi}\eta_r + \vartheta_r}{\tilde{\xi}\eta_a + \vartheta_a}\frac{\sigma_r}{\sigma_a}. \tag{66}$$

Then solving for the fixed point, η_a and η_r, and substituting the result into (65) – (66) implies:

$$\begin{aligned} 1+\frac{\sigma_a^2}{\upsilon_a^{*2}} &= 2^{\kappa}\gamma\frac{1-\tilde{\xi}\left(1+\frac{2^{-\kappa}}{\gamma}\right)}{1-\tilde{\xi}\left(1+\gamma 2^{-\kappa}\right)}, \\ 1+\frac{\sigma_r^2}{\upsilon_r^{*2}} &= \frac{2^{\kappa}}{\gamma}\frac{1-\tilde{\xi}\left(1+\gamma 2^{-\kappa}\right)}{1-\tilde{\xi}\left(1+\frac{2^{-\kappa}}{\gamma}\right)}. \end{aligned}$$

The conditions on parameters for an interior solution are obtained by imposing $\frac{\sigma_j^2}{\upsilon_j^2} \geq 0$, and are given by:

$$\begin{cases} \tilde{\xi} \leq \frac{1-\frac{1}{\gamma}2^{-\kappa}}{1-2^{-2\kappa}} & \text{if} \quad \gamma \leq 1 \\ \tilde{\xi} \leq \frac{1-\gamma 2^{-\kappa}}{1-2^{-2\kappa}} & \text{if} \quad \gamma > 1 \end{cases}$$

Then, the corner solutions for η_a and η_r are:

$$\begin{cases} (\eta_a,\eta_r) = \left(\vartheta_a\frac{1-2^{-2\kappa}}{1-\tilde{\xi}(1-2^{-2\kappa})},0\right) & \text{if } \tilde{\xi} > \frac{1-\gamma 2^{-\kappa}}{1-2^{-2\kappa}} \text{ and } \gamma > 1 \\ (\eta_a,\eta_r) = \left(0,\vartheta_r\frac{1-2^{-2\kappa}}{1-\tilde{\xi}(1-2^{-2\kappa})}\right) & \text{if } \tilde{\xi} > \frac{1-\frac{1}{\gamma}2^{-\kappa}}{1-2^{-2\kappa}} \text{ and } \gamma \leq 1 \end{cases}$$

9.3 Appendix D

The production function Fernald (2007) uses to estimate the TFP growth rate through a Solow residual is:

$$Y_t^{va} = Z_t\left(K_t u_t\right)^{\alpha} L_t^{1-\alpha}$$

where Y_t^{va} is value added output, and the measure of labor takes into account of the quality and effort into hours worked. The implied expression for the Solow residual is:

$$\hat{Z}_t = \hat{Y}_t^{va} - \left(\alpha\hat{K}_t + \alpha\hat{u}_t + (1-\alpha)\hat{L}_t\right).$$

By log-differentiating the aggregate production function in my model on the non-stochastic balanced growth path, I obtain:

$$\hat{A}_t = \hat{Y}_t - (1-\mu)\left(\alpha\hat{K}_t + \alpha\hat{u}_t + (1-\alpha)\hat{L}_t\right) - \mu\hat{X}_t.$$

Then, considering the fact that gross output growth, $\hat{Y}_t$, can be partitioned in intermediate inputs growth, $\hat{X}_t$, and value added output growth, $\hat{Y}_t^{va}$:

$$\hat{Y}_t = \mu\hat{X}_t + (1-\mu)\hat{Y}_t^{va},$$

I obtain the final expression for the Solow residual implied by my model:

$$\frac{\hat{A}_t}{(1-\mu)} = \hat{Y}_t^{va} - \left(\alpha \hat{K}_t + \alpha \hat{u}_t + (1-\alpha)\,\hat{L}_t\right).$$

Therefore $\hat{Z}_t = \frac{\hat{A}_t}{(1-\mu)}$, which implies $\sigma_a = (1-\mu)\,\sigma_z$.

Tables

Table 1- Forecast error decomposition –VAR, U.S., 1959-2:2007:2				
Horizons	1	5	10	20
Technology Shock	0.68	0.52	0.45	0.33
	(0.20)	(0.14)	(0.12)	(0.11)
Monetary policy shock	0.00	0.01	0.05	0.14
	(0.00)	(0.02)	(0.03)	(0.07)

Table 2 - Parameters Calibration	
β	$1.03^{-0.25}$
δ	0.025
φ_l	1
α	0.36
b	0.7
S’’(1)	5
σ_ψ	0.5
θ	4
μ	0.75
γ_a	0.001
ρ_a	0
$\frac{\sigma_a}{\sigma_r}$	1
κ	4

Table 3 - Taylor rule estimation, U.S., 1959:2-2007:2			
c	ρ_r	ϕ_π	ϕ_y
-0.2	0.96	0.12	0.2
(0.07)	(0.03)	(0.04)	(0.04)

$$r_t = c + \rho_r r_{t-1} + \phi_\pi \pi_t + \phi_y y_t + u_t$$

The constant c is scaled on the basis of annualized nominal interest rates, expressed in percent. SSR is 0.64. The R^2 is 0.94.

Figures

Figure 1: IRF to 1 b.p. TFP and FFR shock, Benchmark VAR, U.S. 1959:2-2007:2

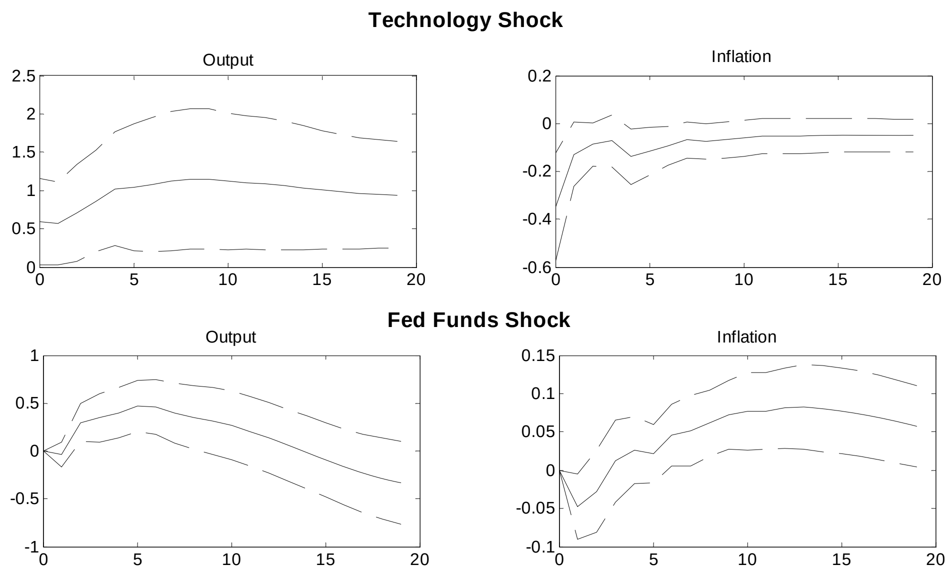


Figure 2: Price responses relative to complete information, Static Model

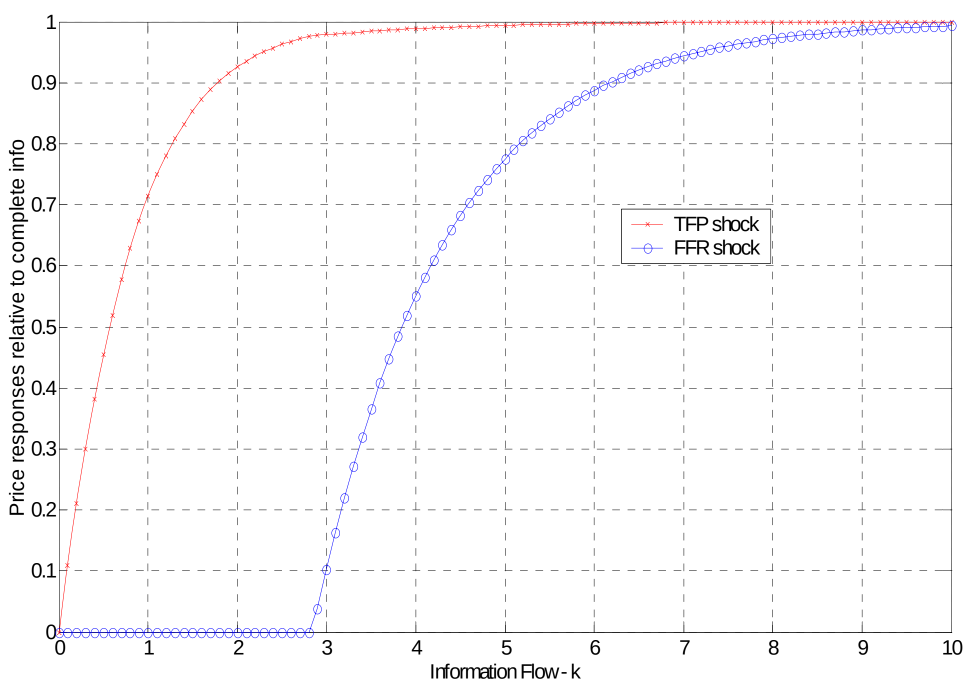


Figure 3: Attention Multiplier as a function of κ, Static Model

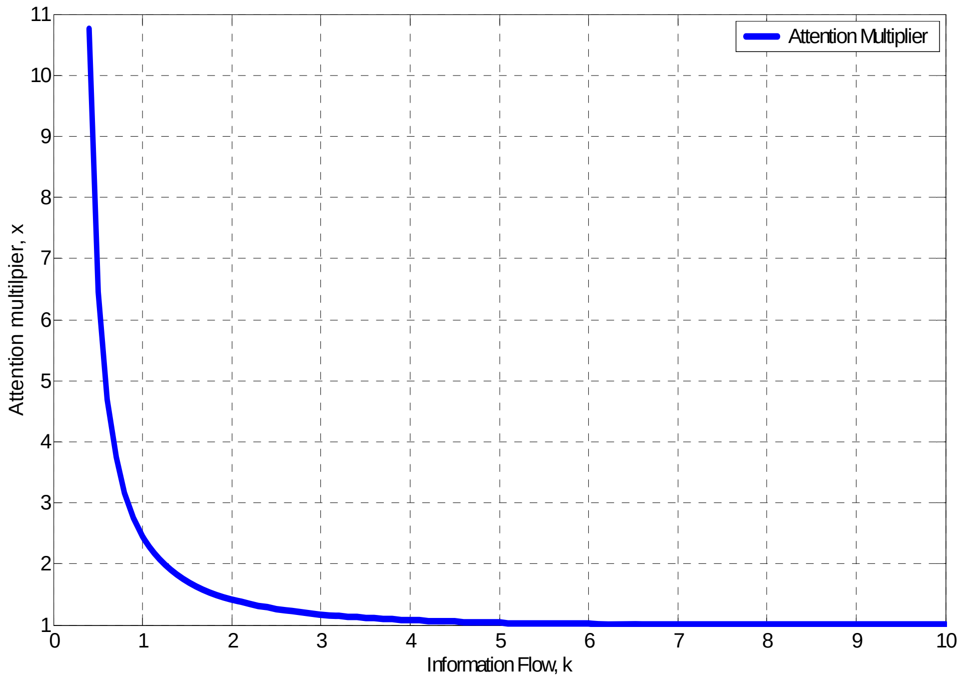


Figure 4: Attention Multiplier as a function of ξ, Static Model

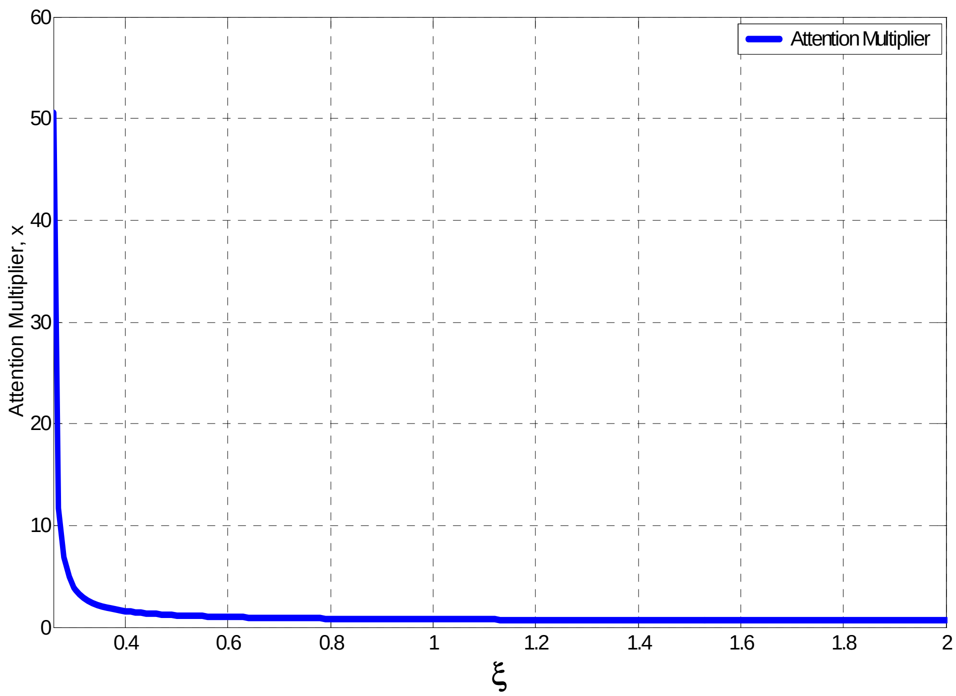


Figure 5: Price responses to ε^r relative to complete information as a function of ξ and κ, Static Model

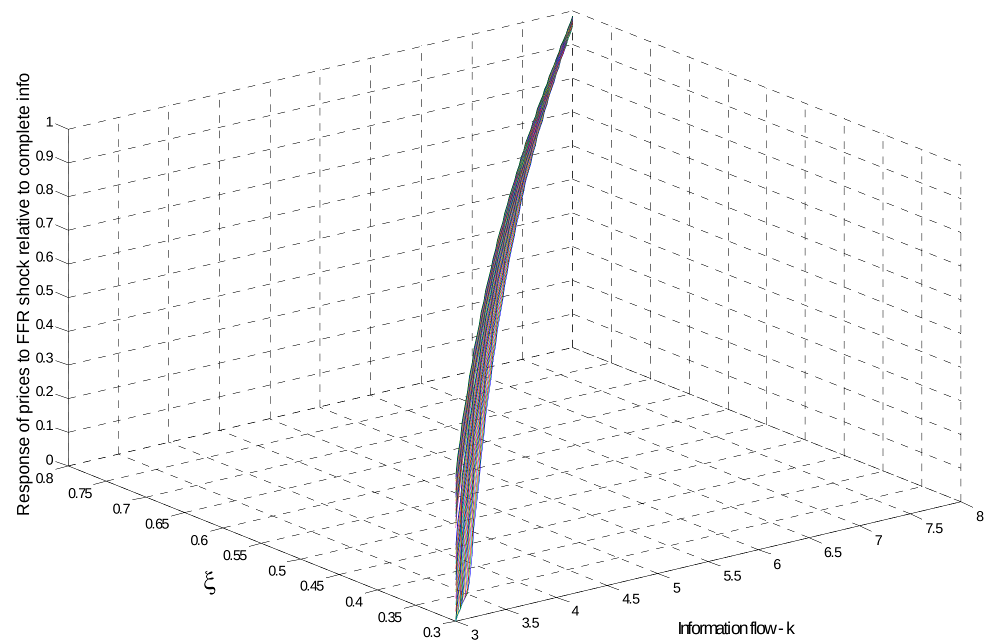


Figure 6: Price responses to ε^a relative to complete information as a function of ξ and κ, Static Model

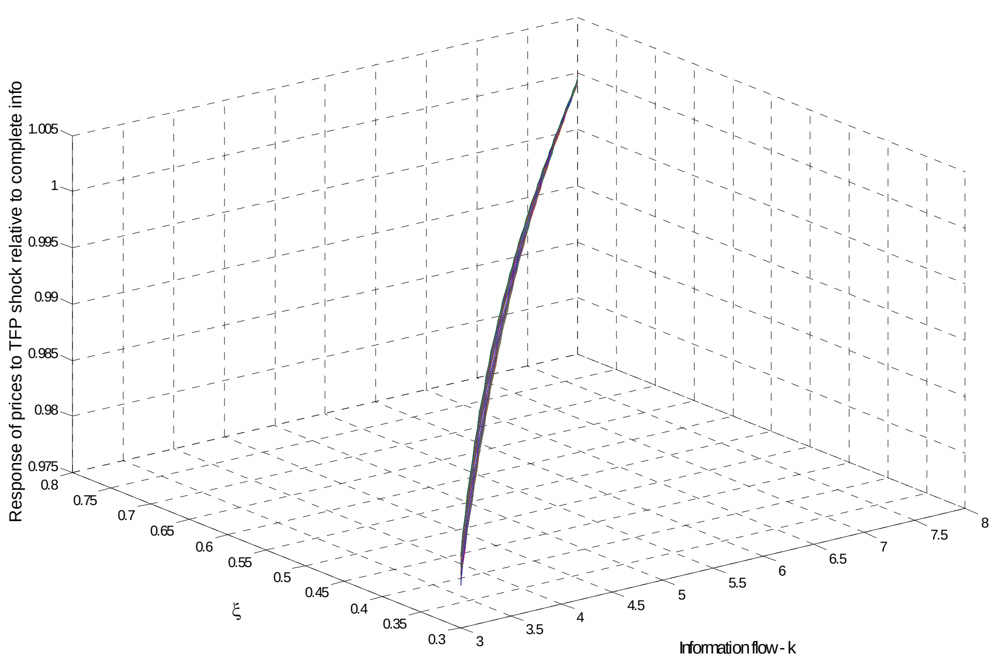


Figure 7: IRF to 1 basis point shock to ε^a and ε^r, Complete information[*]

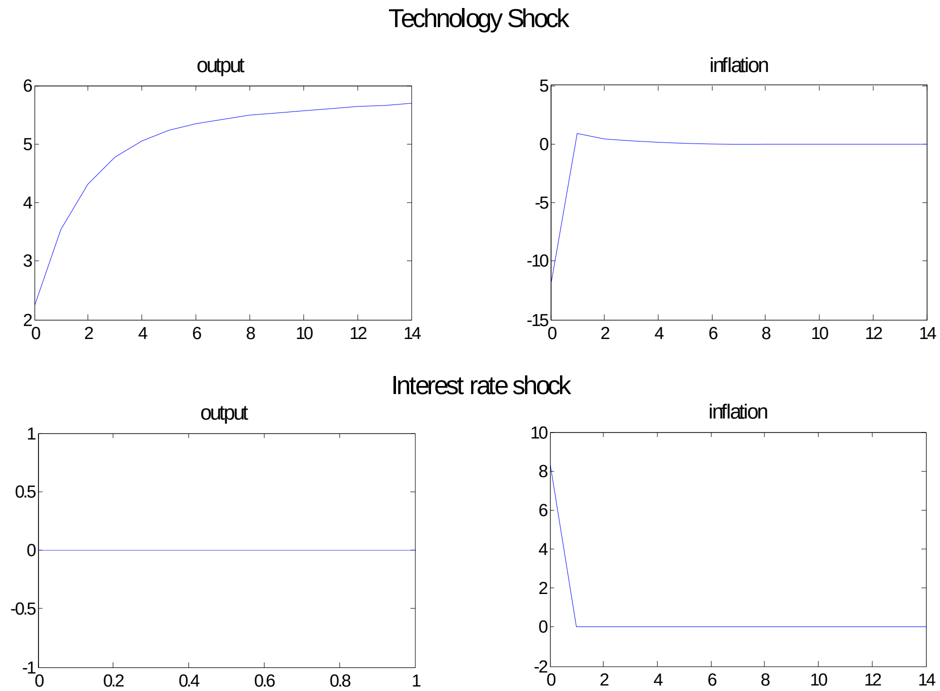
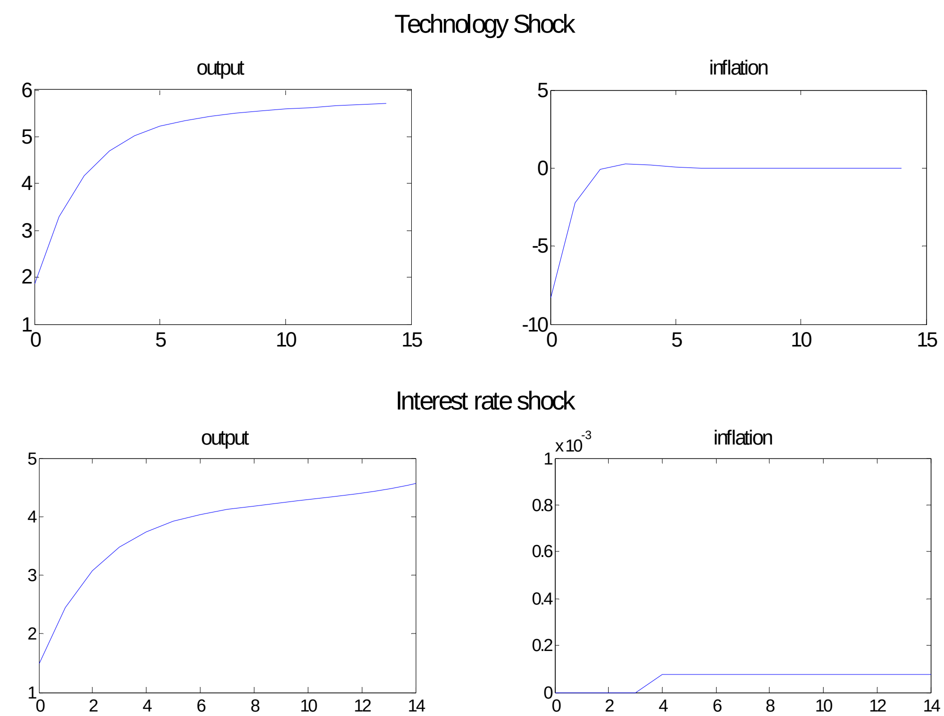


Figure 8: IRF to 1 basis point shock to ε^a and ε^r, Rational Inattention



[*] Notice that a 1 basis point shock to ε^a corresponds to a $1/(1-\mu)$ basis point shock to TFP in the VAR. For more details see Appendix D. All responses are expressed in percent values.

Figure 9: IRF to 1 basis point shock to ε^a and ε^r, Rational Inattention and Complete Info, low strategic complementarities in prices, μ=0.5

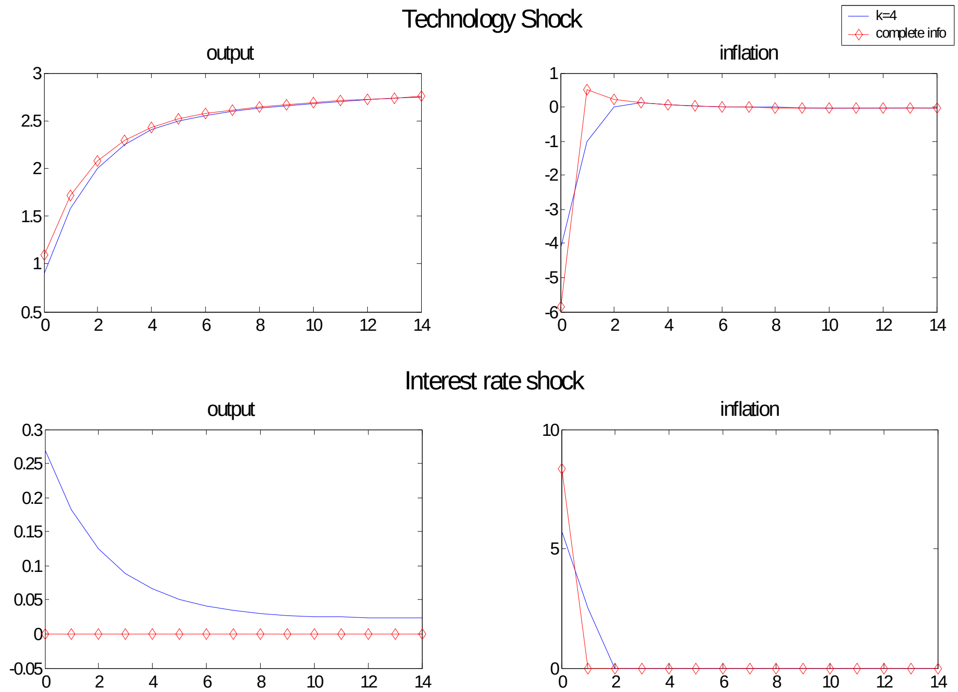


Figure 10: IRF to 1 basis point shock to ε^a and ε^r, Rational Inattention and Complete Info, monetary policy less aggressive on output growth, ϕ_y= 0.1

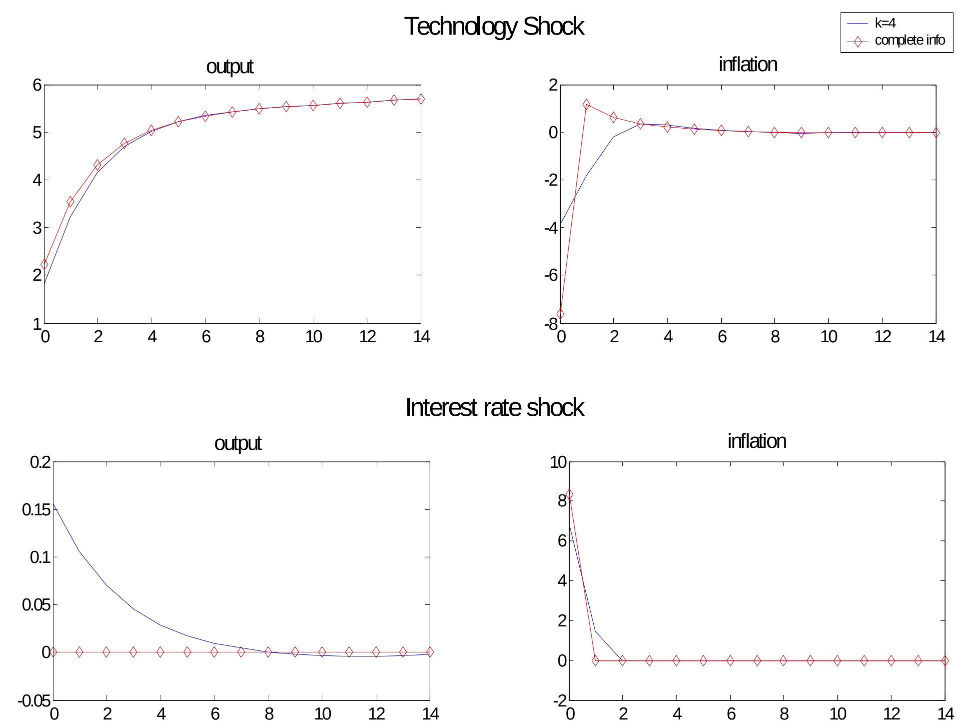


Figure 11: IRF to 1 basis point shock to ε^a and ε^r, Rational Inattention and Complete Info, monetary policy more aggressive on inflation, $\phi_\pi = 0.3$

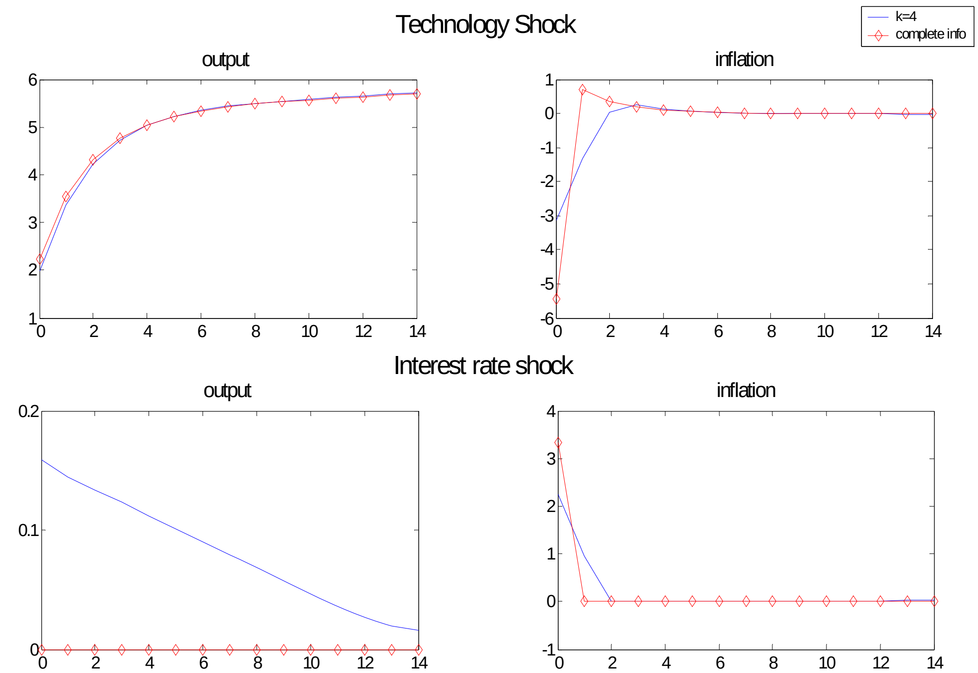


Figure 12: Price responses relative to complete information as a function of κ, Static Model with signals on endogenous variables