Endogenous Monetary Non-Neutrality

Jose Carreño*
University of Oregon

January, 2025

Abstract

This paper develops a general equilibrium model that combines strategic complementarities with boundedly rational expectations to generate short-run non-neutrality of monetary and demand shocks with respect to real variables in line with the evidence. In sharp contrast to the previous literature and leading DSGE models, I do not impose nominal rigidities: neither in the form of a Calvo assumption, menu costs, rational inattention, etc. Instead, price inertia derives as an equilibrium outcome where firms compete for their customer base in horizontally differentiated markets and form level-k beliefs about the future prices of their competitors.

*This paper has benefited from discussions with Martin Eichenbaum, Mark Colas, David Evans, George Evans, and Peio Zuazo-Garin. I started this project while visiting the Federal Reserve Bank of San Francisco, their hospitality is appreciated. All errors are my own. Email: josecarreno.econ@gmail.com.

1 Introduction

There is ample evidence that demand shocks affect output,[1] and a "widely held belief (certaintly among central bankers) in the power of [monetary] policy to influence output and employment developments, at least in the short run" (Galí, 2015).

Real Business Cycle (RBC) theory—which starts with the seminal papers of Kydland and Prescott (1982), and Prescott (1986)—was a methodological revolution in macroeconomics. Large scale macroeconomic models based on behavioral equations and ad hoc assumptions—and hence subject to serious challenges for policy evaluation (Lucas, 1976)—gave way to the use of dynamic stochastic general equilibrium (DSGE) models based on strong microfoundations. However, these models were characterized by a very small or null effect of demand shocks on output and real variables, and most of them directly made no reference to monetary factors at all. Even the versions that introduced a monetary sector (Cooley and Hansen, 1989) generally predicted neutrality (or near neutrality) of monetary policy.

This discrepancy between theoretical predictions and evidence, and between normative implications and policy practice, made way for a new generation of macroeconomic models that brought different exogenous devices to impose price inertia in firms and hence generate outcomes more in line with the evidence. Along these lines we have menu costs and Ss models, where firms have to pay a fee in order to be able to adjust their prices (see, for example, Caplin and Spulber, 1987 and Golosov and Lucas, 2007); the rational inattention models, where agents can only update their information sets and price plans at random dates or they have to pay a fee to do that (see, for example, Mankiw and Reis, 2002 and Reis, 2006); and the New Keynesian framework, which departs from the RBC framework and adds monopolistic competition and exogenous nominal rigidities as in Calvo (1983):[2] firms are not freely able to adjust their own prices if they want to, but there is a fixed probability every period to be allowed to do that (see, for example, Smets and Wouters, 2003, Smets and Wouters, 2007, and Christiano et al., 2005).

The New Keynesian framework rests on the DSGE structure characteristic of RBC models and, in medium and large scale versions, it has been the major workhorse for policy

[1]See Friedman and Schwartz (1963), Eichengreen and Sachs (1985), Mussa (1986), Christiano et al. (1999), Blanchard and Perotti (2002), Romer and Romer (2004), Ramey (2011), Barro and Redlick (2011), Nakamura and Steinsson (2014), Guajardo et al. (2014), Mian and Sufi (2014), Gertler and Karadi (2015), Nakamura and Steinsson (2018).

[2]An alternative is to combine this framework with the assumption in Rotemberg (1982) where firms must pay a proportional adjustment cost to be able to change their prices. Both the Calvo (1983) and Rotemberg (1982) pricing assumptions are equivalent to a first order approximation about a zero inflation steady state and the proper parameterization.

evaluation. However, as in menu costs and rational inattention models, it depends on the pivotal role of the respective exogenous device that imposes nominal rigidities. In other words, price inertia does not derive as an equilibrium outcome but from ad hoc restrictions externally imposed over firms, so it is not explained but assumed. This poses important problems. From a theoretical perspective, firms are forced to accept suboptimal contracts that they would never accept by any rational choice; and, if that was the status quo, firms would even be willing to pay significantly large amounts to get rid of those if they were allowed to, and hence be finally able to change their own price whenever they choose to. From a policy perspective, these models are again subject to Lucas critique (Lucas, 1976) because their main ingredient is an ad hoc assumption lacking in microfoundations, which casts reasonable doubts on their robustness for policy evaluation.

This paper aims to make a contribution towards filling the gap between: (i) an RBC theory that is strongly microfounded but not in line with the evidence on the non-neutrality of monetary and demand shocks with respect to real variables, and (ii) a New Keynesian framework that is in line with such evidence but that rests on the critical role of ad hoc restrictions and is therefore subject to Lucas critique.

The model embeds an industry structure featuring horizontal differentiation in an otherwise standard real business-cycle framework. There is a continuum of industries, each producing a different consumption good. Each industry is a "linear city" or product space of length 1, with consumers uniformly and continuously distributed along the $[0, 1]$ interval and a firm in each of the two extremes of the interval. Consumers have heterogeneous tastes—determined by their location in each industry's product space—and dislike the relative distance between their specific location in the product space (that is, their ideal good) and the location of the firms' goods (that is, the available goods in the market). The level of horizontal differentiation determines the degree of competition in an industry: in the absence of differentiation, we would be back to the Bertrand paradox where whoever charges the smallest price gets the whole market; on the other hand, as the level of horizontal differentiation increases, it becomes relatively more costly for consumers to move across the product space, so the neighboring clientele of a firm becomes more captive or, in other words, firms get more market power.

There is a continuum of households that are identical except for their location in the product space of each industry, which defines part of their preferences. In particular, for each pair industry-household, the household is randomly assigned a location in the industry's product space that is independent and identically drawn from a continuous uniform distribution over the $[0, 1]$ interval. Households consume a final good that is a CES aggregate of the continuum of industry goods, so consumption goods from different industries are imperfect substitutes. For each industry, optimal behavior implies that there is a threshold that depends on the price of each firm, the general level of prices, and the level of horizontal differentiation: households located to the left of that threshold will purchase from the firm at the left extreme of the interval (Firm A) in that particular industry, whereas households located to the right are *sufficiently* close to the other firm (Firm B) and will purchase from it. The market will be equally split when firms charge the same price; however, raising the price would—ceteris paribus—*partially* reduce the customer base of the firm. The firm with the smallest price will get a larger share of the pie, but not all the market because of the horizontal differentiation that arises from consumers' locations in the product space.

Raising—ceteris paribus—the price will be profitable for the firm because it will be charging a higher markup per unit; however, it will be costly not only because its customers will now be purchasing less units, but also because the firm will be completely losing some of its customers, who will be moving away to its industry competitor. Firm's profit maximization with respect to prices then yields a best response function that depends on three objects: the price of the industry competitor, the overall inflation rate, and the marginal cost. Each period, all firms choose their prices simultaneously after observing the realization of the shocks and then inflation is realized. I relax the standard assumption that firms perfectly forecast the actions of their industry competitor, and I instead assume that firms form consistent beliefs about how the competitor's price will react to the shocks and act upon them. In particular, I analyze different degrees of level-k thinking expectations, where level-k thinking for firm A is defined as the belief that firm B will be playing its best response under the belief that firm A will be playing its best response under level-$(k-1)$ thinking. For simplicity, I maintain the rational expectations assumption everywhere else in the model.

A positive monetary shock resulting in a permanent shift in the money supply leads to a rise in consumption and output, together with inflationary pressure, as well as an increase in labor and wages, despite the complete lack of any sort of exogenous nominal rigidity in the model. The main mechanism behind this result is the incomplete marginal cost-to-price pass-through that derives from firms' competition for their customer base. A positive demand shock raises wages, since that is necessary for workers to work more. However, the pass-through from that rise in the marginal cost to the price of the firm depends on how the firm predicts that the industry competitor will react. If the firm rises its price unilaterally, part of its customer base will move away to its industry competitor. Therefore, the firm will be reluctant to rise its price *too much* over the price of the industry competitor, since that would have a detrimental effect on its customer base. Hence, the firm will only be willing to rise its price significantly when the industry competitor is doing something similar, so that this price increase is not as negatively perceived by its customer base. A more standard monetary policy shock to a Taylor rule yields qualitatively similar results.

Lastly, given that the main mechanism behind the results is based on this short-run, incomplete marginal cost-to-price pass-through, I explore how the implications of the model change when we incorporate wage inertia. The result from a standard calibration is that the effects of the shock in the real variables doubles in both size and duration, while it gets much more robust to high levels of thinking.

This paper is, to the best of my knowledge, the first to introduce horizontally differentiated markets in a general equilibrium framework. There is a large literature of general equilibrium models with different sorts of pricing complementarities,[3] as well as a literature introducing boundedly rational expectations in macroeconomic models (for example, Angeletos and Lian (2018) and Farhi and Werning (2019) embed low level-k thinking assumptions in a New Keynesian framework). This paper shares some elements with both literatures, but does not need to include any sort of exogenous nominal rigidities to generate short-run non-neutrality of monetary and demand shocks with respect to real variables. Christiano et al. (2016) endogenizes wage inertia by explicitly modeling wage bargaining while keeping the Calvo (1983) assumption in the goods market.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 carries out the quantitative analysis. Section 4 provides empirical grounds for the main mechanism of the model, discusses the role of the key assumptions, and analyzes the main implications. Concluding remarks appear in section 5.

2 Model

This model embeds an industry structure featuring horizontal differentiation as in Hotelling (1929) in an otherwise standard real business-cycle model. There are a measure-one continuum of industries indexed by $j \in [0, 1]$, each producing a different consumption good. Each industry j consists on a "linear city" of length 1, as displayed in figure 1, with the consumers uniformly and continuously distributed with density 1 along the $[0, 1]$ interval. Each industry has two firms, say Firm A and Firm B, each located at the extremes of the

[3] See, for example, Kimball (1995), Rotemberg and Woodford (1997), Chari et al. (2000), Bergin and Feenstra (2001), Dotsey and King (2006), Burstein and Hellwig (2007), Klenow and Willis (2006), Eichenbaum and Fisher (2007), Atkeson and Burstein (2008), Gertler and Leahy (2008), Nakamura and Steinsson (2010), Altig et al. (2011), Mongey (2021).

[0,1] interval[4].

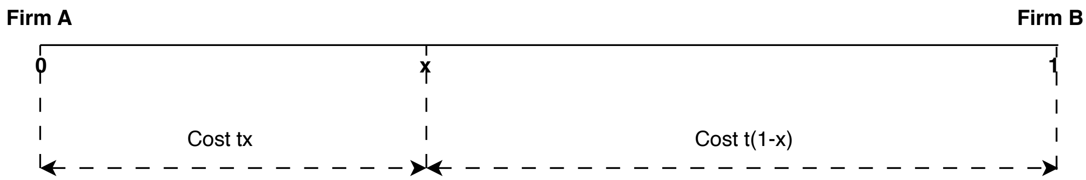

Figure 1: An industry.

A consumer located at point $x \in [0,1]$ and interested in purchasing some units of the industry j good has the option of purchasing from Firm A or B (or both). In both cases, the consumer will incur in a transportation cost $t > 0$ per unit, which is linear and proportional to the distance between the consumer and the firm. There are two ways to think about this framework. The classical interpretation is that the products offered by both firms are identical, but the consumer is physically distant from the firm (or the store) and hence has to pay a transportation cost to get there either in terms of time or money. A modern interpretation is to consider the "linear city" as the product space, along which goods are differentiated by some characteristic. Consumers have heterogeneous tastes and have to pay a utility cost when consuming a product relatively different from their ideal, represented by their location in the product space. While I consider the later interpretation as more appropriate, I model transportation costs as a monetary cost instead of a utility cost, as otherwise the model would become much more convoluted and less tractable without clear conceptual gains.

Notice that Firm A and B's products will be more differentiated the higher the transportation cost. When $t = 0$, we are back to the Bertrand paradox where whoever charges the smallest price gets the whole market and eventually both firms charge a price equal to the marginal cost in equilibrium. As t increases, it becomes more costly for consumers to move across the product space, so the neighboring clientele of a firm becomes more captive or, in other words, firms get more and more market power (so prices are higher in equilibrium). From the modeller's perspective, this way of introducing strategic complementarities is advantageous because, despite its simplicity, it embodies any potential degree of market competition as a function of a single parameter.[5]

[4]For simplicity, I am abstracting from firm location choice, considering the standard "maximal differentiation" location as fixed and predetermined. This allows me to keep linear transportation costs while ensuring tractability both at the industry-level and in the general model.

[5]Another advantage is its theoretical soundness and its relative realism. Bertrand's criticism to Cournot's competition was that such a solution cannot represent an equilibrium because, at any point, anyone could slightly reduce the price and take all his opponent's market, and that by focusing on quantities as an

The model has three types of agents: the households, the firms, and the government.

2.1 Households

There are a measure-one continuum of households indexed by $i \in [0,1]$. For every pair industry j - household i, household i is randomly assigned a location in the industry j product space, denoted by $x(j,i) \in [0,1]$. This $x(j,i)$ location is independent and identically drawn from a continuous uniform $[0,1]$ distribution[6]. Households are identical in all fronts except for their location in the product space of each particular industry $j \in [0,1]$, which is part of their preferences.

A particular household i seeks to maximize the utility function

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t(i)^{1-\sigma}}{1-\sigma} - \Psi \frac{N_t(i)^{1+\eta}}{1+\eta} + \theta \ln \left(\frac{M_t(i)}{P_t} \right) \right] \tag{1}$$

where $C_t(i)$ is consumption of the final good, which is a CES aggregate of the continuum of industry goods

$$C_t(i) = \left(\int_0^1 C_t(j,i)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}}. \tag{2}$$

Here $\epsilon > 2$, and $C_t(j,i) = C_t^A(j,i) + C_t^B(j,i)$ is the sum of the consumption made from the goods produced by each of the firms in industry j. $N_t(i)$ denotes employment or hours worked, while $\frac{M_t(i)}{P_t}$ is the real balance of money held by household i and P_t is the price of the final good in terms of money (numeraire). The period budget constraint is

independent variable instead of prices one was hiding this fallacy. Hotelling's criticism to Bertrand was, in his own words, that "in all his illustrations of competition one merchant can take away his rival's entire business by undercutting his price ever so slightly [...] Quite commonly a tiny increase in price by one seller will send only a few customers to the other".

[6] Notice that an immediate implication of this distributional assumption is the fact that, for a given industry $\tilde{j} \in [0,1]$, the distribution of locations across households $[x(\tilde{j},i)]_{i\in[0,1]}$ will be continuously uniform along the interval $[0,1]$. The same thing happens, given a household $\tilde{i} \in [0,1]$, to the distribution of locations across industries $[x(j,\tilde{i})]_{j\in[0,1]}$.

$$
\begin{aligned}
&\int_0^1 \Big[P_t^A(j)C_t^A(j,i) + P_t^B(j)C_t^B(j,i)\Big]dj + \\
P_t t &\int_0^1 \Big[x(j,i)C_t^A(j,i) + [1-x(j,i)]C_t^B(j,i)\Big]dj + \\
B_{t+1}(i) + M_t(i) = M_{t-1}(i) &+ W_t N_t(i) + \Pi_t + (1+i_{t-1}^n)B_t(i) - P_t T_t,
\end{aligned} \tag{3}
$$

where $B_t(i)$ is nominal bond holdings, i_t^n is the nominal interest rate, $M_t(i)$ is nominal money holdings, W_t is the nominal wage, Π_t is the nominal per-capita profit accruing to households from the ownership of the firms (assumed to be equally distributed across households), and T_t is a lump sum government tax.

The first line of equation (3) is the total cost of the industry goods, where $(P_t^A(j), P_t^B(j))$ are the prices charged by firms A and B of industry j. The second line is the total transportation costs, where $P_t t$ is the nominal transportation cost per unit of length in the product space of industry j, with $t > 0$ governing the magnitude of the cost.

Given a pair industry j - household i, with $x(j,i) \in [0,1]$, the effective cost per unit when purchasing from firm A would be $P_t^A(j) + P_t t x(j,i)$, while it would be $P_t^B(j) + P_t t(1-x(j,i))$ when purchasing from firm B. Cost minimization implies that, when purchasing goods from industry j, household i will choose the firm with the lowest effective cost per unit. In particular, household i will purchase from firm A if, and only if, $P_t^A(j) + P_t t x(j,i) \leq P_t^B(j) + P_t t(1-x(j,i))$ or, equivalently, $x(j,i) \leq \frac{P_t^B(j) - P_t^A(j)}{P_t 2t} + \frac{1}{2}$. This allows us to define the threshold location at industry j as:

$$
x_t^*(j) = \frac{P_t^B(j) - P_t^A(j)}{P_t 2t} + \frac{1}{2}. \tag{4}
$$

Households located to the left of this threshold are *sufficiently* close to firm A and will purchase from it, whereas those located to the right will choose firm B. Notice that the market will be equally split when firms charge the same price; however, raising the price would—ceteris paribus—*partially* reduce the customer base of the firm. Notice also how the price of the competitor directly affects the customer base of each firm. Lastly, I will restrict attention to prices such that $x_t^*(j) \in (0,1)$ for all $j \in [0,1]$ or, in other words, such that the price difference between the two firms is not so high that one of the firms has no demand. Such a thing cannot happen in equilibrium anyways, as otherwise one of the firms would have incentives to lower its price and get positive profits (conditional on wages being sufficiently small, which would be in equilibrium by labor market clearing).

Apart from the consumption/savings and the labor/leisure choice dimensions, the household has to decide how to allocate its consumption expenditures across the continuum of industry goods. Optimal consumer behavior requires $C_t(i)$ to be maximized for any given level of expenditures. First, this requires the household to purchase from firm A in those industries where $x(j,i) \leq x_t^*(j)$, and from B otherwise. Furthermore, and as shown in Appendix A.1, this yields the set of demand equations

$$C_t(j,i) = \left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t(i)} \right)^{-\epsilon} C_t(i) \tag{5}$$

for all $j \in [0,1]$, where I define

$$\tilde{P}_t(j,i) \equiv [P_t^A(j) + P_t t x(j,i)] \mathbb{1}\{x(j,i) \leq x_t^*(j)\} + [P_t^B(j) + P_t t(1 - x(j,i))] \mathbb{1}\{x(j,i) > x_t^*(j)\}$$

as the effective price (or cost) of a unit of industry good j optimally purchased by household i; and where I define

$$\tilde{P}_t(i) \equiv \left(\int_0^1 \tilde{P}_t(j,i)^{1-\epsilon} dj \right)^{\frac{1}{1-\epsilon}}$$

as the *effective* aggregate price index.[7] Moreover, and conditional on such optimal behavior, we have that

$$\begin{aligned} \tilde{P}_t(i) C_t(i) = & \int_0^1 \left[P_t^A(j) C_t^A(j,i) + P_t^B(j) C_t^B(j,i) \right] dj \\ & + P_t t \int_0^1 \left[x(j,i) C_t^A(j,i) + [1 - x(j,i)] C_t^B(j,i) \right] dj, \end{aligned} \tag{6}$$

which allows us to write the budget constraint of the household as

$$\tilde{P}_t(i) C_t(i) + B_{t+1}(i) + M_t(i) = M_{t-1}(i) + W_t N_t(i) + \Pi_t + (1 + i_{t-1}^n) B_t(i) - P_t T_t. \tag{7}$$

Appendix A.2 shows that, as a result of the continuity and random assignment assumptions, the effective aggregate price index is homogeneous across households, meaning that $\tilde{P}_t(i) = \tilde{P}_t(i') = \tilde{P}_t$ for any $i, i' \in [0,1]$. Therefore, the problem of the household reduces to

[7] Notice that the effective price of a unit of industry good j is household specific, as it depends on the specific location of the particular household in the product space of industry j. This makes the effective aggregate price index to be household specific too.

$$\max_{C_t(i),N_t(i),B_{t+1}(i),M_t(i)} \quad \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t(i)^{1-\sigma}}{1-\sigma} - \Psi \frac{N_t(i)^{1+\eta}}{1+\eta} + \theta \ln\left(\frac{M_t(i)}{P_t}\right) \right]$$

$$s.t.\ \tilde{P}_t C_t(i) + B_{t+1}(i) + M_t(i) = M_{t-1}(i) + W_t N_t(i) + \Pi_t + (1+i^n_{t-1})B_t(i) - P_t T_t.$$

However, since all households face the exactly same problem with the same set of prices, they will make the same optimizing decisions. We can then drop the i index and we are back to the representative household framework. The simplicity of this setting is that, even though different households will distribute consumption spending differently across and within industries, they will make the same overall consumption/saving and labor/leisure choices.

The Lagrangian for household i would then be

$$\mathcal{L} = \mathbb{E} \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \Psi \frac{N_t^{1+\eta}}{1+\eta} + \theta \ln\left(\frac{M_t}{P_t}\right) + \right.$$
$$\left. \lambda_t \left(M_{t-1} + W_t N_t + \Pi_t + (1+i^n_{t-1})B_t - P_t T_t - \tilde{P}_t C_t - B_{t+1} - M_t \right) \right],$$

which yields the intra-temporal (labor/leisure) optimality condition

$$\Psi N_t^{\eta} = C_t^{-\sigma} \frac{W_t}{\tilde{P}_t}, \tag{8}$$

the inter-temporal (consumption/saving) optimality condition

$$C_t^{-\sigma} = \beta \mathbb{E}_t \left[C_{t+1}^{-\sigma} (1+i^n_t) \frac{\tilde{P}_t}{\tilde{P}_{t+1}} \right], \tag{9}$$

and the optimality condition for money (money demand)

$$\theta \left(\frac{M_t}{\tilde{P}_t} \right)^{-1} = C_t^{-\sigma} \frac{i^n_t}{1+i^n_t}. \tag{10}$$

2.2 Firms

As introduced above, an industry $j \in [0,1]$ is a "linear city" or product space of length 1. Each industry j has two firms—Firm A and Firm B[8]—located at the extremes of the $[0,1]$ interval and, as stated above, households are uniformly distributed along the interval. Both firms set their price simultaneously, and these prices determine each firm's customer base as illustrated in figure 2. Under optimal consumer behavior, the customer bases of firms A and B would be

$$CB_t^A(j) = \frac{P_t^B(j) - P_t^A(j)}{P_t 2t} + \frac{1}{2} = \frac{\pi_t^B(j) - \pi_t^A(j)}{\pi_t 2t} + \frac{1}{2}$$

and

$$CB_t^B(j) = \frac{1}{2} - \frac{P_t^B(j) - P_t^A(j)}{P_t 2t} = \frac{1}{2} - \frac{\pi_t^B(j) - \pi_t^A(j)}{\pi_t 2t},$$

where $\pi_t \equiv \frac{P_t}{P_{t-1}} - 1$, $\pi_t^A(j) \equiv \frac{P_t^A(j)}{P_{t-1}} - 1$, and $\pi_t^B(j) \equiv \frac{P_t^B(j)}{P_{t-1}} - 1$.

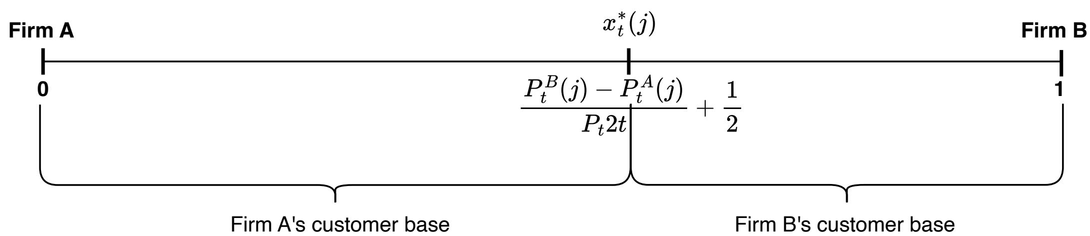

Figure 2: Distribution of the customer base in industry j.

Notice that each firm's customer base will directly depend on the price of the other firm. In particular, the customer base in industry j would be equally split if both firms set the same price; however, the firm with the smallest price would get a larger share of the pie when prices are different, but not all the market because of the horizontal differentiation that arises from the heterogeneity in consumer locations.

As I show below, profit maximization will yield a best response function that will depend on three objects: the price of the industry competitor, the overall inflation rate, and the marginal cost. Each period, both firms choose their price simultaneously after observing the realization of the shocks and then—while this happens across industries—inflation is realized.

[8]A more precise notation would be Firm A (j) and Firm B (j), as firms are industry-specific.

I relax the standard assumption that firms perfectly forecast the action of their competitor, and I instead assume that firms form consistent beliefs about how the competitor's price will react to the shocks and act upon them. In particular, I will analyze different degrees of level-k thinking expectations, where a level-0 thinking would correspond to the belief that the industry competitor is myopic in the sense that it will not react to the shocks; and where level-k thinking for firm A is defined as the belief that firm B will be playing its best response under the belief that firm A will be playing its best response under level-$(k-1)$ thinking.[9] For consistency, and given that we have a continuum of industries and each particular firm is atomistic, I assume that firms keep the same beliefs with respect to the overall inflation.[10] For simplicity, I maintain the rational expectations assumption everywhere else in the model.

To keep things as simple as possible, I assume a constant returns to scale technology in labor with a common productivity shock A_t,[11] so that the production function of firms A and B of industry j are

$$Y_t^A(j) = A_t N_t^A(j) \tag{11}$$

and

$$Y_t^B(j) = A_t N_t^B(j), \tag{12}$$

where $N_t^A(j)$ and $N_t^B(j)$ denote the labor demands of firms A and B from industry j. Labor is homogeneous across industries, there is a unique labor market, and hence all firms face a common wage W_t. It is easy to see that cost minimization implies the marginal cost $\frac{W_t}{A_t}$.

Let us, without loss of generality, take the perspective of firm A from industry j. This firm will set its price $P_t^A(j)$ to maximize expected profits subject to the constraint of producing enough to meet demand. The expression for profits is

[9] We can keep raising the level of sophistication until level-∞ thinking, which corresponds to the belief that the competitor will directly choose the unique Nash equilibrium's price in pure strategies.

[10] This makes sense as all industries are alike and there are no industry-specific shocks, so there is no reason to believe that the price variations of the economy as a whole had to be substantially different from the price variation of my industry competitor. However, placing different beliefs between the overall and individual inflation yields qualitatively similar results.

[11] The results of the paper are of course robust to the presence of capital and more general production functions, the model just becomes slightly more convoluted but still relatively simple and fully tractable.

$$\Pi_t^A(j) = P_t^A(j)Y_t^A(j) - W_t N_t^A(j) = P_t^A(j)Y_t^A(j) - \frac{W_t}{A_t} A_t N_t^A(j) = \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] Y_t^A(j), \tag{13}$$

where $\left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right]$ is the nominal price-cost markup per unit, and where $Y_t^A(j)$ is the production of firm A, which equals the demand for firm A, which will be a function of the customer base of firm A and the amount of goods purchased by the consumers in such a customer base.[12] The profits function of firm A has then three dimensions: raising—ceteris paribus—the price of firm A will be profitable because the firm will be charging a higher markup per unit; however, it will be costly not only because its customers will now be purchasing less units, but also because the firm will be completely losing some of its customers, who will be moving away to its industry competitor.

Appendix A.3 shows that, given prices, we can write the profits function of firm A from industry j as

$$\begin{aligned}\Pi_t^A(j) = \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] Y_t^A(j) &= \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] \int_0^{\frac{P_t^B(j) - P_t^A(j)}{P_t 2t} + \frac{1}{2}} C_t(j, x) dx \\ &= \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] \tilde{P}_t^{\epsilon} C_t \frac{1}{(-\epsilon + 1) P_t t} \\ &\left\{\left[P_t^A(j) + P_t t\left(\frac{P_t^B(j) - P_t^A(j)}{P_t 2t} + \frac{1}{2}\right)\right]^{-\epsilon+1} - [P_t^A(j)]^{-\epsilon+1}\right\}.\end{aligned} \tag{14}$$

We can then take the first order condition to get the best response function as a function of

[12] In fact, in a simplified partial-equilibrium model where we assume that all consumers purchase the same amount regardless of location, we can write the profits expression as $\Pi_t^A = \left[P_t^A - \left(\frac{W_t}{A_t}\right)\right]\left(\frac{P_t^B - P_t^A}{P_t 2t} + \frac{1}{2}\right) C_t(P_t^A)$, where $C_t(P_t^A)$ is a downward-slopping demand function. This simple expression is of course not correct in the general setting.

prices

$$\left\{\left[P_t^A(j)+P_tt\left(\frac{P_t^B(j)-P_t^A(j)}{P_t2t}+\frac{1}{2}\right)\right]^{-\epsilon+1}-[P_t^A(j)]^{-\epsilon+1}\right\}$$
$$=(\epsilon-1)\left[P_t^A(j)-\left(\frac{W_t}{A_t}\right)\right] \tag{15}$$
$$\left\{\left[P_t^A(j)+P_tt\left(\frac{P_t^B(j)-P_t^A(j)}{P_t2t}+\frac{1}{2}\right)\right]^{-\epsilon}\frac{1}{2}-[P_t^A(j)]^{-\epsilon}\right\},$$

which can be written in terms of inflation rates as

$$\left\{\left[(1+\pi_t^A(j))+(1+\pi_t)t\left(\frac{\pi_t^B(j)-\pi_t^A(j)}{(1+\pi_t)2t}+\frac{1}{2}\right)\right]^{-\epsilon+1}-[1+\pi_t^A(j)]^{-\epsilon+1}\right\}$$
$$=(\epsilon-1)\left[(1+\pi_t^A(j))-\left(\frac{w_t}{A_t}(1+\pi_t)\right)\right] \tag{16}$$
$$\left\{\left[(1+\pi_t^A(j))+(1+\pi_t)t\left(\frac{\pi_t^B(j)-\pi_t^A(j)}{(1+\pi_t)2t}+\frac{1}{2}\right)\right]^{-\epsilon}\frac{1}{2}-[1+\pi_t^A(j)]^{-\epsilon}\right\},$$

where the real wage is defined as $w_t \equiv \frac{W_t}{P_t}$, and which implicitly defines the best response function of firm A, $BR^A\left(\pi_t^B(j), \pi_t, \frac{w_t}{A_t}\right)$, as a function of the price inflation of the industry competitor $\pi_t^B(j)$, the overall inflation π_t, and the real marginal cost $\frac{w_t}{A_t}$. Firm A's best response to its beliefs about the price inflation of the competitor and the overall inflation—$\mathcal{B}_{\pi_t^B(j)}$ and $\mathcal{B}_{\pi_t}$—would then be $BR^A\left(\mathcal{B}_{\pi_t^B(j)}, \mathcal{B}_{\pi_t}, \frac{w_t}{A_t}\right)$, implicitly defined as

$$\left\{\left[(1+\pi_t^A(j))+(1+\mathcal{B}_{\pi_t})t\left(\frac{\mathcal{B}_{\pi_t^B(j)}-\pi_t^A(j)}{(1+\mathcal{B}_{\pi_t})2t}+\frac{1}{2}\right)\right]^{-\epsilon+1}-[1+\pi_t^A(j)]^{-\epsilon+1}\right\}$$
$$=(\epsilon-1)\left[(1+\pi_t^A(j))-\left(\frac{w_t}{A_t}(1+\mathcal{B}_{\pi_t})\right)\right] \tag{17}$$
$$\left\{\left[(1+\pi_t^A(j))+(1+\mathcal{B}_{\pi_t})t\left(\frac{\mathcal{B}_{\pi_t^B(j)}-\pi_t^A(j)}{(1+\mathcal{B}_{\pi_t})2t}+\frac{1}{2}\right)\right]^{-\epsilon}\frac{1}{2}-[1+\pi_t^A(j)]^{-\epsilon}\right\},$$

where level-0 thinking beliefs are defined as

$$\mathcal{B}^0_{\pi_t^B(j)} \equiv \pi_{ss}^B(j),$$

and

$$\mathcal{B}^0_{\pi_t} \equiv \pi_{ss},$$

where $\pi_{ss}^B(j)$ and π_{ss} are the steady state levels of $\pi_t^B(j)$ and π_t—which are equal—and where level-k thinking beliefs are defined as

$$\mathcal{B}^k_{\pi_t^B(j)} \equiv \mathcal{B}^k_{\pi_t} \equiv BR^B\left(\mathcal{B}^{k-1}_{\pi_t^A(j)}, \mathcal{B}^{k-1}_{\pi_t}, \frac{w_t}{A_t}\right),$$

where $BR^B\left(\pi_t^A(j), \pi_t, \frac{w_t}{A_t}\right)$ is the best response function of firm B. Following the same argument as before, one can show that firm B's best response to its beliefs about the price inflation of the industry competitor and the overall inflation is implicitly defined from the expression[13]

$$\begin{aligned}\left\{[1+\pi_t^B(j)]^{-\epsilon+1} - \left[(1+\pi_t^B(j)) + (1+\mathcal{B}_{\pi_t})t\left(\frac{1}{2} - \frac{\pi_t^B(j) - \mathcal{B}_{\pi_t^A(j)}}{(1+\mathcal{B}_{\pi_t})2t}\right)\right]^{-\epsilon+1}\right\} \\ = (\epsilon-1)\left[(1+\pi_t^B(j)) - \left(\frac{w_t}{A_t}(1+\mathcal{B}_{\pi_t})\right)\right] \\ \left\{[1+\pi_t^B(j)]^{-\epsilon} - \left[(1+\pi_t^B(j)) + (1+\mathcal{B}_{\pi_t})t\left(\frac{1}{2} - \frac{\pi_t^B(j) - \mathcal{B}_{\pi_t^A(j)}}{(1+\mathcal{B}_{\pi_t})2t}\right)\right]^{-\epsilon}\frac{1}{2}\right\}.\end{aligned} \tag{18}$$

Notice that the best response functions are time- and industry-independent, as the problem of the firm is static and industries are all alike.

[13] These best responses would also arise from the problem of maximizing the *expected* profits of the firm, where such an expectation is taken over a degenerate probability distribution that places all probability mass on these given beliefs. Notice that a more sophisticated probability distribution that considers a set of possible outcomes would not have a continuous support but a discrete one, as firm A knows firm B's best response function but just does not know the degree of level-k thinking that firm B is applying to form its beliefs.

2.3 Government

The government in this economy consists on a monetary policy rule that determines M_t, and a fiscal rule that determines T_t. First, I assume that the money supply follows an $AR(1)$ process in the growth rate

$$\Delta \ln M_t = (1 - \rho_m)\pi_{ss} + \rho_m \Delta \ln M_{t-1} + \epsilon_{m,t}, \tag{19}$$

where $\Delta \ln M_t \equiv \ln M_t - \ln M_{t-1}$, $\rho_m \in (0, 1)$, and $\epsilon_{m,t}$ is a money supply shock.

Second, I assume that the government does no spending, but it just generates revenues with the lump sum tax T_t and by printing money. Therefore, the government's budget constraint in nominal terms is

$$0 \leq P_t T_t + M_t - M_{t-1}. \tag{20}$$

At equality, the lump sum tax must satisfy

$$T_t = -\frac{M_t - M_{t-1}}{P_t}, \tag{21}$$

which defines a fiscal rule that determines T_t. So, if money growth is positive—$M_t > M_{t-1}$— the lump sum tax will be actually negative, meaning that the government will be rebating its seignorage revenues to the households.

2.4 Equilibrium and Aggregation

To close the model we need to specify a process for A_t. I assume that it follows a mean zero $AR(1)$ process in the log

$$\ln A_t = \rho_a \ln A_{t-1} + \epsilon_{a,t}, \tag{22}$$

where $\rho_a \in (0, 1)$, and $\epsilon_{a,t}$ is a productivity shock.

For simplicity, I will restrict to equilibria where all firms apply the same level-k thinking when forming their expectations[14].

[14]The presence of belief heterogeneity would lead to a much more convoluted framework without adding clear conceptual gains for the purposes of this paper. It does, nevertheless, play an important role in learning

Let us define the aggregate price index P_t as

$$P_t Y_t \equiv \int_0^1 P_t^A(j) Y_t^A(j) + P_t^B(j) Y_t^B(j) dj, \tag{23}$$

where aggregate output is

$$Y_t \equiv \left(\int_0^1 Y_t(j)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}}. \tag{24}$$

Integrating the budget constraint across all households and applying the fiscal rule (21), the fact that bonds are in zero net supply, and the results from appendix A.1, we get

$$\int_0^1 \tilde{P}_t(i) C_t(i) di = W_t N_t + \Pi_t,$$

which, given that the effective price index is homogeneous across households, reduces to

$$\tilde{P}_t C_t = W_t N_t + \Pi_t. \tag{25}$$

Total profits are the aggregation of the profits of all firms across industries

$$\begin{aligned} \Pi_t &= \int_0^1 \pi_t^A(j) + \pi_t^B(j) dj \\ &= \int_0^1 [P_t^A(j) Y_t^A(j) - W_t N_t^A(j)] + [P_t^B(j) Y_t^B(j) - W_t N_t^B(j)] dj \\ &= \int_0^1 P_t^A(j) Y_t^A(j) + P_t^B(j) Y_t^B(j) dj - W_t \int_0^1 N_t(j) dj \\ &= \int_0^1 P_t^A(j) Y_t^A(j) + P_t^B(j) Y_t^B(j) dj - W_t N_t, \end{aligned} \tag{26}$$

where $N_t(j) \equiv N_t^A(j) + N_t^B(j)$ for any $j \in [0,1]$, and where labor market clearing implies $N_t = \int_0^1 N_t(j) dj$.

Combining both equations we get

$$\tilde{P}_t C_t = \int_0^1 P_t^A(j) Y_t^A(j) + P_t^B(j) Y_t^B(j) dj, \tag{27}$$

models.

which, from the definition of the aggregate price index, reduces to

$$\tilde{P}_t C_t = P_t Y_t,$$

or

$$\frac{1+\tilde{\pi}_t}{1+\pi_t} C_t = Y_t, \tag{28}$$

which is the resource constraint of this economy. Appendix A.4 shows that, in equilibrium,

$$\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right) = \left[2\frac{\left[1+\frac{t}{2}\right]^{-\epsilon+1} - 1}{(-\epsilon+1)t}\right]^{-\frac{1}{\epsilon}}, \tag{29}$$

which is a function of t, equals 1 as t goes to zero, and has a positive first partial derivative and a negative second partial derivative. This increasing function of t captures aggregate transportation costs, which are proportional to consumption. When $t = 0$ there are no transportation costs, $\frac{1+\tilde{\pi}_t}{1+\pi_t} = 1$, and $C_t = Y_t$. However, as t increases, aggregate transportation costs will be higher. Another way to think about this function is as $\frac{1+\tilde{\pi}_t}{1+\pi_t} = 1 + \Phi(t)$, where $\Phi(t)C_t$ would be the aggregate transportation costs[15].

Appendix A.4 derives the following equilibrium equations for $\tilde{\pi}_t$,

$$\begin{aligned}(1+\tilde{\pi}_t)^{1-\epsilon} = & \frac{\left[(1+\pi_t^A) + (1+\pi_t)t\frac{1}{2}\right]^{2-\epsilon} + \left[(1+\pi_t^B) + (1+\pi_t)t\frac{1}{2}\right]^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t} \\ & - \frac{(1+\pi_t^A)^{2-\epsilon} + (1+\pi_t^B)^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t},\end{aligned} \tag{30}$$

and π_t,

[15]Notice that this resource constraint is similar to the one in models with menu or Rotemberg adjustment costs. The difference is that, while in those models the costs enter additively, here they are effectively a mark-up over consumption. This makes sense as transportation costs at the micro level are per unit purchased/consumed. Notice also that, to be fully rigorous, the production of each individual firm includes not only the consumption from its customer base but also its share in the aggregate transportation costs, paid in terms of final goods. If these shares are distributed as consumption, equation (14) should include the component $\frac{1+\tilde{\pi}_t}{1+\pi_t}$ multiplying the consumption function. However, the reason why I am not doing this is for the sake of simplicity and clarity, as such component is a constant that would cancel out in the first order conditions and not be part of the best response function anyways, since the firm would not be internalizing the effect of its price on the aggregate transportation costs.

$$
\begin{aligned}
(1+\pi_t) = (1+\pi_t^A)\left(\frac{(1+\tilde{\pi}_t)}{(1+\pi_t)}\right)^{\epsilon} & \frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right)+\frac{t}{2}\right]^{-\epsilon+1}-\left(\frac{1+\pi_t^A}{1+\pi_t}\right)^{-\epsilon+1}}{(-\epsilon+1)t} \\
+(1+\pi_t^B)\left(\frac{(1+\tilde{\pi}_t)}{(1+\pi_t)}\right)^{\epsilon} & \frac{\left[\left(\frac{1+\pi_t^B}{1+\pi_t}\right)+\frac{t}{2}\right]^{-\epsilon+1}-\left(\frac{1+\pi_t^B}{1+\pi_t}\right)^{-\epsilon+1}}{(-\epsilon+1)t},
\end{aligned} \tag{31}
$$

where $\pi_t^A = \pi_t^A(j)$ and $\pi_t^B = \pi_t^B(j)$ for all $j \in [0,1]$, as prices are homogeneous across industries in equilibrium because of the fact that all industries are identical.

Appendix A.4 also derives the aggregate production function as

$$
A_t N_t = Y_t, \tag{32}
$$

and shows that, in equilibrium, $\pi_t = \pi_t^A = \pi_t^B$. Notice that, from equation (29), $\tilde{\pi}_t$ is always higher than π_t. However, they are proportional to each other and have the same time dynamics.

3 Quantitative Analysis

This section carries out the quantitative analysis based on a calibrated version of the model.

I set the parameters of the utility function to the standard values in macroeconomics $\sigma = \Psi = \eta = \theta = 1$. Also following the standards in the literature, I set the subjective discount factor equal to $\beta = 0.98$, the elasticity of substitution between differentiated industry goods to $\epsilon = 6$, and I assume a zero inflation steady state, $\pi_{ss} = 0$.

The parameter t is the transportation cost per unit of distance in the product space of a given industry. This is a new parameter that determines the degree of market competition in the economy. I set $t = 0.05$, which implies that purchasing a unit of an industry good one unit of distance away from one's ideal good incurs in a cost or depreciation equivalent to 5% of a unit of the final good. Appendix B shows the robustness of the results to alternative calibrations for t.

I consider a shock to the money supply. I set $\rho_m = 0$ so that nominal money follows a random walk and the shock results in a one time permanent shift in M_t, and I set the standard deviation of the shock to $\sigma_m = 0.01$. I carry out this exercise six times, on six economies that differ in the level of thinking that firms apply to their beliefs: respectively characterized by level-0, level-1, level-2, level-3, level-4, and level-5 thinking. Figure 3 shows the impulse responses from all six exercises together.

In all the six cases, the shock leads to a permanent shift in M_t, since I have set $\rho_m = 0$. Moreover, the nominal interest rate does not move.[16] The permanent, upward shift in M_t—which is then rebated to the households—is a positive demand shock that rises quantities and prices. We have a rise in consumption and output, together with inflationary pressure; and also an increase in labor and wages, since labor demand rises. There is also a temporary rise in the real money balance, m_t.

Why are not prices strongly and inmediately rising to the extend of undoing the effect of the shock as in the RBC model? Consider the case of a given industry j, which has two competitors located at the extremes of the product space. A positive demand shock will rise the marginal cost, since that is necessary to get workers to work more. However, the pass-through from that rise in the marginal cost to the price of the firm will depend on how the firm predicts that the industry competitor will react. If the firm rises the price unilaterally, part of its customer base will move away to its industry competitor. Therefore, the firm will be reluctant to rise its price *too much* over the price of the industry competitor, since that would have a detrimental effect on its customer base. Hence, the firm will only be willing to rise its price significantly when the industry competitor is doing something similar, so that this price increase is not as negatively perceived by its customer base. That is also the reason why the effect on output is lower the higher is the level of thinking that the firm is applying to its beliefs, since those correspond to the belief that the industry competitor would be rising its price by more.[17] Figure B.1, in appendix B, shows that the real effects of the shock slightly decline with t, as a higher t captures a less competitive market where the customer base of the firm is more *captive*, and so the firm is more willing to rise prices. Section 4 explores this incomplete marginal cost-to-price pass-through mechanism in depth, providing strong empirical grounds and discussing the role of the key assumptions.

Taylor Rule

This section substitutes the money supply process for the following Taylor rule:

[16] This is easy to see from the first order conditions of the household's problem. Solving forward the first order condition for money, we get $\lambda_t = \mathbb{E}_t \sum_{j=0}^{\infty} \beta^t \frac{\theta}{M_{t+j}}$. Since M_t follows a random walk, the shock will make both λ_t and λ_{t+1} fall, but by the same amount. From the first order condition for bonds, $\lambda_t = \beta \mathbb{E}_t[\lambda_{t+1}(1 + i_t)]$, this implies that i does not react.

[17] Note how the difference in impulse responses from level-k to level-$k+1$ thinking will be smaller the higher is k.

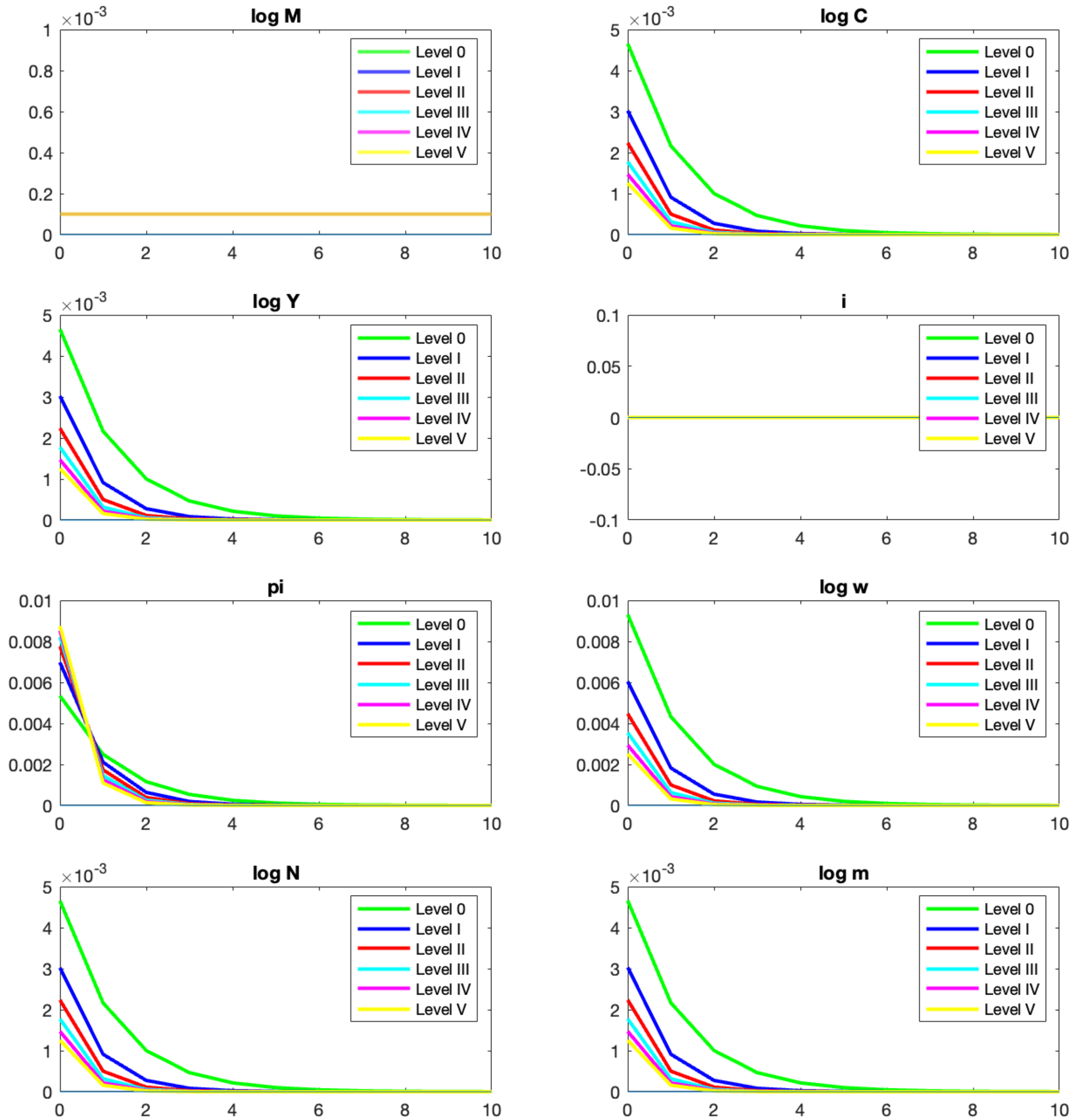

Figure 3: Impulse Responses

$$i_t = (1-\rho_i)i_{ss} + \rho_i i_{t-1} + (1-\rho_i)\phi_\pi(\pi_t - \pi_{ss}) + \epsilon_i, \tag{33}$$

where $i_{ss} = \frac{1+\pi_{ss}}{\beta} - 1$ is the steady state nominal interest rate, and ϵ_i is a monetary policy shock—analogous to the ϵ_m in the money supply process. Notice that there is no mention of money in this policy rule specification: given the chosen nominal interest rate, the central bank will implicitly print the necessary amount of money to meet money demand at that interest rate. Following the standards in the literature, I set $\rho_i = 0.8$ and $\phi_\pi = 1.5$. Next, I set the standard deviation of the shock to $\sigma_i = 0.01$. I consider an expansionary monetary policy shock and, as before, I carry out the exercise six times, on six economies that differ in the beliefs of their firms. Figure 4 shows the impulse responses from all six exercises together.

Wage Inertia

One of the main mechanisms of this model rests on the incomplete marginal cost-to-price pass-through that derives from firms' competition for their customer base. Hence, it is natural to wonder how the model behaves when wages follow more realistic dynamics. An interesting way to do this would be to model the labor market as in Christiano et al. (2016), who derive wage inertia endogenously, based on how firms and workers negotiate wages. Instead—with the aim of maximizing simplicity and clarity—I extend the model introducing Calvo sticky wages as in Erceg et al. (2000). Appendix B.2 goes through the details of this exercise, where I set the elasticity of substitution $\epsilon_w = 10$, and the Calvo parameter $\phi_w = 0.5$. Figure 5 shows the impulse responses of a money supply shock with standard deviation $\sigma_m = 0.01$.

This is exactly the same shock as in figure 3, yet the impulse responses are quantitatively different. First, we can see that the responses of all real variables roughly double both in size and duration. This is due to wage inertia: the money supply shock is effectively a positive demand shock that rises labor demand and puts an upward pressure on wages; however, the dynamics of wages are now much more sluggish as a result of the Calvo assumption, in the sense that the initial reaction of wages is much smaller and the overall response is more prolonged over time. The interaction of the incomplete marginal cost-to-price pass-through mechanism with the sluggish dynamics of wages leads to a much more sluggish behavior of prices too.[18] As a result, the responses of the real variables in the economy are much more

[18] Notice how the wage dynamics follow a bit of a hump-shape, which is also passed to inflation. This is a standard feature of the Calvo assumption.

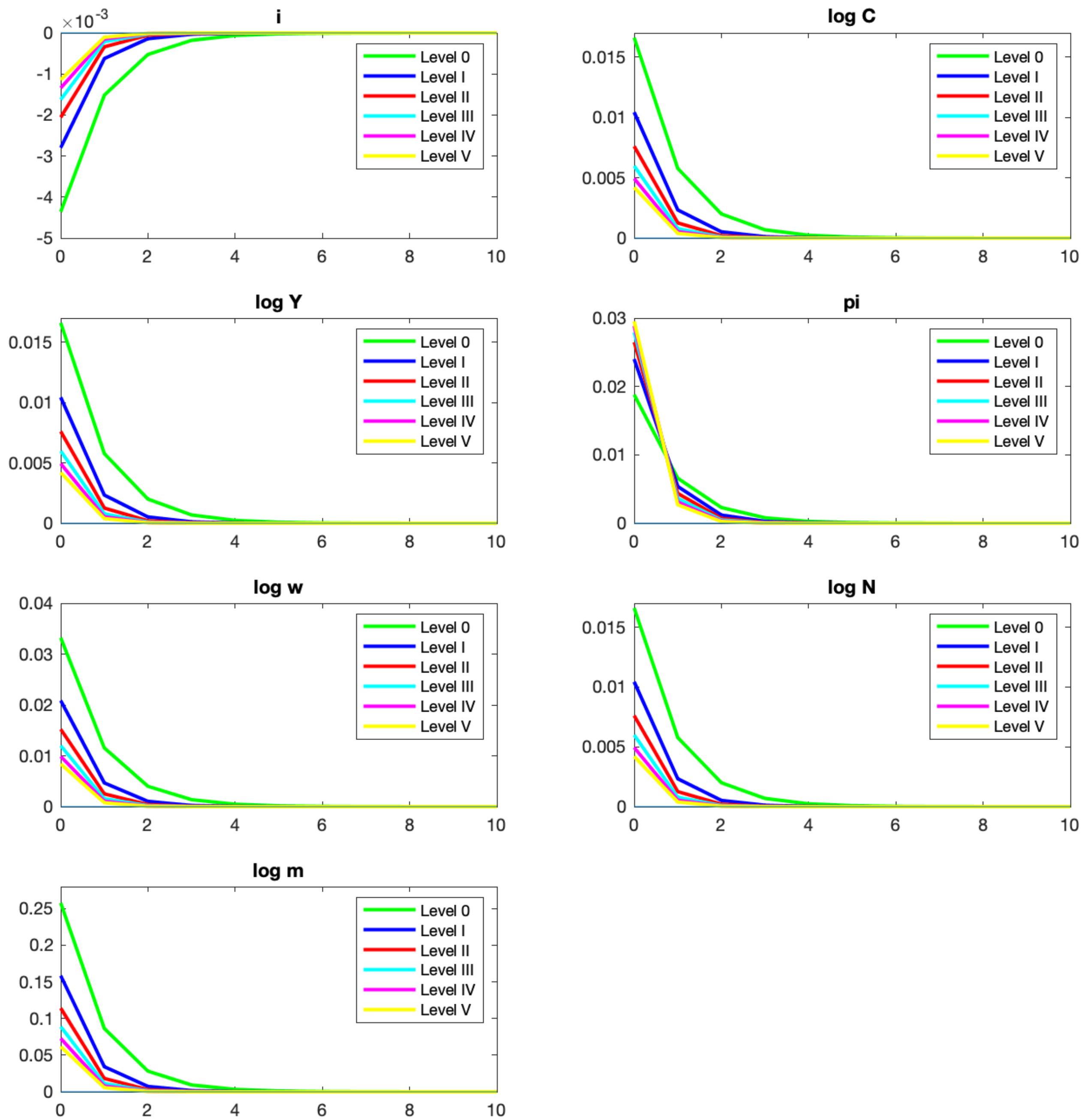

Figure 4: Impulse Responses under a Taylor rule

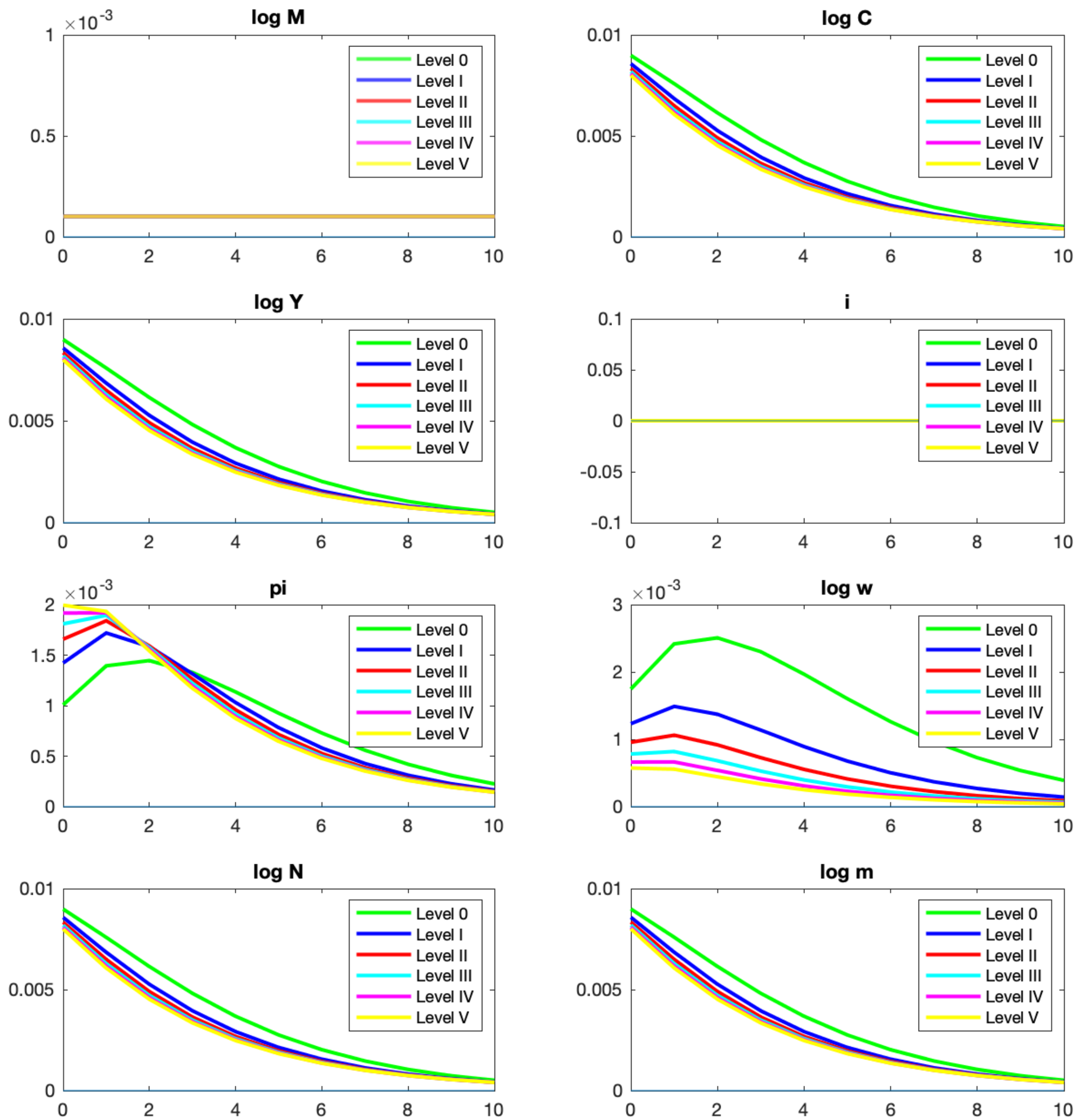

Figure 5: Impulse Responses under sticky wages

sizable and prolonged.

Moreover, we can see that the impulse responses of the real variables across levels of thinking are much more close to each other than in the case of flexible wages.[19] In other words, the pace at which impulse responses get smaller as k gets higher is much slower under sticky wages than it is under flexible wages.

4 Empirical Foundations and Theoretical Robustness

The mechanisms driving the results in the previous section are not only theoretically tractable but also grounded in empirical evidence. This section reviews the empirical support for the model's core assumptions, underscores the robustness of its central mechanism, and outlines the model's main implications—each of which is backed by solid empirical findings.

The central mechanism behind the model's non-neutrality result is the incomplete marginal cost-to-price pass-through that arises in the short run following a demand shock. This mechanism is strongly supported by the extensive empirical literature on pass-through. For example, Nakamura and Zerom (2010) study the U.S. coffee industry and find that prices respond incompletely—and with a lag—to changes in the cost of raw materials. Gorodnichenko and Talavera (2017) find highly incomplete pass-through even in online markets. Ganapati et al. (2020) document incomplete pass-through from energy prices to manufacturing prices in the U.S., while Gopinath et al. (2010) focus on exchange rate pass-through to export prices. Finally, Cavallo et al. (2021) study tariff pass-through using variation from U.S. trade policy.

It is important to note that both strategic complementarities and level-k beliefs are essential to generate the incomplete marginal cost-to-price pass-through and, therefore, the model's non-neutrality result.

On the one hand, could a standard monopolistic competition model combined with level-k thinking deliver a similar outcome? The answer is no. In a monopolistic competition framework—equivalent to the current setting under infinite transportation costs—each industry is served by a monopolist facing a demand curve of the form

$$Y_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} Y_t.$$

Suppose the firm believes the aggregate price level to be $\mathcal{B}_{P_t}$. The firm then maximizes the

[19] As an illustration, with this calibration the impulse responses for output at the time of the shock are around 64% under level-1 thinking than what they are under level-0 thinking in figure 3, while this number rises to 96% in figure 5 under sticky wages. Something similar happens to the duration of the responses.

profit function:

$$\left[P_t(j) - \frac{W_t}{A_t}\right]\left(\frac{P_t(j)}{\mathcal{B}_{P_t}}\right)^{-\epsilon} Y_t.$$

The optimal price that solves this problem remains:

$$P_t(j) = \frac{\epsilon}{\epsilon - 1}\frac{W_t}{A_t},$$

a constant markup over marginal cost. The firm's beliefs about the aggregate price index affect its expected profits, but not its optimal price choice. This is because the demand elasticity, $\frac{\partial \ln Y_t(j)}{\partial \ln P_t(j)} = -\epsilon$, is constant and independent of the level of prices. By contrast, in the horizontally differentiated framework used here, the price elasticity of demand depends on both the firm's own price and the price set by its competitor, as equation (14) makes clear.

Moving away from the standard representative-agent and monopolistic competition setups to an industry structure with horizontal differentiation is not only critical for the model's mechanism, but also enhances its empirical realism. Amiti et al. (2019) empirically investigate a micro panel of Belgian manufacturing firms and find that the typical firm adjusts its price with an elasticity of about 35% in response to competitor prices, and 65% in response to its own cost shocks. Koga et al. (2020) report similar findings using micro data from Japanese firms.

On the other hand, could the model generate non-neutrality by combining horizontal differentiation with level-∞ thinking? The answer is again no. Under such infinite sophistication, firms behave as if endowed with perfect foresight: they fully anticipate competitors' responses and instantaneously coordinate on the unique Nash equilibrium. As a result, the model collapses to the standard neutrality outcome. By contrast, under finite level-k thinking, firms form coherent but boundedly rational beliefs about competitors' actions and best respond accordingly. As shown in Figure 3, this leads to a gradual convergence toward the Nash equilibrium outcome over time, as firms iteratively adjust their strategies and marginal costs evolve.

This assumption of bounded strategic reasoning is grounded in both theory and evidence. Evolutionary game theory suggests that Nash equilibrium emerges only after a process of learning and adaptation, making it an evolutionarily stable outcome rather than an initial condition (Mailath, 1998; Samuelson, 2002; Shapiro et al., 2014). Similarly, a vast experimental literature documents that the play of experimental subjects systematically violates the Nash equilibrium predictions—which can be very complex to calculate as it assumes a high degree of rationality on the part of the subject and also assumes that the subject assumes a high degree of rationality on the part of others—supporting level-k thinking as a simpler and empirically plausible approximation (Nagel, 1995; Bosch-Domenech et al., 2002).

Lastly, the model yields two direct implications that are both testable and policy-relevant:[20]

1. **Monetary and demand shocks have stronger real effects in markets with higher levels of competition.**

   In the model, the strength of real effects declines with market power, which is directly proportional to transportation costs. This implication aligns well with a broad empirical literature. For instance, Duval et al. (2024) combine firm-level data for the U.S. and 14 other advanced economies to show that greater market power dampens firms' output responses to monetary policy shocks. Complementary evidence comes from the literature on pass-through: Hong and Li (2017) find that cost pass-through increases with firm-level market share using retail scanner data, while Pless and Benthem (2019) document similar results in the context of solar subsidies. These findings suggest that firms with greater market power adjust prices more aggressively in response to shocks, muting the real effects of demand disturbances. From a policy perspective, if market power is trending upward over time (Autor et al., 2020; De Loecker et al., 2020), the effectiveness of monetary policy as a stabilization tool may be structurally declining.

2. **The real effects of shocks depend on the inflationary environment—through firms' evolving cognitive sophistication.**

   In the model, real effects diminish with the level-k of thinking: more sophisticated firms coordinate more quickly and offset shocks more effectively. At the same time, experimental economics shows that repeated exposure to similar games may increase players' strategic sophistication (Duffy and Nagel, 1997; Cooper et al., 2024). A natural implication of these two insights is that monetary and demand shocks should be more inflationary in economies with a recent history of high inflation, whereas they should have larger real effects in stable inflation contexts. This asymmetry is intuitive: in low-inflation contexts, price signals are less salient and learning incentives are weaker, whereas high-inflation environments trigger faster adaptation and deeper strategic reasoning.[21] The empirical evidence is consistent with this. Alvarez et al.

[20] Of course beyond the model's prediction of price inertia, which is also strongly supported by empirical evidence, e.g., Bils and Klenow (2004) and Nakamura and Steinsson (2008).

[21] This insight is supported by macroeconomic theory (Sims, 2003; Mackowiak and Wiederholt, 2009),

(2019) and Gagnon (2009) study Argentina and Mexico, respectively, and find that, in high-inflation contexts, firms adjust prices more frequently and by larger amounts. In contrast, in low-inflation settings, pricing behavior is relatively stable. Nakamura et al. (2018) finds similar results for the US during the Great Inflation, except that price changes became more frequent, but not necessarily larger. The policy implication of this is that the effectiveness of monetary policy may be limited in economies with a recent history of high-inflation.

5 Concluding Remarks

This paper develops a general equilibrium model that combines strategic complementarities—based on horizontally differentiated markets—with boundedly rational expectations to generate short-run non-neutrality of monetary and demand shocks with respect to real variables in line with the evidence. In sharp contrast to the previous literature and leading DSGE models, I do not impose any sort of exogenous nominal rigidities or ad hoc restrictions to agents. Instead, price inertia derives as an equilibrium outcome where firms compete for their customer base in horizontally differentiated markets, which results in a short-run, incomplete marginal cost-to-price pass-through.

This paper abstracts from a number of things that go beyond the scope and are left for future research. First, an alternative way to introduce wage inertia would be to model the labor market as in Christiano et al. (2016), and derive wage inertia from the way firms and workers negotiate wages. This would result in a framework where both prices and wages are free from any sort of exogenously imposed nominal rigidity. Second, the model abstracts from capital and investment in order to maximize clarity and simplicity, but it is easy to introduce them. Third, the model abstracts from learning, but it takes the degree of level-k thinking—which is common across firms—as *given*. A natural way to introduce learning would be to have firms that revise their depth of reasoning in response to forecast errors as in Evans et al. (2024). Learning can be slow (Christiano et al., 2024) and hence its influence at business cycle frequencies may be limited; however, it opens the door to a very interesting cross-sectional analysis as it would suggest that the same monetary policy shock would have different effects in economies characterized by different inflation histories.[22] Fourth, the model takes the degree of market competition as given. I find this to be a reasonable

behavioral economics (Bordalo et al., 2013), and the empirical evidence from natural experiments (Alvarez et al., 2019; D'Acunto et al., 2021).

[22] In particular, this framework—equipped with learning—would predict that the same monetary shock would be relatively more inflationary in an economy like Argentina than in economies with a recent history of more stable inflation such as the US or Germany, where it would be relatively more expansionary.

simplification at business cycle frequencies; however, it can be endogenized by assuming instead a circular product space where there is firm entry and a certain entry cost. Last, a natural way to account for the micro-level infrequency of price adjustment in the data would be to introduce some sort of rational inattention in firms. However, the model does not need to match this micro-level feature of the data to generate non-neutrality at the aggregate level, so I consider this to be second-order for the purposes of this paper.

This paper provides microfoundations to the phenomenon of price inertia and short-run non-neutrality—prevalent in both the empirical evidence and the common understanding of policymakers—leading to a framework reasonably robust for policy evaluation.

References

Altig, D., L. Christiano, M. Eichenbaum, and J. Lindé (2011). "Firm-Specific Capital, Nominal Rigidities and the Business Cycle". *Review of Economic Dynamics* 14(2), pp. 225–247.

Alvarez, F., M. Beraja, M. Gonzalez-Rozada, and P. A. Neumeyer (2019). "From Hyperinflation to Stable Prices: Argentina's Evidence on Menu Cost Models". *The Quarterly Journal of Economics* 134(1), pp. 451–505.

Amiti, M., O. Itskhoki, and J. Konings (2019). "International Shocks, Variable Markups and Domestic Prices". *Review of Economic Studies* 86(6), pp. 2356–2402.

Angeletos, G.-M. and C. Lian (2018). "Forward Guidance without Common Knowledge". *American Economic Review* 108(9), pp. 2477–2512.

Atkeson, A. and A. Burstein (2008). "Pricing-to-Market, Trade Costs, and International Relative Prices". *American Economic Review* 98(5), pp. 1998–2031.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020). "The Fall of the Labor Share and the Rise of Superstar Firms". *The Quarterly Journal of Economics* 135(2), pp. 645–709.

Barro, R. and C. Redlick (2011). "Macroeconomic Effects from Government Purchases and Taxes". *The Quarterly Journal of Economics* 126(1), pp. 51–102.

Bergin, P. and R. Feenstra (2001). "Pricint-to-Market, Staggered Contracts, and Real Exchange Rate Persistence". *Journal of International Economics* 54(2), pp. 333–359.

Bils, M. and P. J. Klenow (2004). "Some Evidence on the Importance of Sticky Prices". *Journal of Political Economy* 112.5, pp. 947–985.

Blanchard, O. and R. Perotti (2002). "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output". *The Quarterly Journal of Economics* 117(4), pp. 1329–1368.

Bordalo, P., N. Gennaioli, and A. Shleifer (2013). "Salience and Consumer Choice". *Journal of Political Economy* 121(5).

Bosch-Domenech, A., J. G. Montalvo, R. Nagel, and A. Satorra (2002). "One, Two, (Three), Infinity, ... : Newspaper and Lab Beauty-Contest Experiments". *American Economic Review* 92(5), pp. 1687–1701.

Burstein, A. and C. Hellwig (2007). "Prices and Market Shares in a Menu Cost Model". *NBER Working Paper No. w13455*.

Calvo, G. A. (1983). "Staggered Prices in a Utility-Maximizing Framework". *Journal of Monetary Economics* 12, pp. 383–98.

Caplin, A. S. and D. F. Spulber (1987). "Menu Costs and the Neutrality of Money". *The Quarterly Journal of Economics* 102(4), pp. 703–726.

Cavallo, A., G. Gopinath, B. Neiman, and J. Tang (2021). "Tariff Pass-Through at the Border and at the Store: Evidence from US Trade Policy". *American Economic Review: Insights* 3(1), pp. 19–34.

Chari, V., P. Kehoe, and E. McGrattan (2000). "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?" *Econometrica* 68, pp. 1151–1179.

Christiano, L., M. Eichenbaum, and C. Evans (1999). "Monetary Policy Shocks: What have we learned and to what end?" *Handbook of Macroeconomics* 1, Part A, pp. 65–148.

Christiano, L., M. Eichenbaum, and M. Trabandt (2016). "Unemployment and Business Cycles". *Econometrica* 84(4), pp. 1523–1569.

– (2024). "Slow Learning". *Working Paper*.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005). "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy". *Journal of Political Economy* 113(1), pp. 1–45.

Cooley, T. F. and G. D. Hansen (1989). "Inflation Tax in a Real Business Cycle Model". *American Economic Review* 79, pp. 733–748.

Cooper, D. J., E. Fatas, A. J. Morales, and S. Qi (2024). "Consistent Depth of Reasoning in Level-k Models". *American Economic Journal: Microeconomics* 16(4), pp. 40–76.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber (2021). "Exposure to Grocery Prices and Inflation Expectations". *Journal of Political Economy* 129(5).

De Loecker, J., J. Eeckhout, and G. Unger (2020). "The Rise of Market Power and the Macroeconomic Implications". *The Quarterly Journal of Economics* 135(2), pp. 561–644.

Dotsey, M. and R. King (2006). "Pricing, Production, and Persistence". *Journal of the European Economic Association* 4, pp. 893–928.

Duffy, J. and R. Nagel (1997). "On the Robustness of Behavior in Experimental 'Beauty Contest´ Games". *The Economic Journal* 107(445), pp. 1684–1700.

Duval, R., D. Furceri, R. Lee, and M. Tavares (2024). "Market Power and Monetary Policy Transmission". *Economica, London School of Economics and Political Science* 91(362), pp. 669–700.

Eichenbaum, M. and J. Fisher (2007). "Estimating the Frequency of Price Re-optimization in Calvo-style Models". *Journal of Monetary Economics* 54(7), pp. 2032–2047.

Eichengreen, B. and J. Sachs (1985). "Exchange Rates and Economic Recovery in the 1930s". *The Journal of Economic History* 45(4), pp. 925–946.

Erceg, C. J., D. W. Henderson, and A. T. Levin (2000). "Optimal Monetary Policy with Staggered Wage and Price Contracts". *Journal of Monetary Economics* 46(2), pp. 281–313.

Evans, G., C. Gibbs, and B. McGough (2024). "A Unified Model of Learning to Forecast". *American Economic Journal: Macroeconomics*, forthcoming.

Farhi, E. and I. Werning (2019). "Monetary Policy, Bounded Rationality, and Incomplete Markets". *American Economic Review* 109(11), pp. 3887–3928.

Friedman, M. and A. Schwartz (1963). "A Monetary History of the United States, 1867-1960". *Princeton University Press.*

Gagnon, E. (2009). "Price Setting during Low and High Inflation: Evidence from Mexico". *The Quarterly Journal of Economics* 124(3), pp. 1221–1263.

Galí, J. (2015). "Monetary Policy, Inflation, and the Business Cycle". *Princeton University Press.*

Ganapati, S., J. Saphiro, and R. Walker (2020). "Energy Cost Pass-Through in the U.S. Manufacturing: Estimates and Implications for Carbon Taxes". *American Economic Journal: Applied Economics* 12(2), pp. 303–342.

Gertler, M. and P. Karadi (2015). "Monetary Policy Surprises, Credit Costs, and Economic Activity". *American Economic Journal: Macroeconomics* 7(1), pp. 44–76.

Gertler, M. and J. Leahy (2008). "A Phillips Curve with an Ss Foundation". *Journal of Political Economy* 116(3), pp. 533–572.

Golosov, M. and R. E. Lucas (2007). "Menu Costs and Phillips Curves". *Journal of Political Economy* 115(2), pp. 171–199.

Gopinath, G., O. Itskhoki, and R. Rigobon (2010). "Currency Choice and Exchange Rate Pass-Through". *American Economic Review* 100(1), pp. 304–336.

Gorodnichenko, Y. and O. Talavera (2017). "Price Setting in Online Markets: Basic Facts, International Comparisons, and Cross-Border Integration". *American Economic Review* 107(1), pp. 249–282.

Guajardo, J., D. Leigh, and A. Pexcatori (2014). "Expansionary Austerity? International Evidence". *Journal of the European Economic Association* 12(4), pp. 949–968.

Hong, G. H. and N. Li (2017). "Market Structure and Cost Pass-Through in Retail". *The Review of Economics and Statistics* 99(1), pp. 151–166.

Hotelling, H. (1929). "Stability in Competition". *The Economic Journal* 39.153, pp. 41–57.

Kimball, M. (1995). "The Quantitative Analytics of the Basic Neomonetarist Model". *Journal of Money, Credit and Banking* 27, pp. 1241–77.

Klenow, P. and J. Willis (2006). "Real Rigidities and Nominal Price Changes". *Stanford University and Federal Reserve Bank of Kansas City*, Working paper.

Koga, M., K. Yoshino, and T. Sakata (2020). “Strategic Complementarity and Asymmetric Price Setting Among Firms”. *BIS Papers Chapters, in: Bank for International Settlements (ed.), Inflation dynamics in Asia and the Pacific* 111, pp. 85–97.

Kydland, F. E. and E. C. Prescott (1982). “Time to Build and Aggregate Fluctuations”. *Econometrica* 50, pp. 1345–1370.

Lucas, R. E. (1976). “Econometric Policy Evaluation: A Critique”. *Carnegie-Rochester Conference Series on Public Policy* 1, pp. 19–46.

Mackowiak, B. and M. Wiederholt (2009). “Optimal Sticky Prices under Rational Inattention”. *American Economic Review* 99(3), pp. 769–803.

Mailath, G. J. (1998). “Do People Play Nash Equilibrium? Lessons from Evolutionary Game Theory”. *Journal of Economic Literature* 36(3), pp. 1347–1374.

Mankiw, N. and R. Reis (2002). “Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve”. *The Quarterly Journal of Economics* 117(4), pp. 1295–1328.

Mian, A. and A. Sufi (2014). “What Explains the 2007-2009 Drop in Employment?” *Econometrica* 82(6), pp. 2197–2223.

Mongey, S. (2021). “Market Structure and Monetary Non-neutrality”. *NBER Working Paper No. 29233*.

Mussa, M. (1986). “Nominal Exchange Rate Regimes and the Behavior of Real Exchange Rates: Evidence and Implications”. *Carnegie-Rochester Conference Series on Public Policy* 25, pp. 117–214.

Nagel, R. (1995). “Unraveling in Guessing Games: An Experimental Study”. *American Economic Review* 85(5), pp. 1313–1326.

Nakamura, E. and J. Steinsson (2008). “Five Facts about Prices: A Reevaluation of Menu Cost Models”. *The Quarterly Journal of Economics* 123.4, pp. 1415–1464.

– (2010). “Monetary Non-neutrality in a Multisector Menu Cost Model”. *The Quarterly Journal of Economics* 125(3), pp. 961–1013.

Nakamura, E. and J. Steinsson (2014). "Fiscal Stimulus in a Monetary Union: Evidence from US Regions". *American Economic Review* 104(3), pp. 753–92.

– (2018). "Identification in Macroeconomics". *Journal of Economics Perspectives* 32(3), pp. 59–86.

Nakamura, E., J. Steinsson, P. Sun, and D. Villar (2018). "The Elusive Costs of Inflation: Price Dispersion During the U.S. Great Inflation". *The Quarterly Journal of Economics* 133(4), pp. 1933–1980.

Nakamura, E. and D. Zerom (2010). "Accounting for Incomplete Pass-Through". *Review of Economic Studies* 77, pp. 1192–1230.

Pless, J. and A. A. van Benthem (2019). "Pass-Through as a Test for Market Power: An Application to Solar Subsidies". *American Economic Journal: Applied Economics* 11(4), pp. 367–401.

Prescott, E. C. (1986). "Theory Ahead of Business Cycle Measurement". *Federal Reserve Bank of Minneapolis Quarterly Review* 10, pp. 9–22.

Ramey, V. (2011). "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature* 49(3), pp. 673–85.

Reis, R. (2006). "Inattentive Producers". *The Review of Economic Studies* 73(3), pp. 793–821.

Romer, C. and D. Romer (2004). "A New Measure of Monetary Shocks: Derivation and Implications". *American Economic Review* 94(4), pp. 1055–1084.

Rotemberg, J. and M. Woodford (1997). "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy". *1997 NBER Macroeconomic Annual*, pp. 297–346.

Rotemberg, J. J. (1982). "Sticky Prices in the United States". *Journal of Political Economy* 90, pp. 1187–1211.

Samuelson, L. (2002). "Evolution and Game Theory". *Journal of Economic Perspectives* 16(2), pp. 47–66.

Shapiro, D., X. Shi, and A. Zillante (2014). “Level-k Reasoning in a Generalized Beauty Contest”. *Games and Economic Behavior* 86, pp. 308–329.

Sims, C. A. (2003). “Implications of Rational Inattention”. *Journal of Monetary Economics* 50(3), pp. 665–690.

Smets, F. and R. Wouters (2003). “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area”. *Journal of the European Economic Association* 1(5), pp. 1123–75.

– (2007). “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach”. *American Economic Review* 97(3), pp. 586–606.

Taylor, J. (1980). “Aggregate Dynamics and Staggered Contracts”. *Journal of Political Economy* 88, pp. 1–23.

Appendices

A Technical Appendix

A.1 Optimal Allocation of Consumption Expenditures

For any given household i, optimal behavior requires $C_t(i)$ to be maximized for any given level of expenditures

$$X_t(i) \equiv \int_0^1 \Big[P_t^A(j)C_t^A(j,i)+P_t^B(j)C_t^B(j,i)\Big]dj+P_t t\int_0^1 \Big[x(j,i)C_t^A(j,i)+[1-x(j,i)]C_t^B(j,i)\Big]dj.$$

This first requires the household to optimally choose the firm that minimizes the effective cost per unit (price plus transportation cost) in each industry or, in other words, to purchase from firm A in those industries where $x(j,i) \leq x_t^*(j)$, and from B otherwise.

We can define the effective price of a unit of industry good j as $\tilde{P}_t(j,i) \equiv [P_t^A(j) + P_t t x(j,i)]\mathbb{1}\{x(j,i) \leq x_t^*(j)\}+[P_t^B(j)+P_t t(1-x(j,i))]\mathbb{1}\{x(j,i) > x_t^*(j)\}$, which is the effective cost of a unit of industry good j for household i when behaving optimally. The solution to the previous problem is then the same as the solution to the problem of maximizing $C_t(i)$ subject to

$$\int_0^1 \tilde{P}_t(j,i)C_t(j,i)dj = X_t(i),$$

which can be formalized with the Lagrangian

$$\mathcal{L} = \left(\int_0^1 C_t(j,i)^{\frac{\epsilon-1}{\epsilon}}dj\right)^{\frac{\epsilon}{\epsilon-1}} + \lambda_t\left[X_t(i) - \int_0^1 \tilde{P}_t(j,i)C_t(j,i)dj\right].$$

The implied first order condition is

$$C_t(i)^{\frac{1}{\epsilon}}C_t(j,i)^{-\frac{1}{\epsilon}} = \lambda_t\tilde{P}_t(j,i)$$

so, for any two industries $j, j' \in [0,1]$, we have that

$$C_t(j,i) = C_t(j',i)\left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t(j',i)}\right)^{-\epsilon}.$$

This expression can be plugged in the constraint to get

$$C_t(j,i) = \left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t(i)}\right)^{-\epsilon}\frac{X_t(i)}{\tilde{P}_t(i)},$$

where $\tilde{P}_t(i) \equiv \left(\int_0^1 \tilde{P}_t(j,i)^{1-\epsilon}dj\right)^{\frac{1}{1-\epsilon}}$, which can be substituted into the definition of $C_t(i)$ to get that

$$\tilde{P}_t(i)C_t(i) = X_t(i).$$

Combining the two previous equations we finally get

$$C_t(j,i) = \left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t(i)}\right)^{-\epsilon} C_t(i),$$

which, for a given household i, defines a set of demand equations for all $j \in [0,1]$.

A.2 Homogeneity of the Effective Aggregate Price Index

This appendix shows that $\tilde{P}_t(i) = \tilde{P}_t(i') = \tilde{P}_t$ for any $i, i' \in [0,1]$.

For a given household i notice that, by definition,

$$\tilde{P}_t(i)^{1-\epsilon} = \int_0^1 \tilde{P}_t(j,i)^{1-\epsilon}dj =$$

$$\int_0^1 \left[[P_t^A(j) + P_t t x(j,i)]\mathbb{1}\{x(j,i) \le x_t^*(j)\} + [P_t^B(j) + P_t t(1-x(j,i))]\mathbb{1}\{x(j,i) > x_t^*(j)\}\right]^{1-\epsilon} dj.$$

Let me define $\theta \equiv \int_0^1 \mathbb{1}\{x(j,i) \le x_t^*(j)\}dj$ as the proportion of industries where household i optimally purchases from firm A; and notice that, given that each $x(j,i)$ location is iid drawn from a continuous uniform $[0,1]$ distribution, the distribution of locations—for that particular household—across industries $[x(j,i)]_{j\in[0,1]}$ is also continuously uniform along the interval $[0,1]$. As no particular order matters, we can relabel the j's to break up the integral in the following way:

$$\tilde{P}_t(i)^{1-\epsilon} = \int_0^\theta \left[[P_t^A(j) + P_t t x(j,i)] \right]^{1-\epsilon} dj + \int_\theta^1 \left[[P_t^B(j) + P_t t(1 - x(j,i))] \right]^{1-\epsilon} dj.$$

First, the weak law of large numbers immediately gives us that $\int_0^\theta P_t^A(j)dj = \theta\mathbb{E}[P_t^A(j)|x \leq x_t^*(j)]$ or, in other words, that the integral (sum) of the individual prices of all the industries where household i's location falls to the left of the threshold is proportional to the weighted sum over the entire unit interval, where the weights are the conditional probabilities given the thresholds $x_t^*(j)$ for all $j \in [0,1]$, and where the proportion is equal to the subset of industries where $x(j,i) \leq x_t^*(j)$[23]. Given that the thresholds $x_t^*(j)$'s are common across households, $\mathbb{E}[P_t^A(j)|x \leq x_t^*(j)]$ is also common across households. Moreover, given iidness and the continuum of industries, so is the proportion $\theta \in [0,1]$. The following result—which is a version of the law of large numbers that does not require an identical distribution but just uncorrelation and finite variance—will be very useful for the rest of this section.

Result 1[24] (Chebyshev's L_2 law of large numbers). *Let X_1, X_2, ... be random variables such that $\mathbb{E}|X_i|^2 < \infty$ for every i. If*

$$\frac{1}{n^2} Var(X_1 + ... + X_n) \xrightarrow[n\to\infty]{} 0,$$

then denoting $S_n = X_1 + ... + X_n$,

$$\frac{S_n}{n} - \mathbb{E}\frac{S_n}{n} \xrightarrow[n\to\infty]{L_2} 0.$$

In particular, this holds when the X_i are uncorrelated with bounded variance, that is $Var(X_i) \leq M$ for every i for some M.

Proof. We have

[23] Actually, the different laws of large numbers do not state equality but different notions of convergence. However, since we have a continuum of industries, we are already in the limit.

[24] This result, stated and proved here for convenience, is of course not my own but one of Chebyshev's weak laws of large numbers, taken from the probability lecture notes of Tomasz Tkocz of Carnegie Mellon University.

$$\mathbb{E}\left|\frac{S_n}{n}-\mathbb{E}\frac{S_n}{n}\right|^2=\frac{1}{n^2}\mathbb{E}|S_n-\mathbb{E}S_n|^2=\frac{1}{n^2}Var(X_1+...+X_n)\xrightarrow[n\to\infty]{}0.$$

Since

$$Var(X_1+...+X_n)=\sum_{i=1}^{n}Var(X_i)+2\sum_{1\le i<j\le n}Cov(X_i,X_j),$$

when the X_i are uncorrelated with bounded variance, we have

$$\frac{1}{n^2}Var(X_1+...+X_n)\le\frac{Mn}{n^2}=\frac{M}{n}$$

which goes to 0 as $n\to\infty$. $\square$

Now, for our given household i, consider the set of $x(j,i)$'s across $j\in[0,1]$ such that $x(j,i)\le x_t^*(j)$. We know that, conditional on such a set, $x(j,i)\sim\mathbb{U}[0,x_t(j)]$. Therefore, and again conditional on such a set, $Var(x(j,i))=\frac{1}{12}(x_t^*(j))^2<\frac{1}{12}$. If we randomly selected n variable x(j,i)'s out of that set and call them $X_1,...,X_n$, we would have that

$$Var(X_1+...+X_n)=\sum_{z=1}^{n}Var(X_z)+2\sum_{1\le z<z'\le n}Cov(X_z,X_{z'}).$$

But, as these variables are independent, we would have that

$$Var(X_1+...+X_n)=\sum_{z=1}^{n}Var(X_z)=\sum_{z=1}^{n}\frac{1}{12}(x_t^*(j))^2=\frac{1}{12}\sum_{z=1}^{n}(x_t^*(j))^2<\frac{n}{12}.$$

Therefore,

$$\frac{1}{n^2}Var(X_1+...+X_n)<\frac{1}{n^2}\frac{n}{12}=\frac{1}{12n}\xrightarrow[n\to\infty]{}0.$$

Since convergence in L_2 implies convergence in probability, the same result would apply in this case, so that $\int_0^\theta x(j,i)dj=\theta\mathbb{E}[x(j,i)|x(j,i)\le x_t^*(j)]$, which is again a function of the set of $x_t^*(j)$ for all $j\in[0,1]$ and hence common across households.

The rest of the proof of this section follows from the continuous mapping theorem, which states that, given a sequence of random variables $\{X_n\}_{n=1}^{\infty}$ and X taking values in a metric space $(\mathcal{X}, d)$,

$$X_n \xrightarrow{p} X \implies g(X_n) \xrightarrow{p} g(X),$$

for all bounded continuous functions g.

By the continuous mapping theorem, $\int_0^\theta P_t tx(j,i)dj = \theta \mathbb{E}[P_t tx(j,i)|x(j,i) \leq x_t^*(j)]$ (where $g(x) = cx$ for a constant c). Applying again the continuous mapping theorem yields

$$\int_0^\theta P_t^A(j) + P_t tx(j,i)dj = \theta \mathbb{E}[P_t^A(j) + P_t tx(j,i)|x(j,i) \leq x_t^*(j)]$$

where $g(x,y) = x + y$ for two variables x and y. And finally,

$$\int_0^\theta [P_t^A(j) + P_t tx(j,i)]^{1-\epsilon} dj = \theta \mathbb{E}\left[[P_t^A(j) + P_t tx(j,i)]^{1-\epsilon} | x(j,i) \leq x_t^*(j) \right],$$

which is again common across households. An equivalent argument follows for the B part of the integral, $\int_\theta^1 [P_t^B(j) + P_t t(1 - x(j,i))]^{1-\epsilon} dj$.

A.3 Derivation of Equation (14)

Given prices, we can write the profits function of firm A from industry j as

$$
\begin{aligned}
\Pi_t^A(j) = \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] Y_t^A(j) &= \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] \int_0^{\frac{P_t^B(j)-P_t^A(j)}{P_t 2t}+\frac{1}{2}} C_t(j,x)dx \\
&= \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] \int_0^{\frac{P_t^B(j)-P_t^A(j)}{P_t 2t}+\frac{1}{2}} \left(\frac{\tilde{P}_t(j,x)}{\tilde{P}_t(x)}\right)^{-\epsilon} C_t(x)dx \\
&= \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] \int_0^{\frac{P_t^B(j)-P_t^A(j)}{P_t 2t}+\frac{1}{2}} \left(\frac{\tilde{P}_t(j,x)}{\tilde{P}_t}\right)^{-\epsilon} C_t dx \\
&= \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] \tilde{P}_t^\epsilon C_t \int_0^{\frac{P_t^B(j)-P_t^A(j)}{P_t 2t}+\frac{1}{2}} [\tilde{P}_t(j,x)]^{-\epsilon} dx \\
&= \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] \tilde{P}_t^\epsilon C_t \int_0^{\frac{P_t^B(j)-P_t^A(j)}{P_t 2t}+\frac{1}{2}} [P_t^A(j) + P_t t x]^{-\epsilon} dx \\
&= \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] \tilde{P}_t^\epsilon C_t \left[\frac{[P_t^A(j) + P_t t x]^{-\epsilon+1}}{(-\epsilon+1)P_t t}\Bigg|_0^{\frac{P_t^B(j)-P_t^A(j)}{P_t 2t}+\frac{1}{2}}\right] \\
&= \left[P_t^A(j) - \left(\frac{W_t}{A_t}\right)\right] \tilde{P}_t^\epsilon C_t \frac{1}{(-\epsilon+1)P_t t} \\
&\left\{\left[P_t^A(j) + P_t t\left(\frac{P_t^B(j)-P_t^A(j)}{P_t 2t} + \frac{1}{2}\right)\right]^{-\epsilon+1} - [P_t^A(j)]^{-\epsilon+1}\right\},
\end{aligned} \tag{34}
$$

where the third line follows from the fact that, as shown in Appendix A.2, the effective aggregate price index is homogeneous across households.

A.4 Equilibrium Results

$\tilde{\pi}_t$ equation

From the definition of the effective aggregate price index and Appendix A.2 we have that

$$
\tilde{P}_t^{1-\epsilon} = \tilde{P}_t(i)^{1-\epsilon} = \int_0^1 \tilde{P}_t(j,i)^{1-\epsilon} dj,
$$

for a given $i \in [0,1]$.

Given that we are restricting to equilibria where all firms apply the same level-k thinking to the formation of their beliefs and given that—for any given industry $j \in [0,1]$—both firms A and B face a symmetric problem, $P_t^A(j) = P_t^B(j)$ in equilibrium, so $x^*(j) = \frac{1}{2}$. From random assignment and the continuum of industries, this leads to the fact that household i will optimally choose to buy from firm A in half of the industries, and from B in the other half. Moreover, as locations are drawn from a continuous uniform $[0,1]$ distribution, the distribution of locations—for that particular household i—across industries $\Big[x(j,i)\Big]_{j\in[0,1]}$ is also continuously uniform along the interval $[0,1]$. Since the particular order of the $j \in [0,1]$ index does not matter, one can relabel the j's to be equal to the exact location of household i in such particular industry, and then we can break up the integral of the expression above in the following way

$$\tilde{P}_t^{1-\epsilon} = \tilde{P}_t(i)^{1-\epsilon} = \int_0^{\frac{1}{2}} \Big[P_t^A(j) + P_t tj\Big]^{1-\epsilon} dj + \int_{\frac{1}{2}}^1 \Big[P_t^B(j) + P_t t(1-j)\Big]^{1-\epsilon} dj.$$

Dividing by $P_{t-1}^{1-\epsilon}$ and noticing that—since all industries are identical and beliefs are homogeneous—the optimal prices for firms A and B will be the same across industries in equilibrium, we get that

$$(1+\tilde{\pi}_t)^{1-\epsilon} = \int_0^{\frac{1}{2}} \Big[(1+\pi_t^A) + (1+\pi_t)tj\Big]^{1-\epsilon} dj + \int_{\frac{1}{2}}^1 \Big[(1+\pi_t^B) + (1+\pi_t)t(1-j)\Big]^{1-\epsilon} dj,$$

where $\pi_t^A = \pi_t^A(j)$ and $\pi_t^B = \pi_t^B(j)$ for all $j \in [0,1]$. The two components of this equation can be written as

$$\begin{aligned}\int_0^{\frac{1}{2}} \Big[(1+\pi_t^A) + (1+\pi_t)tj\Big]^{1-\epsilon} dj &= \frac{\Big[(1+\pi_t^A) + (1+\pi_t)tj\Big]^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t}\Bigg|_0^{\frac{1}{2}} \\ &= \frac{\Big[(1+\pi_t^A) + (1+\pi_t)t\frac{1}{2}\Big]^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t} - \frac{(1+\pi_t^A)^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t},\end{aligned}$$

and

$$\int_{\frac{1}{2}}^{1} \Big[(1+\pi_t^B)+(1+\pi_t)t(1-j)\Big]^{1-\epsilon} dj = \frac{\Big[(1+\pi_t^B)+(1+\pi_t)t(1-j)\Big]^{2-\epsilon}}{(2-\epsilon)(-(1+\pi_t)t)}\Bigg|_{\frac{1}{2}}^{1}$$

$$= \frac{\Big[(1+\pi_t^B)+(1+\pi_t)t\frac{1}{2}\Big]^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t} - \frac{(1+\pi_t^B)^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t},$$

so the $\tilde{\pi}_t$ equation reduces to

$$(1+\tilde{\pi}_t)^{1-\epsilon} = \frac{\Big[(1+\pi_t^A)+(1+\pi_t)t\frac{1}{2}\Big]^{2-\epsilon} + \Big[(1+\pi_t^B)+(1+\pi_t)t\frac{1}{2}\Big]^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t} - \frac{(1+\pi_t^A)^{2-\epsilon}+(1+\pi_t^B)^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t}.$$

π_t equation

Recall the definition of the aggregate price index

$$P_tY_t \equiv \int_0^1 P_t^A(j)Y_t^A(j) + P_t^B(j)Y_t^B(j)dj,$$

and consider a particular industry $j \in [0,1]$. Since both firms choose the same price in equilibrium, the market will be equally split and $x_t^*(j) = \frac{1}{2}$. Since households are uniformly distributed across the product space of industry j, and since the particular order of the $i \in [0,1]$ index does not matter, one can relabel the i's to be equal to the exact location of each household in the $[0,1]$ interval of industry j. Then, we can compute $Y_t^A(j)$ and $Y_t^B(j)$ of industry j as

$$
\begin{aligned}
Y_t^A(j) &= \int_0^{\frac{1}{2}} \frac{1+\tilde{\pi}_t}{1+\pi_t} C_t(j,i)di = \int_0^{\frac{1}{2}} \frac{1+\tilde{\pi}_t}{1+\pi_t} \left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t(i)}\right)^{-\epsilon} C_t(i)di \\
&= \int_0^{\frac{1}{2}} \frac{1+\tilde{\pi}_t}{1+\pi_t} \left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t}\right)^{-\epsilon} C_t di = \int_0^{\frac{1}{2}} \left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t}\right)^{-\epsilon} Y_t di \\
= \int_0^{\frac{1}{2}} \tilde{P}_t(j,i)^{-\epsilon} \tilde{P}_t^{\epsilon} Y_t di &= \tilde{P}_t^{\epsilon} Y_t \int_0^{\frac{1}{2}} \tilde{P}_t(j,i)^{-\epsilon} di = \tilde{P}_t^{\epsilon} Y_t \int_0^{\frac{1}{2}} \left[P_t^A + P_t t i\right]^{-\epsilon} di \\
&= \tilde{P}_t^{\epsilon} Y_t \left[\left.\frac{[P_t^A + P_t t i]^{-\epsilon+1}}{(-\epsilon+1)P_t t}\right|_0^{\frac{1}{2}}\right] = \tilde{P}_t^{\epsilon} Y_t \frac{[P_t^A + P_t t \frac{1}{2}]^{-\epsilon+1} - [P_t^A]^{-\epsilon+1}}{(-\epsilon+1)P_t t},
\end{aligned}
$$

and

$$
\begin{aligned}
Y_t^B(j) &= \int_{\frac{1}{2}}^{1} \frac{1+\tilde{\pi}_t}{1+\pi_t} C_t(j,i)di = \int_{\frac{1}{2}}^{1} \frac{1+\tilde{\pi}_t}{1+\pi_t} \left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t(i)}\right)^{-\epsilon} C_t(i)di \\
&= \int_{\frac{1}{2}}^{1} \frac{1+\tilde{\pi}_t}{1+\pi_t} \left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t}\right)^{-\epsilon} C_t di = \int_{\frac{1}{2}}^{1} \left(\frac{\tilde{P}_t(j,i)}{\tilde{P}_t}\right)^{-\epsilon} Y_t di \\
= \int_{\frac{1}{2}}^{1} \tilde{P}_t(j,i)^{-\epsilon} \tilde{P}_t^{\epsilon} Y_t di &= \tilde{P}_t^{\epsilon} Y_t \int_{\frac{1}{2}}^{1} \tilde{P}_t(j,i)^{-\epsilon} di = \tilde{P}_t^{\epsilon} Y_t \int_{\frac{1}{2}}^{1} \left[P_t^B + P_t t(1-i)\right]^{-\epsilon} di \\
&= \tilde{P}_t^{\epsilon} Y_t \left[\left.\frac{[P_t^B + P_t t(1-i)]^{-\epsilon+1}}{(-\epsilon+1)(-P_t t)}\right|_{\frac{1}{2}}^{1}\right] = \tilde{P}_t^{\epsilon} Y_t \frac{[P_t^B + P_t t \frac{1}{2}]^{-\epsilon+1} - [P_t^B]^{-\epsilon+1}}{(-\epsilon+1)P_t t}.
\end{aligned}
$$

Plugging both expressions in the definition of the aggregate price index and noticing that, in equilibrium, both prices and quantities are homogeneous across industries, we get

$$
P_t Y_t = P_t^A \tilde{P}_t^{\epsilon} Y_t \frac{[P_t^A + P_t t \frac{1}{2}]^{-\epsilon+1} - [P_t^A]^{-\epsilon+1}}{(-\epsilon+1)P_t t} + P_t^B \tilde{P}_t^{\epsilon} Y_t \frac{[P_t^B + P_t t \frac{1}{2}]^{-\epsilon+1} - [P_t^B]^{-\epsilon+1}}{(-\epsilon+1)P_t t},
$$

where $P_t^A = P_t^A(j)$ and $P_t^B = P_t^B(j)$ for all $j \in [0,1]$. This equation can be written in terms of inflation rates as

$$(1+\pi_t) = (1+\pi_t^A)\left(\frac{(1+\tilde{\pi}_t)}{(1+\pi_t)}\right)^{\epsilon} \frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right)+\frac{t}{2}\right]^{-\epsilon+1} - \left(\frac{1+\pi_t^A}{1+\pi_t}\right)^{-\epsilon+1}}{(-\epsilon+1)t}$$
$$+(1+\pi_t^B)\left(\frac{(1+\tilde{\pi}_t)}{(1+\pi_t)}\right)^{\epsilon} \frac{\left[\left(\frac{1+\pi_t^B}{1+\pi_t}\right)+\frac{t}{2}\right]^{-\epsilon+1} - \left(\frac{1+\pi_t^B}{1+\pi_t}\right)^{-\epsilon+1}}{(-\epsilon+1)t}.$$

Aggregate Production Function

Recall that the production function of firm A from industry j is $Y_t^A(j) = A_t N_t^A(j)$. From labor market clearing we have that

$$A_t N_t = A_t \int_0^1 N_t(j)dj = \int_0^1 A_t N_t(j)dj = \int_0^1 A_t[N_t^A(j) + N_t^B(j)]dj$$
$$= \int_0^1 A_t N_t^A(j)dj + \int_0^1 A_t N_t^B(j)dj = \int_0^1 Y_t^A(j)dj + \int_0^1 Y_t^B(j)dj.$$

Using the results from the previous section we get that

$$A_t N_t = \tilde{P}_t^{\epsilon} Y_t \left\{ \frac{[P_t^A + P_t t\frac{1}{2}]^{-\epsilon+1} - [P_t^A]^{-\epsilon+1}}{(-\epsilon+1)P_t t} + \frac{[P_t^B + P_t t\frac{1}{2}]^{-\epsilon+1} - [P_t^B]^{-\epsilon+1}}{(-\epsilon+1)P_t t} \right\},$$

which can be written as

$$A_t N_t = Y_t \left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)^{\epsilon} \left\{ \frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right)+\frac{t}{2}\right]^{-\epsilon+1} + \left[\left(\frac{1+\pi_t^B}{1+\pi_t}\right)+\frac{t}{2}\right]^{-\epsilon+1}}{(-\epsilon+1)t} \right.$$
$$\left. - \frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right)\right]^{-\epsilon+1} + \left[\left(\frac{1+\pi_t^B}{1+\pi_t}\right)\right]^{-\epsilon+1}}{(-\epsilon+1)t} \right\}.$$

Claim: $\pi_t = \pi_t^A = \pi_t^B$

I have argued that, in equilibrium, $P_t^A(j) = P_t^B(j)$ for any industry $j \in [0,1]$. Moreover, the fact that all industries are alike leads to $\pi_t^A(j) = \pi_t^A$ and $\pi_t^B(j) = \pi_t^B$ for any $j \in [0,1]$. Therefore, $\pi_t^A = \pi_t^B$. This section shows that $\pi_t = \pi_t^A = \pi_t^B$.

To show this, we can depart from the $\tilde{\pi}_t$ equation, which reduces to

$$(1+\tilde{\pi}_t)^{1-\epsilon} = 2\frac{\left[(1+\pi_t^A) + (1+\pi_t)t\frac{1}{2}\right]^{2-\epsilon} - (1+\pi_t^A)^{2-\epsilon}}{(2-\epsilon)(1+\pi_t)t}$$

after applying the fact that $\pi_t^A = \pi_t^B$. This equation can be written as

$$\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)^{-\epsilon+1} = 2\frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right) + \frac{t}{2}\right]^{2-\epsilon} - \left(\frac{1+\pi_t^A}{1+\pi_t}\right)^{2-\epsilon}}{(2-\epsilon)t},$$

which is true for all $\epsilon > 2$. Let's make the change of variables $\epsilon = y+1$, so

$$\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)^{-y} = 2\frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right) + \frac{t}{2}\right]^{1-y} - \left(\frac{1+\pi_t^A}{1+\pi_t}\right)^{1-y}}{(1-y)t},$$

which is true for all $y > 1$. But this implies that

$$\frac{1}{2} = \left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)^{y}\frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right) + \frac{t}{2}\right]^{1-y} - \left(\frac{1+\pi_t^A}{1+\pi_t}\right)^{1-y}}{(1-y)t}$$

for all $y > 1$. Now, recall the π_t equation

$$\begin{aligned}(1+\pi_t) = (1+\pi_t^A)&\left(\frac{(1+\tilde{\pi}_t)}{(1+\pi_t)}\right)^{\epsilon}\frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right) + \frac{t}{2}\right]^{-\epsilon+1} - \left(\frac{1+\pi_t^A}{1+\pi_t}\right)^{-\epsilon+1}}{(-\epsilon+1)t}\\ +(1+\pi_t^B)&\left(\frac{(1+\tilde{\pi}_t)}{(1+\pi_t)}\right)^{\epsilon}\frac{\left[\left(\frac{1+\pi_t^B}{1+\pi_t}\right) + \frac{t}{2}\right]^{-\epsilon+1} - \left(\frac{1+\pi_t^B}{1+\pi_t}\right)^{-\epsilon+1}}{(-\epsilon+1)t}.\end{aligned}$$

Combining the two previous results we get that

$$(1+\pi_t) = (1+\pi_t^A)\frac{1}{2} + (1+\pi_t^B)\frac{1}{2},$$

so $\pi_t = \pi_t^A = \pi_t^B$. Notice that this is true regardless of the specific beliefs of the firms, as long as these beliefs are homogeneous across firms so that they end up choosing the same price in equilibrium.

Understanding the Resource Constraint

The previous section shows that

$$\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)^{\epsilon} \frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right)+\frac{t}{2}\right]^{1-\epsilon} - \left(\frac{1+\pi_t^A}{1+\pi_t}\right)^{1-\epsilon}}{(1-\epsilon)t} = \frac{1}{2}.$$

Applying the fact that $\pi_t = \pi_t^A = \pi_t^B$, the equation reduces to

$$\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right) = \left[2\frac{\left[1+\frac{t}{2}\right]^{-\epsilon+1} - 1}{(-\epsilon+1)t}\right]^{-\frac{1}{\epsilon}},$$

which is the term that multiplies consumption in the resource constraint and is a function of t. This section explores the properties of this term. In particular, and given that $t > 0$, I analyze what happens as t goes to zero, and I compute the partial derivative with respect to t.

First, let me define $g(t) \equiv 2\frac{[1+\frac{t}{2}]^{-\epsilon+1}-1}{(-\epsilon+1)t}$, and let us calculate the limit of $g(t)$ as t goes to zero with the application of L'Hopital's rule as

$$\lim_{t\to 0} g(t) = \lim_{t\to 0} 2\frac{[1+\frac{t}{2}]^{-\epsilon+1} - 1}{(-\epsilon+1)t} = \lim_{t\to 0} 2\frac{(-\epsilon+1)[1+\frac{t}{2}]^{-\epsilon}\frac{1}{2}}{(-\epsilon+1)} = 1.$$

Therefore,

$$\lim_{t\to 0}\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right) = \lim_{t\to 0}[g(t)]^{-\frac{1}{\epsilon}} = 1.$$

Second, let us analyze the partial derivative of $\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)$ with respect to t

$$\frac{\partial\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)}{\partial t} = \left(-\frac{1}{\epsilon}\right)\left[2\frac{\left[1+\frac{t}{2}\right]^{-\epsilon+1}-1}{(-\epsilon+1)t}\right]^{-\frac{1}{\epsilon}-1} 2$$
$$*\frac{(-\epsilon+1)^2[1+\frac{t}{2}]^{-\epsilon}\frac{t}{2}-(-\epsilon+1)[(1+\frac{t}{2})^{-\epsilon+1}-1]}{(-\epsilon+1)^2t^2}.$$

It is easy to see that this term will be positive if, and only if,

$$(-\epsilon+1)\left[\left(1+\frac{t}{2}\right)^{-\epsilon+1}-1\right] > (-\epsilon+1)^2\left[1+\frac{t}{2}\right]^{-\epsilon}\frac{t}{2}$$
$$\iff \left(1+\frac{t}{2}\right)^{-\epsilon}\left[1+\frac{t}{2}-(-\epsilon+1)\frac{t}{2}\right] < 1.$$

Let me define $f(t) \equiv \left(1+\frac{t}{2}\right)^{-\epsilon}\left[1+\frac{t}{2}-(-\epsilon+1)\frac{t}{2}\right]$. Notice that $f(0)=1$, and

$$\frac{\partial f(t)}{\partial t} = -\epsilon\left(1+\frac{t}{2}\right)^{-\epsilon-1}\left(\frac{1}{2}\right)\left[1+\frac{t}{2}-(-\epsilon+1)\frac{t}{2}\right] + \left(1+\frac{t}{2}\right)^{-\epsilon}\left[\frac{1}{2}-\frac{1}{2}(-\epsilon+1)\right]$$
$$= \left(1+\frac{t}{2}\right)^{-\epsilon}\left(\frac{\epsilon}{2}\right)\frac{(-\epsilon+1)t}{2+t} < 0,$$

as $\epsilon > 2$. Therefore, $f(t) < 1$ for all $t > 0$, so $\frac{\partial\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)}{\partial t} > 0$.

One can also show that the second derivative is negative, although that involves a significant amount of algebra. Given that the expression has only one parameter, ϵ, figure A.1 shows a graphical representation for the particular case where this parameter is, say, $\epsilon = 6$.

Understanding the Aggregate Production Function

The aggregate production function, as calculated above, is

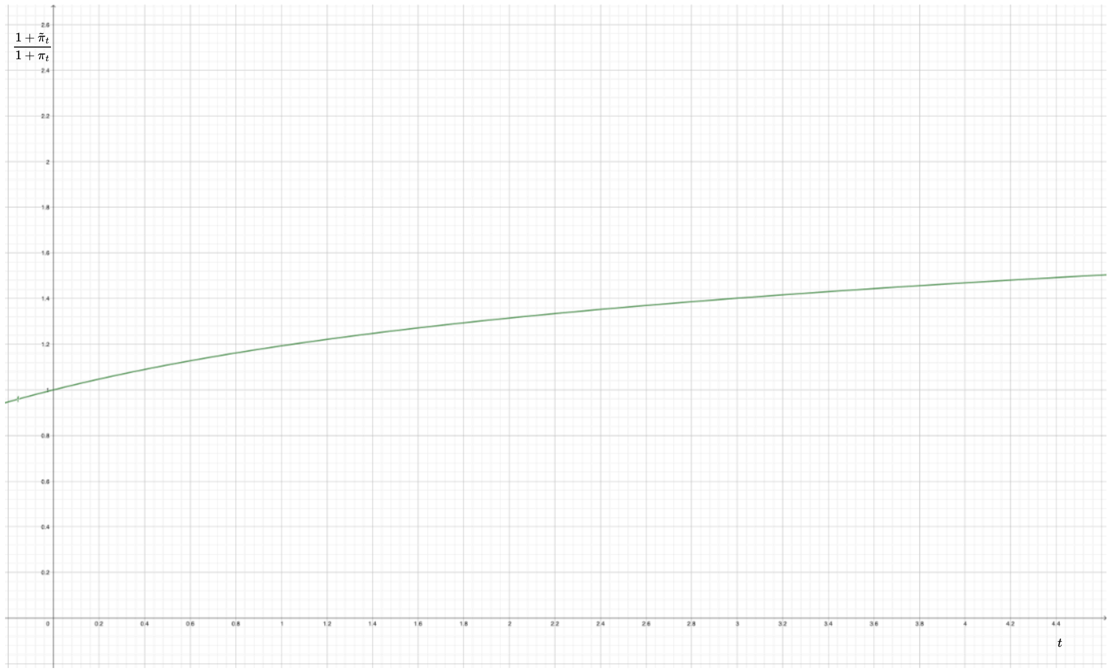

Figure A.1: $\frac{1+\tilde{\pi}_t}{1+\pi_t}$ as a function of t.

$$A_tN_t = Y_t\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)^{\epsilon}\left\{\frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right)+\frac{t}{2}\right]^{-\epsilon+1}+\left[\left(\frac{1+\pi_t^B}{1+\pi_t}\right)+\frac{t}{2}\right]^{-\epsilon+1}}{(-\epsilon+1)t}\right.$$
$$\left.-\frac{\left[\left(\frac{1+\pi_t^A}{1+\pi_t}\right)\right]^{-\epsilon+1}+\left[\left(\frac{1+\pi_t^B}{1+\pi_t}\right)\right]^{-\epsilon+1}}{(-\epsilon+1)t}\right\}.$$

Applying the fact that, in equilibrium, $\pi_t = \pi_t^A = \pi_t^B$ we get

$$A_tN_t = Y_t\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right)^{\epsilon} 2\frac{\left(1+\frac{t}{2}\right)^{-\epsilon+1}-1}{(-\epsilon+1)t},$$

but in the previous section we found that

$$\left(\frac{1+\tilde{\pi}_t}{1+\pi_t}\right) = \left[2\frac{\left[1+\frac{t}{2}\right]^{-\epsilon+1}-1}{(-\epsilon+1)t}\right]^{-\frac{1}{\epsilon}},$$

so the aggregate production function reduces to

$$A_t N_t = Y_t.$$

B Quantitative Appendix

B.1 Alternative Calibrations

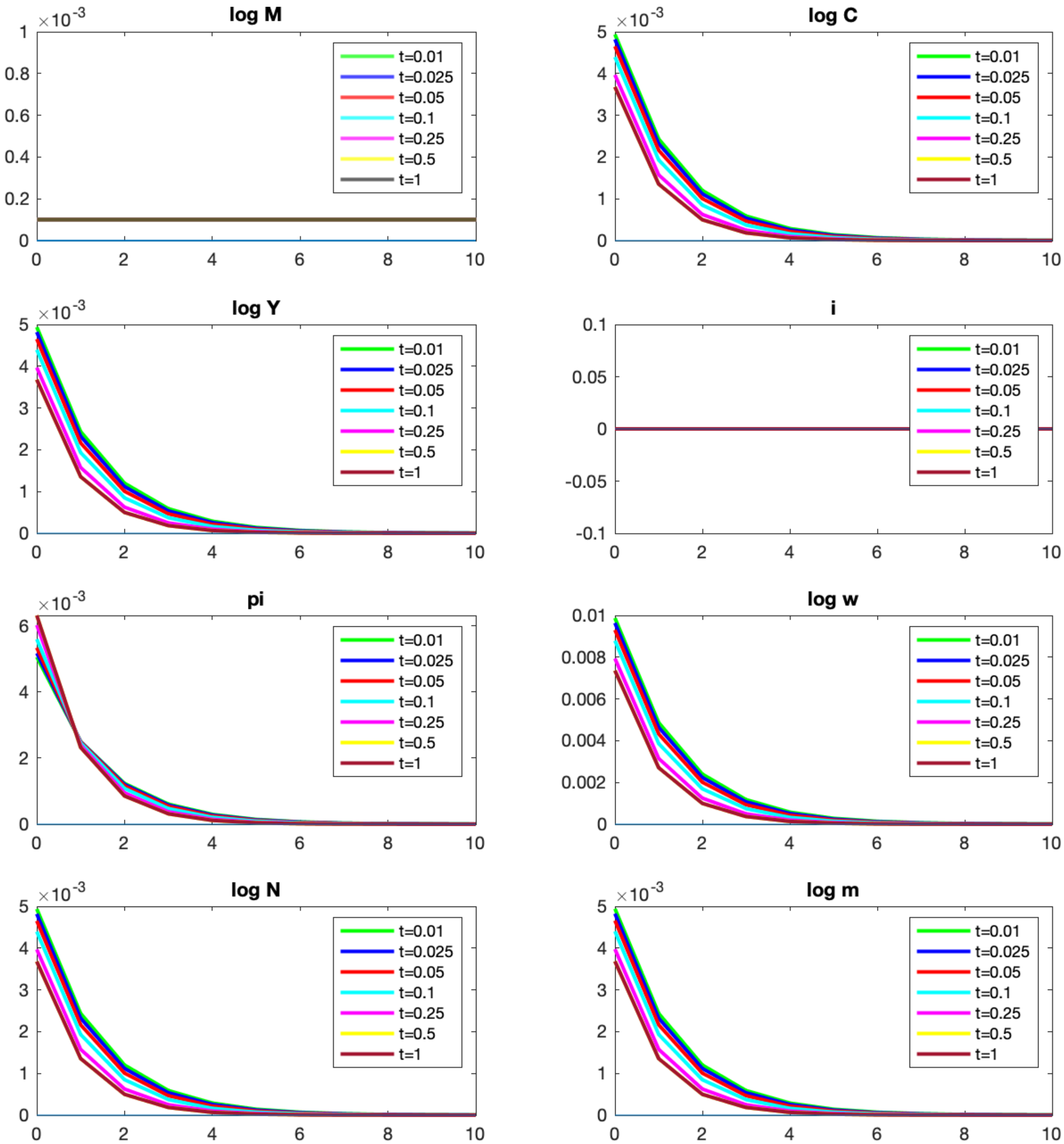

Figure B.1: Impulse Responses for different t calibrations and level-0 thinking.

B.2 Sticky Wages

This section introduces the details about expanding the model with sticky wages as in Erceg et al. (2000). Households supply differenciated labor input that is then "packed" into a bundled labor input that is sold to firms. The profit maximization problem of a *labor packer* or union that aggregates labor with a CES function—with elasticity of substitution $\epsilon_w > 1$—produces the following downward slopping demand for labor

$$N_t(i) = \left(\frac{W_t(i)}{W_t}\right)^{-\epsilon_w} N_t,$$

where

$$N_t = \left(\int_0^1 N_t(i)^{\frac{\epsilon_w - 1}{\epsilon_w}} di\right)^{\frac{\epsilon_w}{\epsilon_w - 1}},$$

and

$$W_t = \left(\int_0^1 W_t(i)^{1-\epsilon_w} di\right)^{\frac{1}{1-\epsilon_w}}.$$

The problem of the households is slightly different because they are now supplying differentiated labor and setting their own wage. The optimal allocation of consumption expenditures result derived in Appendix A.1, and the homogeneity of the effective aggregate price index result derived in Appendix A.2 still hold without changes. Moreover, the fact that preferences are separable in consumption and leisure and markets are complete, allows us to apply the Erceg et al. (2000) result so that households will only differ in the wage they charge and their labor supply.

The rest of derivations are standard and I will omit most details not to make this appendix overly long. The problem of a household that chooses its own wage to maximize utility subject to labor demand, and subject to the Calvo assumption, results in the wage setting equation

$$w_t^{\#, \epsilon_w \eta + 1} = \frac{\epsilon_w}{\epsilon_w - 1} \frac{H_{1,t}}{H_{2,t}},$$

where $w_t^{\#}$ is the reset wage, and where

$$H_{1,t} = \Psi w_t^{\epsilon_w(1+\eta)} N_t^{1+\eta} + (\beta\phi_w)\mathbb{E}_t(1+\pi_{t+1})^{\epsilon_w(1+\eta)} H_{1,t+1},$$

and

$$H_{2,t} = C_t^{-\sigma} \frac{1+\pi_t}{1+\tilde{\pi}_t} w_t^{\epsilon_w} N_t + (\beta\phi_w)\mathbb{E}_t(1+\pi_{t+1})^{\epsilon_w - 1} H_{2,t+1}.$$

Lastly, in equilibrium,

$$w_t^{1-\epsilon_w} = (1-\phi_w) w_t^{\#,1-\epsilon_w} + \phi_w w_{t-1}^{1-\epsilon_w}(1+\pi_t)^{\epsilon_w - 1}.$$

The resource constraint does not change and the rest of equilibrium results also hold.