

Aggregate Hazard Function in Price-Setting: A Bayesian Analysis Using Macro Data

Fang Yao*

Humboldt University of Berlin

First draft: Oct. 20, 2009

This version: March. 26, 2010

Abstract

This paper presents an approach to identify aggregate price reset hazards from the joint dynamic behavior of inflation and macroeconomic aggregates. The identification is possible due to the fact that inflation is composed of current and past reset prices and that the composition depends on the price reset hazard function. The derivation of the generalized NKPC links those composition effects to the hazard function, so that only aggregate data is needed to extract information about the price reset hazard function. The empirical hazard function is generally increasing with the age of prices, but with spikes at the 1st and 4th quarters. The implication of this finding for sticky price modeling is that the pricing decision has both time- and state-dependent aspects.

JEL classification: E12; E31

Key words: Sticky prices, Aggregate hazard function, Bayesian estimation

* I am grateful to Michael Burda, Carlos Carvalho, Harald Uhlig, Lutz Weinke, Alexander Wolman and other seminar participants in Berlin for helpful comments. I acknowledge the support of the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk". All errors are my sole responsibility. Address: Institute for Economic Theory, Humboldt University of Berlin, Spandauer Str. 1, Berlin, Germany +493020935667, email: yaofang@rz.hu-berlin.de.

1 Introduction

In the current generation of monetary models, effects of monetary policy are closely related to the speed of the aggregate price level reacting to a nominal disturbance. The adjustment of aggregate price in turn depends on two factors. One is the optimal reset price an adjusting firm chooses, and the other is the fraction of firms changing their prices. With the exception of a few micro-founded state-dependent models¹, the majority of research on sticky prices is limited to addressing the optimal reset price decision, but leaving the adjustment timing to be exogenously given by some simplified assumptions, e.g. models incorporating the Calvo (1983) or Taylor (1980) approaches. Put in more technical terms, it amounts to restricting the price reset hazard function to a specific shape and studying other issues on the basis of this assumption.

Until recently, the aggregate price reset hazard function remains a largely ignored topic in the macro literature. It begins, however, to draw more attention, because the competing theoretical models of sticky prices deliver clear mappings between specific aggregate hazard functions and implications for macro dynamics and monetary policy. Pioneer work by Wolman (1999) and Kiley (2002) demonstrated that aggregate dynamics should be sensitive to the hazard function underlying different pricing rules. For this reason, the aggregate hazard function provides a new metric used to select theoretical models and identify most relevant propagation mechanisms for monetary policy shocks.

Despite its uses, empirical studies of the aggregate hazard function are rare in the macro literature. By contrast, fast growing evidence from micro data sets becomes available in the recent years². However, I want to argue that it is the aggregate hazard function that of great interest to macroeconomists, and it is important to distinguish between the macro and micro hazard function. The aggregate hazard is defined as the probability of the price adjustment reacting to aggregate shocks. In the theoretical macro models, those hazard rates can be clearly mapped into impulse responses of aggregate variables. By contrast, mapping between micro hazard functions and macro dynamics is much trickier³. For example, Caplin and Spulber (1987) demonstrated that, when the selection effect is present, the aggregate economy is completely immune to price stickiness at the micro level, and thereby has no real effect of monetary policy. Hazard functions estimated from the micro data are therefore not a perfect substitute for the aggregate hazard function defined in the theoretical models. Besides this theoretical consideration, there are also empirical pitfalls that cause for attention in interpreting micro hazard rates. First, micro hazard rates are typically higher than aggregate hazard rates, because individual prices react to both idiosyncratic and aggregate shocks. It is very difficult to disentangle them with a micro data set. Second, evidence of the shape of the hazard function from microeconomic studies is not conclusive⁴. Micro data sets differ substantially in the range of goods included,

¹See: e.g. Caplin and Spulber (1987), Dotsey et al. (1999), Caballero and Engel (2007) and Golosov and Lucas (2007). The strength of those models is to endogenize both the optimal reset price decision and the adjustment timing decision in the same framework. However, due to the complexity of this approach, few of them are actually applied in the policy analysis.

²See: e.g. Bils and Klenow (2004), Alvarez et al. (2006) Midrigan (2007) and Nakamura and Steinsson (2008) among others.

³See: Mackowiak and Smets (2008) for elaboration on this point.

⁴Some find strong support for increasing hazard functions (e.g.: Cecchetti, 1986 and Fougere et al., 2005), while others find evidence in favour of decreasing hazards (e.g.: Campbell and Eden, 2005, Alvarez, 2007 and

the countries and time periods covered, and thereby make it difficult to compare their results; and, even though comprehensive micro data sets have now become available, they are usually short compared to aggregate time series data. Most of the CPI or PPI data sets for the U.S. or Europe are only available from the late 80's⁵. It is reasonable to think that the shape of hazard functions could depend on the underlying economic conditions, and would therefore change over the time periods of the collected data.

The objective of this paper is to estimate the aggregate price reset hazard function directly from the time series data. To do that, I first construct a fully-specified DSGE model featuring nominal rigidity that allows for a flexible hazard function of price setting. I derive the generalized New Keynesian Phillips curve (NKPC hereafter) and then estimate this model with the Bayesian approach. The identification of the aggregate hazard function is possible due to the fact that inflation rate can be decomposed into current and past reset prices and its composition is determined by the aggregate hazard function. The derivation of the generalized NKPC links those composition effects to the hazard function, so that only aggregate data is needed to extract information about the price reset hazard function. The advantages of this identification method is that, first, it is based on a generic assumption of the firm's level pricing behavior, making the mapping between the hazard function and aggregate dynamics robust to the modeling of sticky prices. In addition, this method identifies aggregate hazard function from the fluctuations of aggregate price level, so that effects of idiosyncratic shocks are removed. However, this method is not free from other typical identification problems which prevail in the estimated New Keynesian models⁶, e.g. observational equivalence of the labor supply elasticity. For those poorly identified structural parameters, I conduct various sensitivity tests to check the robustness of hazard function estimates.

I estimate the hazard function using the U.S. quarterly data of inflation, the growth rate of real output and effective federal funds rate from 1955 to 2008. The empirical aggregate hazard function has a U-shape with a spike at the fourth quarter. The interpretation of this finding is that price setting is characterized by both time- and state-dependent aspects. For the time-dependent pricing aspect, one quarter and 4-quarter seem to be the most important frequencies of the aggregate price adjustment. About 34.2% of prices hold for less than one quarter, while, 12.4% of prices have the mean duration of four quarters. Besides the time-dependent pricing pattern, the upward-sloping hazard function indicates that the state-dependent pricing also plays an important role in price decisions, especially when a price becomes more outdated. In fact, this generalized time-dependent model can be viewed as a tractable approximation for the more microfounded state-dependent model, when we consider a relative stable economy. The hazard function of the deviation from the optimum largely coincides with the hazard function of time-since-last-adjustment. The longer a price is fixed, the more likely it deviates significantly from the optimum, and hence its probability of being adjusted rises. Since the annual inflation rates in my data set stay under 2% for most of the sample periods except for 1970's, it is arguable that the time elapsed since last adjustment is a good proxy for the deviation from the optimum, therefore the increasing part of the hazard function gives the pricing decision a state-dependent aspect.

Nakamura and Steinsson, 2008).

⁵For more details see Table 2 in Alvarez (2007)

⁶For the recent discussion on this topic, see Canova and Sala (2009) and Rios-Rull et al. (2009).

This paper is broadly related to progress in developing empirical models of sticky prices based on the New Keynesian framework. The early empirical model of sticky prices was solely based on the NKPC under the Calvo pricing assumption (See, e.g. Gali and Gertler, 1999, Gali et al., 2001 and Sbordone, 2002). These authors estimated the NKPC with GMM, and found a considerable degree of price rigidity in the aggregate data. The empirical price reset hazard rate is around 20% per quarter for the U.S. and 10% for Europe. These results, however, are at odds with micro evidence in two ways. First, recent micro studies generally conclude that the average frequency of price adjustments at the firm's level is not only higher, but also differs substantially across sectors in the economy⁷. Second, the Calvo assumption implies a constant hazard function, meaning that the probability of adjusting prices is independent of the length of the time since the last price revision, and the flat hazard function has been largely rejected by empirical evidence from micro level data (See, e.g.: Cecchetti, 1986, Campbell and Eden, 2005 and Nakamura and Steinsson, 2008). Given these discrepancies between the macro and micro evidence, empirical models allowing for more flexible price durations or hazard functions have become popular in the recent literature. Jadresic (1999) presented a staggered pricing model featuring a flexible distribution over price durations and used VAR approach to demonstrate that the dynamic behavior of aggregate data on inflation and other macroeconomic variables provide information about the disaggregated price dynamics underlying the data. More recently, Sheedy (2007) constructed a generalized Calvo model and parameterized the hazard function in such a way that the resulting NKPC implied intrinsic inflation persistence when the hazard function was upward sloping. Based on this hazard function specification, he estimated the NKPC using GMM and found evidence of an upward-sloping hazard function. Coenen et al. (2007) developed a staggered nominal contracts model with both random and fixed durations, and estimated the generalized NKPC with an indirect inference method. Their results showed that price rigidity is characterized by a very high degree of real rigidity, as opposed to modest nominal rigidity with an average duration of about 2-3 quarters. Carvalho and Dam (2008) estimated a semi-structural multi-price-duration model with the Bayesian approach, and found that allowing for prices to last longer than 4 quarters is crucial to avoid underestimating the relative importance of nominal rigidity.

The remainder of the paper is organized as follows: in section 2, I present the model with generalized time-dependent pricing and derive the New Keynesian Phillips curve; section 3 introduces the empirical method and the data I use to estimate the model. At the end, results regarding the hazard function are presented and discussed; section 4 contains some concluding remarks.

⁷See: e.g. Bils and Klenow (2004), Alvarez et al. (2006) Midrigan (2007) and Nakamura and Steinsson (2008) among others.

2 The Model

In this section, I present a DSGE model of sticky prices due to nominal rigidity. I introduce nominal rigidity by means of a general form of hazard functions⁸. A hazard function of price setting is defined as the probabilities of price adjustment conditional on the spell of time elapsed since the last price change. In this model, the hazard function is a discrete function taking values between zero and one on its time domain. Many well known models of price setting in the literature can be shown to imply hazard functions of one form or another. For example, the most prominent pricing assumption of Calvo (1983) implies a constant hazard function over the infinite horizon.

2.1 Representative Household

A representative, infinitely-lived household derives utility from the composite consumption good C_t , and its labor supply L_t , and it maximizes a discounted sum of utility of the form:

$$\max_{\{C_t, L_t, B_t\}} E_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\delta}}{1-\delta} - \chi_H \frac{L_t^{1+\phi}}{1+\phi} \right) \right].$$

Here C_t denotes an index of the household's consumption of each of the individual goods, $C_t(i)$, following a constant-elasticity-of-substitution aggregator (Dixit and Stiglitz, 1977).

$$C_t \equiv \left[\int_0^1 C_t(i)^{\frac{\eta-1}{\eta}} di \right]^{\frac{\eta}{\eta-1}}, \quad (1)$$

where $\eta > 1$, and it follows that the corresponding cost-minimizing demand for $C_t(i)$ and the welfare-based price index, P_t , are given by

$$C_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\eta} C_t \quad (2)$$

$$P_t = \left[\int_0^1 P_t(i)^{1-\eta} di \right]^{\frac{1}{1-\eta}}. \quad (3)$$

For simplicity, I assume that households supply homogeneous labor units (L_t) in an economy-wide competitive labor market.

The flow budget constraint of the household at the beginning of period t is

$$P_t C_t + \frac{B_t}{R_t} \leq W_t L_t + B_{t-1} + \int_0^1 \pi_t(i) di. \quad (4)$$

Where B_t is a one-period nominal bond and R_t denotes the gross nominal return on the bond. $\pi_t(i)$ represents the nominal profits of a firm that sells good i . I assume that each household owns an equal share of all firms. Finally this sequence of period budget constraints is supplemented with a transversality condition of the form $\lim_{T \rightarrow \infty} E_t \left[\frac{B_T}{\prod_{s=1}^T R_s} \right] \geq 0$.

⁸In the theoretical literature, the general time-dependent pricing model has been first outlined in Wolman (1999), who studied some simple examples and found that inflation dynamics are sensitive to different pricing rules. Similar models have also been studied by Mash (2004) and Yao (2009).

The solution to the household's optimization problem can be expressed in two first order necessary conditions. First, optimal labor supply is related to the real wage:

$$\chi_H L_t^\phi C_t^\delta = \frac{W_t}{P_t}. \quad (5)$$

Second, the Euler equation gives the relationship between the optimal consumption path and asset prices:

$$1 = \beta E_t \left[\left(\frac{C_t}{C_{t+1}} \right)^\delta \frac{R_t P_t}{P_{t+1}} \right]. \quad (6)$$

2.2 Firms in the Economy

2.2.1 Real Marginal Cost

The production side of the economy is composed of a continuum of monopolistic competitive firms, each producing one variety of product i by using labor. Each firm maximizes real profits, subject to the production function

$$Y_t(i) = Z_t L_t(i) \quad (7)$$

where Z_t denotes an aggregate productivity shock. Log deviations of the shock, \hat{z}_t , follow an exogenous AR(1) process $\hat{z}_t = \rho_z \hat{z}_{t-1} + \varepsilon_{z,t}$, and $\varepsilon_{z,t}$ is white noise with $\rho_z \in [0, 1)$. $L_t(i)$ is the demand of labor by firm i .

Following equation (2), demand for intermediate goods is given by

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\eta} Y_t. \quad (8)$$

In each period, firms choose optimal demands for labor inputs to maximize their real profits given nominal wage, market demand (8) and the production technology (7):

$$\max_{L_t(i)} \Pi_t(i) = \frac{P_t(i)}{P_t} Y_t(i) - \frac{W_t}{P_t} L_t(i) \quad (9)$$

And real marginal cost can be derived from this maximization problem in the following form:

$$mc_t = \frac{W_t/P_t}{(1-a)Z_t}.$$

Furthermore, using the production function (7), output demand equation (8), the labor supply condition (5) and the fact that at the equilibrium $C_t = Y_t$, I can express real marginal cost only in terms of aggregate output and technology shock.

$$mc_t = Y_t^{\phi+\delta} Z_t^{-(1+\phi)}. \quad (10)$$

2.2.2 Pricing Decisions under Nominal Rigidity

In this section, I introduce a general form of nominal rigidity, which is characterized by a set of hazard rates depending on the spell of the time since last price adjustment. I assume that monopolistic competitive firms cannot adjust their price whenever they want. Instead, opportunities for re-optimizing prices are dictated by the hazard rates, h_j , where j denotes the time-since-last-adjustment and $j \in \{0, J\}$. J is the maximum number of periods in which a firm's price can be fixed.

Dynamics of the Price-duration Distribution In the economy, firms' prices are heterogeneous with respect to the time since their last price adjustment. Table 1 summarizes key notations concerning the dynamics of the price-duration distribution.

Duration	Hazard Rate	Non-adj. Rate	Survival Rate	Distribution
j	h_j	α_j	S_j	$\theta(j)$
0	0	1	1	$\theta(0)$
1	h_1	$\alpha_1 = 1 - h_1$	$S_1 = \alpha_1$	$\theta(1)$
\vdots	\vdots	\vdots	\vdots	\vdots
j	h_j	$\alpha_j = 1 - h_j$	$S_j = \prod_{i=0}^j \alpha_i$	$\theta(j)$
\vdots	\vdots	\vdots	\vdots	\vdots
J	$h_J = 1$	$\alpha_J = 0$	$S_J = 0$	$\theta(J)$

Table 1: Notations of the Dynamics of Price-vintage-distribution.

Using the notation defined in Table 1, and also denoting the distribution of price durations at the beginning of each period by $\Theta_t = \{\theta_t(0), \theta_t(1) \dots \theta_t(J)\}$, we can derive the ex-post distribution of firms after price adjustments ($\tilde{\Theta}_t$) as

$$\tilde{\theta}_t(j) = \begin{cases} \sum_{i=1}^J h_i \theta_t(i), & \text{when } j = 0 \\ \alpha_j \theta_t(j), & \text{when } j = 1 \dots J. \end{cases} \quad (11)$$

Firms reoptimizing their prices in period t are labeled with 'Duration 0', and the proportion of those firms is given by hazard rates of all duration groups multiplied by their corresponding densities. The firms left in each duration group are the firms that do not adjust their prices. When the period t is over, this ex-post distribution, $\tilde{\Theta}_t$, becomes the ex-ante distribution for the new period, Θ_{t+1} . All price duration groups move to the next one, because all prices age by one period.

As long as the hazard rates lie between zero and one, dynamics of the price-duration distribution can be viewed as a Markov process with an invariant distribution, Θ , and is obtained by solving $\theta_t(j) = \theta_{t+1}(j)$. It yields the stationary price-duration distribution $\theta(j)$:

$$\theta(j) = \frac{S_j}{\sum_{j=0}^{J-1} S_j}, \text{ for } j = 0, 1 \dots J - 1. \quad (12)$$

In a simple example, when $J = 3$, the stationary price-duration distribution

$$\Theta = \left\{ \frac{1}{1 + \alpha_1 + \alpha_1\alpha_2}, \frac{\alpha_1}{1 + \alpha_1 + \alpha_1\alpha_2}, \frac{\alpha_1\alpha_2}{1 + \alpha_1 + \alpha_1\alpha_2} \right\}.$$

I assume the economy converges to this invariant distribution fairly quickly, so that regardless of the initial price-duration distribution, I only consider the economy with the invariant distribution of price durations. This assumption makes aggregation problem of the economy tractable.

The Optimal Pricing under the General Form of Nominal Rigidity Given the general form of nominal rigidity introduced above, the only heterogeneity among firms is the time when they last reset their prices, j . Firms in price duration group j share the same probability of adjusting their prices, h_j , and the distribution of firms across durations is given by $\theta(j)$.

In a given period when a firm is allowed to reoptimize its price, the optimal price chosen should reflect the possibility that it will not be re-adjusted in the near future. Consequently, adjusting firms choose optimal prices that maximize the discounted sum of real profits over the time horizon in which the new price is expected to be fixed. The probability that a new price will be fixed at least for j periods is given by the survival function, S_j , defined in Table 1.

I setup the maximization problem of an adjuster as follows:

$$\max_{P_t^*} E_t \sum_{j=0}^{J-1} S_j Q_{t,t+j} \left[Y_{t+j|t}^d \frac{P_t^*}{P_{t+j}} - \frac{TC_{t+j}}{P_{t+j}} \right].$$

Where E_t denotes the conditional expectation based on the information set in period t , and $Q_{t,t+j}$ is the stochastic discount factor appropriate for discounting real profits from t to $t+j$. An adjusting firm maximizes the profits subject to the demand for its intermediate good in period $t+j$ given that the firm resets the price in period t and can be expressed as.

$$Y_{t+j|t}^d = \left(\frac{P_t^*}{P_{t+j}} \right)^{-\eta} Y_{t+j}.$$

This yields the following first order necessary condition for the optimal price:

$$P_t^* = \frac{\eta}{\eta - 1} \frac{\sum_{j=0}^{J-1} S_j E_t [Q_{t,t+j} Y_{t+j} P_{t+j}^{\eta-1} MC_{t+j}]}{\sum_{j=0}^{J-1} S_j E_t [Q_{t,t+j} Y_{t+j} P_{t+j}^{\eta-1}]}, \quad (13)$$

where MC_t denotes the nominal marginal cost. The optimal price is equal to the markup multiplied by a weighted sum of future marginal costs, whose weights depend on the survival rates. In the Calvo case, where $S_j = \alpha^j$, this equation reduces to the Calvo optimal pricing condition.

Finally, given the stationary distribution, $\theta(j)$, aggregate price can be written as a distributed sum of all optimal prices. I define the optimal price which was set j periods ago as P_{t-j}^* . Following the aggregate price index from equation (3), the aggregate price is then obtained by:

$$P_t = \left(\sum_{j=0}^{J-1} \theta(j) P_{t-j}^{*1-\eta} \right)^{\frac{1}{1-\eta}}. \quad (14)$$

2.3 New Keynesian Phillips Curve

In this section, I derive the New Keynesian Phillips curve for this generalized model. To do that, I first log-linearize equation (13) around the flexible price steady state. The log-linearized optimal price equations are obtained by

$$\hat{p}_t^* = E_t \left[\sum_{j=0}^{J-1} \frac{\beta^j S(j)}{\Omega} (\widehat{mc}_{t+j} + \hat{p}_{t+j}) \right], \quad (15)$$

where :

$$\Omega = \sum_{j=0}^{J-1} \beta^j S(j) \quad \text{and} \quad \widehat{mc}_t = (\delta + \phi)\hat{y}_t - (1 + \phi)\hat{z}_t.$$

In a similar fashion, I derive the log deviation of the aggregate price by log linearizing equation (14).

$$\hat{p}_t = \sum_{k=0}^{J-1} \theta(k) \hat{p}_{t-k}^*. \quad (16)$$

After some algebraic manipulations on equations (15) and (16), I obtain the New Keynesian Phillips curve as follows⁹

$$\begin{aligned} \hat{\pi}_t = & \sum_{k=0}^{J-1} \frac{\theta(k)}{1 - \theta(0)} E_{t-k} \left(\sum_{j=0}^{J-1} \frac{\beta^j S(j)}{\Psi} \widehat{mc}_{t+j-k} + \sum_{i=1}^{J-1} \sum_{j=i}^{J-1} \frac{\beta^j S(j)}{\Psi} \hat{\pi}_{t+i-k} \right) \\ & - \sum_{k=2}^{J-1} \Phi(k) \hat{\pi}_{t-k+1}, \quad \text{where } \Phi(k) = \frac{\sum_{j=k}^{J-1} S(j)}{\sum_{j=1}^{J-1} S(j)}, \quad \Psi = \sum_{k=0}^{J-1} \beta^k S(k). \end{aligned} \quad (17)$$

At the first glance, this Phillips curve is quite different from the one in the Calvo model. It involves not only lagged inflation but also lagged expectations that were built into pricing decisions in the past. All coefficients in the NKPC are derived from structural parameters which are either the hazard function parameters or the preference parameters. When $J = 3$, for example, then the NKPC is of the following form

⁹The detailed derivation of the NKPC can be found in the technical Appendix (A).

$$\begin{aligned}
\hat{\pi}_t = & \frac{1}{(\alpha_1 + \alpha_1\alpha_2)\Psi} \widehat{mc}_t + \frac{\alpha_1}{(\alpha_1 + \alpha_1\alpha_2)\Psi} \widehat{mc}_{t-1} + \frac{\alpha_1\alpha_2}{(\alpha_1 + \alpha_1\alpha_2)\Psi} \widehat{mc}_{t-2} \\
& + \frac{1}{\alpha_1 + \alpha_1\alpha_2} E_t \left(\frac{\beta\alpha_1}{\Psi} \widehat{mc}_{t+1} + \frac{\beta^2\alpha_1\alpha_2}{\Psi} \widehat{mc}_{t+2} + \frac{\beta\alpha_1 + \beta^2\alpha_1\alpha_2}{\Psi} \hat{\pi}_{t+1} + \frac{\beta^2\alpha_1\alpha_2}{\Psi} \hat{\pi}_{t+2} \right) \\
& + \frac{\alpha_1}{\alpha_1 + \alpha_1\alpha_2} E_{t-1} \left(\frac{\beta\alpha_1}{\Psi} \widehat{mc}_t + \frac{\beta^2\alpha_1\alpha_2}{\Psi} \widehat{mc}_{t+1} + \frac{\beta\alpha_1 + \beta^2\alpha_1\alpha_2}{\Psi} \hat{\pi}_t + \frac{\beta^2\alpha_1\alpha_2}{\Psi} \hat{\pi}_{t+1} \right) \\
& + \frac{\alpha_1\alpha_2}{\alpha_1 + \alpha_1\alpha_2} E_{t-2} \left(\frac{\beta\alpha_1}{\Psi} \widehat{mc}_{t-1} + \frac{\beta^2\alpha_1\alpha_2}{\Psi} \widehat{mc}_t + \frac{\beta\alpha_1 + \beta^2\alpha_1\alpha_2}{\Psi} \hat{\pi}_{t-1} + \frac{\beta^2\alpha_1\alpha_2}{\Psi} \hat{\pi}_t \right) \\
& - \frac{\alpha_1\alpha_2}{\alpha_1 + \alpha_1\alpha_2} \hat{\pi}_{t-1}, \tag{18}
\end{aligned}$$

where : $\Psi = 1 + \beta\alpha_1 + \beta^2\alpha_1\alpha_2$.

As we can observe that all coefficients in equation (18) are expressed in terms of non-adjustment rates ($\alpha_j = 1 - h_j$) and the subjective discount factor, β , thereby the coefficients in the generalized NKPC link dynamic effects of reset prices on inflation to the hazard function. As a result, information about the price reset hazard rates can be extracted from the aggregate data through the dynamic structure of the Phillips curve.

The economic reason why those lagged dynamic components should appear in the generalized NKPC but miss in the Calvo model is because they exert two opposing effects on current inflation through \hat{p}_t and \hat{p}_{t-1} respectively. The magnitudes of these effects depend on the price reset hazard function. In the general case, the effect of past optimal prices on current aggregate price \hat{p}_t should be different to those affecting lagged aggregate price \hat{p}_{t-1} . As a result, lagged expectations and lagged inflations should appear in the generalized NKPC. Conversely, in the Calvo case, the constant hazard function leads relevant reset prices to exert the exactly same amount of impact on both \hat{p}_t and \hat{p}_{t-1} , and thereby it causes them to be cancelled out.

2.4 The Final System of Equations

The general equilibrium system consists of equilibrium conditions derived from the optimization problems of economic agents, market clearing conditions and a monetary policy equation. Market clearing conditions require real prices clear the factor and good markets, while monetary policy determines nominal value of the real economy. I choose a Taylor rule to close the model.

$$I_t = I_{t-1}^{\rho_i} \left[\left(\frac{P_t}{P_{t-1}\bar{\pi}} \right)^{\phi_\pi} \left(\frac{Y_t}{Y_{t-1}} \right)^{\phi_y} \right]^{1-\rho_i} e^{q_t}. \tag{19}$$

Equation (19) is motivated by the interest rate smoothing specification for the Taylor rule¹⁰, which specifies a policy rule that the central bank uses to determine the nominal interest rate in the economy, where ϕ_π and ϕ_y denote short-run responses of the monetary authority to log deviations of inflation and the output growth rate, and q_t is a sequence of *i.i.d.* white noise with zero mean and a finite variance $(0, \sigma_q^2)$.

¹⁰See: the empirical work by Clarida et al. (2000)

After log-linearizing those equilibrium equations around the flexible-price steady state, log-linearized general equilibrium system consists of the generalized NKPC (20), real marginal cost (21), the household's intertemporal optimization condition (22), the Taylor rule (19) and exogenous stochastic processes. In the IS curve, I add an exogenous shock, d_t , to represent real aggregate demand disturbances¹¹

$$\hat{\pi}_t = \sum_{k=0}^{J-1} W_1(k) E_{t-k} \left(\sum_{j=0}^{J-1} W_2(j) \widehat{m}c_{t+j-k} + \sum_{i=1}^{J-1} W_3(i) \hat{\pi}_{t+i-k} \right) - \sum_{k=2}^{J-1} W_4(k) \hat{\pi}_{t-k+1}, \quad (20)$$

$$\widehat{m}c_t = (\delta + \phi) \hat{y}_t - (1 + \phi) \hat{z}_t, \quad (21)$$

$$\delta E_t [\hat{y}_{t+1}] = \delta \hat{y}_t + (\hat{i}_t - E_t [\hat{\pi}_{t+1}]) + d_t, \quad (22)$$

$$\hat{i}_t = (1 - \rho_i) (\phi_\pi \hat{\pi}_t + \phi_y (\hat{y}_t - \hat{y}_{t-1})) + \rho_i \hat{i}_{t-1} + q_t, \quad (23)$$

$$\hat{z}_t = \rho_z * \hat{z}_{t-1} + \epsilon_t \quad \text{where } \epsilon_t \sim N(0, \sigma_z^2), \quad (24)$$

$$d_t = \rho_d * d_{t-1} + \varepsilon_t \quad \text{where } \varepsilon_t \sim N(0, \sigma_d^2), \quad (25)$$

$$q_t \sim N(0, \sigma_q^2), \quad (26)$$

where weights (W_1, W_2, W_3, W_4) in the generalized NKPC are defined in equation (17). I collect the structure parameters into a vector $\mu = (\beta, \delta, \phi, \eta, \phi_\pi, \phi_y, \rho_i, \alpha_j s, \rho_z, \rho_d, \rho_i, \sigma_z^2, \sigma_d^2, \sigma_q^2)$. In the empirical study, I am interested in estimating values for those structural parameters by using the Bayesian approach.

3 Estimation

In this section, I solve and estimate the New Keynesian model described above by using the Bayesian approach. The full information Bayesian method has some appealing features in comparison to methods employed in the literature. As pointed out by An and Schorfheide (2007), this method is system-based, meaning that it fits the DSGE model to a vector of aggregate time series. Through a full characterization of the data generating process, it provides a formal framework for evaluating misspecified models on the basis of the data density. In addition, the Bayesian approach also provides a consistent method for dealing with rational expectations – one of the central elements in the DSGE models.

3.1 Bayesian Inference

I apply the Bayesian approach, set forth by DeJong et al. (2000), Schorfheide (2000) and Fernandez-Villaverde and Rubio-Ramirez (2004) among others, to estimate the structural parameters of the DSGE model. The Bayesian estimation is based on combining information gained from maximizing likelihood of the data and additional information about the parameters (the prior distribution). The main steps of this approach are as follows:

First, the linear rational expectation model is solved by using standard numerical methods (See: e.g. Uhlig, 1998 and Sims, 2002) to obtain the reduced form equations in its predetermined and exogenous variables.

¹¹Introducing this shock is not necessary for the theoretical model, but, in the Bayesian estimation, due to the singularity problem I need three shocks for three observables.

For example, the linearized DSGE model can be written as a rational expectations system in the form

$$\Upsilon_0(\mu)S_t = \Upsilon_1(\mu)S_{t-1} + \Upsilon_\epsilon(\mu)\epsilon_t + \Upsilon_\omega(\mu)\omega_t. \quad (27)$$

Here, S_t is a vector of all endogenous variables in the model, such as \hat{y}_t , $\hat{\pi}_t$, \hat{i}_t , *etc.* The vector ϵ_t stacks the innovations of the exogenous processes and ω_t is composed of one-period-ahead rational expectations forecast errors. Entries of $\Upsilon(\mu)$ matrices are functions of structural parameters in the model. The solution to (27) can be expressed as

$$S_t = \Psi_1(\mu)S_{t-1} + \Psi_\epsilon(\mu)\epsilon_t. \quad (28)$$

The second step involves writing the model in state space form. This is to augment the solution equation (28) with a measurement equation, which relates the theoretical variables to a vector of observables Y_obs_t .

$$Y_obs_t = A(\mu) + BS_t + CV_t. \quad (29)$$

Where $A(\mu)$ is a vector of constants, capturing the mean of S_t , and V_t is a set of shocks to the observables, including measurement errors.

Third, when we assume that all shocks in the state space form are normally distributed, we can use the Kalman filter (Sargent, 1989) to evaluate the likelihood function of the observables $\mathcal{L}(\mu|Y_obs^T)$. In contrast to other maximum likelihood methods, the Bayesian approach combines the likelihood function with prior densities $p(\mu)$, which includes all extra information about the parameters of interest. The posterior density $p(\mu|Y_obs^T)$ can be obtained by applying Bayes' theorem

$$p(\mu|Y_obs^T) \propto \mathcal{L}(\mu|Y_obs^T) p(\mu). \quad (30)$$

In the last step, μ is estimated by maximizing the likelihood function given data $\mathcal{L}(\mu|Y_obs^T)$ reweighted by the prior density $p(\mu)$, in that numerical optimization methods are used to find the posterior mode for μ and the inverse Hessian matrix. Finally, the posterior distribution is generated by using a random-walk Metropolis sampling algorithm¹².

3.2 Data and Priors

According to the empirical framework and research questions to be addressed in this paper, I choose the following three observables: growth rate of real GDP per capita, annualized inflation rate calculated from the consumer price index (CPI) and nominal interest rate for the U.S. over the period 1955.Q1 - 2008.Q4¹³. The output growth rate and inflation are detrended by the Hodrick-Prescott filter. Based on the definition of the model's variables and the observables, the measurement equations are defined as follows:

¹²I implement the Bayesian estimation procedure discussed above by using the MATLAB based package DYNARE, which is available at: <http://www.ceprenmap.cnrs.fr/dynare/>

¹³Details on the construction of the data set are provided in Appendix (B).

$$\begin{aligned}
y_obs &= \hat{y}_t - \hat{y}_{t-1} \\
\pi_obs &= \hat{\pi}_t \\
i_obs &= \hat{i}_t.
\end{aligned}$$

The priors I choose are in line with the mainstream values used in the Bayesian literature (e.g. Smets and Wouters, 2007 and Lubik and Schorfheide, 2005). They are centered around the average value of estimates of micro and macro data with fairly loose standard deviations.

I fix two parameters in advance. The discount factor β is equal to 0.99, implying an annual steady state real interest rate of 4%. The elasticity of substitution between intermediate goods is set to be 10, implying an average mark-up of around 11%. Both values are common in the literature.

The key structural parameters in this model are the hazard rates or the non-adjustment rates with respect to time-since-last-adjustment, α_j . I choose the priors for these parameters based on micro evidence on the mean frequency of price adjustments, reported by Bils and Klenow (2004). They find that the U.S. sectoral prices on average last only 2 quarters, which implies the hazard rate is equal to 0.5. Because the main goal of this study is to find out what shape of hazard function fits best to the macro data, I set all non-adjustment rates with the same mean of 2 quarters and a very loose standard deviation of 0.28. This prior leads to a 95% inter-quantile-range basically covering the whole interval between 0 and 1 quite evenly, except for the two extremes. In addition, same priors for all α_j reflect the view of a pricing model using a constant-hazard assumption. By choosing a large standard deviation for the prior, I allow the data to speak out quite freely about the shape of hazard rates over the time horizon, so that I can evaluate theoretical models from the point of view of a hazard function.

Moving to the other structural parameters, the prior for the relative risk aversion, δ , is set to follow a gamma distribution with mean 1.5 and a standard error of 0.375. This prior covers a wide range of values from the experimental and macro literature. The inverse elasticity of labour supply, ϕ , is difficult to calibrate, because there is a wide range of evidence in the literature. I choose the prior for this parameter to be normally distributed around the mean of 1.5. A mean of 1.5 is commonly estimated in the micro-labor studies (See: e.g. Blundell and Macurdy, 1999). I set a large standard error of 1.0. In the sensitivity analysis, I also check the robustness of my result to the other values of the prior mean.

Proceeding with priors for parameters in the Taylor rule, the priors for ϕ_π and ϕ_y are centered at the values commonly associated with a Taylor rule. Following Smets and Wouters (2007), I set a prior for the response coefficient to deviation of annualized inflation ϕ_π to be centered around 1.5 with a standard error of 0.1, and a prior for response coefficient to output growth rate ϕ_y to be centered around 0.5 with a standard error of 0.1. This rule also allows for interest rate smoothing with a prior mean of 0.5 and a standard deviation of 0.1.

Finally, I assume that the standard errors of the innovations follow an inverse-gamma distribution with a mean of 0.1 and two degrees of freedom. The persistence of the AR(1) process of the productivity shock is beta-distributed with a mean of 0.8 and the standard deviation of 0.1, and the persistence of the AR(1) process of the aggregate demand shock is beta-distributed with a mean of 0.5 and the standard deviation of 0.1.

3.3 Results

By applying the methodology described above, I proceed to gauge the degree of nominal rigidity in terms of the estimated structural parameters based on those prior distributions discussed above. The posterior modes of parameters are calculated by maximizing the log likelihood function of the data, and then the posterior distributions are simulated using the “Metropolis-Hastings” algorithm. The results presented here are based on 500,000 out of 1 million total draws and the average acceptance rate is around 0.31. With the simulation, I obtain convergence and relative stability in all measures of the parameter moments. The posterior mode, mean and 5%, 95% quantiles of the 17 estimated parameters are reported in Table 2 and the prior-posterior distributions are plotted in the figure appendix. The data provides strong information about most of the structural parameters, except for the inverse of the elasticity of labor supply and one of Taylor rule parameters. In those cases, I conduct various sensitivity tests to check the robustness of estimates of hazard function to changes in those poorly identified structural parameters.

Parameters	Prior			Posterior (M-H 500,000)			
	Dist.	Mean	S.D.	Mode	Mean	5%	95%
δ	gamma	1.5	0.375	4.311	4.149	3.379	4.994
ϕ	normal	1.5	1.0	1.459	1.282	-0.072	2.553
ϕ_π	normal	1.5	0.1	1.914	1.938	1.796	2.083
ϕ_y	normal	0.5	0.1	0.745	0.740	0.579	0.899
ρ_i	beta	0.5	0.1	0.634	0.622	0.572	0.674
α_1	beta	0.5	0.28	0.403	0.454	0.334	0.571
α_2	beta	0.5	0.28	0.941	0.855	0.717	0.998
α_3	beta	0.5	0.28	0.991	0.931	0.848	0.999
α_4	beta	0.5	0.28	0.663	0.676	0.431	0.962
α_5	beta	0.5	0.28	0.978	0.833	0.648	0.995
α_6	beta	0.5	0.28	0.975	0.801	0.620	0.983
α_7	beta	0.5	0.28	0.641	0.590	0.247	0.994
ρ_z	beta	0.8	0.1	0.992	0.988	0.978	0.998
ρ_d	beta	0.5	0.1	0.850	0.849	0.811	0.887
σ_z	invgam	0.1	2.0	1.663	1.859	1.142	2.602
σ_m	invgam	0.1	2.0	1.176	1.211	1.069	1.347
σ_d	invgam	0.1	2.0	0.721	0.735	0.591	0.873

Table 2: Posterior Distributions of Parameters (U.S.83-08)

Estimate for the relative risk aversion is high (4.149), but well in line with the benchmark values for macroeconomic studies, while the inverse of the elasticity of labor supply is not well identified in this model. Prior and posterior distributions are very close to each other, indicating that data does not provide information on this parameter under the current identification scheme.

The estimated monetary policy reaction function is consistent with the common view of the Taylor rule. Monetary policy responds strongly to the deviation of inflation (1.938), but not as much as to the output growth rate (0.74). There is a considerable degree of interest rate

smoothing, as the posterior mean of ρ_i is around 0.62. But the response of the nominal interest rate to output growth rate is also not well identified.

I turn now to the nominal rigidity that is represented by the estimates of non-adjustment rates, α_j . Contrary to the prior distributions, which are motivated by the Calvo model where all hazard rates are constant over time, the estimates reveal that the hazard function changes shape over time and the data strongly advocates a non-constant hazard function. The mean frequency of price adjustment is 32% per quarter which implies a mean price duration of 9.2 months. This result is consistent with findings by Nakamura and Steinsson (2008) using micro data. More importantly, price reset hazards vary substantially around this mean, depending on how long the price has been fixed. I will discuss the hazard function in more detail after the sensitivity analysis.

3.4 Robustness Tests

In this section, I test the robustness of the structural parameter estimates, especially those for the hazard function, to alternative priors, different model setups and data using different detrending methods. Table (3) and (4) report results of the posterior modes. In Table (3), I summarize results using the Hodrick-Prescott filter detrended data. I conduct the sensitivity analysis in the following steps: first, I check the prior sensitivity by altering the prior mean of ϕ , from 1.5 to 0.5. I choose to check this parameter because the estimation result shows that the inverse of labor supply elasticity is poorly identified. In addition, there is no consensus about the calibration value for this parameter in the literature. The first three columns in the table compare the results from the three alternative priors. 0.5 is a typical value motivated in the macro literature, while $\phi = 1.5$ is commonly estimated in the micro-labor studies (See: e.g. Blundell and Macurdy, 1999). I also check the value $\phi = 1$, which can be often found in the RBC literature. I find that changing the prior for ϕ mainly affects the posterior mode of ϕ itself, leaving estimates of other parameters for preference and monetary policy qualitatively unchanged. These results manifest an observational equivalence problem commonly found in estimating DSGE models. The log likelihood function is mostly flat on the choices of priors for ϕ . Posterior estimates are mainly driven by the prior instead of data. As for the non-adjustment rates, the choice of the prior for labor elasticity affects magnitudes of non-adjustment rates at all frequencies. Interestingly, it shows that making labor supply more elastic, decreasing in the value of ϕ , leads to more frequent price adjustments estimated. Non-adjustment rates are significantly lower at all frequencies except for the 7th. Despite changes in the magnitude, the general pattern of the hazard function remains the same.

In the next two columns, I change the model setup. When adopting the standard Taylor rule with output gap instead of output growth rate, it results in a large change in the estimated ϕ_y , which becomes very small, indicating that central bank reacts less to output gap in the monetary policy decision making, because it is an unobservable in the economy. Estimates for the non-adjustment rates, however, are almost identical as in the benchmark case. I also change the model setup by fixing the hazard rates to the value of 0.5, implying an average duration of 2 quarters¹⁴. As seen in the last column, fixing the hazard rates has implications for the

¹⁴I call it the pseudo-Calvo model because, in this case, I truncate the hazard function at the 7th quarter. As a result, it is not exactly equivalent to the Calvo model, which implies an infinite horizon for the hazard function.

Tests	H-P filter Detrended Data				
Parameter	$\phi = 1.5$	$\phi = 1$	$\phi = 0.5$	<i>TaylorRule</i>	<i>Calvo</i>
δ	4.311 (0.322)	4.012 (0.062)	4.097 (0.063)	4.790 (0.332)	2.949 (0.247)
ϕ	1.46 (0.867)	1.063 (0.647)	0.618 (0.918)	1.687 (0.433)	0.315 (0.309)
ϕ_π	1.914 (0.091)	1.864 (0.087)	1.859 (0.087)	1.956 (0.084)	2.052 (0.083)
ϕ_y	0.745 (0.097)	0.794 (0.097)	0.793 (0.097)	0.111 (0.069)	0.704 (0.097)
α_1	0.403 (0.076)	0.050 (0.034)	0.051 (0.034)	0.469 (0.085)	0.5
α_2	0.941 (0.138)	0.772 (0.148)	0.781 (0.148)	0.918 (0.126)	0.5
α_3	0.991 (0.028)	0.968 (0.153)	0.968 (0.152)	0.992 (0.072)	0.5
α_4	0.663 (0.151)	0.408 (0.157)	0.407 (0.156)	0.624 (0.145)	0.5
α_5	0.980 (0.067)	0.914 (0.232)	0.916 (0.231)	0.973 (0.147)	0.5
α_6	0.980 (0.064)	0.926 (0.242)	0.926 (0.241)	0.978 (0.137)	0.5
α_7	0.641 (0.262)	0.737 (0.388)	0.741 (0.386)	0.586 (0.252)	0.5
Log Margin. Likeli.	-907.32	-914.91	-912.63	-935.57	-918.04

Table 3: Sensitivity Check for H-P Detrended Data

Tests	Linearly detrended Data				
Parameter	$\phi = 1.5$	$\phi = 1$	$\phi = 0.5$	<i>TaylorRule</i>	<i>Calvo</i>
δ	4.352 (0.640)	4.357 (0.627)	4.364 (0.621)	5.226 (0.473)	3.671 (0.489)
ϕ	2.10 (0.886)	1.72 (0.854)	1.373 (0.816)	2.305 (0.878)	1.942 (0.797)
ϕ_π	1.912 (0.081)	1.911 (0.081)	1.911 (0.081)	1.908 (0.097)	1.952 (0.077)
ϕ_y	0.669 (0.099)	0.669 (0.099)	0.670 (0.099)	0.258 (0.395)	0.662 (0.098)
α_1	0.504 (0.075)	0.493 (0.075)	0.482 (0.075)	0.5632 (0.077)	0.5
α_2	0.716 (0.126)	0.710 (0.126)	0.705 (0.128)	0.789 (0.138)	0.5
α_3	0.980 (0.063)	0.979 (0.064)	0.979 (0.065)	0.969 (0.187)	0.5
α_4	0.143 (0.218)	0.124 (0.216)	0.106 (0.215)	0.397 (0.135)	0.5
α_5	0.193 (0.645)	0.182 (0.604)	0.180 (0.608)	0.601 (0.389)	0.5
α_6	0.462 (1.216)	0.446 (1.180)	0.437 (1.163)	0.785 (0.415)	0.5
α_7	0.329 (1.149)	0.359 (1.147)	0.379 (1.147)	0.221 (0.460)	0.5
Log Margin. Likeli.	-792.12	-792.44	-792.79	-837.09	-797.88

Table 4: Sensitivity Check for Linearly Detrended Data

estimates of other structural parameter. For example, it leads to a much lower estimate for the intertemporal elasticity of substitution and inverse of labor elasticity. In addition, in terms of log marginal likelihood, both the output-gap-Taylor-rule model and the fixed-hazard setup are clearly less favored by the data. In the last row of the table, I report the log marginal likelihood of the data for each model. It shows that changing priors of ϕ only marginally affects the data density, but the data gives strong support of the flexible hazard model as opposed to the fixed-hazard model and the output gap Taylor rule. The Bayes factor in favor of the flexible hazard model is approximately in the order of 10^5 .

I conduct the same tests by using linearly detrended data again, which is reported in Table (4). All results from previous exercises are confirmed, but the drawback of using linearly detrended data is that they do not deliver accurate information about the hazard function. As seen in the table, non-adjustment rates are much different to what we have from the HP-detrended data and those after the 3rd quarter are all statistically insignificant. The reason for this could be that the linearly detrending mixes macro dynamics at the business cycle frequencies with those from other frequencies, so that it biases the estimates and reduces the efficiency of the estimation too.

3.5 Aggregate Hazard Function and Implications for Macro Modeling

In this part, I evaluate new evidence on the aggregate price reset hazard rates obtained from my empirical analysis and also discuss its implication for macro modeling of sticky prices.

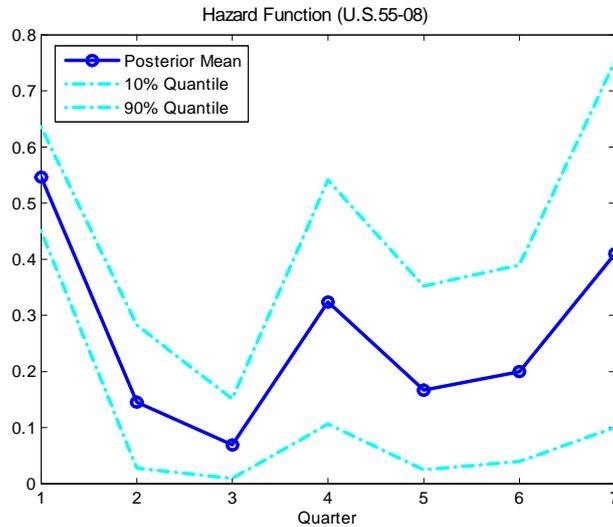


Figure 1: The Price Reset Hazard Function

I plot the estimates of hazard rates in Figure (1). The hazard rate is high one quarter after the last price adjustment (55%), and drops in the next two consecutive quarters to around 10% and rise again in the 4th quarter. Afterwards, hazard rates are largely increasing with

This pseudo-Calvo can be view as an approximation of the real Calvo model. I estimate also the pseudo-Calvo model with longer horizons, but it does not change the main conclusion.

the age of the price. Overall, the hazard function has a U-shape with a spike at the fourth quarter. I also calculate the distribution of price durations from the estimated hazard rates by using formula (12). It yields that around 34.2% of prices last for less than one quarter. From micro data studies, we learn that prices of apparel, unprocessed food, energy and travel are the most frequently adjusted prices, whose median durations last less than three months. Moreover, 12.4% of prices have the mean duration of four quarters. Examples for those prices are services, such as hairdressers or public transportation.

This finding has important bearing on the macro modeling of sticky prices. Overall, I find new evidence can not be explained by any single theory of sticky prices. For the first half of the hazard function (from the first to the fourth quarters), it appears that the pricing decision is mainly characterized by either the flexible price setting or by a time-dependent aspect (e.g. Taylor, 1980). The survey evidence has shown that many firms conduct yearly price revisions due to costly information. This kind of behavior can also be motivated by the theory of customer markets, which indicates that long-term customer relationships are an important consideration in pricing decisions (See: e.g. Rotemberg, 2005). On the other hand, the upward-sloping part of the hazard function indicates that the state-dependent pricing also plays an important role in the pricing decision, the more outdated a price becomes. In fact, this generalized time-dependent model can be viewed as a proxy of the more microfounded state-dependent model. More microfounded pricing models, such as Dotsey et al. (1999), show that the state-dependent pricing behavior implies an increasing hazard function. If we consider a relative stable economy, the hazard function of the deviation from the optimum largely coincides with the hazard function of time-since-last-adjustment. The longer a price is fixed, the more likely it deviates significantly from the optimum, and hence its probability of being adjusted rises. Since in my data set, the annual inflation rates stay under 2% for most of the sample periods except for the turbulent decade between early 1970s and early 1980s, it is arguable that the time elapsed since last adjustment is a good proxy for the deviation from the optimum, therefore the increasing part of the hazard function gives the pricing decision a state-dependent aspect.

To summarize these results, the new evidence of the aggregate hazard function reveals that, for the less sticky prices ranging in duration from one to four quarters, time-depend pricing plays a major role, while, for stickier prices with a duration longer than 4 quarters, the state-dependent pricing dominates.

Conclusion

In this paper, I document new evidence on the shape of the aggregate hazard function. I construct a DSGE model featuring nominal rigidity that allows for a flexible hazard function of price setting. The generalized NKPC possesses a richer dynamic structure, with which I can infer the shape of the hazard function underlying aggregate dynamics. Identifying the hazard function from aggregate data is a useful exercise, because, first, estimating hazard rates directly from a DSGE model provides the most consistent way to compare the theoretical concept with the empirical evidence. Second, it overcomes some weaknesses of estimates using micro data, such as contamination by the idiosyncratic effects and the limited availability of the long time series data. As a result, this study delivers some useful insights for macroeconomists, which can be readily used to guide macro modeling. At last, a caveat should be note that the identification method relies on scrutinizing effects of past reset prices on current inflation, and those effects decay over time. Consequently, the information contents of aggregate data for hazard rates deteriorate fast with the length of the time since last adjustment. After the fourth quarter, estimated hazard rates become imprecise.

A Deviation of the New Keynesian Phillips Curve

Starting from 15

$$\hat{p}_t^* = E_t \left[\sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} (\widehat{mc}_{t+j} + \hat{p}_{t+j}) \right] \quad (31)$$

$$= E_t \left[\sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \widehat{mc}_{t+j} \right] + E_t \left[\sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{p}_{t+j} \right] \quad (32)$$

The last term can be further expressed in terms of future rates of inflation

$$\begin{aligned} \sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{p}_{t+j} &= \frac{1}{\Psi} \hat{p}_t + \frac{\beta S_1}{\Psi} \hat{p}_{t+1} + \dots + \frac{\beta^{J-1} S_{J-1}}{\Psi} \hat{p}_{t+J-1} \\ &= \frac{1}{\Psi} \hat{p}_t + \frac{\beta S_1}{\Psi} \hat{p}_t + \frac{\beta S_1}{\Psi} (\hat{p}_{t+1} - \hat{p}_t) + \dots + \frac{\beta^{J-1} S_{J-1}}{\Psi} \hat{p}_{t+J-1} \\ &= \left(\frac{1}{\Psi} + \frac{\beta S_1}{\Psi} \right) \hat{p}_t + \sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+j} + \frac{\beta^2 S_2}{\Psi} \hat{p}_{t+1} + \dots + \frac{\beta^{J-1} S_{J-1}}{\Psi} \hat{p}_{t+J-2} \\ &= \left(\frac{1}{\Psi} + \frac{\beta S_1}{\Psi} + \frac{\beta^2 S_2}{\Psi} \right) \hat{p}_t + \sum_{j=1}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+j} + \sum_{j=2}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+j-1} \\ &\quad + \frac{\beta^3 S_3}{\Psi} \hat{p}_{t+1} + \dots + \frac{\beta^{J-1} S_{J-1}}{\Psi} \hat{p}_{t+J-2} \\ &\quad \vdots \\ &= \left(\frac{1}{\Psi} + \frac{\beta S_1}{\Psi} + \dots + \frac{\beta^{J-1} S_{J-1}}{\Psi} \right) \hat{p}_t + \sum_{j=1}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+j} \\ &\quad + \sum_{j=2}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+j-1} + \dots + \frac{\beta^{J-1} S_{J-1}}{\Psi} \hat{\pi}_{t+1} \\ &= \hat{p}_t + \sum_{i=1}^{J-1} \sum_{j=i}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+i} \end{aligned}$$

The optimal price can be expressed in terms of inflation rates, real marginal cost and aggregate prices.

$$\hat{p}_t^* = \hat{p}_t + E_t \left[\sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \widehat{mc}_{t+j} \right] + E_t \left[\sum_{i=1}^{J-1} \sum_{j=i}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+i} \right] \quad (33)$$

Next, I derive the aggregate price equation as the sum of past optimal prices. I lag equation

33 and substitute it for each \hat{p}_{t-j}^* into equation 16

$$\begin{aligned}
\hat{p}_t &= \theta(0) \hat{p}_t^* + \theta(1) \hat{p}_{t-1}^* + \dots + \theta(J-1) \hat{p}_{t-J+1}^* \\
&= \theta(0) \left[\hat{p}_t + E_t \left(\sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \widehat{m}c_{t+j} \right) + E_t \left(\sum_{i=1}^{J-1} \sum_{j=i}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+i} \right) \right] \\
&+ \theta(1) \left[\hat{p}_{t-1} + E_{t-1} \left(\sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \widehat{m}c_{t+j-1} \right) + E_{t-1} \left(\sum_{i=1}^{J-1} \sum_{j=i}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+i-1} \right) \right] \\
&\vdots \\
&+ \theta(J-1) \left[\hat{p}_{t-J+1} + E_{t-J+1} \left(\sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \widehat{m}c_{t+j-J+1} \right) + E_{t-J+1} \left(\sum_{i=1}^{J-1} \sum_{j=i}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+i-J+1} \right) \right] \\
\hat{p}_t &= \sum_{k=0}^{J-1} \theta(k) \left[\hat{p}_{t-k} + E_{t-k} \left(\underbrace{\sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \widehat{m}c_{t+j-k} + \sum_{i=1}^{J-1} \sum_{j=i}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+i-k}}_{F_{t-k}} \right) \right] \tag{34}
\end{aligned}$$

Where F_t summarizes all current and lagged expectations formed at period t . Finally, we derive the New Keynesian Phillips curve from equation 34.

$$\begin{aligned}
\hat{p}_t &= \sum_{k=0}^{J-1} \theta(k) \hat{p}_{t-k} + \underbrace{\sum_{k=0}^{J-1} \theta(k) F_{t-k}}_{Q_t} \\
\hat{\pi}_t &= \sum_{k=0}^{J-1} \theta(k) \hat{p}_{t-k} - \hat{p}_{t-1} + Q_t \\
&= \theta(0) (\hat{p}_t - \hat{p}_{t-1}) + \theta(0) \hat{p}_{t-1} + \theta(1) \hat{p}_{t-1} + \dots + \theta(J-1) \hat{p}_{t-J+1} - \hat{p}_{t-1} + Q_t \\
&= \theta(0) (\hat{p}_t - \hat{p}_{t-1}) + (\theta(0) + \theta(1)) \hat{p}_{t-1} + \theta(2) \hat{p}_{t-2} + \dots + \theta(J-1) \hat{p}_{t-J+1} - \hat{p}_{t-1} + Q_t \\
&= \underbrace{\theta(0) \hat{\pi}_t}_{W(0)} + \underbrace{(\theta(0) + \theta(1)) \hat{\pi}_{t-1} + (\theta(0) + \theta(1) + \theta(2)) \hat{p}_{t-2} + \dots + \theta(J-1) \hat{p}_{t-J+1} - \hat{p}_{t-1}}_{W(1)} + Q_t \\
&\vdots \\
&= W(0) \hat{\pi}_t + W(1) \hat{\pi}_{t-1} + \dots + W(J-2) \hat{\pi}_{t-J+2} + \underbrace{W(J-1) \hat{p}_{t-J+1} - \hat{p}_{t-1}}_{=1} + Q_t \\
&= W(0) \hat{\pi}_t + \dots + W(J-2) \hat{\pi}_{t-J+2} + \underbrace{\hat{p}_{t-J+1} - \hat{p}_{t-J+2}}_{-\hat{\pi}_{t-J+2}} + \hat{p}_{t-J+2} - \dots + \underbrace{\hat{p}_{t-2} - \hat{p}_{t-1}}_{-\hat{\pi}_{t-1}} + Q_t \\
(1 - W(0)) \hat{\pi}_t &= -(1 - W(2)) \hat{\pi}_{t-1} - \dots - (1 - W(J-1)) \hat{\pi}_{t-J+2} + Q_t \\
\hat{\pi}_t &= - \sum_{k=2}^{J-1} \frac{1 - W(k)}{1 - \theta(0)} \hat{\pi}_{t-k+1} + \sum_{k=0}^{J-1} \frac{\theta(k)}{1 - \theta(0)} F_{t-k}
\end{aligned}$$

The generalized New Keynesian Phillips curve is:

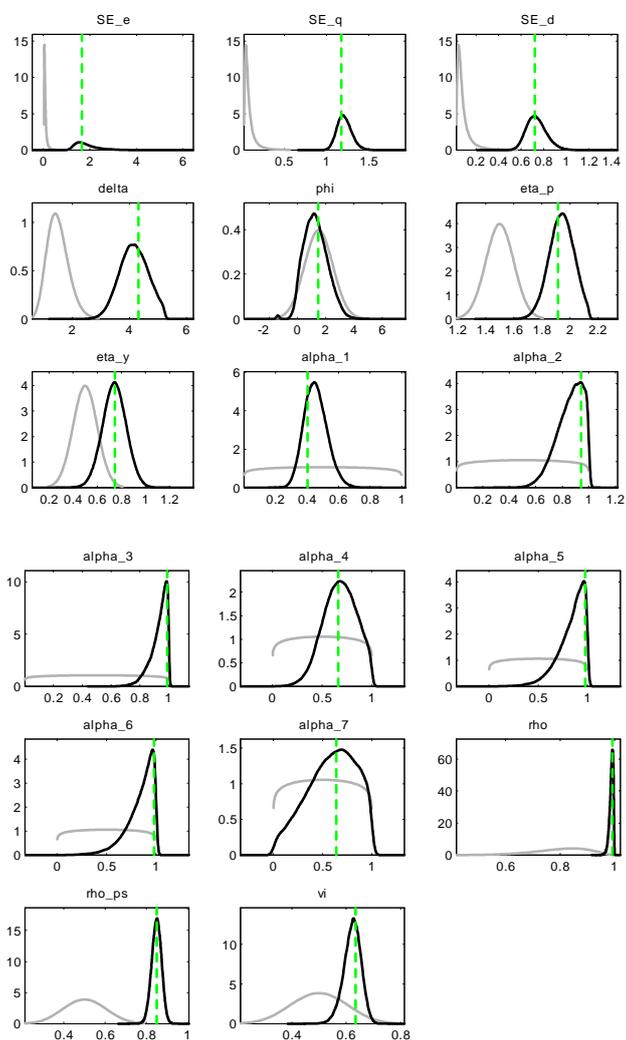
$$\begin{aligned}
\hat{\pi}_t &= \sum_{k=0}^{J-1} \frac{\theta(k)}{1-\theta(0)} E_{t-k} \left(\sum_{j=0}^{J-1} \frac{\beta^j S_j}{\Psi} \widehat{m}c_{t+j-k} + \sum_{i=1}^{J-1} \sum_{j=i}^{J-1} \frac{\beta^j S_j}{\Psi} \hat{\pi}_{t+i-k} \right) \\
&\quad - \sum_{k=2}^{J-1} \Phi(k) \hat{\pi}_{t-k+1}, \quad \text{where } \Phi(k) = \frac{\sum_{j=k}^{J-1} S_j}{\sum_{j=1}^{J-1} S_j}, \quad \Psi = \sum_{j=0}^{J-1} \beta^j S_j
\end{aligned} \tag{35}$$

B Data

The data used in this paper is taken from the FRED (Federal Reserve Economic Data) maintained by the Federal Reserve Bank of St. Louis.

- Growth rate of real GDP per capita: is based on the Real Gross Domestic Product (Series: GDPC1). They are in the unit of billions of chained 2005 dollars, quarterly frequency and seasonally adjusted. To construct per capita GDP, I use the Civilian Noninstitutional Population (Series: CNP16OV) from the Bureau of Labor Statistics, U.S. Department of Labor. The monthly data is converted into quarterly frequency by arithmetic averaging. Per capita real output growth is defined as: $100 \times [\ln(GDP_t/POP_t) - \ln(GDP_{t-1}/POP_{t-1})]$. Finally the data is detrended by the Hodrick-Prescott filter.
- Inflation rate: is calculated by using Consumer Price Index for all urban consumers: all items (Series: CPIAUCSL), seasonally adjusted. The monthly data is converted into quarterly frequency by arithmetic averaging. Annualized Inflation rate is defined as $400 \times \ln(P_t/P_{t-1})$. Finally the data is detrended by the Hodrick-Prescott filter.
- Nominal interest rate: is the Effective Federal Funds Rate (Series: FEDFUNDS). The monthly data is converted into quarterly frequency by arithmetic averaging. The data is detrended with the trend inflation calculated by using the Hodrick-Prescott filter and then mean adjusted.

C Figure



Prior and Posterior Distributions for data
1955-2008

References

- Alvarez, L. J.** (2007), What do micro price data tell us on the validity of the new keynesian phillips curve?, Kiel working papers, Kiel Institute for the World Economy.
- Alvarez, L. J., E. Dhyne, M. Hoeberichts, C. Kwapil, H. L. Bihan, P. Lünnemann, F. Martins, R. Sabbatini, H. Stahl, P. Vermeulen, and J. Vilmunen** (2006), Sticky prices in the euro area: A summary of new micro-evidence, *Journal of the European Economic Association*, 4(2-3), 575–584.
- An, S. and F. Schorfheide** (2007), Bayesian analysis of dsge models, *Econometric Reviews*, 26(2-4), 113–172.
- Bils, M. and P. J. Klenow** (2004), Some evidence on the importance of sticky prices, *Journal of Political Economy*, 112(5), 947–985.
- Blundell, R. and T. Macurdy** (1999), Labor supply: A review of alternative approaches, in: O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*.
- Caballero, R. J. and E. M. Engel** (2007), Price stickiness in ss models: New interpretations of old results, Cowles Foundation Discussion Papers 1603, Cowles Foundation, Yale University.
- Calvo, G. A.** (1983), Staggered prices in a utility-maximizing framework, *Journal of Monetary Economics*, 12(3), 383–98.
- Campbell, J. R. and B. Eden** (2005), Rigid prices: evidence from u.s. scanner data, Working paper series, Federal Reserve Bank of Chicago.
- Canova, F. and L. Sala** (2009), Back to square one: Identification issues in dsge models, *Journal of Monetary Economics*, 56(4), 431–449.
- Caplin, A. S. and D. F. Spulber** (1987), Menu costs and the neutrality of money, *The Quarterly Journal of Economics*, 102(4), 703–25.
- Carvalho, C. and N. A. Dam** (2008), Estimating the cross-sectional distribution of price stickiness from aggregate data, Ssrn elibrary working papers, SSRN.
- Cecchetti, S. G.** (1986), The frequency of price adjustment : A study of the newsstand prices of magazines, *Journal of Econometrics*, 31(3), 255–274.
- Clarida, R., J. Galí, and M. Gertler** (2000), Monetary policy rules and macroeconomic stability: Evidence and some theory, *The Quarterly Journal of Economics*, 115(1), 147–180.
- Coenen, G., A. T. Levin, and K. Christoffel** (2007), Identifying the influences of nominal and real rigidities in aggregate price-setting behavior, *Journal of Monetary Economics*, 54(8), 2439–2466.
- DeJong, D. N., B. F. Ingram, and C. H. Whiteman** (2000), A bayesian approach to dynamic macroeconomics, *Journal of Econometrics*, 98(2), 203–223.

- Dixit, A. K. and J. E. Stiglitz** (1977), Monopolistic competition and optimum product diversity, *American Economic Review*, 67(3), 297–308.
- Dotsey, M., R. G. King, and A. L. Wolman** (1999), State-dependent pricing and the general equilibrium dynamics of money and output, *The Quarterly Journal of Economics*, 114(2), 655–690.
- Fernandez-Villaverde, J. and J. Rubio-Ramirez** (2004), Comparing dynamic equilibrium models to data: a bayesian approach, *Journal of Econometrics*, 123(1), 153–187.
- Fougere, D., H. L. Bihan, and P. Sevestre** (2005), Heterogeneity in consumer price stickiness - a microeconomic investigation, Working paper series, European Central Bank.
- Gali, J. and M. Gertler** (1999), Inflation dynamics: A structural econometric analysis, *Journal of Monetary Economics*, 44(2), 195–222.
- Gali, J., M. Gertler, and J. D. Lopez-Salido** (2001), European inflation dynamics, *European Economic Review*, 45(7), 1237–1270.
- Golosov, M. and R. E. Lucas** (2007), Menu costs and phillips curves, *Journal of Political Economy*, 115, 171–199.
- Jadresic, E.** (1999), Sticky prices - an empirical assessment of alternative models, Imf working papers, International Monetary Fund.
- Kiley, M. T.** (2002), Partial adjustment and staggered price setting, *Journal of Money, Credit and Banking*, 34(2), 283–98.
- Lubik, T. and F. Schorfheide** (2005), A bayesian look at new open economy macroeconomics, Economics working paper archive, The Johns Hopkins University, Department of Economics.
- Mackowiak, B. and F. Smets** (2008), On implications of micro price data for macro models, Working Paper Series 960, European Central Bank.
- Mash, R.** (2004), Optimising microfoundations for inflation persistence, Tech. rep.
- Midrigan, V.** (2007), Menu costs, multi-product firms, and aggregate fluctuations, CFS Working Paper Series 2007/13, Center for Financial Studies.
- Nakamura, E. and J. Steinsson** (2008), Five facts about prices: A reevaluation of menu cost models, *The Quarterly Journal of Economics*, 123(4), 1415–1464.
- Rios-Rull, J.-V., F. Schorfheide, C. Fuentes-Albero, M. Kryshko, and R. S. Llopis** (2009), Methods versus substance: Measuring the effects of technology shocks on hours, NBER Working Papers 15375, National Bureau of Economic Research, Inc.
- Rotemberg, J. J.** (2005), Customer anger at price increases, changes in the frequency of price adjustment and monetary policy, *Journal of Monetary Economics*, 52(4), 829–852.
- Sargent, T. J.** (1989), Two models of measurements and the investment accelerator, *Journal of Political Economy*, 97(2), 251–87.

- Sbordone, A. M.** (2002), Prices and unit labor costs: a new test of price stickiness, *Journal of Monetary Economics*, 49(2), 265–292.
- Schorfheide, F.** (2000), Loss function-based evaluation of dsge models, *Journal of Applied Econometrics*, 15(6), 645–670.
- Sheedy, K. D.** (2007), Intrinsic inflation persistence, CEP Discussion Papers dp0837, Centre for Economic Performance, LSE.
- Sims, C. A.** (2002), Solving linear rational expectations models, *Computational Economics*, 20(1-2), 1–20.
- Smets, F. and R. Wouters** (2007), Shocks and frictions in us business cycles: A bayesian dsge approach, *American Economic Review*, 97(3), 586–606.
- Taylor, J. B.** (1980), Aggregate dynamics and staggered contracts, *Journal of Political Economy*, 88(1), 1–23.
- Uhlig, H.** (1998), A toolkit for analysing nonlinear dynamic stochastic models easily.
- Wolman, A. L.** (1999), Sticky prices, marginal cost, and the behavior of inflation, *Economic Quarterly*, (Fall), 29–48.
- Yao, F.** (2009), The cost of tractability and the calvo pricing assumption, SFB 649 Discussion Papers 042, Humboldt University, Berlin, Germany.