

Time Variation in the Sticky Information Phillips Curve

Evan Hart

May 3, 2024

Abstract

I analyze the empirical evidence of variation in the structural parameter of the sticky information model. This involves scrutinizing both the statistical significance of the variation and its economic implications. Upon examination, I discover a systematic trend in firms' attention to relevant macroeconomic conditions, indicating a decline in attention over time. This decline in attention is found to have significant economic implications, particularly in relation to the expected volatility of prices. By incorporating models proposed by Ball et al. (2003), I demonstrate that this trend in attention is associated with a decrease in the expected volatility of prices. Comparisons between fixed attention models and time-varying attention models reveal substantial differences in parameter estimates. A Bayesian model comparison favors a fixed parameter model but does not eliminate the time varying parameter model altogether. These findings underscore the importance of considering time-varying attention in economic modeling and policymaking, as it provides valuable insights into the evolving nature of decision-making processes in response to macroeconomic information.

1 Introduction

In the Sticky Information Phillips curve, attention is used as a nominal rigidity. Firms strive to set a profit-maximizing price path based on the current aggregate price and output, but they do not have access to future data. Therefore, they must rely on forecasts of these variables to set future prices in their path. Attention is modeled as a probability that they are given a current history of relevant variables to use in their forecasts. Empirical research on attention to economic conditions holds relevance for both macroeconomic policy and firm strategy. Identifying this aggregate attention presents an opportunity for the Federal Reserve to enact policies that are more effective. In my analysis of attentiveness, I identify variations that are important economically and cannot be ruled out statistically.

Research into structural models employs various forms of imperfect knowledge. The first is Lucas (1972), where agents possess limited knowledge about changes in prices. A more recent perspective introduced by Sims (2003) posits that agents suffer from limited attention, preventing them from processing all available information. In my study, I employ the Phillips Curve developed by Mankiw and Reis (2002), where the likelihood of setting a fixed price is substituted with the probability of receiving new information. (Henceforth referred to as SIPC.)

We may anticipate that the rate of attention changes over time. Branch et al. (2009) proposed a model where the attention rate is an equilibrium value determined by the current state of the economy. For instance, the attention rate declines if the costs associated with acquiring and processing information increase. These variations are significant because the level of attention influences the economy's responsiveness to external shocks. Therefore, I estimate a model that permits the rate of attention to follow a random walk process. This approach enables the rate to change in every period of the model, allowing the data to dictate the path instead of imposing a more structured pattern such as a structural break.

Shifts in attentiveness have been explored through various frameworks. Coibion and Gorodnichenko (2015) utilize a method that links ex post mean forecast errors to ex ante revisions in the average forecast among surveyed professionals, providing a benchmark for evaluating potential deviations by economic agents. On the other hand, Carroll (2003) constructs a model that compares a survey of household expectations to the forecasts from the Survey of Professional Forecasters.

The empirical approach employed in this paper closely resembles that implemented by Coibion (2010). The expectations of firms will be represented using data from the Survey of Professional Forecasters, replacing the conventional rational expectations theory, wherein his-

torical values of the variable of interest would be used. In the SIPC, adopting historical values would result in an error process highly correlated with the regressors and instruments. In addition, this model uses an error term that is assumed to be additive to the inflation equation and drawn from an independent and identically distributed (IID) process. This approach deviates from the AR(1) process used to introduce uncertainty in the models of Ball et al. (2003) and Branch et al. (2009).

Using Bayesian estimation, I identify a downward trend in the rate of attention over the sample period. To determine if the changes are economically significant, I turn to a general equilibrium model developed by Ball et al. (2003). This enriched model uses a similar sticky information Phillips curve but creates a general equilibrium model by specifying a demand curve and an objective function by which monetary policy can be derived. Using this enriched model, I have uncovered economically significant changes. Specifically, the shift in attention translates to a substantial decrease in the expected volatility of aggregate price, ranging from 5.5% to 12.7%. These findings highlight the potential impact of implementing the proposed adjustments, suggesting a notable improvement in the stability of aggregate price movements.

Bayesian model comparison allows us to gauge how credible two models are based on our data. Chib and Jeliazkov (2001) provide the framework for estimating the marginal likelihood when implementing a Markov chain Monte Carlo estimation. The Bayesian model comparison places a 16% probability on the TVP model. This means it cannot be ruled out statistically.

Enriched attention models can inform us of the causes of these behaviors. In Branch et al. (2005), firms will decrease their attention rate in response to an increase in the cost of absorbing aggregate information. The increasing volume of available data is usually associated with increased information costs. In Hart (2023), I extended the Branch et al. (2005) model by introducing a bounded rationality assumption. In response to large shocks, firms will be temporarily more attentive. This response is caused by the belief that the economic shocks are drawn from a distribution with a higher standard deviation.

The rest of the paper has the following structure. First, I develop a TVP version of the SIGE. Then, I explain what data will be used and how it is structured. I go on to explain the estimation method and priors. Finally, I present and explore the results of the estimation.

2 Empirical Approach

2.1 Model

The assumption developed by Mankiw and Reis 2002 is used to model inflation. Firms use outdated information until they are randomly “selected” to receive new information. The parameter that measures firms’ attention is $(1-\lambda)$, the probability of being selected. Aggregate price is assumed to be a weighted average over firms’ expectation of the optimal price:

$$\pi_t = \frac{(1-\lambda)}{\lambda} \alpha x_t + \sum_{j=0}^{J-1} (1-\lambda) * \lambda^j E_{t-1-j}(\pi_t + \alpha \Delta x_t) + \nu_t.$$

The weighting system is formed by the fixed probability of receiving new information. The proportion of firms using the newest information set I_{t-1} is $(1-\lambda)$. The proportion of firms using the information set I_{t-2} is $(1-\lambda)\lambda$, the probability of receiving new information last period multiplied by the probability of staying in that information set. As an information set becomes more outdated it will shrink by the probability of staying in that information set each period.

If we allow the probability of new information to change over time the weighting system is formed by the history of probabilities of receiving new information. The proportion of firms using the newest information set I_{t-1} is $(1-\lambda_t)$. The proportion of firms using the information set I_{t-2} is $(1-\lambda_{t-1})\lambda_t$, the probability of receiving new information last period multiplied by the probability of staying in that information set. This time-varying parameter model has the following form of the SIPC:

$$\pi_t = \frac{(1-\lambda_t)}{\lambda_t} \alpha x_t + \sum_{j=0}^{J-1} a_{t,j+1} E_{t-1-j}(\pi_t + \alpha \Delta x_t) + \nu_t, \quad (1)$$

s.t

$$a_{t,1} = (1-\lambda_t)$$

$$a_{t,2} = \lambda_t * (1-\lambda_{t-1})$$

$$a_{t,3} = \lambda_{t-1} * \lambda_t * (1-\lambda_{t-2})$$

$$a_{t,4} = \lambda_{t-2} * \lambda_{t-1} * \lambda_t * (1-\lambda_{t-3})$$

$$a_{t,5} = \lambda_{t-3} * \lambda_{t-2} * \lambda_{t-1} * \lambda_t * (1-\lambda_{t-4}).$$

$$x_t = \delta Z_t + \xi_t, \quad (2)$$

$$\lambda_t = \lambda_{t-1} + \omega_t. \quad (3)$$

The transition equation is modeled as a random walk. The selection of a random walk stems from uncertainty regarding the underlying process. Prior to this study, there has been no estimation of an SIPC with a time-varying parameter. The lack of structure placed on the random walk assumption allows the data to inform us on the underlying form.

2.2 Data

The data is naturally divided into two groups. First, we have realized times series data on inflation, output and the natural rate of output. Second, we have firm's expectation. Expectation data of firms will be proxied using expectation data from the survey of professional forecasters.

The GDP price index forecasts are the median from the survey of Professional Forecasters. Collected from the Philadelphia Federal Reserve website. This data set runs from 1968-10-01 to 2022-10-01. The file includes 9 columns, two are dedicated to the date, two are annual average forecasts. The other 6 contain our expectation data. For a row labeled time t , we have the expectations using information set $t - 1$. Using this information set, the median forecast is provided starting with P_{t-1} ending with P_{t+4} . Forecasts for the quarterly level of the chain-weighted GDP price index can be used to construct an expectation of inflation using the following formula:

$$E_{t-1}\pi_{t+i} = \log(E_{t-1}P_{t+i}) - \log(E_{t-1}P_{t+i-1}) \text{ for } i \in \{0, 1, 2, 3, 4\}$$

The Real GDP forecasts are the taken from the same source and have the same format. No expectation data on the change in natural output is proxied with the actual change. The expected change in the output gap can be broken into two pieces given our simplifying assumption. The expectation data will form the first part. While the second part will use the actual change in the output gap. Therefore, the expected change in the output gap has the following form:

$$E_{t-1}\alpha\Delta x_{t+i} = \alpha * (E_{t-1}\Delta y_{t+i} - \Delta y_{t+i}^N)$$

s.t.

$$E_{t-1}\Delta y_{t+i} = \log(E_{t-1}(Y_{t+i})) - \log(E_{t-1}(Y_{t-1+i})),$$

$$\Delta y_{t+i}^N = \log(Y_{t+i}^N) - \log(Y_{t+i-1}^N) \text{ for } i \in \{0, 1, 2, 3, 4\}.$$

Inflation is measured using the implicit GDP price deflator. The output gap is the log

difference between real gross domestic product and the CBO measure of potential output.

2.3 Estimation Method

A particle filter will be used to evaluate the likelihood of the model, and the Metropolis-Hastings algorithm used to simulate from the posterior. The highly nonlinear nature of the measurement equation means that traditional estimation methods cannot be used. The Metropolis-Hastings algorithm is a simple algorithm for producing samples from the distribution that otherwise would be hard to characterize.

The MH sampler can be used to take a sample of parameter values to characterize the posterior distribution. The prior beliefs of this model use a gamma distribution to characterize the standard deviation parameters and a normal distribution for α and λ_0 . There are 1,000 burn in draws and 2,000 post convergence draws. Before beginning the sampler, we must choose a θ_0 and the variance-covariance matrix R, used to create variation from the previously accepted value of theta.

The stochastic terms are assumed to be normally distributed. Standard deviations of the stochastic terms in model equations are assumed to have a Gamma distribution with hyperparameters α and β .

$$\sigma_\nu^{-1} \sim \text{Gamma}(\alpha_\nu, \beta_\nu) \quad st \quad \alpha_\nu = 100 \quad \beta_\nu = 4$$

$$\sigma_\omega^{-1} \sim \text{Gamma}(\alpha_\omega, \beta_\omega) \quad st \quad \alpha_\omega = 80 \quad \beta_\omega = 1$$

$$\sigma_\xi^{-1} \sim \text{Gamma}(\alpha_\xi, \beta_\xi) \quad st \quad \alpha_\xi = 45 \quad \beta_\xi = 2$$

All the coefficients are assumed to be drawn from independent normal distributions. Hyperparameters for Normal distribution describing prior for conditional mean parameters:

$$\alpha \sim N(\mu_\alpha, \sigma_\alpha^2) \quad st \quad \mu_\alpha = 0.1 \quad \sigma_\alpha^2 = .001$$

$$\lambda_0 \sim N(\mu_{\lambda_0}, \sigma_{\lambda_0}^2) \quad st \quad \mu_{\lambda_0} = 0.7 \quad \sigma_{\lambda_0}^2 = 1$$

The priors for the conditional mean parameters in the instrumental variable equation are

established in this section. A diffuse prior is employed, wherein the hyperparameters dictate a mean of 0 and a variance of 1. This choice aims to minimize the influence of the prior on our posterior estimates.

$$\delta_i \sim N(\mu_{\delta_i}, \sigma_{\delta_i}^2) \quad st \quad \mu_{\delta_i} = 0 \quad \sigma_{\delta_i}^2 = 1 \quad \forall i$$

The covariance between the instrumental error: ξ and the measurement error: ν is also assumed to be drawn from a normal distribution:

$$\sigma_{\nu, \xi} \sim N(\mu_{\sigma}, \sigma_{\sigma}^2) \quad st \quad \mu_{\sigma} = 0 \quad \sigma_{\sigma}^2 = 1$$

The assessment for weak instruments typically employs an F-test. The leading evaluation for weak instruments in time series models is Olea and Pflueger (2013) approach. Nonetheless, this technique hinges on the linear Ordinary Least Squares (OLS) estimates of the structural equation. An alternative test, Stock and Yogo (2005), has been completed for this model by Coibion (2010) and stands as the next viable option.

3 Results

Estimating the SIPC with variation in the rate of attention gives us parameter estimates and a time series of firms' attentiveness to information relevant to price setting. The Metropolis Hasting sampler drew 2000 samples of our parameters. These samples can be used to understand an estimate of the parameter and the accuracy of the estimates.

3.1 TVP Estimation

The results of modeling time-varying attention to macroeconomic information are as follows. The final acceptance rate of the draws stood at 0.44, surpassing the recommended threshold in the literature, yet remaining within an acceptable range. The estimated values for α and $1 - \lambda_0$ are 0.06 and 0.46, respectively. The ensuing figures illustrate the generated samples. Each sample of the static parameter models gave rise to a corresponding sample path of λ . By aggregating these sample paths using a time-wise mean, a composite graph was constructed. Subsequently, the 95 % confidence bounds per period were determined.

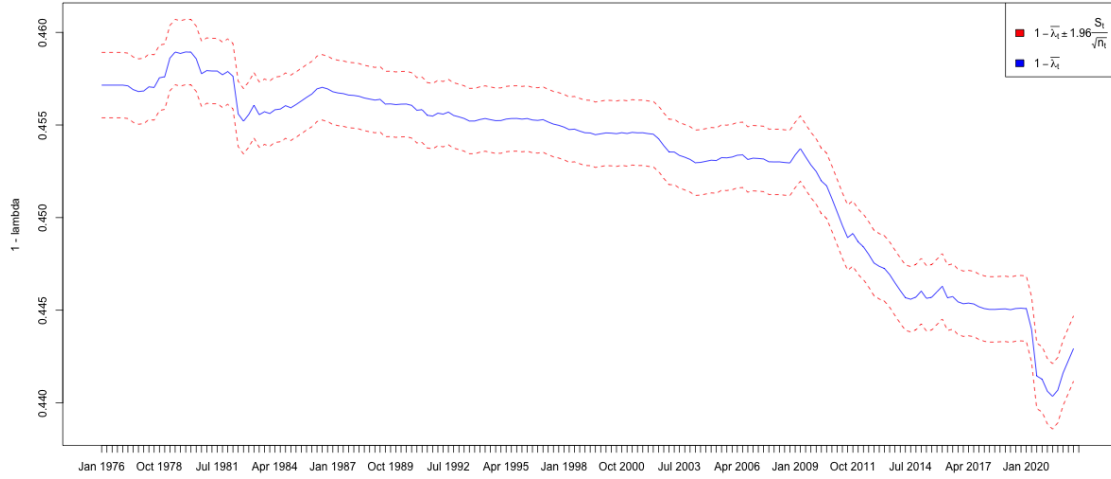


Figure 1: Proportion of firms using the newest information over time.

	Alpha	Lambda0	HV	HW	HE	Covariance
Mean	0.06	0.54	329.01	77.92	90.32	0.00
SD	0.01	0.03	17.88	8.97	5.07	0.00

Table 1: Measure Parameters in TVP Model

	Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
Mean	-0.00	0.81	0.20	-0.14	-0.11
SD	0.00	0.05	0.04	0.02	0.05

Table 2: IV Parameters in TVP Model

The time-varying parameter estimation reveals a consistent trend of declining attention over the analyzed period. Commencing from January 1976, the estimated value of $1 - \lambda$ stands at 0.457, progressively decreasing to 0.443 by April 2022. To gauge the economic implications of this trend, I will refer to the model introduced by Ball et al. (2003).

The changes in attention may be driven by shifts in other parameters. The standard deviation of the error used in the measurement equation could be shifting over time. To test for this possibility, another model could be estimated and compared to the ones presented here.

Prior research conducted by Coibion and Gorodnichenko (2015) found that the frequency of forecasts by professional forecasters surveyed does not change over time. Therefore, the shifts in attention are isolated to the model presented and not the data.

3.2 Economic Implications

To analyze the economic impact of the identified shifts in the rate of attention, I use the model by Ball et al. (2003). A similar model that incorporates a demand and policy equation to form a general equilibrium model. The demand equation is a simple quantity equation in which nominal spending is proportional to the money supply. The policy rule is derived from maximizing the Woodford and Woodford and Walsh (2005) approximation of utility. In the rational expectations equilibrium, the variance of price can be expressed in terms of the model parameters and the rate of attention. The equation for price variance is given here:

$$\sigma_p^2 = Var(\varepsilon_t) \sum_{j=0}^{\infty} \phi_j^2$$

$$\phi_j = \frac{\rho^j}{\alpha^2 \omega + \frac{(1-\lambda)^{j+1}}{1-(1-\lambda)^{j+1}}}$$

Using the parameterization in Branch et al. (2009) , the initial volatility registers at 1.66, tapering down to 1.57 towards the end of the analyzed period. A discernible shift in volatility is observed, a decrease of 5.5% in the variance of price. An alternative parameterization found in Ball et al. (2003) increases the magnitude of the shift to -12.7%. This suggests that policymaking may need adjustments in response to the time-varying estimates of attention. The parameterization taken from that model is $\alpha = 0.1$, $\rho = 0.8$, $\omega = 1$, and $\sigma_\varepsilon^2 = 1$. The parameterization from Branch et al. (2009) is $\alpha = 0.1$, $\rho = 0.8$, $\omega = 20$, and $\sigma_\varepsilon^2 = 0.1$.

3.3 Fixed Parameter Estimation

Fixed estimation is done using a Metropolis Hasting Sampler and a joint likelihood function equivalent to the particle filter with zero variation. This likelihood method was chosen to provide an equivalent measure of the marginal likelihood to compare to the time varying model. The results of modeling fixed attention to macroeconomic information are as follows: the final acceptance rate of the draws stood at 0.33, within the recommended threshold in the literature. The estimated values for α and $1 - \lambda$ are 0.07 and 0.45, respectively. The estimated value of λ implies that most firms are expected to update their information twice a year.

Comparing the results for the two model types using the values $1 - \lambda$, I find that the mean value of the TVP estimate is 0.45, similar to the fixed estimate. The fixed estimate of 0.55 is significantly different to the value found by Coibion (2010), but similar to that found by Reis (2009). A Bayesian model comparison will shed some light on which estimate is best.

	Alpha	Lambda	HV	HE	Covariance
Mean	0.07	0.55	325.01	91.61	0.00
SD	0.01	0.03	17.56	4.55	0.00

Table 3: Measure Parameters in Fixed Model

	Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
Mean	-0.00	0.87	-0.03	0.02	0.03
SD	0.00	0.05	0.05	0.02	0.05

Table 4: IV Parameters in Fixed Model

3.4 Bayesian Model Comparison

Bayesian model comparison can provide the relative plausibility of two models given our data. To calculate the probability of a time-varying attention model, I find the ratio between the marginal likelihood of the TVP model and the total marginal likelihood of both the TVP model and the fixed parameter model. Chib and Jeliazkov (2001) provide the framework for estimating the marginal likelihood when implementing a Markov chain Monte Carlo estimation.

The fundamental model comparison equation is the posterior odds ratio:

$$\frac{p(M_1|Y)}{p(M_2|Y)} = \frac{p(Y|M_1)p(M_1)}{p(Y|M_2)p(M_2)}$$

The probability placed on each model: $p(M_i)$ will be set equal for i equals one and two. The posterior odds ration is simplified to be a function of the marginal likelihood of each model:

$$p(Y|M_i) = \frac{p(Y|\theta)p(\theta)}{p(\theta|Y)}$$

The Metropolis-Hasting Sampler is used to sample from the posterior distribution, creating N samples. I calculate the likelihood function and the prior density at the median of the samples. Then to estimate the posterior ordinate we use the following equation:

$$\widehat{p(\tilde{\theta}|Y)} = \frac{\frac{1}{g} \sum_{g=1}^G \alpha(\theta^{[g]}, \tilde{\theta}) q(\tilde{\theta}|\theta^{[g]})}{\frac{1}{J} \sum_{j=1}^J \alpha(\tilde{\theta}, \theta^{[j]})},$$

in which $\alpha(\theta^{[g]}, \tilde{\theta})$ is the probability of accepting the median value instead of sample g and $q(\tilde{\theta}|\theta^{[g]})$ is the likelihood of the variation between the median and the sample. J draws are taken by random variation from the median, then $\alpha(\tilde{\theta}, \theta^{[j]})$ is the probability of accepting that draw instead of the median. The probabilities are then averaged to get the numerator and

denominator to get and estimate of the ordinate.

A Bayesian model comparison between the fixed model and the TVP model favors the fixed parameter model. The probability placed on the fixed parameter model is 0.84. This analysis favors a model that excludes a time-varying parameter for attention but it cannot be ruled out statistically.

4 Conclusion

In conclusion, the estimation of the sticky information Phillips curve model with variation in the rate of attention provides valuable insights into parameter estimates and the dynamics of firms' attentiveness to macroeconomic information relevant to price setting.

The results from modeling time-varying attention reveal a declining trend in attention over the analyzed period, indicating potential economic implications for inflation volatility. Incorporating models proposed by Ball et al. (2003) suggests that policymakers may need to adjust their strategies in response to these time-varying estimates of attention.

Overall, these findings underscore the importance of considering time-varying attention in economic modeling and policymaking, as it provides valuable insights into the evolving nature of decision-making processes in response to macroeconomic information.

References

- Ball, L., Mankiw, N. G., and Reis, R. (2003). Monetary policy for inattentive economies. Working Paper 9491, National Bureau of Economic Research.
- Branch, W. A., Carlson, J., Evans, G. W., and McGough, B. (2009). Monetary policy, endogenous inattention and the volatility trade-off. *The Economic Journal*, 119(534):123–157.
- Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of Economics*, 118(1):269–298.
- Chib, S. and Jeliazkov, I. (2001). Marginal likelihood from the metropolis-hastings output. *Journal of the American Statistical Association*, 96(453):270–281.
- Coibion, O. (2010). Testing the sticky information phillips curve. *The Review of Economics and Statistics*, 92(1):87–101.
- Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–78.
- Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, 4(2):103–124.
- Mankiw, N. G. and Reis, R. (2002). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve*. *The Quarterly Journal of Economics*, 117(4):1295–1328.
- Olea, J. L. M. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business and Economic Statistics*, 31(3):358–369.
- Reis, R. (2009). A sticky-information general-equilibrium model for policy analysis. Working Paper 14732, National Bureau of Economic Research.
- Sims, C. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.
- Stock, J. and Yogo, M. (2005). *Testing for Weak Instruments in Linear IV Regression*, pages 80–108. Cambridge University Press, New York.
- Woodford, M. and Walsh, C. E. (2005). Interest and prices: Foundations of a theory of monetary policy. *Macroeconomic Dynamics*, 9(3):462–468.