Universidade de São Paulo
Faculdade de Economia, Administração, Contabilidade e Atuária
Departamento de Economia
Programa de Pós-Graduação em Economia

Henrique Danyi da Silveira Correia

STICKY-INFORMATION PHILLIPS CURVE WITH TIME-VARYING INATTENTION

CURVA DE PHILLIPS DE INFORMAÇÃO RÍGIDA COM INATENÇÃO VARIANTE NO TEMPO

São Paulo

2023

Henrique Danyi da Silveira Correia

STICKY-INFORMATION PHILLIPS CURVE WITH TIME-VARYING INATTENTION

Dissertação apresentada ao programa de Pós-Graduação em Economia do Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, como requisito parcial para obtenção do título de Mestre em Ciências.

Advisor: Rodrigo de Losso da Silveira Bueno
Co-Advisor: Prof. Ricardo Dias de Oliveira Brito

Versão corrigida

São Paulo

2023

Correia, Henrique Danyi da Silveira
STICKY-INFORMATION PHILLIPS CURVE WITH TIME-VARYING INATTENTION/ Henrique Danyi da Silveira Correia. – São Paulo, 2023-
55 p.

Dissertação (Mestrado) - Universidade de São Paulo, 2023.
Orientador: Rodrigo de Losso da Silveira Bueno
Co-orientador: Ricardo Dias de Oliveira Brito

1. Curva de Phillips. 2. Inatenção. 3. Informação Rígida. I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

ACKNOWLEDGMENTS

To my mother, father, and sister, Márcia, Jorge, and Juliana, for their unending and unconditional love and support. I am deeply indebted to this family and will be forever grateful for their company and inspiration.

To my aunt and uncle, Marina and Mauro, who propelled me towards my aspirations.

To my partner, Beatriz, whose love and care gave me the encouragement and helped me face the challenges of the recent years. This endeavor would not have been possible without her.

To my friends Amanda, Bruno, Carolina, Gabriela, Pedro, and many others who are now fellow economists, I am grateful for the journey so far. Regardless of what lies ahead, I am thankful for the wonderful and insightful people who have come into my life.

To my advisor and my co-advisor, Rodrigo Bueno and Ricardo Brito. I was very lucky to have such a great pair to guide me in this journey and could not have undertaken it without them.

This study was financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

"Adaptation seems to be, to a substantial extent,
a process of reallocating your attention."
Daniel Kahneman

"Time changes everything
except something within us
which is always surprised by change."
Thomas Hardy

ABSTRACT

Inflation persistence is a pivotal topic to macroeconomic policy discussions. The Sticky Information Phillips Curve (SIPC) incorporates inflation persistence in a microfounded approach as a result of inattention from firms. Previous empirical analyses of the SIPC estimated inattention as constant in time, but this need not be the case. We review the canonical model of the SIPC to allow for time-varying inattention and present empirical evidence to support it. We also find that inattention in the SIPC is heavily associated with inflation volatility, as predicted by the Reis (2006a).

Key words: Sticky Information, Time-Varying Inattention, Inflation Volatility.

JEL Classification: D8, E1, E3, D84, E31

RESUMO

A persistência inflacionária é um tópico de grande relevância para as discussões de política macroeconômica. A Curva de Phillips de Informação Rígida (SIPC) incorpora a persistência inflacionária de maneira microfundamentada como resultado de inatenção por parte das firmas. Análises empíricas anteriores da SIPC estimaram a inatenção como constante no tempo, mas esse não é necessariamente o caso. Nós revisamos o modelo canônico da SIPC para permitir uma inatenção variante no tempo e apresentamos evidência empírica que apoia essa característica. Nós também encontramos que a inatenção no modelo da SIPC está fortemente associada à volatilidade da inflação, como previsto por Reis (2006a).

Palavras-chave: Informação Rígida, Inatenção Variante no Tempo, Volatilidade da Inflação.

Código JEL: D8, E1, E3, D84, E31

SUMMARY

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Inflation persistence is a pivotal topic to macroeconomic policy discussions (see, e.g., Woodford (2007) and Sbordone (2007)). The degree to which inflation expectations are anchored to inflation history can greatly alter the form of the optimal policy rule of the monetary authority. One of the proposals that incorporates inflation persistence in a macroeconomic model with microfoundations is the *Sticky Information Phillips Curve* (SIPC) from Mankiw and Reis (2002), which found it to be a consequence of slow information diffusion, or inattention.

On the one hand, Granville and Zeng (2019) presented evidence of time variation in inflation persistence. On the other, the state-dependent nature of inattention was well-documented in the rational inattention literature (see, e.g., Maćkowiak and Wiederholt (2009) and Maćkowiak Matějka and Wiederholt (2021)) and in Sticky-Information models (see Reis (2006a)). Yet, previous empirical analyses of the SIPC estimated the degree of inattention as constant in time (see, e.g., Khan and Zhu (2006), D̈opke et al. (2008), Gillitzer (2016), and Hung and Kwan (2022)).

This paper presents evidence of time variation in inattention using the SIPC model and rewrites it to account for this feature. This new model is an extension of the canonical SIPC of Mankiw and Reis (2002) with time-varying parameters. In contrast to the original model, this extension allows for varying degrees of inflation persistence in time and is also able to empirically verify the implications of Reis (2006a) on inattention's determinants.

Following Lucas' (1976) Critique: it is unlikely that agents' reactions are invariant to changes in policy regimes; thus, more generally, one should not expect the degree of inattention to remain constant in time. As in Reis (2006a), inattention must be determined by the marginal costs and benefits of information gathering, which need not be constant in time. To clarify this, we present two examples below:

- Changes in the dynamics of variables relevant for the pricesetting procedure can alter the benefits associated with information updates, e.g., assuming strategic complementarities, a lower and steadier inflation would imply in higher inattention than otherwise.

- Advances in information technology in the last decades have reduced the costs of accessing information, which, *ceteris paribus*, should decrease the degree of inattention. (REIS, 2006a, p. 812).

We estimate what would be the average inattention at different periods in time through nonlinear regressions for the US. This approach presents evidence of time-varying average inattention, thus supporting that inattention itself cannot be considered time-invariant.

Together with such evidence, we find that the dynamics of average inattention are strongly associated with inflation volatility, as predicted in Reis (2006a). Our first empirical approach divides the whole sample into sub-samples according to macroeconomic regimes and in quinquennia. Here, we find that inattention is significantly lower in periods with high inflation volatility, such as the Great Inflation period or the Great Financial Crisis (GFC) period, when compared to the inattention of the Great Moderation or Post-GFC periods. In further detail, Table 1 below summarises our estimates for macroeconomic regimes. Next, we estimate 10-year rolling regressions, presenting what can be interpreted as a 10-year moving average of the degree of inattention in the SIPC. The results validate our findings from the first exercise and reinforce what was predicted by the literature: a time-varying estimate of inattention is heavily associated with a time-varying estimate of inflation volatility.

Table 1 – Median Estimates and 95% Confidence Interval of Inattention in Macroeconomic Regimes.

Regime	Median Estimate (95% Confidence Interval)
Great Inflation	0.5156 (0.4110,0.5963)
Great Moderation	0.2306 (0.1410,0.3864)
Great Financial Crisis	0.8279 (0.0284,0.8775)
Post-Crisis	0.1031 (0.0635,0.1680)

This thesis is structured as follows. In the next chapter, we present a literature review covering the relevant developments on sticky information and inattentiveness. In the Contribution and Methodology chapter, we propose a new time-varying-attention SIPC and delve into the details of our estimation process. The Results chapter presents our estimates. Finally, chapter 5 concludes.

2 THEORETICAL FOUNDATIONS AND PREVIOUS ESTIMATES

Since its emergence, the macroeconomic literature on imperfect information resulting from optimal behavior has been developed in two categories: rational-inattention models and sticky-information models. Overall, their motivation is the same: the introduction of a source of rigidity in the economy brought by the limitation of individuals in acquiring and processing information. In this section, we will review in further detail the literature on sticky-information models.

We will start by revisiting the canonical *Sticky-Information Phillips Curve* (SIPC) model from Mankiw and Reis (2002), discussing its buildup and overall advantages over the *New Keynesian Phillips Curve* (NKPC) in section 2.1. Section 2.2 revisits Reis (2006a) microfounded approach to the SIPC where we can first take note on the determinants of inattention. Section 2.3 covers a few contributions from the rational-inattention literature which add to the discussion of inattention determinants. Finally, section 2.4 reviews the empirical literature on the matter, focusing on estimations of the inattention parameter in the SIPC model.

2.1 Mankiw and Reis (2002)

Mankiw and Reis (2002) introduced a model of optimal pricesetting that forgoes the sticky-prices hypothesis from the NKPC and takes into account the costs and benefits of acquiring and/or processing information thus being able to include slow information diffusion in the economy: the sticky-information model.

In the original proposal, Mankiw and Reis (2002) advocate the substitution of the sticky-prices tradition for the sticky-information model based on closer resemblance of its predictions to actual data and no loss of microfoundations. There are strong arguments in favor of this approach. First and foremost, as emphasized by Mankiw and Reis (2010), when trying to adapt the sticky-prices models to account for trend inflation and to avoid inflation jumps in response to news today about future circumstances—features empirically verified—, the literature assumed backward-looking behavior in ways that contrast micro data, renouncing 'the enterprise of microfoundations'. Additionally, the SIPC is in line with a number of stylized facts on inflation dynamics which the NKPC alone is not:

1. The full impact of monetary policy changes on inflation is lagged by an amount of six to eight quarters;

2. Inflation exhibits high serial correlation;

3. Disinflation policies have contractionary effects; and

4. The maximum effect of monetary policy actions on cyclical output occurs before its maximum effect on inflation.

They implement their idea in a similar fashion to that of Calvo's (1983) sticky-prices method: In every period, instead of a fraction of firms being randomly selected to optimally adjust their prices, in Mankiw and Reis (2002) a fraction of firms are randomly selected to update their information set and all firms are allowed to optimally choose their prices in all periods based on their information set. As a result, in each period, some firms will have their prices optimally set based on current information, but others will have their prices optimally set based on previous expectations of current conditions. The result is what Mankiw and Reis (2002) coined as the *Sticky Information Phillips Curve* (SIPC).

As in the sticky prices case, the key parameter in such an environment would be the fraction of firms that update their information set in one period, which Mankiw and Reis (2002) define as λ. Unsurprisingly, λ is restricted to values between 0 and 1 and can also be interpreted as the information diffusion or information stickiness parameter and hence is a particular estimate of the degree of attention of an economy. Since the original proposal, a number of estimates for λ emerged and will be covered in section 2.4.

In Appendix A, we revisit the step-by-step procedure used by Mankiw and Reis (2002) to develop the SIPC. The resulting Phillips Curve, as shown there, is:

$$\pi_t = \frac{\lambda\alpha}{1-\lambda} y_t + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-1-j} \left[\pi_t + \alpha \Delta y_t \right], \Delta y_t = y_t - y_{t-1}, \tag{1}$$

where π_t is the inflation rate and y_t is the output gap.

Hence, in Mankiw and Reis' (2002) canonical model, the kind of inflation and output gap expectations that determine the short-run Phillips curve are past expectations of current economic conditions, rather than current expectations regarding future inflation alone as in the NKPC. This modifies the dynamic behavior of prices and output in response to monetary policy, in particular creating a lagged, inertial response of inflation to policy changes, as observed in data and which the standard sticky-price model is unable to deliver without the introduction of further frictions.

Further insights are possible in equation 1. Assuming a Poisson distribution for the arrival of information update decision periods[1], one can interpret the inverse of λ as the average duration of information stickiness, i.e., $D = \frac{1}{\lambda}$. In words, it takes on average $\frac{1}{\lambda}$ periods for a firm to update its information set and revise its optimal price trajectory for the future accordingly in this environment. Thus, from an sticky-information point of view, $\frac{1}{\lambda}$ is the degree of inattention or information rigidity in an economy.

[1] This assumption was later proved to be verified under more general conditions by Reis (2006a).

2.2 Reis (2006a)

Since its original proposal, the sticky-information model has received much attention and has been duly extended. Reis (2006b) considered the equivalent of the sticky-information pricesetting problem for the consumers and was able to provide an explanation for the excess sensitivity and excess smoothness puzzles. Alternatively, Reis (2006a) pursued the endeavor of microfoundations for the SIPC model and presented its continuous-time equivalent. Reis' (2006a) microfounded version of the SIPC will inherently expose the main determinants of inattentiveness. The importance of this should be clear since we can easily infer that if the determinants of inattention are not constant, there should be no *a priori* reason to expect attention itself to be constant. In this subsection, we will thus focus on Reis (2006a).

Reis (2006a) indentified two major factors that should define the degree of attention in the SIPC model. First, he found that the larger the costs of planning are, the longer is the duration of inattentiveness. Second, he found that the faster the losses from being inattentive accumulate, the shorter is the duration of inattentiveness. Since these findings come from a microfounded model, both are thus associated with the net marginal benefits of information from the firms perspective:

1. '*How much will I benefit if I bear the cost of updating my information set?*' - This question is directly associated with the costs associated with acquiring and processing information;

2. '*How much will updating information set facilitate my predictability capacity? Is it really that necessary?*' - These questions are directly associated with the volatility in supply and demand conditions (staying inattentive is more costly in a world that is rapidly changing).

Note that neither the costs of acquiring and processing information nor overall economic volatility are necessarily time-invariant.

Reis (2006a) then goes even further and elaborates the proposition below:

Proposition 1 *Reis (2006a) proves that under certain conditions, λ is proportional to the volatility of the output gap (and hence inflation), and inversely proportional to the cost of information.*

To prove Proposition 1, we present a simplified version of Reis' (2006a) analysis below.

In Reis' (2006a), producers are inattentive to a set of state variables indexed by time s_t. Time is continuous and firms choose the set of information-update periods (D) in order to maximize:

$$J(s_0, D) = E\left\{\sum_{i=0}^{\infty}\left(\int_{D_i}^{D_{i+1}} e^{-rt}\Pi\left(s_{D_i}, t - D_i\right)dt - e^{-rD_{i+1}}K\left(s_{D_{i+1}}\right)\right)\right\},$$

which is the sum of the present value of the profit function $\Pi(\cdot)$ (continuous in all arguments) between update periods D_i and D_{i+1} for all i, subtracted by the present value of the cost function associated with information updates $K(\cdot)$ (continuous in all arguments), evaluated in the future periods of information update. Recursively, this problem can be rewritten as:

$$V(s) = \sup_d\left\{\int_0^d e^{-rt}\Pi(s,t)dt + e^{-rd}E\left[-K\left(s_d\right) + V\left(s_d\right)\right],\ \text{s.t. } s_d = \Psi\left(s, u^d\right)\right\},$$

where d is the time length between information-update periods and, thus, the control variable.

Reis (2006a) then proceeds by introducing the function $G(s,t)$ (continuous in all arguments), which is the expected difference between profits earned with full information and profits earned while following a pre-chosen plan. He also defines $K(s_t) = \kappa^2\tilde{K}(s_t)$ and rewrites the problem as:

$$V(s) = \max_d\left\{-\int_0^d e^{-rt}G(s,t)dt + e^{-rd}E\left[-\kappa^2\tilde{K}(s_t) + V\left(s_d\right)\right],\ \text{s.t. } s_d = \Psi\left(s, u^d\right)\right\}$$

The reason advocated for the form of $K(s_t)$ is to facilitate the evaluation of the system of equations resulting from the maximum principle around $\kappa = 0$ since "analytically, in general, the optimal inattentiveness is a complicated function of the state of the economy". (REIS, 2006a, p. 802).

Optimality conditions are:

$$-G(s,d) + r\kappa^2\mathbb{E}[\tilde{K}(s_d) = \mathbb{E}\left[rV(s_d) + (\kappa^2\tilde{K}_s(s_d) - V_s(s_d))\Psi(s,u^d)d\right]$$

$$V_j(s) = -\int_0^d e^{-rt}G_j(s,t)dt + e^{-rd}\mathbb{E}[(-\kappa^2\tilde{K}_s(s_d) + V_s(s_d))\Psi_j(s,u^d)]$$

$$V_\kappa(s) = e^{-rd}\mathbb{E}[-2\kappa\tilde{K}(s_d) + V_\kappa(s_d)]$$

These, together with the value function, define the optimum. As anticipated, the analytical solution for inattentiveness is non-trivial, so he perturbs these equations around the point where the costs of planning are 0 ($\kappa = 0$). In such environment, $d^* = 0$, $V(s) = 0$, $V_s(s) = 0$ $G(s) = 0$, and $G_j(s) = 0$. The system is:

$$\begin{aligned} V_\kappa &= V_\kappa \\ -G_t d_\kappa &= rV_\kappa - \frac{d}{d\kappa}\left[\frac{1}{dt}\mathbb{E}(dV)\right] d_\kappa \\ V_{j\kappa} &= V_{s\kappa}\Psi_j \\ 0 &= -2\tilde{K} - rV_\kappa d_\kappa + \frac{d}{d\kappa}\left[\frac{1}{dt}\mathbb{E}(dV)\right] d_\kappa \end{aligned}$$

Combining the second and third equations:

$$d_\kappa = \sqrt{\frac{2\tilde{K}}{G_t}} \qquad \Longrightarrow \qquad d^* \approx d_\kappa\sqrt{\kappa} = \sqrt{\frac{2K(s)}{G_t(s,0)}}$$

Reis (2006a) proceeds to make a second-order approximation of the difference between profits with full information and profits earned while following a pre-chosen plan. He finds that (y is the output gap):

$$G(s,t) = -\frac{\pi_{pp}\alpha^2\mathbb{E}[Var(y_t)]}{2}$$

Reis (2006a) thus finds that:

$$d^* = \frac{2}{\alpha}\sqrt{\frac{K(s)}{-\pi_{pp}\sigma_y^2}}$$

,

Reis (2006a) then proves that the only stationary equilibrium distribution of inattentiveness (d) is the exponential distribution with parameter λ. This is the same λ that Markiw and Reis (2002) worked with.

Remember that, if $d \sim Exp(\lambda)$, $\mathbb{E}[d] = \frac{1}{\lambda}$. Therefore, Reis (2006a) proves that under certain conditions, λ is proportional to the volatility of the output gap (and hence inflation) and inversely proportional to the cost of information.

Proposition 1 can thus provide us with incentive for our objective. There is no reason to expect the cost of information and inflation volatility to be constant in time. Hence, there is no reason to expect the inattention in the SIPC model to be constant in time. It also provides us with a direction to look for evidence of time-varying attention in the SIPC model which we will use in the estimation exercises of section 4.1.1.

2.3 Other Contributions on Inattention Determinants

Other researchers used different models to study parameters similar or equivalent to that of the SIPC canonical model. In this subsection, we analyse two alternative approaches.

Begg and Imperato's (2001) monopoly model included the decision of thinking in a scenario where thinking is costly but provides benefits in the form of improved forecasts (and higher profits). The implication is that optimal information updates occur only at discrete points in time and the length between these periods—in other words, the degree of inattention, equivalent to our D or d—

> *"(...) depends on the cost of decision-making itself, the variance of the stochastic variable the agent is trying to forecast, its speed of convergence towards its long-run mean, the discount rate which applies to future mistakes, and the effect of forecast errors on the loss function" (BEGG; IMPERATO, 2001, p. 250).*

Therefore, Begg and Imperato's (2001) model also provides insights into the dynamics of attention. We explore their model in further detail in Appendix B.

As mentioned in the beginning of this section, the macroeconomic literature on imperfect information developed in another category different from the sticky-information models: the rational inattention framework. The rational inattention literature started with Woodford (2001) and Sims (2003). In this type of model, agents continuously update their information sets but are never capable of observing the true state of the economy. To form and update beliefs, agents enact in a signal extraction problem, weighting prior beliefs and the new information received. In this structure, the weights to prior beliefs are, thus, the degree of information stickiness. This was followed by several further developments. Most notably, Sims (2010) introduced rational inattention into dynamic optimization problems and analyzed its implications for monetary policy. Later, Hébert and Woodford (2017, 2021) associated rational inattention models with the problem of optimal dynamic evidence accumulation. Maćowiak, Matějka and Wiederhold (2021) provide a summary of this literature.

In the rational inattention framework, Maćkowiak and Wiederholt (2009) were responsible for anticipating the discussion on state-dependent attention as well. They developed a model building on Sims' (2003) proposal and had an interesting finding for our research. In their model, firms face a trade-off between paying attention to aggregate conditions and paying attention to idiosyncratic conditions. Their calibration for the US implies that firms allocate almost all attention to idiosyncratic conditions. But, "when the variance of nominal aggregate demand increases, firms shift attention toward aggregate conditions and away from idiosyncratic conditions" (MAĆKOWIAK; WIEDERHOLT, 2009, p. 770).

Thus, Maćkowiak and Wiederholt (2009) found that attention to aggregate conditions is associated to the variance of nominal aggregate demand. Later, Maćkowiak, Matějka, and Wiederholt (2021) associated attention dynamics to inflation dynamics: "In a rational inattention model, tracking the evolving state of the economy requires scarce attention [...] When inflation is high and volatile, paying attention to monetary policy becomes more important to agents. "(MAĆKOWIAK; MATĚJKA; WIEDERHOLT, 2021, p. 1-2). In Coibion and Gorodnichenko (2015) these results were coined as state-dependent attention. Coibion and Gorodnichenko (2015) also provided empirical evidence showing that price-setters increase their attention levels as the economy migrates from low to high volatility states.

In short, previous research have generally associated attention to aggregate variables with the volatility of these variables and the cost of accessing/processing the information regarding these variables.

When comparing their model to that of Mankiw and Reis (2002) and to Reis (2006a), Maćkowiak and Wiederholt (2009) argue that, in the sticky-information framework, prices respond with equal speed to all disturbances, while in their model, prices respond quickly to some shocks and slowly to others. However, note that this is only true if one assumes λ be constant. If one does not make such assumption, then, λ becomes a measure of overall attention to aggregate conditions at a specific point in time. Prices will then not respond with equal speed to all disturbances every time; this will depend on the λ in each period, that is, on the degree of attention to aggregate variables in each period.

2.4 Empirical Time-Invariant Attention Estimates

Since Mankiw and Reis' (2002) original proposal, given the desirable features of the model, a substantial empirical literature was developed. A number of estimates for the parameter emerged, in particular for the US economy. In this section, we cover some of the research of this literature.

First, Carroll (2003) used Michigan's Survey of Consumers and the Survey of Professional Forecasters data from the third quarter of 1981 to the second quarter of 2000 to estimate the inattentiveness parameter through a very simple model which he considers to be an alternative version of Mankiw and Reis' (2002) SIPC model with microfoundations. His model is structured as follows: households gather information from professional forecasters' news reports to only then form their expectations; this imply in some inattention which, in turn, implies in aggregate-expectations stickiness. In his estimations, Carroll (2003) finds a value of 0.27 for λ and also concludes that the dynamics of expectations are well captured by a his model. The estimates of λ in this paper are very similar to the value assumed by Mankiw and Reis (2002): a quarter of firms adjust their prices in each period on average.

Departing from Mankiw and Reis (2002), Khan and Zhu (2006) develop an estimation process different from Carroll's (2003) using US data from the first quarter of 1969 to the last quarter of 2000 based on

$$\pi_t = \frac{\lambda\alpha}{1-\lambda} y_t + \lambda \sum_{j=1}^{T} (1-\lambda)^{j-1} E_{t-j} \left[\pi_t + \alpha \Delta y_t\right]. \quad (2)$$

The only difference from this equation to equation 1 is the inclusion of a truncation point in the summation index, which is necessary for estimation purposes. Khan and Zhu (2006) set $T = 20$ based on sensitivity analysis. Due to the need of data on past expectations of current inflation and the output gap, they compare two different approaches for out-of-sample forecasts:

1. Simple autoregressive models with lags defined based on the smallest root-of-mean-squared forecasting error; and

2. VAR-based methodology that uses other economic variables to increase forecasting power.

Finally, they estimate bootstrapped confidence intervals to avoid the "generated regressors" problem. Khan and Zhu (2006) deliver estimates for average duration of information stickiness in the US that range from three ($\lambda = 0.3$) to seven quarters ($\lambda = 0.14$).

For reference purposes, we replicate Khan and Zhu (2006) estimations using our data (described in detail in section 3.1). First, we restrict our sample to match theirs (1Q1980-4Q2000) and obtain a similar point estimate to theirs, 0.211. Next, we use the entire available sample (1Q1960-4Q2019) and obtain a point estimate of 0.3656, a considerably higher result. This result supports our hypothesis that λ should not be treated as constant. The difference between these estimates suggests that λ varies over time.

Other estimates for the inattentiveness parameter in the SIPC were made: Döpke, et al. (2008) estimated the SIPC for Germany, France, Italy, and the United Kingdom; Gillitzer (2016) for Australia; Reid and Du Rand (2015) for South Africa; and more recently, Hung and Kwan (2022) for Hong Kong. But, to the best of our knowledge, no empirical exercise considered the potential time-varying characteristic of this parameter highlighted in sections 2.2 and 2.3.

3 CONTRIBUTION AND METHODOLOGY

Our first contribution is to review Mankiw and Reis' (2002) model and adapt it to allow for changes in λ through time. To do this, we index time with t, and assume that history starts at time t = 0. Through a series of algebraic transformations available in Appendix C, we reach the equations presented below:

$$\pi_t = \left(\frac{\lambda_t}{1-\lambda_t}\right)\alpha y_t + \lambda_{t-1}E_{t-1}\left[\Delta p_t^o\right] + \sum_{i=1}^{t-2}\lambda_{t-i-1}E_{t-i-1}\left[\Delta p_t^o\right]\prod_{j=1}^{i}\left(1-\lambda_{t-j}\right) + E_0\left[\Delta p_t^o\right]\prod_{j=1}^{t-1}\left(1-\lambda_{t-j}\right); \text{ and} \quad (3)$$

$$\pi_t = \lambda_t\left(\pi_t + \alpha y_t\right) + \sum_{i=0}^{t-2}\lambda_{t-i-1}E_{t-i-1}\left[\Delta p_t^o\right]\prod_{j=0}^{i}\left(1-\lambda_{t-j}\right) + E_0\left[\Delta p_t^o\right]\prod_{j=0}^{t-1}\left(1-\lambda_{t-j}\right). \quad (4)$$

Ideally, one would approach the dynamic estimation of λ_t using either equations 3 or 4. Alas, the parametrization of both equations are non-trivial, since they are non-Markovian and nonlinear in the λ coefficient. Due to lack of conditional independence of observations in both equations 3 and 4, no state-space models can be applied to estimate the time-varying parameter. To circumvent this issue and still fulfill our objective, we return to equation 2 and explore alternative approaches to the estimation process using standard Nonlinear Least Squares (NLS) regressions.

In fact, with an additional hypothesis, we can make a small alteration in equation 2 which reduces the number of possible solutions associated with it. Note that by truncating equation 2 at T, we have the possibility to slightly alter it. The modified version of it is exposed below.

$$\pi_t = \frac{\lambda\alpha}{1-\lambda}y_t + \lambda\sum_{j=1}^{T-1}(1-\lambda)^{j-1}E_{t-j}\left[\pi_t + \alpha\Delta y_t\right] + (1-\lambda)^{T-1}E_{t-T}\left[\pi_t + \alpha\Delta y_t\right] \quad (5)$$

This modification is possible if one assumes that past expectations of current conditions converge if made in every period before $t-T$. We can then eliminate the multiplying λ in the last term[2]. To check if this is empirically acceptable, we estimate both equations 2 and 5[3]. For brevity, we report only the estimated results from equation 5; no significant disparities emerges from both estimations and we can thus safely affirm that using equation 5 is acceptable.

[2] Appendix E explains this modification in further detail and walks through the necessary algebra.

[3] In our estimations, we have calibrated $\alpha = 0.1$, as Mankiw and Reis (2002) and Khan and Zhu (2006). We have replicated all estimations for different values of α, but still assuming strategic complementarities ($\alpha < 1$), and found no significant changes in the final result.

3.1 Data

As one can already notice, we need data for inflation and the output gap. We use quarterly data covering the period from 2Q1947 to 4Q2019 made available by the Federal Reserve Bank of Saint Louis through their database 'Federal Reserve Economic Data' (FRED). Specifically, we will use GDP data from the Bureau of Economic Analysis (BEA) and CPI data from the Bureau of Labor Statistics (BLS)[4].

Regarding the expectations component, we will use forecasts generated by employing simple 8-year rolling AR(4) estimations for all variables. That means that in each period, the estimations of λ is fed with current data and past forecasts for the current period.

At this stage, it is important to include a disclaimer. While we are using data from 2Q1947 to 4Q2019, it should be noted that our estimates do not cover this entire period. This is because we generate expectations using this same data, which means that we lose some initial observations before we start estimating λ.

We loose 32 quarters to start generating forecasts using the 8-years rolling AR(4). As mentioned previously, we will assume a truncation in the summation of the equation, stopping at T. Sadly, the higher the T, the less available data we have for the expectations necessary in the estimation exercise. After a sensibility analysis for T[5], we found that $T = 20$ is conservative enough for convergence. Thus, we loose a total of 51 quarters to have enough data for the first estimation, making estimations possible for 1Q1960 to 4Q2019.

3.2 Median and Confidence Interval Estimation

As we feed estimated forecasts as expectations data for the estimation of λ as in Khan and Zhu (2006), we suffer from the generated regressors problem pointed by Pagan (1984). Therefore, standard errors and confidence intervals (CIs) are incorrect. To address this issue, one would refer to bootstrap methods.

We have adapted an algorithm from the bootstrapping method used in Khan and Zhu (2006) and have taken inspiration from the technique applied in Hansen (1999). Our approach involves simulating 1000 possible expectations trajectories using the expectations generating process. For each simulation, we estimate the relevant λ using NLS regressions. We then run a kernel regression on these λ's to estimate the distribution of the estimated λ's. Finally, we select the 50th, 2.5th, and 97.5th quantiles to obtain the median estimate and the 95% confidence interval (CI) of this distribution, respectively.

Our method summarizes several possible values of λ, taking into account the

[4] Output gap is estimated using an HP-Filter with smoothing parameter set to 1600, as is done for quarterly data in Hodrick Prescott (1981).

[5] All estimations were replicated for $T = 12$ and $T = 16$.

frequency of each possibility. As a result, it contains sufficient information on the possible values for the estimated λ.

We chose not to use standard bootstrap procedures due to the possibility of multiple solutions associated with the non-linearity of the reference equation. Standard bootstrapping procedures result in excessively wide intervals due to this non-linearity. Despite not being a standard bootstrap procedure, our algorithm results in fairly wide CIs in some cases, but they are considerably smaller than the CIs of standard bootstrap methods. For the sake of completeness, we provide the standard bootstrap median and confidence intervals in Appendix D.

4 RESULTS

This chapter is divided into 3 sections. In sections 4.1 and 4.2, we explore 2 alternative approaches to the estimation process, sub-sampling and rolling regressions respectively. Here, we are able to provide evidence of time-varying attention in the SIPC model and identify a major determinant in such a parameter. Finally, section 4.3 closes chapter 4, covering some caveats associated with our results and future steps to be developed on this topic.

4.1 Sub-sampled Estimates

Our first objective is to check for evidence of variation in λ in time. To do this, we run NLS estimations of equations 2 and 5, but now using selected sub-samples.

4.1.1 Macroeconomic Regimes

The first selection criteria for sub-samples we apply is to differentiate periods according to their macroeconomic regimes: the Great Inflation, the Great Moderation, the Great Financial Crisis, and the post-Crisis periods. Sadly, there is little consensus on which specific years/quarters that distinguish these periods. For the sake of brevity, we adopt the division used by the Federal Reserve History website from the Federal Reserve Bank of St.Louis, summarized in table 2.

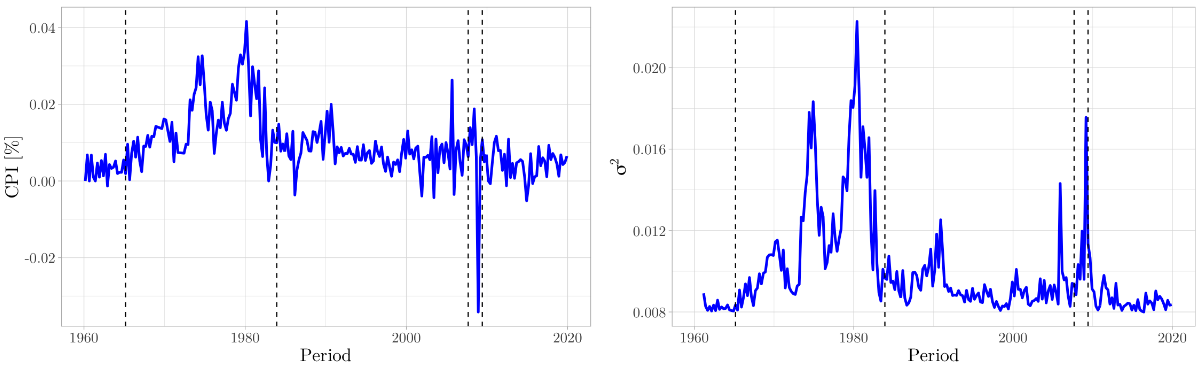
Table 2 – Classification of Periods in Macroeconomic Regimes.

Regime	Period
Great Inflation	1965-1983
Great Moderation	1984-2007Q2
Great Financial Crisis	2007Q3-2009Q2
Post-Crisis	2009Q3-

Fortunately, there is overall consensus regarding the characteristics of these periods in the US. The Great Inflation was characterised by high and volatile inflation. Next, the Great Moderation period stabilized inflation at lower levels. The Great Financial Crisis brought high uncertainty and volatility, but the years thereafter, until 2019, brought back the stability with low inflation. These characteristics can be seen in Figure 1, which present quarterly CPI and its fitted values of a GARCH(4,1) model.

Figure 1 – US CPI Data divided by Macroeconomic Regimes.

Source: own elaboration, BLS.



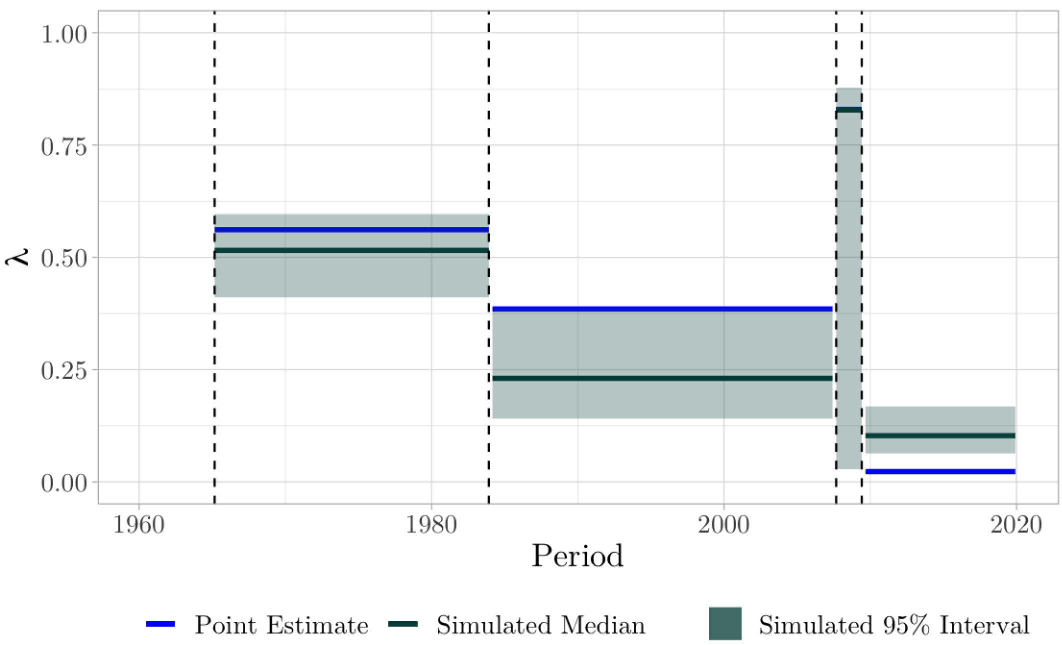
(a) Quarterly CPI. (b) Quarterly CPI Fitted GARCH(4,1).

Following the predictions made in the Literature Review section, in particular from Proposition 1, inflation volatility should a major determinant of the attention levels in the SIPC model. Hence, these contrasting characteristics between these periods can serve as a starting point in checking whether or not there is dynamics in λ.

In Figure 2, we present the point estimates along with the simulated confidence intervals and median estimates following the procedure detailed in section 3.2.

Figure 2 – Point Estimate of Average λ for Macroeconomic Regimes and Simulated Confidence Interval and Median.

Source: own elaboration, BEA, BLS.



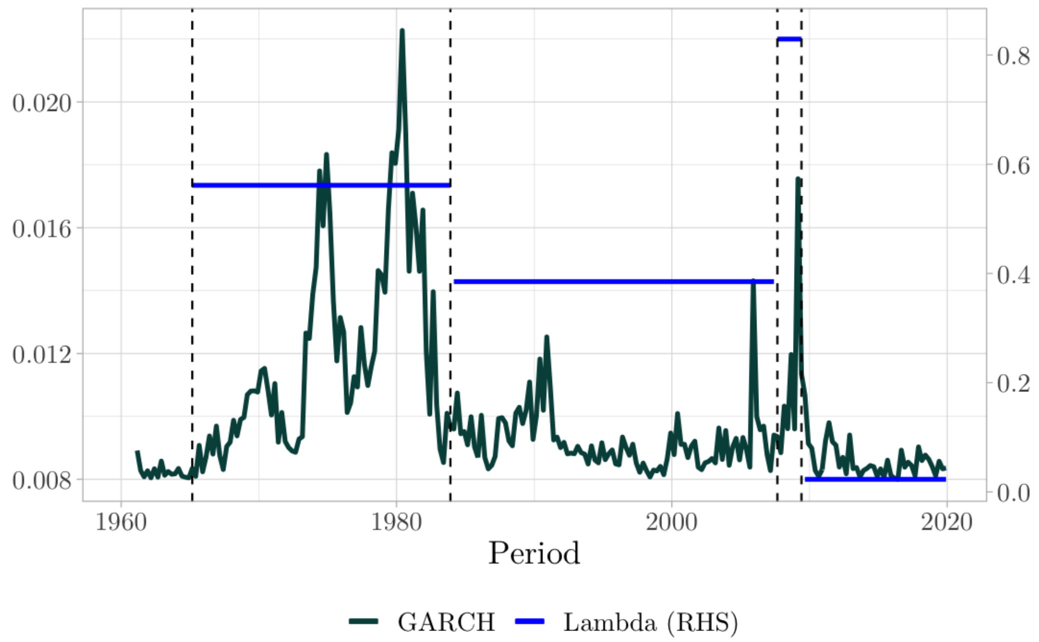
Our estimates reveal a significant finding that supports the main hypothesis of this project: attention levels in the SIPC model should not be treated as constant. Specifically, our results show that the average λ estimate for the entire sample can differ significantly from that of sub-samples within the same sample, as seen in section 4.1.1. This holds true even when considering multiple trajectories of the relevant expectations, as indicated by the presented confidence intervals.

An undesirable result in Figure 2 is the wide size of the confidence interval during the GFC period. Note that equation 5 can have up to 19 possible solutions. We believe that the wide confidence interval is due to this characteristic combined with the high volatility of the period. The slight perturbation made in the expectations during the bootstrap procedure may have caused another root within the (0,1) interval to have lower squared errors, leading to the exaggerated size of the confidence interval.

Next, we present the same point estimates as in Figure 2 combined with the fitted GARCH(4,1) from Figure 1.

Figure 3 – Estimated Average λ for Macroeconomic Regimes and Inflation Fitted GARCH(4,1)

Source: own elaboration, BEA, BLS.



In Figure 3, we can see that the time-variation of the average λ is heavily associated with inflation variance, in line with the analyzed predictions. The Great Moderation period was characterised by a smaller λ than its predecessor. The Great Financial Crisis presented an considerable increase in inflation volatility and, likewise, a much higher average λ compared to the period that followed. This is consistent with Granville an Zeng's (2019) results: Using the SIPC model, we have also found a decline in the degree of inflation persistence during the periods following the Great Inflation.

The intuition behind these results is straightforward:

- In times of higher inflation, it is costly to stay inattentive due to the strategic complementarity in pricing decisions.

- In times of higher inflation volatility, forecasting inflation is harder. Hence, a higher frequency of information updates (attention) is desirable.

4.1.2 5-year intervals

As a complement to our previous exercise, we now break down the whole sample into quinquennia and present the results in Figures 4 and 5.

Figure 4 – Point Estimate of Average λ for Quinquennia and Simulated Confidence Interval and Median.

Source: own elaboration, BEA, BLS.

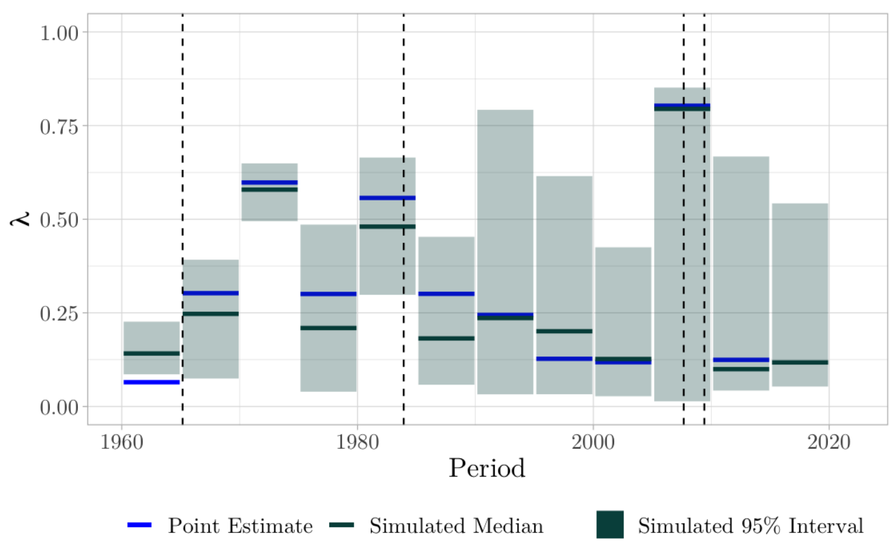
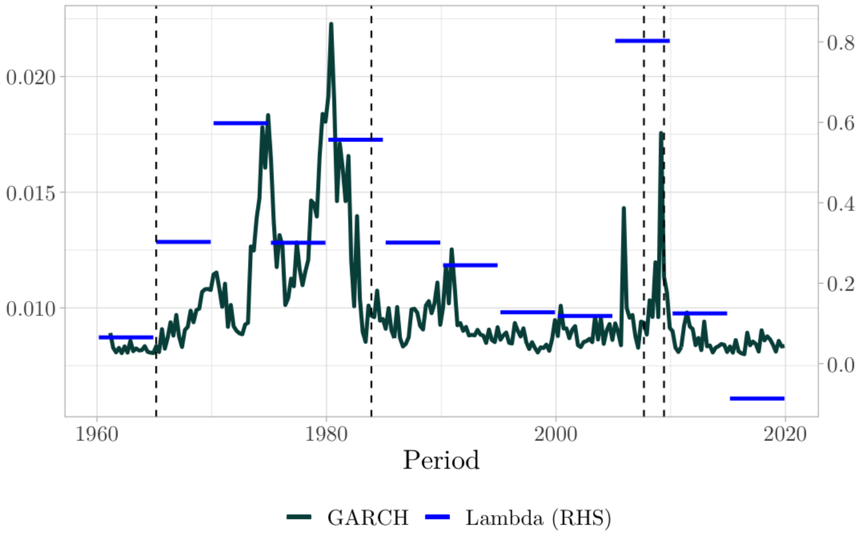


Figure 5 – Estimated Average λ for Quinquennia and Fitted Inflation GARCH(4,1)

Source: own elaboration, BEA, BLS.



The results from Figures 4 and 5 are consistent with those from the previous exercise and support our hypothesis of dynamic attention levels in the SIPC model. Breaking down the sample into quinquennia shows that the point estimates of λ for sub-periods exhibit dynamics even within macroeconomic regimes. However, the issue of confidence interval size is further exacerbated in this exercise because reducing the data in the sub-samples can make different solutions more attractive in terms of squared errors.

4.2 Rolling nonlinear regressions

The next logical step is to present time-varying estimates of λ for the US following a sequence of rolling NLS estimations. Specifically, we will estimate equation 3 for rolling subsets of 10-year periods for the US[6]. This approach should allow us to gather a series for the average λ in time and understand its dynamics.

Figure 6 presents point estimates together with a simulated confidence interval of 95% and its median. Figure 7 presents the resulting series and compare it to a rolling estimate of inflation standard deviation for 10 years.

6 We adopt the convention of presenting the estimate of the sub-sampled period at the last date of that sub-sample. Hence, the first estimate of λ we can present is for 4Q1970, which used data starting from 1Q1961.

Figure 6 – Rolling Estimates of Average λ and its Simulated Confidence Interval and Median.

Source: own elaboration, BEA, BLS.

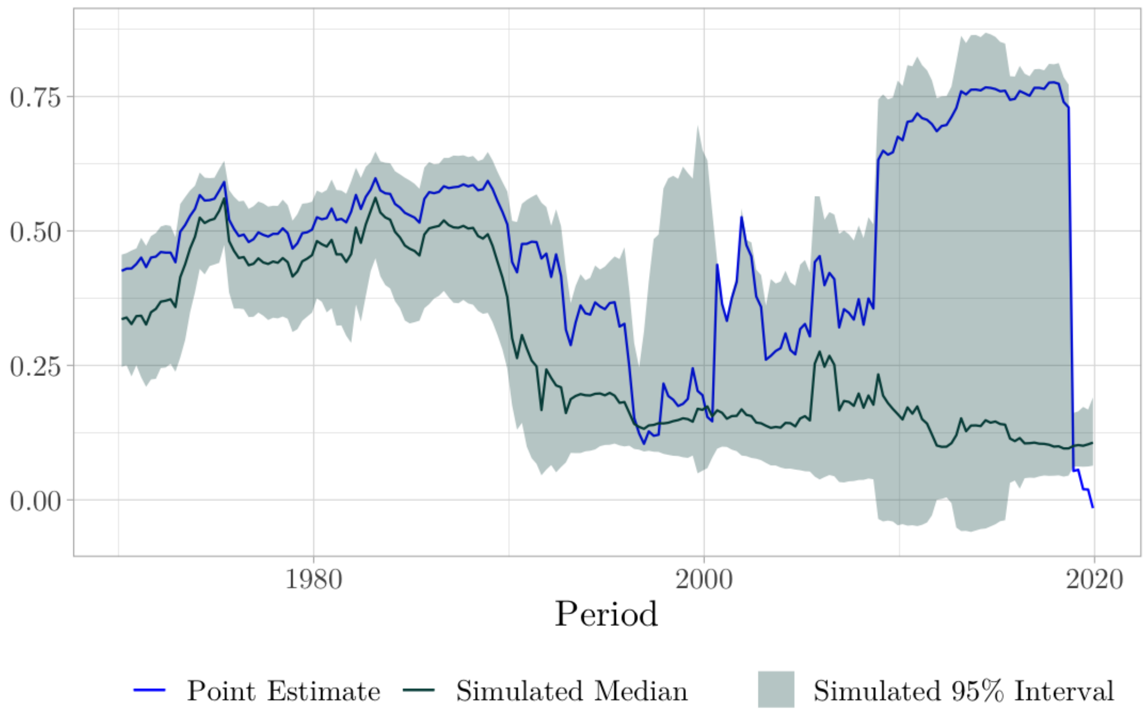
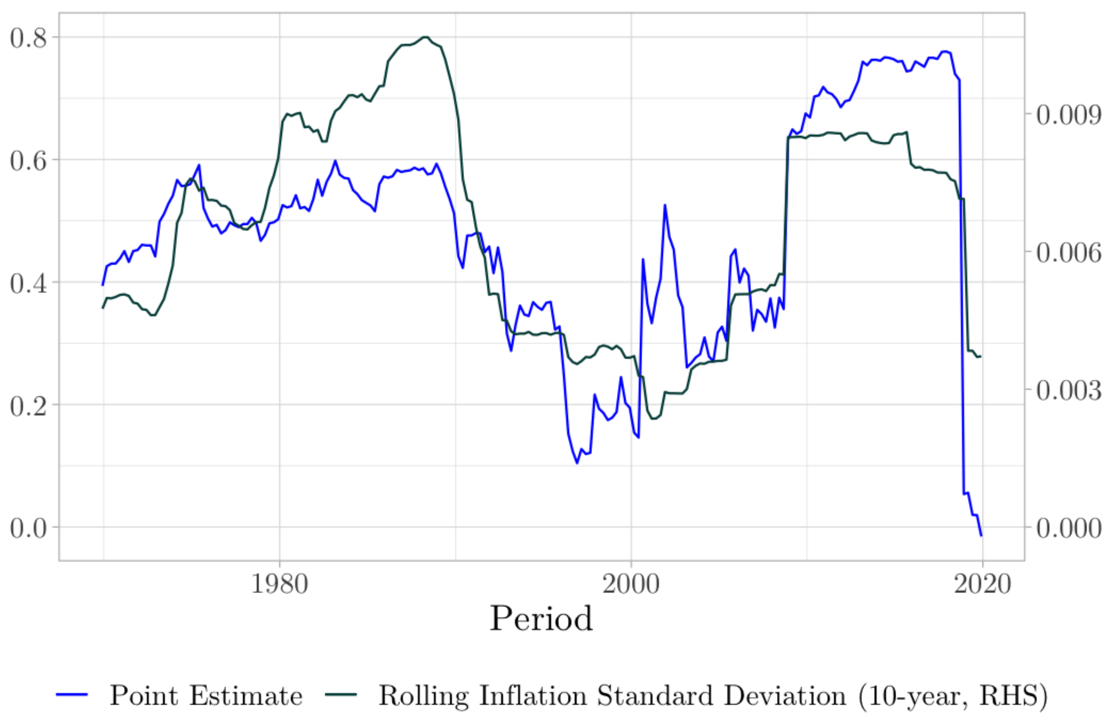


Figure 7 – Estimated λ and 10-year Rolling Inflation Standard Deviation

Source: own elaboration, BEA, BLS.



As before, the first noticeable evidence is in support of a non-constant λ. Since the 1980s until the end of the 1990s, there was a downward trend in λ. This overall decrease in λ can be interpreted as an increase in the average time taken for firms to adjust their prices, i.e., an increase in the inattention levels of the economy. This is most likely due to the increased perceived predictability associated with the Great Moderation period in comparison to its predecessor. This trend is consistent with the results found in Figures 3 and 5.

Since our estimations of λ indicate the the average value over the last 10 years, we compare it with a 10-year rolling estimate of CPI volatility. As the literature suggests[7], the fit is incredibly good, with a correlation of 0.7424. Again, the increased predictability associated with low volatility should be associated with lower benefits of updating the information, *ceteris paribus*. If economic conditions are more volatile, we should have higher benefits of updating the information set to create better expectations. This is consistent with the state-dependent attention formulation presented from Reis (2006a) in Proposition 1, from the results of Begg and Imperato (2001) and from Maćkowiak and Wiederholt (2009) in chapter 2.

This also explains why λ reached higher estimated values after the Great Moderation. Note that here the estimated λ for this period significantly contrasts with the results in the previous exercises. That is because the Great Financial Crisis period is affecting the result. It is possible to observe that the estimated high λ completely vanishes after precisely 10 years of the end of the GFC period. That 'elevated attention' is in fact solely due to the 10-years window size chosen for the rolling exercise. The same issue occurs when one estimates the 10-year rolling inflation volatility. Thus, this contrasting result is solely a consequence of the size of the window in the rolling procedure. For this reason, we believe the median estimate presented to be a better estimate of *lambda*. It is in line with the sub-sampling results and with what was expected by the literature.

We have therefore provided sufficient evidence in favor of dynamic attention in the SIPC model. In addition, we have shown that attention dynamics in the SIPC model are heavily in line with predictions, being strongly associated with inflation and its variability.

4.3 Caveats and Future Steps

The first caveat associated with the our results is the use of forecasted variables as expectations data. Ideally, one would use expectations directly collected from firms regarding prices and the output gap as inputs. However, to the best of our knowledge, no database containing historical expectations of firm for every quarter up to 5 years ahead is yet available. Khan and Zhu (2006) experimented with AR and VAR forecasts and no significant changes emerged. The bootstrapping procedures used (detailed in section 3.2

[7] See sections 2.2 and 2.3.

and appendix D) were capable of considering multiple potential trends for the expectations, thus making our results robust to this critique.

The next issue worth noting is the possibility of multiple solutions of equation 5 with small standard errors. As previously mentioned, this is the primary reason behind the size of the confidence intervals in both bootstrapping procedures used. In the sub-samples and rolling results, we have found that the median estimates presented are vary much in line with the period's implicit CPI volatility estimated through a GARCH(4,1). This is heavily in line with the literature, in particular with Proposition 1 adapted from Reis (2006a), and with the results of Begg and Imperato (2001) and Maćkowiak and Wiederholt (2009). Thus, using this factor as a signal, we believe that our median estimates perform better in capturing the true value of λ."

As mentioned in chapter 3, we would ideally present an estimation of the altered version of the SIPC, equations 3 and 4. Sadly, our attempts thus far have not provided us with adequate results. The non-Markovian nature of λ in equations 3 and 4 is the real challenge, as it rules out any state-space methods by violating conditional independence of observations.

Applying any linearization methods in equations (3) and (4) would: (1) not solve the non-Markovian issue; and (2) imply in an identification problem, since all λ_t's would either emerge as constants or be dropped completely.

There is thus considerable room for future developments in this topic. Exploring alternative non-Markovian estimators should be the main area of focus. With an empirically estimated series for λ_t, one could use it to improve inflation forecasts by incorporating time-varying attention in the SIPC model and comparing its fit quality with that of the standard SIPC and NKPC forecasts. Further exploration of the trends of λ and the implicit information cost is also warranted.

5 CONCLUDING REMARKS

As highlighted by Reis (2021), the inattentiveness framework in macroeconomics, composed by sticky-information models and rational inattention models, presents itself as a well-structured and comparatively simple solution to the empirical puzzles related to the rational expectations hypothesis, in particular inflation persistence. According to the author, this framework allows for a new class of models with only one additional parameter to calibrate and that is “as easy to solve as models with rational-expectations” (REIS, 2021, p. 99).

Even tough the SIPC with constant inattention degree presents a great improvement in fitting with stylized facts and consensual beliefs when compared to the standard NKPC, there is no *a priori* reason to expect a time-invariant inattention. In fact, when considering the determinants of inattention found in the literature and some stylized facts regarding inflation and output gap data, a time-varying inattention is expected.

Our empirical exercises with the standard SIPC have shown evidence supporting our hypothesis of a non-constant inattention. We were also able to show that inflation volatility is a good indicator for the inattention degree, as predicted by the literature. Sadly, when revisiting Mankiw and Reis’ (2002) canonical model to account for this time-varying inattention, the resulting SIPC loses its desired tractability.

Looking ahead, the first step to be further developed on this research is to find a viable estimation process for the time-varying parameters in our version of the SIPC. With a series for the SIPC inattention parameter, one could verify the increase in fit quality of our version of the SIPC and compare it to its original counterpart, in particular comparing the fit of inflation and expected inflation.

REFERENCES

BEGG, D.; IMPERATO, I. The rationality of information gathering monopoly. *The Manchester School*, v. 69, n. 3, p. 237–252, Jun. 2001.

CALVO, G. A. Staggered prices in a utility maximizing framework. *Journal of Monetary Economics*, v. 12, p. 383–398, 1983.

CARROLL, C. D. Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of Economics*, v. 118, n. 1, p. 269–298, Feb. 2003.

COIBION, O.; GORODNICHENKO, Y. Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, v. 105, n. 8, p. 2644–2678, Aug. 2015.

DÖPKE, J. et al. Sticky information phillips curve: European evidence. *Journal of Money, Credit and Banking*, v. 40, n. 7, p. 1513–1519, Oct. 2008.

GALÍ, J. *Monetary Policy, Inflation, and the Business Cycle – An Introduction to the New Keynesian Framework*. 1. ed. [S.l.]: Princeton University Press, 2008.

GILLITZER, C. The sticky information phillips curve: Evidence for australia. *Economic Record*, v. 92, n. 299, p. 548–567, Dec. 2016.

GRANVILLE, B.; ZENG, N. Time variation in inflation persistence: New evidence from modelling us inflation. *Economic Modelling*, v. 81, p. 30–39, 2019. ISSN 0264-9993. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S0264999318304462>.

HANSEN, B. The grid bootstrap and the autoregressive model. *The Review of Economics and Statistics*, v. 81, n. 4, p. 594–607, Nov. 1999.

HODRICK, R.; PRESCOTT, E. *The Center for Mathematical Studies in Economics and Management Sciences - Discussion Paper*, n. 451, p. 1–24, May 1981.

HUNG, T. H.; KWAN, Y. K. Hong kong's new keynesian phillips curve: Sticky information or sticky price? *Pacific Economic Review*, v. 27, n. 1, p. 42–55, Feb. 2022.

HéBERT, B.; WOODFORD, M. Rational inattention and sequential information sampling. *NBER Working Paper Series*, n. 23787, p. 1–68, Sep. 2017.

HéBERT, B.; WOODFORD, M. Rational inattention when decisions take time. *NBER Working Paper Series*, n. 26415, p. 1–68, Aug. 2021.

JR., R. L. Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, v. 1, p. 19–46, 1976.

KHAN, H.; ZHU, Z. Estimates of the sticky-information phillips curve for the united states. *Journal of Money, Credit and Banking*, v. 38, n. 1, p. 195–207, Feb. 2006.

MAĆKOWIAK, B.; MATĚJKA, F.; WIEDERHOLT, M. Rational inattention: a review. *ECB Working Paper Series*, n. 2570, p. 1–74, Jun. 2021.

MAĆKOWIAK, B.; WIEDERHOLT, M. Optimal sticky prices under rational inattention. *The American Economic Review*, v. 99, n. 3, p. 769–803, Jun. 2009.

MANKIW, G.; REIS, R. Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, v. 117, n. 4, p. 1295–1328, Nov. 2002.

MANKIW, G.; REIS, R. Imperfect information and aggregate supply. In: FRIEDMAN, B. M.; WOODFORD, M. (Ed.). *Handbook of Monetary Economics*. 1. ed. [S.l.]: Elsevier, 2010. v. 3, cap. 5, p. 183–229.

PAGAN, A. Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, v. 25, n. 1, p. 221–247, Feb. 1984.

REID, M.; RAND, G. D. A sticky information phillips curve for south africa. *South African Journal of Economics*, v. 83, n. 4, p. 506–526, Dec. 2015.

REIS, R. Inattentive producers. *Review of Economic Studies*, v. 73, p. 793–821, 2006a.

REIS, R. Inattentive consumers. *Journal of Monetary Economics*, v. 53, p. 1761–1800, Oct. 2006b.

REIS, R. Imperfect macroeconomic expectations: Yes, but, we disagree. In: EICHENBAUM, M.; HURST, E.; PARKER, J. A. (Ed.). *NBER Macroeconomics Annual 2020*. [S.l.]: The University of Chicago Press, 2021. v. 35, cap. 1 - Comment, p. 99–111.

SBORDONE, A. Inflation persistence: Alternative interpretations and policy implications. *Journal of Monetary Economics*, v. 54, n. 5, p. 1311–1339, Jul. 2007.

SIMS, C. Rational inattention and monetary economics. In: FRIEDMAN, B. M.; WOODFORD, M. (Ed.). *Handbook of Monetary Economics*. 1. ed. Elsevier, 2010. v. 3, cap. 04, p. 155–181. Available from Internet: <https://EconPapers.repec.org/RePEc:eee:monchp:3-04>.

SIMS, C. A. Implications of rational inattention. *Journal of Monetary Economics*, v. 50, p. 665–690, 2003.

WOODFORD, M. Imperfect common knowledge and the effects of monetary policy. In: AGHION, P. et al. (Ed.). *Knowledge, Information, and Expectations in Modern Macroeconomics*. [S.l.]: Princeton University Press, 2001.

WOODFORD, M. Interpreting inflation persistence: Comments on the conference on “quantitative evidence on price determination”. *Journal of Money, Credit and Banking*, v. 39, n. 1, p. 19–46, Feb. 2007.

APPENDIX A - STICKY INFORMATION PHILLIPS CURVE

To develop the SIPC, Mankiw and Reis (2002) start by considering that a firm's optimal behavior is given by the price decision that maximizes profit, which follows:

$$p_t^* = p_t + \alpha y_t,$$

where p_t^* is the firm's desired price, p_t is the overall price level, y_t is the output gap, and α is the degree of strategic complementarity, which is calibrated with a value of $\alpha = 0.1$.

In contrast to models *a la* Calvo (1983), in the SIPC framework

> *"[...] every firm sets its price every period, but firms gather information and recompute optimal prices slowly over time. In each period, a fraction λ of firms obtains new information about the state of the economy and computes a new path of optimal prices. Other firms continue to set prices based on old plans and outdated information. We make an assumption about information arrival that is analogous to the adjustment assumption in the Calvo model: each firm has the same probability of being one of the firms updating their pricing plans, regardless of how long it has been since its last update." (MANKIW; REIS, 2002, p. 1299).*

With the above argument in mind, a firm that last updated its plans j periods ago sets the price:

$$x_t^j = E_{t-j} p_t^*.$$

Considering that the aggregate price level is the average of the prices of all firms in the economy, Mankiw and Reis (2002) define:

$$p_t = \lambda \sum_{j=0}^{\infty} .(1-\lambda)^j x_t^j,$$

We can thus, combine all previous equations to conclude that:

$$p_t = \lambda \sum_{j=0}^{\infty} E_{t-j}(p_t + \alpha y_t).$$

Using further algebra available in their appendix, Mankiw and Reis (2002) obtain the main structure of the SIPC:

$$\pi_t = \frac{\lambda\alpha}{1-\lambda} y_t + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-1-j}\left[\pi_t + \alpha \Delta y_t\right], \Delta y_t = y_t - y_{t-1}, \tag{6}$$

where π_t is the inflation rate. This is equation 1.

APPENDIX B - BEGG AND IMPERATO'S (2001) MODEL - THE INFORMATION GATHERING PROBLEM OF A MONOPOLY

I will present here the main characteristics of the model proposed by Begg and Imperato (2001). This appendix will serve as a summary of their elaboration, indicating the main equations, assumptions and results. However, for the sake of brevity, the step-by-step construction of the model will be omitted.

Below, I present their demand equation:

$$p_t = \alpha_t - \beta q_t,$$

where p is the price of the good and q its quantity. The intercept α is stochastic and follows the mean-reverting Markov process:

$$d\alpha = k(\mu - \alpha)dt + \sigma dz$$

where μ is the long-run mean, $k > 0$ is the speed of adjustment, σ is a diffusion coefficient, and z is a standard Wiener process. Hence, the unconditional mean is μ and the unconditional variance is $\frac{\sigma^2}{2k}$.

Normalizing productions costs to zero, the optimal output (that maximize revenue) is:

$$q_t^* = \frac{\alpha_t}{2\beta} \tag{7}$$

The model presented until this point can be seen as the frictionless case. Now, I will present the alterations in the model if a fixed cost to thinking (b) is imposed. Before, I review the informational characteristics of the environment they introduce:

> *we assume that both the structure of the model and the stochastic process for the evolution of the intercept of the demand function are known: we rule out model uncertainty. Moreover, we assume that the information about the value taken by contemporaneous α is freely available, but that using it to make an output choice entails a fixed sunk cost b, the cost of thinking. (BEGG; IMPERATO, 2001, p. 240)*

The cost of thinking can be interpreted in many ways, e.g., the cost of accessing the available information that will help with the decision process. If there are thinking costs, two possibilities emerge. Either the monopolist incurs such costs and define output as in equation 5, or, the monopolist does not incur those costs and rely on some old information acquired at $\tau < t$ and produce:

$$q_t = \frac{E[\alpha_t|\alpha_\tau]}{2\beta}. \tag{8}$$

The monopolist will then decide at every time whether to think and get some termination payoff or to continue to rely on past information. Begg and Imperato (2001) then show that the solution to this problem is equivalent to choosing a unique optimal length of periods to update their information set, an unique optimal thinking lag Δ^*.

According to their results, Δ^* is implicitly given by:

$$b = \frac{1}{4\beta^2}\frac{\sigma^2}{2k\rho}[1 - \exp(-2k\Delta^*)]$$

where ρ is the discount rate.

Assuming $b > 0$ and $\Delta^* > 0$, which are the conditions of interest to us, we have the following relations:

b	σ^2	β	ρ	k
+	-	+	+	+

Table 3 – Relations between Δ^* and the relevant variables

Note that, by definition, Δ^* is equivalent to the definition of $D = \frac{1}{\lambda}$, the periods between information updates. Hence, with the previous assumptions, we have the following relations:

b	σ^2	β	ρ	k
-	+	-	-	-

Table 4 – Relations between λ and the relevant variables

As the cost of processing information b increases, the level of attention decreases as it gets more expensive to be attentive. If the variance of the forecast variable σ^2 increases, the level of attention also increases, since it is more beneficial to be more attentive under more uncertain conditions. If the discount rate ρ increases, the attention degree decreases as it gets more expensive to update the information set in present value terms, postponing the cost becomes more attractive. If the degree of mean reversion of the forecast variable increases, the level of attention decreases because faster adjustment makes expected errors smaller.

APPENDIX C - ALGEBRAIC DERIVATIONS FOR EQUATIONS 3 AND 4

Assume that history starts at time $t = 0$. Then, the price at $t = 1$ is:

$$p_1 = \lambda_1 p_1^o + (1 - \lambda_1) E_0 [p_1^o] ;$$

where the desired price level, p_1^o:

$$p_1^o = p_1 + \alpha y_1,$$

with p_1 for the overall price level and y_1 for the output.

At $t = 2$ is:

$$p_2 = \lambda_2 p_2^o + (1 - \lambda_2) \{\lambda_1 E_1 [p_2^o] + (1 - \lambda_1) E_0 [p_2^o]\} ;$$

and in general:

$$p_t = \lambda_t p_t^o + (1 - \lambda_t) \left\{ \begin{array}{c} \lambda_{t-1} E_{t-1} [p_t^o] + \\ (1 - \lambda_{t-1}) \left\{ \begin{array}{c} \lambda_{t-2} E_{t-2} [p_t^o] + \\ (1 - \lambda_{t-2}) \left\{ \begin{array}{c} \lambda_{t-3} E_{t-3} [p_t^o] + \\ (1 - \lambda_{t-3}) \left\{ \begin{array}{c} ...+ \\ (1 - \lambda_1) E_0 [p_t^o] \end{array} \right\} \end{array} \right\} \end{array} \right\} \end{array} \right\} .$$

Rearranging p_t :

$$\begin{aligned} p_t &= \lambda_t p_t^o + (1 - \lambda_t) \lambda_{t-1} E_{t-1} [p_t^o] + (1 - \lambda_t) (1 - \lambda_{t-1}) \lambda_{t-2} E_{t-2} [p_t^o] \\ &\quad + (1 - \lambda_t) (1 - \lambda_{t-1}) (1 - \lambda_{t-2}) \lambda_{t-3} E_{t-3} [p_t^o] \\ &\quad +... + \lambda_1 E_1 [p_t^o] \prod_{j=0}^{t-2} (1 - \lambda_{t-j}) + E_0 [p_t^o] \prod_{j=0}^{t-1} (1 - \lambda_{t-j}) ; \end{aligned}$$

or

$$\begin{aligned} p_t &= \lambda_t p_t^o + \lambda_{t-1} E_{t-1} [p_t^o] \prod_{j=0}^{0} (1 - \lambda_{t-j}) + \lambda_{t-2} E_{t-2} [p_t^o] \prod_{j=0}^{1} (1 - \lambda_{t-j}) \\ &\quad + \lambda_{t-3} E_{t-3} [p_t^o] \prod_{j=0}^{2} (1 - \lambda_{t-j}) + ... + \lambda_1 E_1 [p_t^o] \prod_{j=0}^{t-2} (1 - \lambda_{t-j}) + E_0 [p_t^o] \prod_{j=0}^{t-1} (1 - \lambda_{t-j}) ; \end{aligned}$$

or

$$p_t = \lambda_t p_t^o + \sum_{i=1}^{t-1} \lambda_{t-i} E_{t-i} [p_t^o] \prod_{j=0}^{i-1} (1 - \lambda_{t-j}) + E_0 [p_t^o] \prod_{j=0}^{t-1} (1 - \lambda_{t-j}) .$$

Taking out the first term and redefining the summation index:

$$p_t = \lambda_t p_t^o + \lambda_{t-1} E_{t-1}\left[p_t^o\right](1-\lambda_t) + \sum_{i=1}^{t-2} \lambda_{t-i-1} E_{t-i-1}\left[p_t^o\right] \prod_{j=0}^{i}(1-\lambda_{t-j}) + E_0\left[p_t^o\right] \prod_{j=0}^{t-1}(1-\lambda_{t-j}).$$

The previous period's price can be written as:

$$p_{t-1} = \lambda_{t-1} p_{t-1}^o + \sum_{i=1}^{t-2} \lambda_{t-i-1} E_{t-i-1}\left[p_{t-1}^o\right] \prod_{j=1}^{i}(1-\lambda_{t-j}) + E_0\left[p_{t-1}^o\right] \prod_{j=1}^{t-1}(1-\lambda_{t-j}).$$

Subtracting p_{t-1} from p_t :

$$\begin{aligned} \pi_t &= \lambda_t p_t^o + \lambda_{t-1} E_{t-1}\left[\Delta p_t^o\right] + \sum_{i=1}^{t-2} \lambda_{t-i-1} E_{t-i-1}\left[\Delta p_t^o\right] \prod_{j=1}^{i}(1-\lambda_{t-j}) + E_0\left[\Delta p_t^o\right] \prod_{j=1}^{t-1}(1-\lambda_{t-j}) \\ &\quad -\lambda_t \left\{\lambda_{t-1} E_{t-1}\left[p_t^o\right] + \sum_{i=1}^{t-2} \lambda_{t-i-1} E_{t-i-1}\left[p_t^o\right] \prod_{j=1}^{i}(1-\lambda_{t-j}) + E_0\left[p_t^o\right] \prod_{j=1}^{t-1}(1-\lambda_{t-j})\right\}, \end{aligned}$$

where:

$$\Delta p_t^o = \Delta p_t + \alpha \Delta y_t = \pi_t + \alpha \Delta y_t.$$

Given:

$$p_t - \frac{\lambda_t}{1-\lambda_t}\alpha y_t = \lambda_{t-1} E_{t-1}\left[p_t^o\right] + \sum_{i=1}^{t-2} \lambda_{t-i-1} E_{t-i-1}\left[p_t^o\right] \prod_{j=1}^{i}(1-\lambda_{t-j}) + E_0\left[p_t^o\right] \prod_{j=1}^{t-1}(1-\lambda_{t-j}),$$

we can rewrite:

$$\begin{aligned} \pi_t &= \lambda_t\left(p_t + \alpha y_t\right) + \lambda_{t-1} E_{t-1}\left[\Delta p_t^o\right] + \sum_{i=1}^{t-2} \lambda_{t-i-1} E_{t-i-1}\left[\Delta p_t^o\right] \prod_{j=1}^{i}(1-\lambda_{t-j}) \\ &\quad + E_0\left[\Delta p_t^o\right] \prod_{j=1}^{t-1}(1-\lambda_{t-j}) - \lambda_t\left\{p_t - \frac{\lambda_t}{1-\lambda_t}\alpha y_t\right\}, \end{aligned}$$

and

$$\begin{aligned} \pi_t &= \left(\frac{\lambda_t}{1-\lambda_t}\right)\alpha y_t + \lambda_{t-1} E_{t-1}\left[\Delta p_t^o\right] + \sum_{i=1}^{t-2} \lambda_{t-i-1} E_{t-i-1}\left[\Delta p_t^o\right] \prod_{j=1}^{i}(1-\lambda_{t-j}) \\ &\quad + E_0\left[\Delta p_t^o\right] \prod_{j=1}^{t-1}(1-\lambda_{t-j}). \end{aligned}$$

Manipulating:

$$(1-\lambda_t)\,\pi_t = \lambda_t \alpha y_t + \sum_{i=0}^{t-2} \lambda_{t-i-1} E_{t-i-1}\left[\Delta p_t^o\right] \prod_{j=0}^{i} \left(1-\lambda_{t-j}\right) + E_0\left[\Delta p_t^o\right] \prod_{j=0}^{t-1} \left(1-\lambda_{t-j}\right),$$

and finally:

$$\pi_t = \lambda_t\left(\pi_t + \alpha y_t\right) + \sum_{i=0}^{t-2} \lambda_{t-i-1} E_{t-i-1}\left[\Delta p_t^o\right] \prod_{j=0}^{i} \left(1-\lambda_{t-j}\right) + E_0\left[\Delta p_t^o\right] \prod_{j=0}^{t-1} \left(1-\lambda_{t-j}\right).$$

APPENDIX D - BOOTSTRAPPED CONFIDENCE INTERVALS FOR ALL ESTIMATES

In this appendix, we present the same graphs available in the main text together with the bootstrapped confidence intervals for all estimates. As mentioned, we opted to present these results in this appendix for 2 main reasons: (1) the multiple estimates intervals in the main text is already capable of summarizing multiple possibilities for the same estimate of a λ considering numerous possibilities for expectations; and (2) due to the high order of the reference equation, the problem of multiple solutions near the (0,1) interval is exacerbated in the construction of bootstrapped confidence intervals. We first explain the process used for the bootstrapping exercise to then follow with the results in the same order as presented in the main text.

We apply the same method as in Khan and Zhu (2006) for the bootstrapping exercises. For each inflation and output gap expectations, we estimate a rolling AR(4), as explained in section 3.1, and simulate 1000 possible futures. With the simulated expectations data, we repeat the estimation exercises done in the text, storing $\hat{\lambda}_t^i$ and $\hat{\sigma}_t^i$, the simulated estimate $i = \{1, 2, \cdots, 1000\}$ and its standard error. We then calculate the t-statistic $t_t^i = \frac{\hat{\lambda}_t^i - \hat{\lambda}_t}{\hat{\sigma}_t^i}$ (where $\hat{\lambda}_t$ is the point estimate for period t), and order it in an increasing manner, creating the sequence we call $\tilde{t}_t$. Next, we select the 500th element of this sequence of t-statistics for the bootstrapped median estimate, calculated as $\hat{\lambda}_t + \tilde{t}_t^{500}\hat{\sigma}_t$ and 25th and 975th elements of the sequence 5% bootstrapped confidence intervals, $\tilde{t}_t^{25}$, $\tilde{t}_t^{975}$, $[\hat{\lambda}_t + \tilde{t}_t^{975}\hat{\sigma}_t, \quad \hat{\lambda}_t + \tilde{t}_t^{25}\hat{\sigma}_t]$, where $\hat{\sigma}_t$ is the standard error of $\hat{\lambda}_t$.

As anticipated, the resulting confidence intervals are considerably wider, in particular in estimations around the Great Moderation period for the sub-samples exercises and since 1990 for the rolling estimations exercise.

For the estimate of λ as constant for the whole sampled period, we recall that the resulting point estimate of λ for the whole sample is 0.5464. Using the bootstrapping procedure described above, the median bootstrapped estimate is 0.3701 and the bootstrapped confidence interval is [0.2502, 0.4839]. As for the Khan and Zhu (2006) sample period (1Q1980-4Q2000), our point estimate is 0.211, the bootstrapped median estimate is 0.4197 and the bootstrapped confidence interval is [0.2788, 0.5390].

Finally, below we revisit Figures 2, 4, and 6 substituting the CIs from the procedure of section 3.2 for the bootstrapping procedure described above.

Figure 8 – Estimated Average λ for Macroeconomic Regimes and Bootstrapped Confidence Interval and Median

Source: own elaboration, BEA, BLS.

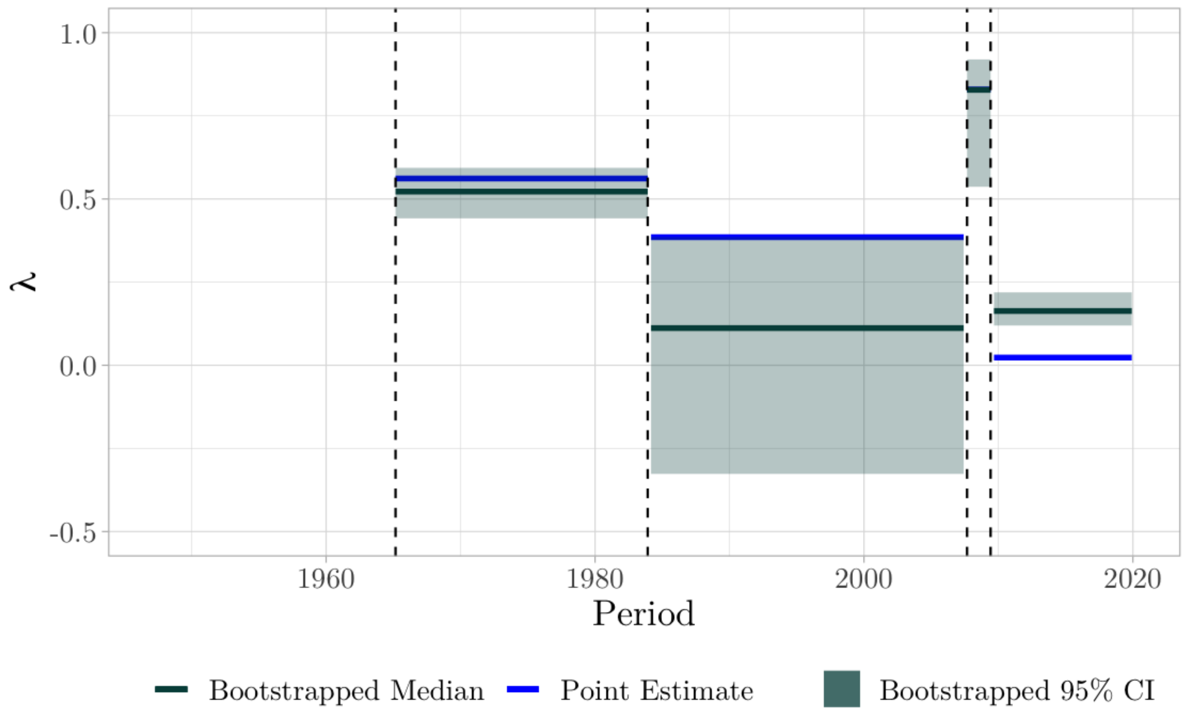


Figure 9 – Estimated Average λ for Quinquennia and Bootstrapped Confidence Interval and Median

Source: own elaboration, BEA, BLS.

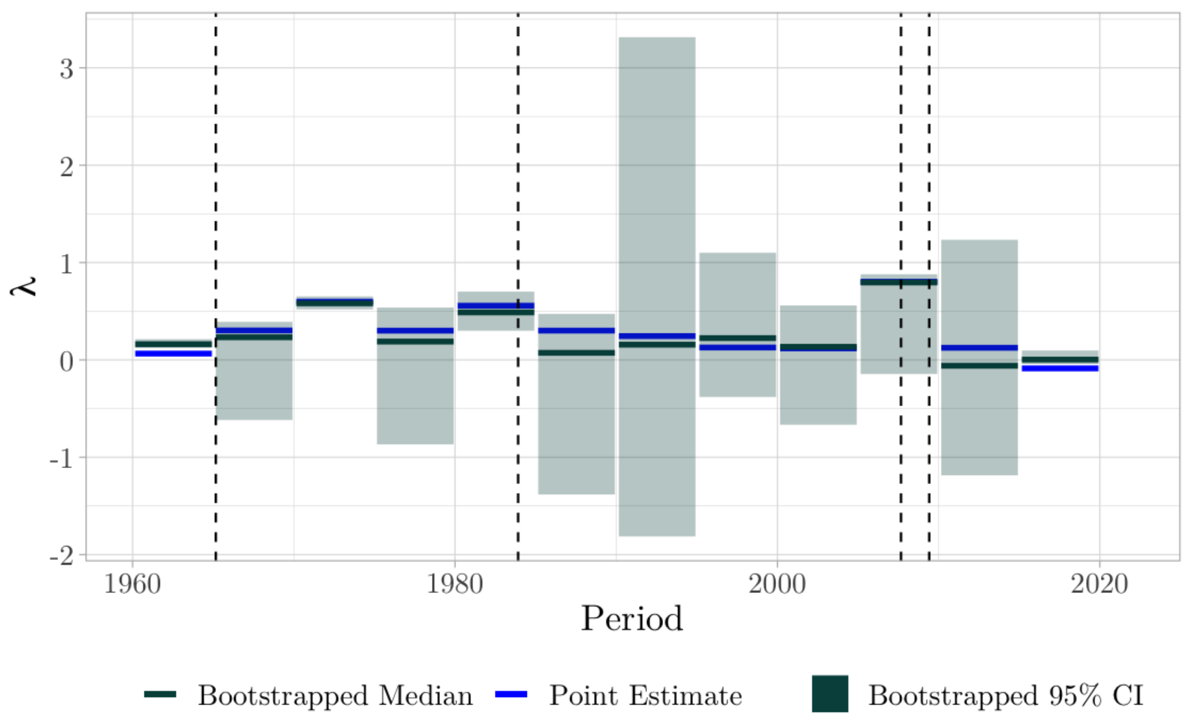
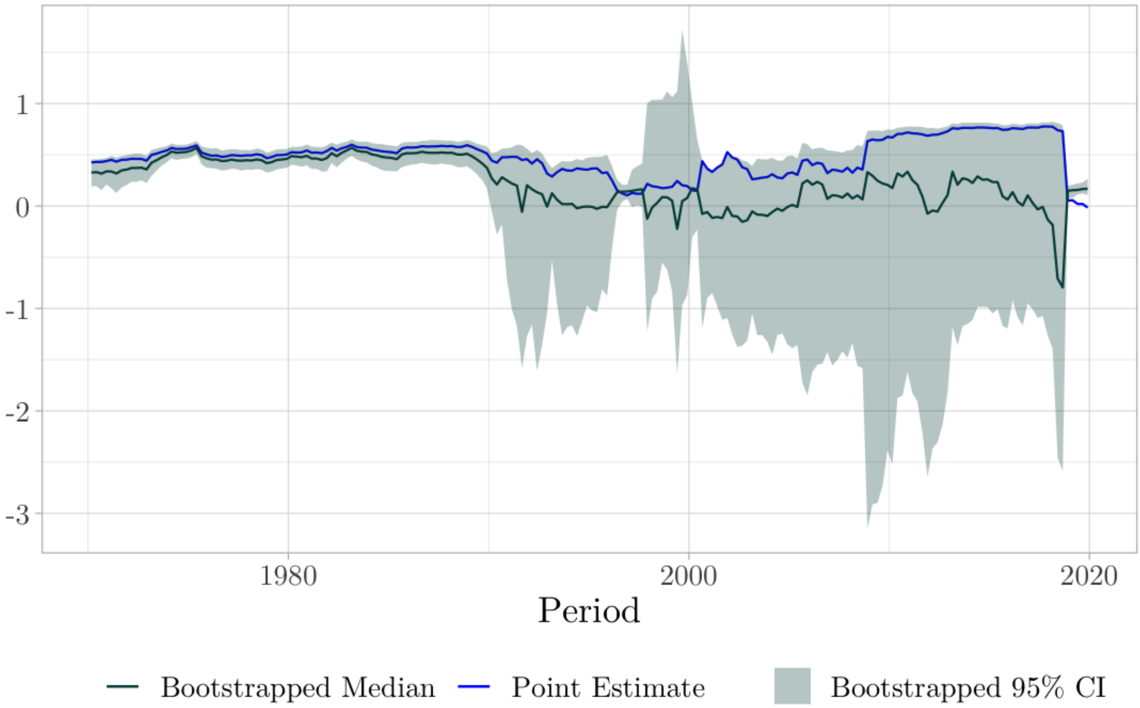


Figure 10 – Rolling Estimates of Average λ and its Bootstrapped Confidence Interval and Median

Source: own elaboration, BEA, BLS.



APPENDIX E - ALGEBRAIC DERIVATIONS FOR EQUATION 5

Departing from equation 1:

$$\pi_t = \frac{\lambda\alpha}{1-\lambda} y_t + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-1-j}\left[\pi_t + \alpha\Delta y_t\right],$$

if we define $T < t$ we can rewrite it as:

$$\pi_t = \frac{\lambda\alpha}{1-\lambda} y_t + \lambda \sum_{j=1}^{T-1} (1-\lambda)^{j-1} E_{t-j}\left[\pi_t + \alpha\Delta y_t\right] + \lambda \sum_{j=T}^{\infty} (1-\lambda)^{j-1} E_{t-j}\left[\pi_t + \alpha\Delta y_t\right].$$

If we assume that $E_{t-j}\left[\pi_t + \alpha\Delta y_t\right] = E_{t-T}\left[\pi_t + \alpha\Delta y_t\right] \quad \forall j > T$, the equation then becomes:

$$\pi_t = \frac{\lambda\alpha}{1-\lambda} y_t + \lambda \sum_{j=1}^{T-1} (1-\lambda)^{j-1} E_{t-j}\left[\pi_t + \alpha\Delta y_t\right] + \lambda E_{t-T}\left[\pi_t + \alpha\Delta y_t\right] \sum_{j=0}^{\infty} (1-\lambda)^{j+T-1},$$

or

$$\pi_t = \frac{\lambda\alpha}{1-\lambda} y_t + \lambda \sum_{j=1}^{T-1} (1-\lambda)^{j-1} E_{t-j}\left[\pi_t + \alpha\Delta y_t\right] + \lambda (1-\lambda)^{T-1} E_{t-T}\left[\pi_t + \alpha\Delta y_t\right] \sum_{j=0}^{\infty} (1-\lambda)^{j}.$$

Since $\lambda \in (0,1)$, we have equation 5:

$$\pi_t = \frac{\lambda\alpha}{1-\lambda} y_t + \lambda \sum_{j=1}^{T-1} (1-\lambda)^{j-1} E_{t-j}\left[\pi_t + \alpha\Delta y_t\right] + (1-\lambda)^{T-1} E_{t-T}\left[\pi_t + \alpha\Delta y_t\right].$$