Department of Economics, University of Warwick

Monash Business School, Monash University

as part of

Monash Warwick Alliance

Application method of rational inattention hypothesis, and Rational Inattention New Keynesian Philips Curve creation

Shun Tonami

Warwick-Monash Economics Student Papers

March 2023

No: 2023/53

ISSN 2754-3129 (Online)

https://warwick.ac.uk/fac/soc/economics/research/wmesp/

The Warwick Monash Economics Student Papers (WM-ESP) gather the best Undergraduate and Masters dissertations by Economics students from the University of Warwick and Monash University. This bi-annual paper series showcases research undertaken by our students on a varied range of topics. Papers range in length from 5,000 to 8,000 words depending on whether the student is an undergraduate or postgraduate, and the university they attend. The papers included in the series are carefully selected based on their quality and originality. WM-ESP aims to disseminate research in Economics as well as acknowledge the students for their exemplary work, contributing to the research environment in both departments.

"*We are very happy to introduce the Warwick Monash Economics Student Papers (WM-ESP). The Department of Economics of the University of Warwick and the Economics Department at Monash University are very proud of their long history of collaboration with international partner universities, and the Monash Warwick Alliance reflects the belief in both Universities that the future will rely on strong links between peer Universities, reflected in faculty, student, and research linkages. This paper series reflects the first step in allowing our Undergraduate, Honours, and Masters students to learn from and interact with peers within the Alliance.*"

Ben Lockwood (Head of the Department of Economics, University of Warwick) and Michael Ward (Head of the Department of Economics, Monash University)

Recommended citation: Tonami, S. (2023). Application Method of the Rational Inattention Hypothesis, and the Rational Inattention New Keynesian Philips Curve Creation. *Warwick Monash Economics Student Papers* 2023/53

[1] Warwick Economics would like to thank Lory Barile, Gianna Boero, and Caroline Elliott for their contributions towards the selection process.

Application Method of Rational Inattention Hypothesis and Rational Inattention New Keynesian Phillips Curve Creation

Shun Tonami*

Abstract

This paper gives a fresh perspective on the New Keynesian Philips Curve (NKPC) when combining the rational inattention hypothesis, which will give a new insight into backward-looking evidence. A further contribution of this paper is to give a unique application method of the development of the rational inattention hypothesis in other economic fields. This study provides the viewpoint that OLS (ordinary least squares) estimators have an imperfect information bias under the noise information model where economic agents optimize their behaviors through the rational inattention hypothesis. Specifically, the information flow constraint is redefined as the regression coefficient constraint by information theory extension. As a result, if the rational inattention hypothesis is expected to hold in economic fields, the Ridge

---

* Contact: shun.tonami.01@icloud.com

I would like to give my special regards to my supervisor Ao Wang for his supportive advice and thoughtful guidance throughout this dissertation.

regression can provide optimal estimations for economists in their fields by allowing them to consider the imperfect information bias through penalizing the coefficients. In empirical work, the rational inattention model complements the NKPC by including the imperfect information bias error in the regression analysis. The estimation reveals that the imperfect information bias error can play a backward-looking role. The rational inattention NKPC introduces the backward-looking element theoretically and demonstrates inflation inertia well.

Keywords: Rational Inattention, Ridge Regression, Information Flow Constraint, New Keynesian Philips Curve

JEL classification codes: C51, D8, E31

Introduction

1.

In order to give a new insight into monetary-policy operation, this paper aims to develop a novel method to study inflation dynamics by means of data science tools. The data science tools are specialized to enhance fittings of models to data sets, although the theoretical context for the reason of the model improvement is ambiguous. By contrast, economic theory is constructed on a theoretical framework. Thus, the theory facilitates the model interpretation in the empirical analysis, whereas some economic theories show an inconsistent performance with the theoretical interpretation there. What data science tools make a better fitting to the data sets and what is missing in the economic theory? If economic theory explains the economic agent behaviors perfectly well, the empirical performance based on the economic model is not expected to be improved significantly by the data science tools. This paper studies the mathematical structure of the data science tools and attempts to incorporate the structure into the empirical study of inflation dynamics.

It is generally observed that inflation has inertia. However, the baseline NKPC cannot theoretically demonstrate the empirical evidence. The NKPC assumes full information rational expectation (FIRE). The assumption allows the firms' expected inflation to be estimated by the econometrics forecast predicted from current information sets. In a short-term forecast, it is known that the simple econometric forecast is likely to be more accurate than the structural economic model. Therefore, the data science method may show a better model which complements something missing in the expected inflation forecast.

To think of the inconsistency of the NKPC with the empirical evidence as FIRE inappropriateness, this paper studied expected inflation theory in terms of rational inattention hypothesis. Mackowiak and Wiederholt (2009) showed an optimization problem for firm price setting when firms are limited in their information processing capacity. This model uses information flow constraint in the optimization problem to represent a firm's limited information processing capacity. To apply the data science tool structure, I attempted to redefine the information flow constraint of the optimization problem. As a result, it was found that the information flow constraint transforms into the regression coefficient constraint. When the rational inattention hypothesis holds, OLS is not a proper estimation to be considered as the econometric forecast that firms are expected to have. The information processing capacity of the firms imposes a constraint upon the size of the coefficients when estimating the expected inflation. This implies a Ridge regression model, which is the method that will give an OLS estimation on the regression coefficients' constraint upon size. The estimation method can control the bias associated with the limited information processing capacity existing in the regression coefficients. The bias is labeled imperfect information bias in this paper. These reasons provided the theoretical evidence for the model improvement by the data science tool on the expected inflation estimation.

For the empirical analysis, this paper built NKPC with the rational inattention hypothesis. The baseline NKPC employs generalized method of moments (GMM). The estimation method requires an independent relation between the current information sets and the forecast error. It is mathematically shown that the relation does not hold with the rational inattention hypothesis. Due to the information flow constraint, an imperfect information bias appears in the estimation. The bias is correlated with the current information sets. Changes in the current information sets have an impact on the forecast error through the change in size of the imperfect information bias. The GMM does not satisfy the consistency. Simultaneously, it is proposed that it is possible to satisfy the condition and have an unbiased estimator if the NKPC incorporates the error term that has arisen from the imperfect information bias into the estimation model as a control variable.

The NKPC controlling the imperfect information bias denotes rational inattention NKPC (RINKPC) in this paper. RINKPC has a backward-looking element because the information lags lead to inflation lags. The RINKPC has an imperfect information bias error term in the estimation model. The error term serves as the path of the shock from the past to the future, as representing the shock dynamics through the information process structure in which firms take time to perfectly remove the noise in the information they received. The information noise and the information processing capacity of the firms lead to the inflation lags. Mackowiak and Wiederholt (2009) gave a proof of the same structure as this model to explain the inflation inertia and confirmed a hump-shaped response of the inflation to the shock in their theoretical analysis.

The results of the empirical analysis were consistent with the implication from the econometric structure model. The estimator of the purely forward-looking NKPC is biased. Comparing RINKPC with the purely forward-looking NKPC, the forward-looking parameter reduces the size, while the coefficients of the real marginal cost increases in size. These indicate that the purely forward-looking NKPC is potentially biased and cannot control the bias associated with the information flow constraint. The RINKPC can obtain an unbiased estimator on the ground that the rational inattention hypothesis holds.

Overall, the purpose of this paper is to attempt the development of a novel method through the application of data science tools to inflation dynamics. By demonstrating the method of the rational inattention application and the creation of the RINKPC, it was possible to achieve a new insight into monetary policy operation in this paper.

The remaining sections proceed as follows. Section 2 briefly overviews the prior research. Sections 3.1 and 3.2 are theoretical frameworks. These explain how the Ridge regression is applied to rational inattention and the RINKPC construction. Section 4.1 and 4.2 specifically show the empirical model structure, the data, and the methodology. The result of the analysis appears in Section 5 and the paper is concluded in Section 6.

Literature Review

2.

The NKPC prevails in the field, benefiting from the microeconomic background providing the central banks with a guideline for their policy operation. However, prior research generally confirms that the empirical performance of the NKPC is open to dispute. For instance, a drawback is that it cannot prove inflation inertia theoretically. Gali and Gertler (1999) and Gali, Gertler and Lopez Saildo (2005) estimate a hybrid NKPC including a backward-looking element. They concluded that, a purely forward-looking model is inconsistent with the data, and does not change the conclusion because the estimated coefficient of the backward-looking parameter is small in size in comparison to the coefficients used in the forward-looking parameter. However, the GMM employed in these papers pointed out that the NKPC is likely to overestimate due to weak instruments and weak identification. Nason (2008) built a model considering these problems and implemented an estimation.

These disputes are currently ongoing. Many researchers have attempted to explain the empirical evidence from various perspectives. One of the research branches presents the position that there is noise information in the model. In this paper, the inflation dynamics in the context of the rational inattention hypothesis is studied. According to Sims (2003), rational inattention theory benefits economic agents by optimizing their attention allocation under their limited information processing capacity. The economic agents are able to recognize the noise in the information and minimize their potential loss under the information flow constraint. Mackowiak and Wiederholt (2009) employs the rational inattention hypothesis with inflation dynamics and builds a microeconomic foundation on inflation inertia. They proved the hump shape of inflation by the rational inattention and optimal attention allocation rules.

As for the rational inattention and Philips curve, the theoretical analysis is developing. Afrouzi and Yang (2020) built an inflation dynamic model under rational inattention and formed an understanding of the relation between the Philips curve slope and the attention optimization. As similar papers, Sims (2003), Mackowiak and Wiederholt (2018) and Fulton (2018) are introduced in their paper (quoted in Afrouzi and Yang, 2020:5).

Looking at the prior papers on the empirical model of the rational inattention hypothesis, many of them are survey-based studies. Coibion and Gorodnichenko (2015) proposes a test model to examine whether FIRE is supported or not. They constructed a model of the relationship between the forecast error and the current information set by rational expectation, sticky information, and a noise information model. The forecast error is correlated with current information in the US survey data and observes information rigidity. Therefore, they support the sticky or noise information assumption. Lena et al. (2019) studied firms' expected inflation using a survey. They regressed the expected inflation on the macro variable and survey variable. The expectations of wages and prices were affected by different variables. The difference may suggest bounded rationality or rational inattention. In addition, Coibion and Gorodnichenko (2015) also tested whether the firms' expectations are rational or not, building a regression equation for forecast error. The results showed that the expected inflation is not rational. Their study supports rational inattention

According to Okuda (2018), the development of imperfect information is attributed to the inappropriateness of FIRE. In fact, Carrol (2003), Kiley (2007) and Coibion and Gorodnichenko (2012) and (2015b), obtained results which did not support FIRE (quoted in Okuda:2018:71), and thus FIRE appears to be a strong assumption.

With respect to the data science application toward inflation, the Bank of Japan Research and Statistical Bureau Economic Analysis Group (2017) employs the statistical learning method for the firms' expected inflation based on survey data. This paper analyzed the big survey data by random forests, which is a data science tool. They did not attempt to combine the data science tool with economic theory. The paper specifies what features about firms' expected inflation can be extracted from the data through the random forest.

Overall, there is much research into rational inattention fields. However, there are few papers about the application of rational attention to the NKPC through the data science method. In the empirical study of rational inattention, survey data is popular. Moreover, the application of the rational inattention hypothesis is rapidly developing into a structural model of the inflation dynamic. My unique point is an attempt to construct a regression model based on rational inattention and empirically applying rational inattention to the NKPC.

Theoretical Model

3.1 The imperfect information bias in OLS estimators under noise information

In this section, I present a mathematical model that shows that an imperfect information bias exists in the OLS estimators. My model is based upon information theory, data science and economics. I will set an optimization problem for the firm price setting under the rational inattention hypothesis by specifying the information flow constraint in terms of information theory. In consequence, a Ridge regression is indicated for a method of estimation built on the rational inattention hypothesis. This regression model can control the imperfect information bias in the estimation. In the subsequent section, the Ridge regression is explained and the Ridge regression model applied to the NKPC for empirical analysis.

I extended the Mackowiak and Wiederholt (2009) model of firm price setting under the rational inattention hypothesis. To shed light on price setting under decision makers' limited attention, prices are fully flexible and in every period can be changed at no cost. Due to information noise, a decision maker cannot perfectly observe the profit-maximizing aggregate price level denoted p_t. A small letter denotes the log-deviation of each variable from its non-stochastic value. Rational inattention argues that the aggregate price level is set to minimize the loss function, which is subject to the information flow constraint. The optimization problem is given.

$$\min_{p_t^*} E\left[\left(p_t - p_t^*\left(I(p_t;s_t)\right)\right)^2\right] \quad s.t. \quad I(p_t;s_t) = \kappa \tag{1}$$

$$where \quad \lim_{I\to\infty} p_t^*\left(I(p_t;s_t)\right) = p_t, \qquad {p_t^*}' > 0, \qquad {p_t^*}'' < 0$$

p_t	:	The profit maximizing aggregate price level
κ	:	Average information flow for aggregate conditions
$p_t^*\left(I(p_t;s_t)\right)$	:	Price expectation function.
$I(p_t;s_t)$	:	Mutual information[1]
$s_t = \{S_1, S_2, \cdots, S_n\}$	:	A signal set containing information about p_t.

[1] (See Appendix) Amount of information that a random variable contains about another. A large amount of such information exists when two variables are strongly dependent on each other.

S_i : A signal for observing p_t. A random variable.

Above, I reestablish the information constraint by signal, not signal set. Information theory defines that mutual information holds with

$$I(X; Y_1, Y_2, \cdots, Y_n) = I(X; Y_1) + I(X; Y_2|Y_1) + I(X; Y_3|Y_1, Y_2,) + \cdots + I(X; Y_n|Y_1, Y_2, \cdots, Y_{n-1},)$$

$I(X; Y_1, Y_2, \cdots, Y_n)$: Mutual information between X and $\{Y_1, Y_2, \ldots, Y_n\}$.

X : A random variable. Information we want.

Y_n : A random variable of signal n. A signal for X.

This expansion is called the chain rule for mutual information[2]. The amount of information X can be expressed as the sum of the conditional mutual information. This rule applies to my model. A firm receives the signals for observing a profit maximizing p_t. This gives

$$I(p_t; s_t) = I(p_t; S_{1t}, S_{2t}, \cdots, S_{nt})$$

[2] (See Appendix)

$$= I(p_t; S_{1t}) + I(p_t; S_{2t}|S_{1t}) + I(p_t; S_{3t}|S_{1t}, S_{2t},) + \cdots$$
$$+ I\left(p_t; S_{nt}|S_{1t}, S_{2t}, \cdots, S_{(n-1)t},\right)$$

If the signals follow a normal distribution with mean zero and variance $\sigma_{S_i}^2$ and are mutually independent, the mutual information constraint[3] becomes

$$I(p_t; s_t) = \frac{1}{2}\log\left(\frac{\sigma_p^2}{\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2}\right) + \frac{1}{2}\log\frac{\left(\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2\right)}{\left(\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2 - \sigma_{S_2}^2\beta_{S_2}^2\right)} + \cdots$$
$$+ \frac{1}{2}\log\frac{\left(\sigma_p^2 - \sum_{i=1}^{n-1}\sigma_{S_i}^2\beta_{S_i}^2\right)}{\left(\sigma_p^2 - \sum_{i=1}^{n}\sigma_{S_i}^2\beta_{S_i}^2\right)}$$
$$= \frac{1}{2}\log\frac{\sigma_p^2}{\left(\sigma_p^2 - \sum_{i=1}^{n}\sigma_{S_i}^2\beta_{S_i}^2\right)}$$

In line with the information flow constraint in the optimization problem, we obtain (2).

$$I(p_t; s_t) = \kappa$$
$$\sum_{i=1}^{n}\frac{\sigma_{S_i}^2}{\sigma_p^2}\beta_{S_i}^2 = (1 - 2^{-2\kappa}), \qquad p_t \sim N\left(0, \sigma_p^2\right), \qquad S_i \sim N\left(0, \sigma_{S_i}^2\right) \tag{2}$$

This shows that the information flow constraint can be represented by the regression coefficients of linear function if the signals and p_t follow normal

[3] (See Appendix) The log base equals two, following information theory manner.

distribution with the mean equal to zero and the signals are mutually independent. I assume that p_t is a linear function of aggregate variables. Therefore, p_t^* is estimated by a linear function model under the information flow constraint.

Here, I derive the regression model built on the rational inattention hypothesis. The firm minimizes the loss function by observing the signals, while the information flow constraint exists. As a result, the size of the regression coefficients is constrained.

$$\min_{\beta} E\left[\sum_{t=1}(p_t - p_t^*)^2\right], \qquad s.t.\ \sum_{i=1}^{n} \frac{\sigma_{S_i}^2}{\sigma_p^2} \beta_{S_i}^2 = (1 - 2^{-2\kappa}) \tag{3}$$

That is, the aggregate price level estimation model yields (4).

$$\min_{\beta} E\left[\sum_{t=1}\left(p_t - \sum_{i=1}^{n} \beta_i S_{it}\right)^2\right], \qquad s.t.\ \sum_{i=1}^{n} \frac{\sigma_{S_i}^2}{\sigma_p^2} \beta_i^2 = (1 - 2^{-2\kappa}) \tag{4}$$

By vector form,

$$p = S\beta + \varepsilon$$

$$p = [p_1 \quad \cdots \quad p_t]^T$$

$$\beta = [\beta_1 \quad \cdots \quad \beta_n]^T$$

$$S = \begin{bmatrix} S_{11} & \cdots & S_{n1} \\ \vdots & \vdots & \vdots \\ S_{1t} & \cdots & S_{nt} \end{bmatrix}$$

$$\sigma^2 = \begin{bmatrix} \frac{\sigma_{S_1}^2}{\sigma_p^2} & 0 & \cdots & 0 \\ 0 & \frac{\sigma_{S_2}^2}{\sigma_p^2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \frac{\sigma_{S_n}^2}{\sigma_p^2} \end{bmatrix}$$

The optimization problem is rewritten as:

$$\min_{\beta}(p - S\beta)^T(p - S\beta), \qquad s.t. \quad \beta^T\sigma^2\beta = K \tag{5}$$

$$L(\beta, \lambda) = (p - S\beta)^T(p - S\beta) + \lambda(\beta^T\sigma^2\beta \ - K)$$

$$\frac{\partial L}{\partial \beta} = -2S^T(p - S\beta) + 2\lambda\sigma^2\beta = 0$$

$$\beta = (\lambda\sigma^2 + S^T S)^{-1} S^T p \qquad (6)$$

$$\frac{\partial L}{\partial \lambda} = \beta^T \sigma^2 \beta - K = 0$$

$$K = p^T \mathrm{S}(\lambda\sigma^2 + S^T S)^{-1} \sigma^2 (\lambda\sigma^2 + S^T S)^{-1} S^T \qquad (7)$$

S_{it} : A signal (such as a GDP gap)

λ : A tuning parameter.

n : Number of signals

K : $K = (1 - 2^{-2\kappa})$

The estimators are biased because of the constraint in the information flow. The size of the regression coefficients changes in accordance with K. When K is small, that is information processing capacity κ is small, λ becomes large and vice versa. In addition, when the variable coefficient has a relatively large variance, it is more biased. Highly fluctuating variables require much more capacity to process the information.

Mackowiak and Wiederholt (2009) argues that firms allocate their limited attentions to a more volatile condition between an aggregate and firm idiosyncratic condition. Through this principle, firms attempt to minimize their loss attributed to the error caused by information noise. This model seeks to extend their findings that the noise trade-off exists even in the signal set, to which trade-off can be also applied by the regression model through the imperfect information bias. In the data science field, the penalized regression model is known as Ridge regression.

Furthermore, this paper provides evidence that when the economy fluctuates considerably, the bias is expected to be heavy. The regression coefficients shrink more when shocks occur to the aggregate conditions than in a stable period, and the gap between the OLS solution and the noise model solution expands. This will have a significant negative impact on the firm's forecast and the economy as a whole. This model proposes that it is beneficial to the economy that the central banks maintain a stable aggregate condition, as it can reduce losses due to misestimation of the economic forecast.

Practically, the Ridge regression is estimated below.

$$\underset{\beta}{\operatorname{argmin}} L = (p - S\beta)^T(p - S\beta) + \lambda(\beta^T\sigma^2\beta)$$

$$\frac{\partial L}{\partial \beta} = -2S^T(p - S\beta) + 2\lambda\sigma^2\beta = 0$$

$$\beta = (\lambda\sigma^2 + S^TS)^{-1}S^Tp$$

The estimated coefficients shrink according to the size of the tuning parameter, which I label as the imperfect information bias which arises from the constraint placed upon the information flow. If there is no information constraint, then λ equals zero, and the OLS solution is optimal. By contrast, when the information flow constraint, K, is small, then λ gives a large value.

In conclusion, rational inattention highlights that firms do not necessarily have an OLS solution due to the limitation of information flow. If the firms face the information flow constraint, the OLS estimators are biased and are less reliable. It is only optimal when there is no information constraint. Moreover, due to variance in the variable, the magnitude of the biases will vary. This model proposes that the stability of the aggregate condition plays a pivotal role in the economy.

3.2. Rational inattention in the New Keynesian Phillips Curve

This section extends the NKPC and relaxes FIRE, by introducing the rational inattention framework. The NKPC has two main features. The first is the forward-looking nature of inflation, and the second is real marginal cost as its determinant. This paper assesses the first feature of the NKPC.

The forward-looking nature builds on FIRE. That is, firms maximize their expected profits of discounted present value through use of their available information. While the model extensively contributes to the theoretical analysis, it tends not to be supported empirically. In general, the NKPC does not fit well into practical data. It is generally accepted that backward-looking elements have a significant effect on inflation in empirical analysis and therefore a flaw of the NKPC is that it does allow for backward-looking evidence to be incorporated into its model, which produces inaccuracies.

Therefore, my paper gives a proposed solution for the potential flaw illustrated above. Employing Calvo's price-setting model and following Gali (1999) for the baseline model, I will reveal the theoretical background of my perspective.

The baseline NKPC is as follow,

$$\pi_t = \beta E[\pi_{t+1}] + \alpha mc_t$$

β : A subjective discount factor

α : $= \frac{(1-\theta)(1-\beta\theta)}{\theta}$, θ is the frequency of price adjustment.

mc_t : Real marginal cost

In the estimation of the NKPC, it is known that there is an issue with the OLS, as the expected inflation cannot be observed. If we substitute $E[\pi_{t+1}]$ with π_{t+1}, then the estimation is not consistent. There is correlation between a leading indicator as a dependent variable and a regression error. Therefore, instrument variables (IV) are the standard method, considering the economic agents' behaviors as FIRE. To show IV validity, the expected inflation error is defined as:

$$\beta(E[\pi_{t+1}] - \pi_{t+1}) = \eta_{t+1}$$

The NKPC is rewritten as:

$$\pi_t = \beta\pi_{t+1} + \alpha mc_t + \eta_{t+1}$$

We need to construct an econometric forecast of the expected inflation. FIRE argues that firms form the expected inflation using all available information at t, and that is there is no information flow constraint.

$$E[\pi_{t+1}|I_t]$$

Our forecast derives from a subset of the information set, z_t, the IV, and can replace these as the representative information for the firm's expectation estimation by the law of iterated expectation.

$$E[E[\pi_{t+1}|I_t]|z_t] = E[\pi_{t+1}|z_t]$$

If FIRE holds true, the information set at t does not correlate with η_{t+1} and the orthogonal term holds.

$$E[I_t\eta_{t+1}] = 0$$
$$E[z_t\eta_{t+1}] = 0, \qquad z_t \in I_t$$

Consequently, the IV satisfies the consistency on the ground that the information vector and expected inflation error will not be correlated. However, if rational inattention holds, then they should show some correlation[4].

The expected inflation error is redefined in the light of the case when firms' behaviors proxy for the rational inattention.

$$\beta(E[\pi_{t+1}] - \pi^*_{t+1}) = \beta(E[\pi_{t+1}] - \pi_{t+1} + E[\pi_{t+1} - \pi^*_{t+1}]) \tag{8}$$

$$= \beta\left(\underbrace{E[\pi_{t+1}] - \pi_{t+1}}_{\eta_{t+1}} + \underbrace{E[\pi_{t+1} - \pi^*_{t+1}]}_{\varepsilon_{b_{t+1}}}\right)$$

$$= \eta_{t+1} + \varepsilon_{b_{t+1}} = \eta^{RI}_{t+1} \tag{9}$$

π^*_{t+1} : Biased inflation by information flow constraint.

π_{t+1} : Unbiased inflation by controlling the information flow constraint influence.

---

4 Coibion and Gorodnichenko (2015) showed that the noise information makes a positive correlation between the forecast error and the updated information from the preceding period.

η_{t+1} : Irreducible error.

$\varepsilon_{b_{t+1}}$: Error from imperfect information bias.

η_{t+1}^{RI} : Forecast error where the rational inattention hypothesis holds.

When the firm's information flow is not constrained, the error term has same value as the IV estimation under FIRE. Since full information has:

$$\pi_{t+1} = \pi_{t+1}^{*}$$

Then,

$$\beta(E[\pi_{t+1}] - \pi_{t+1}^{*}) = \eta_{t+1} = \eta_{t+1}^{RI}$$

Under the rational inattention hypothesis, the error term is composed of two elements. One is the irreducible error η_{t+1}, which is independent of the current information. The other is that the imperfect information bias error arises from the information flow constraint, which has correlation with the information at t. Here, we can see how it works:

$$E[\pi_{t+1} - \pi^*_{t+1}] = (\beta_{z_1} - \beta^*_{z_1})z_1 + (\beta_{z_2} - \beta^*_{z_2})z_2 + \cdots + (\beta_{z_n} - \beta^*_{z_n})z_n = \varepsilon_{b_{t+1}} \quad (10)$$

The value of $\varepsilon_{b_{t+1}}$ depends on which variables fluctuate. If a variable with a heavily biased coefficient fluctuates, then a significant error will result. The error of the expected inflation is projected to show a change in accordance with the information vector if the variables are not constant during the period. To be concise, the information observed until t has some correlation with the expected inflation error under rational inattention.

Thus, rational inattention implies that:

$$E[z_t \eta^{RI}_{t+1}] \neq 0$$

If the firms are limited in their information processing capacity, the standard IV method is not appropriate because it cannot satisfy the consistency. The estimation is not supportive of the empirical inflation mechanism interpretation. The lack of fit to data is likely to be attributed to the imperfection of the information. Therefore, this paper derives a combined NKPC model with rational inattention, which allows for the imperfection bias error measured from the Ridge regression.

The baseline NKPC is expected to be (11)

$$\pi_t = \beta\pi_{t+1}^{*} + \alpha mc_t + \eta_{t+1}^{RI} \tag{11}$$

To satisfy consistency, we need to obtain a controlled error.

$$\eta_{t+1} = E\left[\eta_{t+1}^{RI}\middle|\varepsilon_{b_{t+1}}\right]$$
$$E[z_t\eta_{t+1}] = E\left[z_t E\left(\eta_{t+1}^{RI}\middle|\varepsilon_{b_{t+1}}\right)\right] = 0$$

The controlled error can retain the orthogonal term. For this reason, the controlled NKPC gives an estimation with consistency. I denote rational inattention NKPC as “RINKPC”. This then becomes:

$$\pi_t = \beta\pi_{t+1}^{*} + \alpha mc_t + \varepsilon_{b_{t+1}} + \eta_{t+1} \tag{12}$$

Note that the imperfect information error serves as backward-looking, having a long run effect on inflation. For instance, Gali, Gertler and Lopez Salido’s (2005) NKPC model constructs a reduced form by four lags of inflation, two lags of marginal cost, detrended real output, and nominal wage inflation. The effect of instrument variables on inflation can be seen through the imperfect information bias error,

$$\varepsilon_{b_{t+1}} = \sum_{L=1}^{4} b_{\pi_L}\pi_{t-L} + \sum_{L=1}^{2}(b_{rmac_L}mc_{t-L} + b_{\tilde{x}_L}\tilde{x}_{t-L}+b_{w_L}w_{t-L}) \qquad (13)$$

$$where \quad b_{\pi_L} = \left(\beta_{\pi_L} - \beta^{*}_{\pi_L}\right)$$

The one period ahead from the one period behind and the two periods ahead and the one period behind influence inflation throughout the imperfect information bias error term.

$$\frac{\partial\varepsilon_{b_{t+1}}}{\partial\pi_{t-1}} = \frac{\partial\left(\beta_{\pi_1} - \beta^{*}_{\pi_1}\right)\pi_{t-1}}{\partial\pi_{t-1}} = b_{\pi_1}$$

$$\frac{\partial\varepsilon_{b_{t+2}}}{\partial\pi_{t-1}} = \frac{\partial\left(\beta_{\pi_2} - \beta^{*}_{\pi_2}\right)\pi_{t-1}}{\partial\pi_{t-1}} = b_{\pi_2}$$

These coefficients are the lag weights. It is shown that RINKPC has the backward-looking element. Throughout the imperfect information bias error term, the total change in π_t gives:

$$\Delta\varepsilon_{b_{t+1}} = \sum_{L=1}^{4} b_{\pi_L} + \sum_{L=1}^{2} (b_{mc_L} + b_{\tilde{x}_L} + b_{w_L}) \tag{14}$$

Inflation is affected by the backward-looking element due to the expected inflation error arising from the imperfect information bias. The information lags generate the inflation lags. If the economist includes the one period lag inflation in the estimation, then the NKPC becomes:

$$\pi_t = b_{\pi_1}\pi_{t-1} + \beta\pi^*_{t+1} + \alpha mc_t + \eta^{RI}_{t+1} \tag{15}$$

The backward-looking parameter represents the imperfect information bias error for the one period lag inflation. From the rational inattention hypothesis perspective, the hybrid NKPC only controls the imperfect information bias error with respect to the one period lag inflation. Including only the one lag inflation in the estimation as the backward-looking element cannot completely control the influence of all the past information, elements of the inflation inertia, so the hybrid NKPC is likely to have a biased estimator. Furthermore, the purely looking-forward NKPC is expected to be more biased than the hybrid NKPC, since that model does not control the error at all.

Mackowiak and Wiederholt (2009) theoretically prove that the information flow constraint creates a divergence between the perfect information estimation and the imperfect information estimation, which causes inflation inertia. My model explains the inflation inertia based on the microeconomic foundation of Mackowiak and Wiederholt (2009). Information noise prevents inflation from reflecting the shock instantly. The information lags lead to the inflation lags. The imperfect information bias represents such economic agents' norms and ends in creation of the shock path for the future from the past.

In the next section, I compare the forward-looking parameter coefficient between Gali, Gertler and Lopez Salido's (2005) NKPC model and the RINKPC model.

Empirical model

4.1 The construction of the RINKPC

Here, we see how to obtain the imperfect information bias. When conducting the Ridge regression, this gives an optimal coefficient set associated with λ and then extracts an optimal λ for comparison of fitting in with all sizes of it to the data[5].

$$\underset{\beta}{\operatorname{argmin}} E \sum_{t=1} [\pi_{t+1} - \hat{\pi}_{t+1}]^2 + \lambda \left(\sum_{L=1}^{4} \beta_{\pi_L}^2 + \sum_{L=1}^{2} \beta_{mc_L}^2 + \beta_{\tilde{x}_L}^2 + \beta_{w_L}^2 \right) \tag{16}$$

Gali, Gertler and Lopez Salido (2005) implemented an econometric forecast of π_{t+1} under FIRE. They employed four lags of inflation, and two lags of marginal cost, detrended real output, and nominal wage inflation for the π_{t+1} forecast. In my model, I have conducted the regression analysis both under FIRE and the information flow constraint. In line with the theoretical model, the expected inflation forecast uses OLS estimates. In addition, considering the information

---

[5] This study chose the tuning parameter λ by a cross validation set approach. The approach just extracts the optimal λ where mean squared error is minimized in the test data set. In this paper, the dataset is not large, therefore the estimation is likely to be overfitting. Extending the discussion on how to choose λ requires further research.

availability, my model employs the information observed at $t-1$ for the forecast.

Using two regression parameters, the OLS and the penalized OLS, the quarterly bias is then calculated:

$$\varepsilon_{b_{t+1}} = \sum_{L=1}^{4} b_{\pi_L}\pi_{t-L} + \sum_{L=1}^{2} b_{mc_L}mc_{t-L} + b_{\tilde{x}_L}\tilde{x}_{t-L} + b_{w_L}w_{t-L} \tag{17}$$

$$where \quad b_{\pi_L} = \left(\beta_{\pi_L}(\lambda = \lambda^*) - \beta^*_{\pi_L}(\lambda = 0)\right)$$

$\beta_{\pi_L}(\lambda = \lambda^*)$: penalized OLS solution

$\beta^*_{\pi_L}(\lambda = 0)$: OLS solution

The RINKPC is estimated by GMM with the imperfect information error term.

$$\pi^*_{t+1} = \sum_{L=1}^{4} \beta^*_{\pi_L}\pi_{t-L} + \sum_{L=1}^{2} \beta^*_{mc_L}mc_{t-L} + \beta^*_{\tilde{x}_L}\tilde{x}_{t-L} + \beta^*_{w_L}w_{t-L} \tag{18}$$

$$\pi_t = \beta\pi^*_{t+1} + \alpha mc_t + b\varepsilon_{b_{t+1}} + \eta_{t+1} \tag{19}$$

4.2. Data and Method

Following Gali and Gertler (1999) and Gali, Gertler and Lopez Salido (2005), the data period is quarterly between 1960:1 and 1997:4 in the US. These are obtained from the Federal Reserve Economic Data. The instrumental variables employ four lags of inflation, measured by GDP deflator, two lags of real unit labor costs as marginal cost, real output gap, and nominal wage inflation at one lag. This adopts the Hodrick-Prescott filter for detrending to the latter three variables and sets $\lambda = 1600$, as the data is quarterly observed.

Variable	Definition, Code (from FRED)
π_t	Implicit GDP deflator, [GDPDEF/FRED]
mc_t	Log real unit labor cost, [UCLNFB, GDPDEF/FRED] $\ln\left(\frac{ULC}{GDPDEF}\right)$
$\tilde{x}_t$	Log real GDP gap, [GDPC1/FRED]
w_t	Log nominal wage inflation, [COMPNFB/FRED]

Table 1 Data resource and variable definition

For the estimation, my paper uses a two-step GMM. The HAC weight matrix chooses Newey and West's with 12 lags based on those authors' past papers.

As for the imperfect information error, it is measured by the Ridge regression. The error transforms into an absolute value and takes the natural log. The implication is shown through how the percentage size of the change in the error affects the NKPC. The error estimation uses the adjusted variables at that NKPC estimation. The missing values need to be removed, so that data period is 1961:2 and 1997:3.

Empirical analysis results

5.

	$\boldsymbol{\beta}$	$\boldsymbol{\lambda}$	$\boldsymbol{b}$
NKPC	$\mathbf{0.972}^{***}$ $(\mathbf{0.010})$	$\mathbf{0.028}^{**}$ $(\mathbf{0.014})$	
RINKPC	$\mathbf{0.885}^{***}$ $(\mathbf{0.012})$	$\mathbf{0.043}^{***}$ $(\mathbf{0.013})$	$\mathbf{-0.000}^{***}$ $(\mathbf{0.000})$
Standards errors are shown in brackets. P-value: $* - \mathbf{10\%}$, $** - \mathbf{5\%}$, $*** - \mathbf{1\%}$			

Table 2 Estimates of NKPC and RINKPC

Overall, the imperfect information bias error significantly influences the NKPC. The first line estimates a purely forward-looking aspect. The parameter is consistent with the literature used for the same period as in the US. In the RINKPC model, the size of the forward-looking parameter decreases to approximately 0.087, while the size of the real marginal cost coefficient simultaneously increases to about 0.015. These results remove the biases from the forward-looking parameter and the real marginal cost, and support the need to control the imperfect information bias. FIRE is not consisted with the result. The information set at t has a significant correlation with the expected inflation error. The NKPC is biased on GMM if it does not control the imperfect information bias error.

Looking at b, it has a negative but almost zero impact on the inflation. The size of the error in itself should not influence the inflation. This is consistent with the theoretical perspective. This empirical analysis result indicates that the error has a significant impact on the inflation through other variables, since the expected inflation and the real marginal cost parameter changed in size, while the error term is estimated at zero. The backward-looking parameter potentially leads to an underestimation of the expected inflation, which leads to an underestimation of inflation. This mechanism is considered to make a hump-shaped response of inflation. The backward-looking element is not determinant of inflation, but just information lags.

In conclusion, when any shock hits the economy, and the variance of the aggregate variables grow exponentially, the amount of information flow increases accordingly. As a result, through this mechanism, the volatile state of the economy requires high information processing capacity for economic agents to grasp the true state. However, since our processing capacity is considered to remain unchanged, the imperfect information bias error is estimated to increase significantly. This means that the penalized OLS estimation of the expected inflation is much smaller than the stable period of that estimation. In case of the stable period, the same mechanism reduces the imperfect information error. In short, inflation indirectly depends on the past economic state, technically the past bias size by variable, through the inflation forecasting error arising from the imperfect information bias. Thus, inflation has the effect path that stems from the past economic state.

Changes in the economy cannot be processed instantly by economic agents due to the information flow constraint. This approach is expected to describe the reason why the real economy does not seem to reflect the true effect that an economic model based on FIRE indicates. The baseline NKPC and the inflation hump-shaped response relation is a typical example.

The imperfect information error has a significant effect on inflation. These results cast doubt on the validity of the stabilization policy implemented when the economy becomes unstable. For instance, despite implementation of an aggressive stabilization policy, the amount of the effect in the economy is presumed to be absorbed by the imperfect information error. However, a dynamic model has to be built to discuss this more technically.

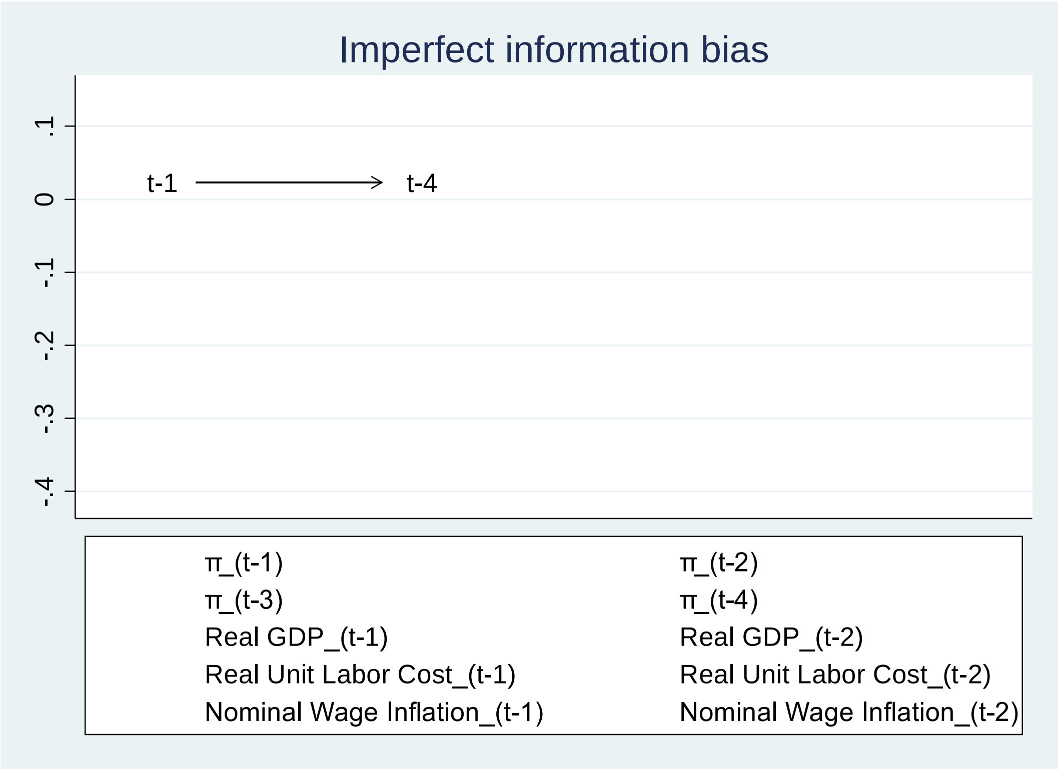

Figure 1 Imperfect information bias by variable (The left bar within a group shows t-1)

This graph illustrates the coefficients in (17). In the econometric estimation of the expected inflation, the real GDP and the real unit labor cost at the one period lag are under the OLS estimation value. In response, firms make adjustments to the gap at the next period, and they overestimate it to recover the underestimation. The behavior can be seen in the nominal wage inflation data with the same rationale but with the opposite bias coefficients. Inflation seems to exhibit a different trend from other criteria (such as real GDP). As time passes, the bias is reduced in size. These estimates simply continue to be underestimated compared with OLS until they finally converge to zero. This might reflect the firms' adjustment behavior mentioned above. Firms seem to adjust their misestimation at the next period, which reduces the noise in the inflation. However, this insight is not confirmed by the mathematical model. The imperfect information bias implication needs to be studied further with a technical mathematical model.

Conclusion

6.

This paper delivers two main conclusions. Firstly, noise information biases the OLS estimators. The Ridge regression has the potential to control the imperfect information bias in the estimation. That is, conducting the regression model allows the analysist to predict to a reasonable degree the implication of the noise information assumption. The method may be applicable to other economic fields. By implementing the penalized OLS, it is possible to compare the results of perfect information with noise information. This allows the user of the model to improve its predictability and consistency of the economic theory.

The regression model selects the most optimal coefficient set which fits the data. The discipline firms have with respect to the noise trade-off is not mentioned in this paper. According to the Mackowiak and Wiederholt (2009) model, firms allocate more attention to volatile conditions than the stable condition, as the allocation rule minimizes the expected loss. The rule is likely to be consistent with my model. The bias to variables with a relatively large variance is more than with other stable variables. A variable with a large variance expands the gap from the utility maximized level. Therefore, the model in this paper is consistent with the model in Mackowiak and Wiederholt (2009). Firms obey the same allocation attention rule as these authors if the rational inattention hypothesis holds. Further empirical evidence is needed to capture the economic agent's unique background in relation to the noise bias allocation.

The second conclusive assertion from this paper is that it provides evidence that the central bank should attempt to maintain and be responsible for maintaining a stable economy, because the size of the imperfect information bias depends on the state of the economy. A volatile condition requires the capacity for high information flow, which yields a large bias, and the economic agents consistently underestimate the impact of economic shock. As a result, they cannot reflect the whole economic policy effect. At an empirical level, the facts are confirmed. This paper constructed an RINKPC model, which controlled the imperfect information bias. The result does not support FIRE and suggests an amendment to factor in the noise information assumption.

RINKPC seems to be supported in the empirical result. The model theoretically incorporates the backward-looking element into inflation. RINKPC has the path of the effect from the past to the present by the imperfect information bias error term. The error term represents the past information imperfectness caused by the information flow constraint. The size of the error depends on the past economic state. If any shock hits the economy, the information flow increases, although the information processing capacity is constant in the economy as a whole. Ultimately, there is much noise in the information employed by firms. This mechanism creates a divergence between the baseline NKPC value and RINKPC. The baseline NKPC responds to the shock instantly, while the RINKPC responds gradually because of the noise error term. The expected inflation parameter in RINKPC changes in accordance with the size of the imperfect information bias error. In other words, the imperfect information error term alternatively absorbs the shock effect on the expected inflation. Thus, the backward-looking element, as an error term, has an indirect effect on inflation. This brings us to the conclusion that the backward-looking element is not determinant of inflation.

The NKPC used in this paper is a simple model. It will need to be consistently refined in order for it to achieve consensus, and to see the performance and consistency of the RINKPC, it needs to be extended using different data, instrumental variable sets, and methods. Until that stage is reached, the model will be less supportive.

The aim of this study was to apply data science tools to economics. Simply using data science tools, it falls in step with the Lucas critique. However, the method of data science can enhance data fitting in some cases, which implies that economic theory is likely to be missing something. There would appear to be value in developing economics through an approach utilizing data science tool structure.

References

Afrouzi, Hasson and Yang, Choongryul (2020) Dynamic Rational Inattention and the Philips Curve. CEFifo Working paper, Np.8840

Bank of Japan Research and Statistical Bureau Economic analysis group (2017) New fact regarding to firm expected inflation formation: Part 2 – machine learning approach-. Bank of Japan: Working paper series, No.17-J-4

Boneva, Lena Cloyne, James, Weale, Martin and Wieladek, Tomasz (2019) Firm Price, Cost and Activity Expectations: Evidence From Micro Data. The Economic Journal, 130:555-586

Coibion, Oliver and Gorodnichenko, Yuriy (2015) Information rigidity and the expectations formation process: a simple framework and new facts. American Economic Review, 105(8):2644-2678

Coibion, Oliver, Gorodnichenko, Yuriy and Kamdar, Rupal (2018) The Formation of Expectations, Inflation, and the Philips Curve. Journal of Economic Literature 56(4):1447-1491

Cover, Thomas M. and Thomas, Joy A. (2006) Differential Entropy. In: John Wiley & Sons, Inc. Elements of Information theory, Second Edition. Hodoken, N.J: 243 259

Fetzer, Thiemo (2022) Shrinkage, non-linear models and feature expansion. Application of Data science. University of Warwick: Topic 3

Gali, Jordi and Gertler, Mark (1999) Inflation dynamics: A structural econometric analysis. Journal of Monetary Economics 44(1999):195-222

Gali,Jordi Gertler, Mark and Lopez-Salido , J David (2005) Robustness of the estimates of the hybrid New Keynesian Philips curve. Journal of Monetary Economics 52(2005):1107-1118

Inamura, Kouki, Hiyama, Kasunori and Shiotani, Kyosuke (2017) Inflation outlook and business conditions of firms: Evidence from the tankan survey. IFC Bulletin 43(session 4 B):842

James, Gareth, Witten, Daniela, Hastie, Trevor and Tibshirani, Robert (2017) Tree-Based Methods. In: Springer Texts in Statistics. An Introduction to Statistical Learning: with application in R. New York: 303:335

Kitayama, Tomiyuki and Tanaka, Masaki (2019) Implication for expected inflation of rational inattention and sticky information. Bank of Japan: Working paper series, No. 19-J-10.

Mackowiak, Bartosz and Wiederholt, Mriko (2009) Optimal Sticky Prices under Rational Inattention. American Economic Review, 99(3):769-803

Mankiw, N. Gregory and Reis, Ricardo (2002) Sticky Information versus Sticky prices: A proposal to replace the New Keynesian Philips Curve. The Quarterly Journal of Economics, 117(4):1295-1328

Nason, James M. and Smith, Gregor W. (2006) Identifying the New Keynesian Philips Curve. Federal reserve of Atlanta. Working Paper, No 2005-1

Nason, James M. and Smith, Gregor W. (2008) The New Keynesian Philips Curve: Lessons From Single-Equation Econometric Estimation. Economic Quarterly, 94(4):361-395

Okuda, Tatushi (2018) Noise information mode and inflation dynamics. Kinyu Kenkyu 37(4):69-120

Sims, Christopher A (2003) Implications of Rational Inattention. Journal of Monetary Economics, 50(3):665-690

Uno, Yousuke, Naganuma, Saori and Hara, Naoko (2017) New fact regarding to firm expected inflation formation: Part 1- Review sticky information model-, Bank of Japan: Working paper series, No.17-J-3

Appendix

Here, I introduce some basic concepts from information theory. For a more technical proof, see the “Elements of Information Theory (2nd Edition)” (2006)

Entropy

Information theory defines the quantity of information as entropy.

The concept is a measure of uncertainty of a random variable.

$$h(X) = -\int f(x) \log_2 f(x) dx$$

$f(x)$: Probability density function. $\int_{-\infty}^{\infty} f(x) dx = 1$

Normal distribution $X \sim N(0, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-x^2}{2\sigma^2}}$$

$$\begin{aligned}
h(x) &= -\int f(x) \ln f(x) dx \\
&= \frac{1}{2\sigma^2} \int x^2 f(x) dx + \frac{1}{2} \ln 2\pi\sigma^2 \\
&= \frac{1}{2} + \frac{1}{2} \ln 2\pi\sigma^2 \\
&= \frac{1}{2} \ln 2\pi e\sigma^2 \\
&= \frac{1}{2} \log 2\pi e\sigma^2
\end{aligned}$$

In a multivariate normal distribution $X \sim N(\mu, K)$, K is a covariance matrix.

Probability density function of $X_1, X_2, \dots, X_n$

$$f(x) = \frac{1}{(\sqrt{2\pi})^n |K|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T K^{-1} (x-\mu)}$$

The entropy of a multivariate normal distribution is given as:

$$h(X_1 X_2, \dots, X_n) = \frac{1}{2} \log(2\pi e)^n |K|$$

Conditional entropy

$$
\begin{aligned}
h(X|Y) &= -\int f(y)f(x|y)\log f(x|y)dxdy \\
&= -\int f(x,y)\log f(x|y)dxdy \\
&= h(X,Y) - h(Y)
\end{aligned}
$$

Mutual information

A random variable can reduce the uncertainty of another random variable. Mutual information is considered to be the amount of information a random variable knows about another random variable. It is defined as

$$
\begin{aligned}
I(X;Y) &= \int f(x,y)\log\frac{f(x,y)}{f(x)f(y)}dxdy \\
&= h(X) - h(X|Y) \\
&= h(X) + h(Y) - h(X,Y)
\end{aligned}
$$

Conditional mutual information

This is mutual information between X and Z given Y. How much reduction in the uncertainty of X arises from Z given Y can be seen.

$$
\begin{aligned}
I(X;Z|Y) &= \iiint f(x,y,z) log \frac{f(x,z|y)}{f(x|y)f(z|y)} dxdydz \\
&= \iiint f(x,y,z) log \frac{1}{f(x|y)} \frac{f(x,y,z)}{f(y)} \frac{f(y)}{f(z|y)} dxdydz \\
&= -\iint f(x,y) log f(x|y) dxdy + \iiint f(x,y,z) log f(x|y,z)\, dxdydz \\
&= h(X|Y) - h(X|Y,Z)
\end{aligned}
$$

Chain rule for Mutual information

$$
\begin{aligned}
I(X;Z|Y) &= h(X|Y) - h(X|Y,Z) \\
&= -\big(h(X) - h(X|Y)\big) + h(X) - h(X|Y,Z) \\
&= -I(X;Y) + I(X;Y,Z) \\
I(X;Y,Z) &= I(X;Y) + I(X;Z|Y)
\end{aligned}
$$

For n random variables,

$$
\begin{aligned}
&I(X;Y_1,Y_2,\cdots,Y_n) \\
&\quad = I(X;Y_1) + I(X;Y_2|Y_1) + I(X;Y_3|Y_1,Y_2,) + I(X;Y_n|Y_1,Y_2,\cdots,Y_{n-1},)
\end{aligned}
$$

Proof of the mutual information constraint

$$
\begin{aligned}
(p_t;s_t) &= I(p_t;S_1,S_2,\cdots,S_n) \\
&= I(p_t;S_1) + I(p_t;S_2|S_1) + I(p_t;S_3|S_1,S_2,) + \cdots + I(p_t;S_n|S_1,S_2,\cdots,S_{n-1},)
\end{aligned}
$$

The signals follow normal distribution with mean zero and variance $\sigma_{S_i}^2$ and are mutually independent,

$$
\begin{aligned}
I(p_t; S_1) &= h(p_t) + h(S_1) - h(p_t, S_1) \\
h(p_t) &= \frac{1}{2}\log(2\pi e)\,\sigma_p^2 \\
h(S_1) &= \frac{1}{2}\log(2\pi e)\,\sigma_{S_1}^2 \\
h(p_t, S_1) &= \frac{1}{2}\log\,(2\pi e)^2 \begin{vmatrix} \sigma_p^2 & \beta_{S_1}\sigma_{S_1}^2 \\ \beta_{S_1}\sigma_{S_1}^2 & \sigma_{S_1}^2 \end{vmatrix}
\end{aligned}
$$

Thus, the mutual information is:

$$
I(p_t; S_1) = -\frac{1}{2}\log\left(\frac{\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2}{\sigma_p^2}\right)
$$

Other mutual information,

$$
\begin{aligned}
I(p_t; S_2|S_1) &= h(p_t|S_1) - h(p_t|S_1, S_2) \\
&= -h(p_t) + h(p_t|S_1) + h(p_t) - h(p_t|S_1, S_2) \\
&= h(p_t) + h(S_1, S_2) - h(p_t, S_1, S_2) - I(p_t; S_1) \\
&= \frac{1}{2}\log(2\pi e)\sigma_p^2 + \frac{1}{2}\log\,(2\pi e)^2 \begin{vmatrix} \sigma_{S_1}^2 & 0 \\ 0 & \sigma_{S_2}^2 \end{vmatrix} \\
&\quad - \frac{1}{2}\log\,(2\pi e)^3 \begin{vmatrix} \sigma_p^2 & \beta_{S_1}\sigma_{S_1}^2 & \beta_{S_2}\sigma_{S_2}^2 \\ \beta_{S_1}\sigma_{S_1}^2 & \sigma_{S_1}^2 & 0 \\ \beta_{S_2}\sigma_{S_2}^2 & 0 & \sigma_{S_2}^2 \end{vmatrix} - I(p_t; S_1)
\end{aligned}
$$

$$= \frac{1}{2}\log\left(\frac{\sigma_p^2}{\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2 - \sigma_{S_2}^2\beta_{S_2}^2}\right) + \frac{1}{2}\log\left(\frac{\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2}{\sigma_p^2}\right)$$

$$= \frac{1}{2}\log\frac{\left(\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2\right)}{\left(\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2 - \sigma_{S_2}^2\beta_{S_2}^2\right)}$$

Hence,

$$I(p_t; s_t) = \frac{1}{2}\log\left(\frac{\sigma_p^2}{\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2}\right) + \frac{1}{2}\log\frac{\left(\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2\right)}{\left(\sigma_p^2 - \sigma_{S_1}^2\beta_{S_1}^2 - \sigma_{S_2}^2\beta_{S_2}^2\right)} + \cdots$$

$$+ \frac{1}{2}\log\frac{\left(\sigma_p^2 - \sum_{i=1}^{n-1}\sigma_{S_i}^2\beta_{S_i}^2\right)}{\left(\sigma_p^2 - \sum_{i=1}^{n}\sigma_{S_i}^2\beta_{S_i}^2\right)}$$