# Rational Inattention

(prepared for The New Palgrave Dictionary of Economics)

Mirko Wiederholt

Northwestern University

August 2010

**Abstract**

Economists have studied for a long time how decision-makers allocate scarce resources. The recent literature on rational inattention studies how decision-makers allocate the scarce resource attention. The idea is that decision-makers have a limited amount of attention and have to decide how to allocate it. The literature on rational inattention argues that the optimal allocation of attention by decision-makers can explain important features of economic data.

The idea of rational inattention is that individuals have a limited amount of attention and therefore have to decide how to allocate their attention.

There is a vast amount of information that is, in principle, available to decision-makers (e.g., all information published in books, magazines, newspapers, and scientific articles; all information available on the internet; knowledge available through colleagues, friends and family), but due to limited attention it is simply impossible to attend to all of this information. Therefore, decision-makers have to choose which information to attend to carefully, which information to attend to less carefully, and which information to ignore. According to the theory of rational inattention, decision-makers take this decision optimally. The literature on rational inattention argues that the optimal allocation of attention by decision-makers can explain important features of economic data.

# 1 Modeling limited attention

Christopher A. Sims proposed to model attention as an information flow and to model limited attention as a bound on information flow. See Sims (1998, 2003). To implement this idea, one has to quantify information flows. Sims (1998, 2003) suggested following a large literature in information theory by quantifying information as reduction in uncertainty, where uncertainty is measured by entropy.

Let us illustrate these concepts with a simple example. Entropy is a measure of uncertainty. The entropy of a normally distributed random variable $X$ equals

$$H(X) = \frac{1}{2}\log_2\left(2\pi e\sigma_X^2\right),$$

where $\sigma_X^2$ denotes the variance of $X$. Conditional entropy is a measure of conditional uncertainty. The conditional entropy of $X$ given $S$, when $X$ and $S$ have a multivariate normal distribution, equals

$$H(X|S) = \frac{1}{2}\log_2\left(2\pi e\sigma_{X|S}^2\right),$$

where $\sigma_{X|S}^2$ denotes the conditional variance of $X$ given $S$. Equipped with measures of uncertainty and conditional uncertainty, one can quantify the information that one random variable contains about another random variable as reduction in uncertainty. For example, the amount of information that $S$ contains about $X$ equals

$$I(X;S) = H(X) - H(X|S).$$

Think of $X$ as a variable that a decision-maker may be interested in. Paying attention to the variable $X$ can be modeled as receiving a signal $S = X + \varepsilon$ where the noise $\varepsilon$ is interpreted as coming from the decision-maker's limited attention and is assumed to be independent of $X$ and normally distributed with mean zero and variance $\sigma_\varepsilon^2$. The idea is that limited attention leads to a noisy perception of the true realization of $X$. Limited attention can be modeled as a bound on information flow

$$I(X;S) \leq \kappa.$$

The constraint on information flow implies

$$\frac{\sigma_X^2}{\sigma_{X|S}^2} \leq 2^{2\kappa},$$

or equivalently

$$\frac{\sigma_X^2}{\sigma_\varepsilon^2} \leq 2^{2\kappa} - 1.$$

In this simple example, limited attention simply imposes a bound on variance reduction. Equivalently, limited attention simply imposes a bound on the signal-to-noise ratio in the signal concerning $X$. The noise in the signal is interpreted as arising from the decision-maker's own nervous system.

One advantage of measuring uncertainty by entropy is that entropy also summarizes in a single number the uncertainty associated with a multivariate distribution. For example, the entropy of a $n$-dimensional random vector $X$ that has a multivariate normal distribution equals

$$H(X) = \frac{1}{2}\log_2\left[(2\pi e)^n \det \Omega_X\right],$$

where $\det \Omega_X$ denotes the determinant of the covariance matrix of $X$. Think of $X$ as a vector of variables that a decision-maker may be interested in. Paying attention to the vector $X$ can be modeled as receiving an $m$-dimensional signal $S$ with the property that $X$

and $S$ have a multivariate normal distribution. Limited attention can again be formalized as a bound on information flow

$$H(X) - H(X|S) \leq \kappa.$$

The constraint on information flow now implies

$$\frac{\det \Omega_X}{\det \Omega_{X|S}} \leq 2^{2\kappa},$$

where $\Omega_{X|S}$ denotes the conditional covariance matrix of $X$ given $S$. One can then ask how the decision-maker would want to use the available information flow. What is the optimal dimension of $S$? Which elements of $X$ would the decision-maker want to learn about? Would the decision-maker want to learn about linear combinations of elements of $X$?

Moreover, entropy can be computed for discrete and continuous distributions. Let $X$ be a discrete random variable with support $\mathcal{X}$ and probability mass function $p(x) = \Pr\{X = x\}$. Then the entropy of $X$ equals

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x).$$

Paying attention can again be modeled as receiving a signal $S$ and limited attention can be modeled as a bound on information flow

$$H(X) - H(X|S) \leq \kappa.$$

One can once more ask how the decision-maker would want to use the available information flow. Is it optimal to find out only whether $X$ is above or below a certain level? What is the optimal form of the joint distribution of $X$ and $S$?

For a different approach to modeling attention, see Reis (2006). In his work, paying attention is modeled as incurring a fixed cost and then learning everything perfectly.

## 2 Rational inattention

Let us now study a rational inattention problem. In particular, let us study the problem of a decision-maker who is responsible for setting a price and has to decide how to allocate his or her attention. Let $p_i$ denote the price of good $i$. Setting a price $p_i$ that differs from the profit-maximizing price $p_i^*$ causes a profit loss equal to $\frac{\omega}{2}\left(p_i - p_i^*\right)^2$. The profit-maximizing price equals $p_i^* = \phi x$ where $x$ is a normally distributed random variable with mean zero and variance $\sigma_x^2$. Here $\omega$ and $\phi$ are parameters. Paying attention to the variable $x$ is modeled as receiving a signal $s_i = x + \varepsilon_i$ where the noise $\varepsilon_i$ is independent of $x$ and normally distributed with mean zero and variance $\sigma_\varepsilon^2$. The decision-maker chooses the amount of attention $\kappa$ devoted to the variable $x$. The decision-maker faces a marginal cost of attention $\mu > 0$. This cost can be interpreted as the opportunity cost of devoting some of the scarce resource attention to the variable $x$. Formally, the decision-maker solves

$$\min_{\kappa \geq 0} \left\{ \frac{\omega}{2} E\left[\left(p_i - p_i^*\right)^2\right] + \mu\kappa \right\},$$

subject to $p_i^* = \phi x$, $s_i = x + \varepsilon_i$, $p_i = E\left[p_i^* | s_i\right]$, and

$$\frac{1}{2}\log_2\left(\frac{\sigma_x^2}{\sigma_{x|s}^2}\right) \leq \kappa.$$

The optimal attention devoted to the variable $x$ equals

$$\kappa^* = \begin{cases} \frac{1}{2}\log_2\left(\frac{\omega\phi^2\sigma_x^2\ln(2)}{\mu}\right) & \text{if } \frac{\omega\phi^2\sigma_x^2\ln(2)}{\mu} \geq 1 \\ 0 & \text{otherwise} \end{cases};$$

and the price set by the decision-maker equals

$$p_i = \left(1 - 2^{-2\kappa^*}\right)\phi\left(x + \varepsilon_i\right).$$

The ratio $\frac{\omega\phi^2\sigma_x^2\ln(2)}{\mu}$ is the marginal benefit of devoting attention to the variable $x$ at $\kappa = 0$ divided by the marginal cost of devoting attention to the variable $x$. If this ratio exceeds one, the decision-maker devotes some attention to the variable $x$. The larger the cost of a price setting mistake and the larger the variance of the profit-maximizing price due to the variable $x$ (i.e., the larger $\omega$ and $\phi^2\sigma_x^2$), the more attention is devoted to the variable $x$ and the stronger is the response of the price $p_i$ to changes in $x$.

The example given above is static. One can make the example dynamic by introducing time and by specifying the stochastic process $\{x_t\}_{t=0}^{\infty}$. One can then solve for the optimal signal process $\{s_{i,t}\}_{t=0}^{\infty}$, subject to a constraint on information flow. If the variable $x_t$ follows a stationary first-order autoregressive process, it is optimal to receive a signal of the form $s_{i,t} = x_t + \varepsilon_{i,t}$, where the noise term $\varepsilon_{i,t}$ is independent across time. See Maćkowiak and

Wiederholt (2009), Propositions 3 and 4. The attention devoted to the variable $x_t$ is again increasing in $\phi^2\sigma_x^2$. Furthermore, the more attention is devoted to the variable $x_t$, the smaller is the variance of noise in the signal and the faster is the response of the price $p_{i,t}$ to changes in $x_t$. Maćkowiak and Wiederholt (2009) argue that this can explain why prices respond quickly to idiosyncratic shocks and slowly to aggregate shocks. If there is a trade-off between attending to idiosyncratic conditions and attending to aggregate conditions and idiosyncratic conditions are more important or more volatile, then price setters devote more attention to idiosyncratic conditions and prices respond faster to idiosyncratic shocks.

In the example given above, the signal $s_i$ and the external state $x$ are assumed to have a multivariate normal distribution, implying that the action $p_i$ and the external state $x$ have a multivariate normal distribution. Sims (2006) argues that an agent with rational inattention will also choose the optimal form of the joint distribution of the action and the external state, subject to the constraint on information flow. Sims (2006) and Tutino (2009) study this question in the case of consumption saving problems. Maćkowiak and Wiederholt (2009), Woodford (2009) and Matejka (2010) study this question in the case of price setting problems. When the decision-maker's objective is quadratic and the external state has a normal distribution, a normally distributed signal turns out to be optimal.

Sims' idea of rational inattention has been applied to a variety of different decision problems, not only to price setting problems. Sims (2003, 2006), Luo (2008) and Tutino (2009) study consumption saving problems with a constant interest rate. Maćkowiak and Wiederholt (2010) study a consumption saving problem with a variable interest rate. Van Nieuwerburgh and Veldkamp (2009) and Mondria (2010) study portfolio choice under rational inattention. Maćkowiak and Wiederholt (2010) solve a dynamic stochastic general equilibrium model with rational inattention on the side of decision-makers in firms and households.

Much remains to be done. There are numerous potential applications of the idea of rational inattention, but so far the theory of rational inattention has only been applied to price setting, consumption, portfolio choice, and wage setting. Furthermore, it is interesting to test models of rational inattention empirically. See Maćkowiak, Moench and Wiederholt (2009) and Kacperczyk, Van Nieuwerburgh and Veldkamp (2010) for empirical tests of rational inattention theories of price setting and of portfolio choice. Finally, it seems interesting to study policy implications of models with rational inattention. Maćkowiak and Wiederholt (2010) and Paciello (2010) conduct monetary policy experiments in models with rational inattention, but fiscal policy, for example, has not been studied yet in models with rational inattention.

## References


[1] Kacperczyk, M., S. Van Nieuwerburgh and L. Veldkamp. 2010. Attention Allocation over the Business Cycle. Manuscript.

[2] Luo, Y. 2008. Consumption Dynamics under Information Processing Constraints. Review of Economic Dynamics 11(2), 366-385.

[3] Maćkowiak, B., E. Moench and M. Wiederholt. 2009. Sectoral Price Data and Models of Price Setting. Journal of Monetary Economics 56(S), 78-99.

[4] Maćkowiak, B. and M. Wiederholt. 2009. Optimal Sticky Prices under Rational Inattention. American Economic Review 99(3), 769-803.

[5] Maćkowiak, B. and M. Wiederholt. 2010. Business Cycle Dynamics under Rational Inattention. CEPR Discussion Paper 7691.

[6] Matejka, F. 2010. Rationally Inattentive Seller: Sales and Discrete Pricing. Manuscript.

[7] Mondria, J. 2010. Portfolio Choice, Attention Allocation, and Price Comovement. Manuscript.

[8] Paciello, L. 2010. Monetary Policy Activism and Price Responsiveness to Aggregate Shocks under Rational Inattention. Manuscript.

[9] Reis, R. 2006. Inattentive Producers. Review of Economic Studies 73(3), 793-821.

[10] Sims, C.A. 1998. Stickiness. Carnegie-Rochester Conference Series on Public Policy 49, 317-356.

[11] Sims, C.A. 2003. Implications of Rational Inattention. Journal of Monetary Economics 50(3), 665-690.

[12] Sims, C.A. 2006. Rational Inattention: Beyond the Linear-Quadratic Case. American Economic Review Papers and Proceedings 96(2), 158-163.

[13] Tutino, A. 2009. The Rigidity of Choice. Lifecycle Savings with Information Processing Limits. Manuscript.

[14] Van Nieuwerburgh, S. and L. Veldkamp. 2009. Information Immobility and the Home Bias Puzzle. Journal of Finance 64(3), 1187-1215.

[15] Woodford, M. 2009. Information-Constrained State-Dependent Pricing. Journal of Monetary Economics 56(S), 100-124.