A Theory of Deflation: Can Expectations Be Influenced by a Central Bank?

Harashima, Taiji

Kanazawa Seiryo University

15 May 2016

A Theory of Deflation:
Can Expectations Be Influenced by a Central Bank?

Taiji HARASHIMA*

May 2016

Abstract

This paper examines how to reverse deflation to inflation. Once deflation takes root, it is not easy to reverse because of the zero lower bound in nominal interest rates. My model indicates that there are two steady states where both inflation/deflation (i.e., changes in prices) and real activity (i.e., quantities) remain unchanged: that is, there are inflationary and deflationary steady states. The model indicates that, to switch a deflationary steady state to an inflationary steady state, a central bank needs to influence the time preference rates of the government and the representative household. It is not easy, however, to do so, and the best way of switching deflation to inflation may be to wait for a lucky event (i.e., an exogenous shock).

JEL Classification code: E31, E52, E58
Keywords: Deflation; The zero lower bound; Monetary policies; Quantitative easing; Time preference

---

*Correspondence: Taiji HARASHIMA, Kanazawa Seiryo University, 10-1 Goshomachi-Ushi, Kanazawa-shi, Ishikawa, 920-8620, Japan.
Email: harashim@seiryo-u.ac.jp or t-harashima@mve.biglobe.ne.jp.

1 INTRODUCTION

Reversing deflation to inflation has been an important policy issue, especially since the 1990s, because some economies have experienced deflation (although temporary) and faced the zero lower bound (ZLB) in nominal interest rates. Although inflation is now tamed in most developed countries, deflation remains a concern. For example, deflation has haunted Japan since the 1990s, although the Bank of Japan has repeatedly tried to reverse the course of deflation, even by using unconventional monetary policies. Once deflation takes root, however, it is not easy to reverse because of ZLB. If nominal interest rates are stuck at or near zero, conventional monetary policies (i.e., manipulations of nominal interest rates by a central bank) are little effective.

An alternative tool to reverse deflation is needed. An important prospective alternative tool is to influence households' expectations (e.g., Eggertsson and Woodford, 2003; Bernanke and Reinhart, 2004; Bernanke et al., 2004; Blinder et al., 2008). However, theoretically, the effectiveness of this measure is ambiguous. In this paper, I examine the feasibility and effectiveness of manipulating expectations on the basis of the inflation/deflation model shown by Harashima (2004b, 2008, 2015a), as well as conventional inflation models—in particular, new-Keynesian Phillips curve (NKPC) models.

The inflation/deflation model of Harashima (2004b, 2008, 2015a) is based on a micro-foundation of trend inflation. It indicates that trend inflation is generated by the difference between the rate of time preference (RTP) of the government and that of the representative household (RTP RH). In addition, both the RTP of the government and the RTP RH are intrinsically temporally variable. Hence, the expectations of both RTPs' future values must be generated by households. The feasibility and effectiveness of expectation manipulation therefore depends on whether a central bank can influence the RTP expectations of both the government and the RH.

Even if nominal interest rates are bounded by ZLB, an economy can be stable if the absolute value of the deflation rate equals the real rate of interest, because the Fisher equation is satisfied. If households generate expectations of the government's RTP and the RTP RH that are consistent with this deflation rate, the economy will be stable. To reverse deflation to inflation, therefore, households' expectations of deflation need to be changed. There are several possible ways to influence households' expectations: verbal intervention, quantitative easing (QE), renouncing central bank independence, raising the nominal interest rate, and depreciating the exchange rate. In addition, imposing taxes on money may be used to reverse deflation, although the main aim in that case is not to influence households' expectations but to recover the ability to manipulate nominal interest rates. In this paper, I examine and evaluate the feasibilities and effectiveness of these measures. The results of examinations indicate that it is not easy for a central bank to reverse deflation to inflation by influencing households' expectations. This conclusion suggests that as inflation shows signs of changing to deflation, it is important for a central bank to act quickly and drastically, for example, by increasing the target rate of inflation and lowering nominal interest rates far more than it usually would (e.g., Williams, 2009).

2 A MECHANISM OF DEFLATION

2.1 The law of motion for inflation/deflation

2.1.1 The law of motion

The model constructed by Harashima (2004b, 2008, 2015a) is used as the inflation/deflation model in this paper. The details of the model are explained in Appendix A. The difference between it and conventional inflation models—particularly NKPC models—is discussed in

Section 2.1.5. The model indicates that the law of motion for inflation/deflation is described by

$$\int_{t-1}^{t}\int_{s}^{s+1}\pi_{\upsilon}\,d\upsilon\,ds=\pi_t+\theta_G-\theta_P \tag{1}$$

where π_t is the inflation/deflation rate at time t, θ_G is the RTP of government, and θ_P is RTP RH. θ_G and θ_P are not necessarily identical. Equation (1) is the same as equation (A19) in Appendix A. It indicates that inflation/deflation accelerates or decelerates as a result of the government and the representative household reconciling the contradiction in their heterogeneous RTPs. A solution of the integral in equation (1) for given θ_G and θ_P is

$$\pi_t=\pi_0+6(\theta_G-\theta_P)t^2 \quad . \tag{2}$$

Generally, the path of inflation/deflation that satisfies equation (1) for $0\leq t$ is expressed as

$$\pi_t=\pi_0+6(\theta_G-\theta_P)\exp[z_t\ln(t)] \quad ,$$

where z_t is a time-dependent variable. The stream of z_t varies depending on the boundary condition. However, if π_t satisfies equation (1) for $0\leq t$, and $-\infty<\pi_t<\infty$ for $-1<t\leq 1$, then

$$\lim_{t\to\infty} z_t=2 \quad .$$

The proof is shown by Harashima (2008). Any inflation/deflation path that satisfies equation (1) for $0\leq t$ therefore asymptotically approaches the path of equation (2).

Equation (2) indicates a trend component in inflation/deflation. In addition to this trend element, actual inflation/deflation will be influenced by output gaps and various disturbances in the short run; thus, an aggregate supply equation (a Phillips curve) consists of those elements as well as the trend component. The model of inflation therefore consists of an aggregate supply equation (which consists of equation [2] as the trend component, output gaps, and various disturbances), an aggregate demand equation, and an instrument rule for the central bank's manipulation of nominal interest rates (e.g., a Taylor rule). See Harashima (2008) for the detailed model structure.

If a central bank is sufficiently independent, it can force the government to change θ_G and achieve $\theta_G = \theta_P$. If $\theta_G = \theta_P$ is kept, inflation/deflation neither accelerates nor decelerates by equation (1). In other words, to stabilize inflation, an independent central bank will punish the government (i.e., force it to change its preference θ_G) if θ_G deviates from θ_P.

2.1.2 Stationarity

Output gaps and disturbances are basically stationary processes. Therefore, if $\theta_G = \theta_P$ is kept, $\pi_t - \pi_0$ is basically a stationary process with a mean of 0; that is, inflation/deflation becomes a stationary process with mean π_0 because, if $\theta_G = \theta_P$, the trend component disappears (i.e., $\pi_t = 0$ in equation [2]). For example, when $\theta_G = \theta_P$ begins to be kept, π_t is negative (e.g., –0.8% or $\pi_0 = -0.008$). An average deflation rate of –0.8% then will continue because $\theta_G = \theta_P$ is kept. The deflation rate may frequently temporarily deviate from –0.8% because of various shocks, but it will soon return to –0.8%.

Trends or unit roots in inflation were clearly observed during the great inflation in the 1960s and 1970s, but they are not clear in the current periods of low inflation. The reason for this lack of clarity (in other words, stationarity) is that central banks currently are sufficiently independent and $\theta_G = \theta_P$ is always kept.

2.1.3 Necessity of generating an expected θ_G

All households behave (i.e., choose their optimal paths) on the basis of the expectation of future inflation/deflation. The model shown in Section 2.1.1 indicates that, to expect future inflation/deflation, households must know the future value of the government's RTP (θ_G). There is, however, no guarantee that θ_G and RTP RH will be constant across time; rather, the RTPs of the government and households will be intrinsically temporally variable. However, households cannot directly know even the current value of θ_G, because households and the government are different entities and do not inherently know each other's preferences. Therefore, households must somehow generate expectations of the future values of θ_G by calculating them using a structural model of the government's RTP, but they first must construct such a model. A model of the government' RTP and the various problems that are created when generating the expected RTP of government are presented by Harashima (2015b) and also in Appendix B.

Equation (2) indicates that π_t depends on θ_G. Therefore, households need to generate an expected inflation/deflation by generating an expected θ_G. That is, the need to generate an expected inflation/deflation necessitates the expectation of θ_G. This means that, if a central bank can influence the expected θ_G, it can also influence the expected inflation/deflation.

2.1.4 Necessity of generating an expected θ_P

As Becker (1980) and Harashima (2014a, b) indicate, it is not possible to assume that the representative household is the same as the average household in dynamic models. Harashima (2014a, b) shows an alternative definition of the representative household such that the behavior of the representative household is defined as the collective behavior of all households under sustainable heterogeneity. The reason why this alternative definition is needed, and the nature of sustainable heterogeneity, are shown in detail in Appendix C. Unlike the case in which the representative household is assumed to be the average household, this alternatively defined representative household reaches a steady state in which all households satisfy all of their optimality conditions in dynamic models, even if the households are heterogeneous. In addition, the alternatively defined representative household has an RTP that is equal to the average RTP as shown in equations (C7) and (C8) in Appendix C. This alternatively defined representative household requires that each household must generate its expected RTP RH *ex ante* for it to behave optimally, as shown in Appendix C (see also Harashima, 2014a, b).

Equation (2) indicates that π_t depends not only on θ_G but also on θ_P. Therefore, households need to generate their expected levels of inflation/deflation by generating not only an expected θ_G but also an expected θ_P. Similar to the case of θ_G, if a central bank can influence the expected θ_P, it can also influence the expected inflation/deflation.

2.1.5 Comparison with conventional inflation models

A typical hybrid NKPC (e.g., Galí et al., 2005), is

$$\pi_t = \alpha_{\pi+1}\pi_{t+1|t} + \alpha_{\pi-1}\pi_{t-1} + \alpha_x x_t + \varepsilon_t \tag{3}$$

where x_t is the output gap and ε_t is an i.i.d. shock with a zero mean at time t, $\pi_{t+1|t}$ π_t is the rate of inflation at time t + 1 expected at time t, and $\alpha_{\pi+1}$, $\alpha_{\pi-1}$, and α_x are constant coefficients. Hybrid NKPC inflation models consist of an aggregate supply equation (a Phillips curve) such as that expressed in equation (3), an aggregate demand equation, and an instrument rule for the central bank's manipulation of nominal interest rates (e.g., a Taylor rule).

An important difference between the model shown in Section 2.1.1 and the hybrid NKPC inflation model shown here is whether or not a mechanism that generates a trend or unit root is explicitly incorporated on the basis of a micro-foundation. In the model in Section 2.1.1, a trend or unit root is generated naturally if $\theta_G \neq \theta_P$ by the law of motion for inflation/deflation. In the hybrid NKPC model, however, an *ad hoc* inclusion of a backward-looking element (π_{t-1}) is needed to generate a trend or unit root.

Nevertheless, if the backward-looking element (π_{t-1}) is excluded from equation (3) (i.e., the hybrid NKPC is reduced to a pure NKPC), inflation also shows stationarity when $\hat{\theta}_G = \hat{\theta}_P$, where $\hat{\theta}_G$ is the expected θ_G and $\hat{\theta}_P$ is the expected θ_P. Hence, if $\hat{\theta}_G = \hat{\theta}_P$, the performances of the model in Section 2.1.1 and the pure NKPC model will be almost the same. However, there is still an important difference between the two models. The expected inflation/deflation depends on the expectations of θ_G and θ_P in the model in Section 2.1.1, whereas the expected inflation/deflation depends on expected future disturbances in the pure NKPC model.

2.2 The zero lower bound (ZLB)

Assume that the central bank is sufficiently independent and capable. Therefore, the central bank always keeps $\hat{\theta}_G = \hat{\theta}_P$ and inflation/deflation is a stationary process with mean π_0. The assumption of a sufficiently independent central bank is natural in most developed countries at the present time.

Suppose that there is a downward shock on expected θ_P and then $\hat{\theta}_G > \hat{\theta}_P$ in equation (2). By the law of motion for inflation/deflation, inflation begins to accelerate (or deflation begins to decelerate). To restore $\hat{\theta}_G = \hat{\theta}_P$, the central bank forces the government to lower θ_G by increasing nominal interest rates, and the government has no choice but to lower θ_G. As a result, $\hat{\theta}_G = \hat{\theta}_P$ is soon attained. However, because of the shock on the expected θ_P and the consequent change of θ_G, π_0 changes. To achieve $\pi_0 = \pi^*$, where π^* is the central bank's target rate of inflation, therefore, the central bank needs to further manipulate nominal interest rates and θ_G. After achieving $\pi_0 = \pi^*$, the central bank keeps $\hat{\theta}_G = \hat{\theta}_P$.

Now suppose that there is an upward shock on expected θ_P, and then $\hat{\theta}_G < \hat{\theta}_P$ in equation (2). By the law of motion for inflation/deflation, inflation begins to decelerate (or deflation begins to accelerate). To restore $\hat{\theta}_G = \hat{\theta}_P$, the central bank forces the government to raise θ_G by decreasing nominal interest rates. The government has no choice but to raise θ_G, but the recovery of $\hat{\theta}_G = \hat{\theta}_P$ is not necessarily guaranteed, because the central bank cannot lower nominal interest rates below the ZLB.[1] If nominal interest rates are bounded by the ZLB, the central bank no longer has the power to force the government to raise θ_G by manipulating nominal interest rates. Unlike the case of a downward shock on the expected θ_P, the capability of the central bank is constrained by the ZLB. As a result, $\hat{\theta}_G = \hat{\theta}_P$ is not necessarily restored and $\hat{\theta}_G < \hat{\theta}_P$ may continue. In this case, inflation eventually changes to deflation and deflation accelerates.

2.3 Households' expectation of $\theta_G = \theta_P$

How do households think the government behaves when nominal interest rates are bounded by the ZLB? They may perceive that, because the central bank is now powerless, the government will freely choose θ_G, and that it will behave on the basis of its intrinsic θ_G. They may also think

[1] Technically, central banks can make nominal interest rates slightly negative as the Bank of Japan and the European Central Bank did. Of course, however, these slightly negative interest rates do not indicate that ZLB does not exist.

that θ_G is still influenced by the central bank in the sense that, when deflation eventually changes to inflation, the central bank will certainly and immediately resume control of θ_G by manipulating nominal interest rates. For example, if the government behaves on the basis of intrinsic θ_G, deflation soon changes to inflation because $\theta_G > \theta_P$ intrinsically. Therefore, manipulations of nominal interest rates become effective again and θ_G can be controlled by the central bank as it was before. This second view indicates that the ZLB does not mean that the government can freely act on its own intrinsic RTP forever, but that it can merely temporarily escape from the discipline of the central bank. Which view is correct?

2.3.1 Habituation

People dislike changing their own preferences. Hence, people and governments feel psychological pain and disutility if their preferences are forced to change. On the other hand, studies on habituation in psychology (e.g., Thompson and Spencer, 1966; Groves and Thompson, 1970; Rankin et al., 2009) imply that once a preference is changed, the initial psychological pain will gradually subside as the person acclimates to the change: that is, the psychological pain will persist but dwindle. According to Rankin et al. (2009), one of the common characteristics of habituation is that repeated application of a stimulus results in a progressive decrease in some parameter of a response to an asymptotic level. Although the psychological pain or a feeling of wrongness may never disappear completely, the level of psychological pain will gradually recede as time passes. Huge initial psychological pains and later acclimation will have an important effect on the discipline of the central bank, because the government will decide its behavior on the basis of its own expectations.

2.3.2 Inhibitory effect

Suppose that a government is considering whether it would be better to behave on its intrinsic θ_G (i.e., $\bar{\theta}_G$) or on the θ_G that is equal to θ_P when nominal interest rates are zero. Unlike in "usual" inflation periods, the government has a choice, because nominal interest rates are zero. If it chooses to behave on $\bar{\theta}_G$, it can relieve the persisting but dwindling psychological pain (disutility) caused by the last forced change in θ_G (i.e., by the last punishment), but it will suffer great psychological pain in the near future because behaving on $\bar{\theta}_G$ indicates the reversal of deflation to inflation (because $\bar{\theta}_G > \theta_P$ by nature) in the near future and the consequent resumption of enforcement (or punishment) by the independent central bank to change θ_G from $\bar{\theta}_G$ to θ_P. The government can enjoy its intrinsic preference only for a short period and soon will have to endure great psychological pains again. On the other hand, if the government continues to behave on $\theta_G = \theta_P$, it will not suffer great psychological pains in the near future even though it has to continue feeling the relatively small and subsiding psychological pains caused by the last forced change in θ_G (i.e., by the last punishment).

Which of the two options a government chooses depends on the level of the initial psychological pain and that of the dwindling psychological pain as time passes. If the initial psychological pain is far larger than the dwindling psychological pain, the government will not change to behaving on $\bar{\theta}_G$ and will keep $\theta_G = \theta_P$. Let $Pain_{s,t}$ be the disutility of a forced change in θ_G in period t where the last forced change in θ_G is undertaken in period s. A larger (positive) value of $Pain_{s,t}$ indicates a larger magnitude of disutility. It is assumed that $\frac{dPain_{s,t}}{dt} < 0$ for $t \geq s$ because the initial psychological pain gradually subsides.

Suppose that deflation sets in and nominal interest rates become zero in period v. Suppose also that, if the government changes θ_G from θ_P to $\bar{\theta}_G$ in period v, deflation reverses to inflation and the central bank resumes forcing the government to change θ_G from $\bar{\theta}_G$ to θ_P in period w where $s < v < w$. The expected disutility of the government generated in period v if it changes θ_G from θ_P to $\bar{\theta}_G$ in period v is therefore

$$
\begin{aligned}
&-\int_v^w \exp\left[-\theta_{G,t}(t-v)\right] pain_{s,t}\,dt + \int_w^\infty \exp\left[-\theta_{G,t}(t-w)\right] pain_{w,t}\,dt \\
&= -\int_v^w \exp\left[-\bar{\theta}_G(t-v)\right] pain_{s,t}\,dt + \int_w^\infty \exp\left[-\theta_P(t-w)\right] pain_{w,t}\,dt
\end{aligned}
$$

where $\theta_{G,t} = \bar{\theta}_G$ for $v \le t < w$ and $\theta_{G,t} = \theta_P$ for $w \le t$, because the government enjoys its intrinsic preference during $v \le t < w$. On the other hand, if the government does not change θ_G from θ_P to $\bar{\theta}_G$, its expected disutility generated in period v is

$$
\begin{aligned}
&\int_v^\infty \exp\left[-\theta_{G,t}(t-v)\right] pain_{s,t}\,dt \\
&= \int_v^\infty \exp\left[-\theta_P(t-v)\right] pain_{s,t}\,dt
\end{aligned}
$$

where $\theta_{G,t} = \theta_P$ for any t ($\ge v$) because the government continues to obey the central bank.

If the disutility in the former case is larger than that in the latter case—that is, if

$$
\begin{aligned}
&-\int_v^w \exp\left[-\bar{\theta}_G(t-v)\right] pain_{s,t}\,dt + \int_w^\infty \exp\left[-\theta_P(t-w)\right] pain_{w,t}\,dt \\
&> \int_v^\infty \exp\left[-\theta_P(t-v)\right] pain_{s,t}\,dt\,, \qquad (4)
\end{aligned}
$$

—then the government will not change θ_G from θ_P to $\bar{\theta}_G$ and thus $\theta_G = \theta_P$ is basically kept even although nominal interest rates are zero and the central bank cannot directly deter the government from changing θ_G from θ_P to $\bar{\theta}_G$. Because $\frac{dPain_{s,t}}{dt} < 0$, then $pain_{w,t} > pain_{s,t}$ for $t > v$, and therefore the probability that inequality (4) holds will not be low. If the value of $\left|\frac{dPain_t(s)}{dt}\right|$ is relatively large—that is, if the initial psychological pain soon subsides—then inequality (4) will usually hold and the change of θ_G from θ_P to $\bar{\theta}_G$ will be always inhibited. Because people will acclimate to the psychological pain, as noted in Section 2.3.1, it is likely that the inhibitory effect usually influences the government's behavior.

The independence of the central bank (in other words, presumable punishments by the central bank) therefore will possess an inhibitory effect, because even if the central bank cannot manipulate nominal interest rates and directly deter the government from changing θ_G from θ_P to $\bar{\theta}_G$ because of ZLB, the government can be still nearly completely under the control of the central bank.

2.3.3 $\hat{\theta}_G = \hat{\theta}_P$, even during deflation

If households firmly believe that the central bank is sufficiently independent and the inhibitory effect is important in controlling the government's behavior, households will expect that the government will continue to keep $\theta_G = \theta_P$, even if nominal interest rates are zero. Therefore, households will basically generate $\hat{\theta}_G = \hat{\theta}_P$ even during deflation if the central bank is sufficiently independent. In the early periods after an upward shock on $\hat{\theta}_P$, households may temporarily generate $\hat{\theta}_G < \hat{\theta}_P$, but they will soon return to $\hat{\theta}_G = \hat{\theta}_P$. This expectation of $\theta_G = \theta_P$ even during deflation is an important factor that makes controlling deflation difficult, as will be shown in later in this section.

2.4 Inflationary and deflationary steady states

2.4.1 Two steady states

A steady state in which both real activity (quantities) and inflation/deflation (changes in prices) stay unchanged requires two conditions, $\theta_P = \bar{r}$ and $\hat{\theta}_G = \hat{\theta}_P$, where $\bar{r}$ is the real rate of interest at steady state. As is well known, $\theta_P = \bar{r}$ is the condition for a steady state of quantities (e.g., Fisher, 1930). $\hat{\theta}_G = \hat{\theta}_P$ is the condition for a steady state of inflation/deflation according to the law of motion for inflation/deflation. As shown in Section 2.3, $\hat{\theta}_G = \hat{\theta}_P$ will be kept even during deflation if the central bank is sufficiently independent.

For both $\theta_P = \bar{r}$ and $\hat{\theta}_G = \hat{\theta}_P$ to be simultaneously satisfied, π_0 in equation (2) needs to take an appropriate value. Among the various possible values, the state that satisfies

(a) $\pi_0 = \pi^*$

is a steady state, which I call an inflationary steady state. Inflation is stabilized at the target rate of the central bank (π^*). However, if deflation is also considered, another π_0 that is consistent with both $\hat{\theta}_G = \hat{\theta}_P$ and $\theta_P = \bar{r}$ can exist, such that

(b) $\pi_0 = -\bar{r}$,

which I call a deflationary steady state. The two steady states are identical except for the inflation/deflation rate.

In addition to the nature that $\hat{\theta}_G = \hat{\theta}_P$ even during deflation, the existence of a deflationary steady state (b) is another important factor that makes controlling deflation difficult. A deflationary steady state (b) can be chosen only when the nominal interest rate is zero (in other worlds, during deflation) because, if nominal interest rates are above zero, the central bank can manipulate nominal interest rates to achieve state (a). On the other hand, $\theta_P = \bar{r}$ cannot necessarily be satisfied for any value of π_0. If, and only if, state (b) is chosen (i.e., if and only if $\pi_0 = -\bar{r}$), the condition $\theta_P = \bar{r}$ is satisfied when the nominal interest rates are zero. Therefore, a deflationary steady state (b) can compete with an inflationary steady state (a) as the steady state once nominal interest rates are stuck at ZLB.

Note that the two steady states are also the only possible steady states in NKPC models. If nominal interest rates are positive, the central bank keeps inflation at the target rate, and therefore an inflationary steady state (a) will be always realized. If nominal interest rates are stuck at ZLB, the instrument rule for the central bank's manipulation of nominal interest rates is useless; thus, inflation depends mostly on households' expected inflation. If households do not wish for the economy to collapse (or reach a non-optimal state), $\theta_P = \bar{r}$ is indispensable and only the expected inflation that is consistent with state (a) or (b) can be generated.

2.4.2 The choice between inflationary and deflationary steady states

The value of π_0 is not given exogenously. It is determined by households in the process of them generating $\hat{\theta}_G$ and $\hat{\theta}_P$, and it varies depending on how and when households generate (or change) $\hat{\theta}_G$ after a shock on $\hat{\theta}_P$. How do households determine the value of π_0? Households are rational and will not select a future path that results in collapse of the economy (or a non-optimal state) due to $\theta_P \neq \bar{r}$. Hence, households will select the value of π_0 that is consistent with $\theta_P = \bar{r}$. In addition, as discussed in Section 2.3, households will basically always generate the expectation of $\theta_G = \theta_P$ under a sufficiently independent central bank. Therefore, households select the value of π_0 that satisfies both $\hat{\theta}_G = \hat{\theta}_P$ and $\theta_P = \bar{r}$. The only states where π_0 satisfies the both conditions are states (a) and (b); thus, households generate only $\hat{\theta}_G$ that is consistent with either state (a) or (b). In other words, households have to choose the value of π_0 from either π^* or $-\bar{r}$ in the process of generating $\hat{\theta}_G$ and $\hat{\theta}_P$.

Of course, the central bank prefers an inflationary steady state (a) and does not want a deflationary steady state (b) because π_0 at state (b) is not the target rate π^*. The central bank therefore will want to make households choose state (a). However, if nominal interest rates are stuck at ZLB, the central bank cannot force households to choose state (a) by manipulating nominal interest rates. It must therefore find other tools to force households to choose state (a). The question arises, however, of whether such a useful and effective tool exists.

3 DIFFICULTY OF ENDING DEFLATION

3.1 Forward-looking information

To switch from state (b) to (a), households' expectation of θ_G or θ_P must be changed because $\pi_0(=-\bar{r})$ is otherwise not changed. The value of π_0 can be changed only in the process of generating $\hat{\theta}_G$ or $\hat{\theta}_P$. In addition, after a change in $\hat{\theta}_G$ or $\hat{\theta}_P$, the condition $\hat{\theta}_G = \hat{\theta}_P$ must be soon be restored because inflation/deflation otherwise accelerates. As discussed in Section 2.3, households will soon restore the condition $\hat{\theta}_G = \hat{\theta}_P$ because the central bank is sufficiently independent. Therefore, if the central bank can influence households' expectation of θ_G or θ_P and make them change $\hat{\theta}_G$ or $\hat{\theta}_P$, it may be able to force households to switch from state (b) to (a).

Households will change $\hat{\theta}_G$ or $\hat{\theta}_P$ if they obtain new important forward-looking information that is related to future θ_G or θ_P. Hence, the ability of the central bank to force a change in $\hat{\theta}_G$ or $\hat{\theta}_P$ depends on whether it can deliver meaningful new forward-looking information that is related to $\hat{\theta}_G$ or $\hat{\theta}_P$ and can make households believe this new information.

3.2 Verbal intervention

One way for the central bank to deliver forward-looking information is through verbal intervention. Forward-looking guidance on the future path of interest rates can be regarded as a kind of verbal intervention in a broad sense. If households change their expectations because of statements from the central bank, the central bank may successfully force the switch from state (b) to (a). Nevertheless, households are not so naïve as to literally believe all the statements of the central bank. The statements are therefore meaningless unless households believe that they contain true forward-looking information. If households suspect that the statements disseminated are deceptive or untrue, the verbal intervention is useless. To succeed, the central bank must be perceived as sincere, honest, and capable.

Can even a sincere, honest, and capable central bank deliver a statement that will make households change their expectations and switch from state (b) to (a)? The central bank can ask, or even beg, households to change their expectations, but it is most likely difficult to persuade households that a deflationary steady state (b) is very harmful to them, because $\theta_P = \bar{r}$ is satisfied at state (b) and the economy proceeds as "normally" as it does at state (a). Therefore, it will not be easy to make households change their expectations of θ_G or θ_P by verbal intervention alone. If the central bank delivers "false" or deceptive information about θ_G or θ_P, households may temporarily change their expectations, but eventually the justification and credibility of the central bank will be questioned and damaged. Verbal intervention therefore will not be sufficiently effective to force a switch from state (b) to (a).

Verbal intervention is also predicted to be ineffective by NKPC models for almost the same reasons. Unless a deceptive statement is delivered by the central banks, households will not feel the need to change their expectations.

3.3 Quantitative easing (QE)

In the Great Recession after the subprime mortgage crisis, some central banks that faced near-zero nominal interest rates adopted QE as a monetary policy to stimulate the economy by increasing the quantity of money in the economy. If QE is effective in changing households' expectation of future θ_G or θ_P, it can be used as a monetary policy tool when nominal interest rates are stuck at ZLB.

QE is adopted on the basis of a strict interpretation of the quantity theory of money such that inflation/deflation is fundamentally governed by the growth rate of the money supply, which is exogenously given. However, the law of motion for inflation/deflation shown in Section 2 indicates that the quantity of money is irrelevant to inflation/deflation as shown in equation (2). The quantity of money will be determined endogenously after the rate of inflation/deflation is determined: that is, the direction of causality is from the rate of inflation/deflation to the quantity of money. Hence, a change in the quantity of money cannot directly affect inflation/deflation. Therefore, it will be difficult, if not impossible, to directly affect the expectation of θ_G or θ_P through the use of QE.

Nevertheless, if the use of QE delivers meaningful forward-looking information about future θ_G or θ_P, it could have a possible indirect effect on households' expectations. By observing QE, households will perceive that the central bank wants them to change their expectations, but households determine their behavior on the basis of their own levels of optimality. If their optimality is not changed by QE, it is unlikely that households will change their expectations. Therefore, unless inflation/deflation and households' optimality are directly affected by QE, $\hat{\theta}_G$ and $\hat{\theta}_P$ will not be affected—even indirectly.

The quantity of money is not usually included in NKPC models (e.g., Ugai, 2007; Woodford, 2012), so it is doubtful, even in these models, whether QE would be able to directly affect households' choices to switch from state (b) to (a). However, if QE influences households' inflation/deflation expectations, it could be effective. The manner in which household's inflation/deflation expectations are generated theoretically in NKPC models is unclear, however, when nominal interest rates are stuck at ZLB. Therefore, it is difficult to arrive at any clear theoretical conclusion about the effectiveness of QE on inflation/deflation channeled through households' expectations of inflation/deflation.

3.4 Renouncing independence

The reason why only steady states (a) and (b) can be chosen is that the condition $\hat{\theta}_G = \hat{\theta}_P$ must be satisfied: that is, the central bank must be sufficiently independent, as discussed in Section 2.

Even during deflation, $\hat{\theta}_G = \hat{\theta}_P$ will hold because of the inhibitory effect resulting from the central bank's independence. Conversely, if the central bank is not independent, a deflationary steady state (b) will not be chosen because the inhibitory effect does not exist, and it is likely that $\theta_G > \theta_P$. Therefore, if the central bank renounces its independence, a switch from deflation to inflation may be possible. Adopting the measure of "helicopter money" as permanent and irreversible QE may be a kind of renouncement of independence. This, however, is an extraordinary and unconventional monetary policy.

Preferences are hard to control solely by oneself. As discussed in Appendix A, even though a government is fully rational and is not weak, foolish, or untruthful, it is still difficult for it to self-regulate its preferences. An independent central bank is therefore essential to control the government's preference (i.e., θ_G), which it has difficulty controlling by itself. Hence, renouncing the central bank's independence indicates that the government will behave on the basis of its own intrinsic preferences—particularly on $\overline{\theta}_G$ —and $\theta_G > \theta_P$ will prevail.

To have any lasting effect, the central bank would have to renounce its authority truthfully and indefinitely. At the least, households would need to firmly believe that the central bank has done so. There is another very serious problem with this solution. The renouncement will be accompanied by high or hyperinflation. If the bank's independence is actually and indefinitely renounced, the deflationary steady state (b) will change, but not necessarily to the inflationary steady state (a) because the central bank is no longer independent. High or hyperinflation will be generated as a byproduct or side effect. Although the reversal of deflation may be successfully achieved, price stability will not. It seems unlikely, therefore, that either the central bank or households would support this measure.

A different conclusion may be drawn in NKPC models, but it is not theoretically clear how the independence of the central bank affects inflation/deflation in these models. In other words, the reason why the central bank and not the government should manipulate nominal interest rates according to a pre-determined instrument rule (e.g., a Taylor rule) is theoretically ambiguous. In NKPC models, it may be implicitly assumed that there is some difference in preferences between the government and the central bank, but this difference is not explicitly modeled. Therefore, it is unclear what would happen if the central bank renounced its independence in NKPC models.

3.5 Raising nominal interest rates

There is another extraordinary and unconventional monetary policy by which a central bank may be able to force a switch from state (b) to state (a): increasing nominal interest rates. Conventionally, when a central bank wants to raise inflation, it decreases nominal interest rates. By this measure, the government has to raise θ_G as the central bank desires because it cannot otherwise achieve optimality and, as a result, inflation increases. However, the same logic can be applied even if the central bank increases nominal interest rates until the government raises θ_G sufficiently. Hence, if nominal interest rates are increased, inflation may be also increased.

However, increasing nominal interest rates is very risky because θ_P may also be affected. Increasing nominal interest rates will generate a temporary recession, and households may feel increased levels of future uncertainty and raise $\hat{\theta}_P$. If $\hat{\theta}_P$ increases as much as θ_G increases, the effect of the higher θ_G will be cancelled out and inflation/deflation will not change. If that occurs, the act of raising nominal interest rates will have created a recession without solving the underlying problem of deflation. Because of this risk, this unconventional monetary policy almost certainly will not actually be used.

This measure may be considered to be effective in NKPC models. If nominal interest rates are stuck at ZLB, the instrument rule for the central bank's manipulation of nominal interest rates is useless, and inflation depends primarily on households' expected inflation/deflation. Usually, θ_P is exogenously given and constant in NKPC models; thus, when nominal interest rates are raised by the central bank, the economy will collapse (or reach a non-optimal state) owing to the permanent condition $\theta_P \neq \bar{r}$ unless the households' inflation expectation is raised. If households strongly want to avoid an economic collapse, they will raise their expected inflation level so that this measure could possibly succeed. However, the conclusions will differ greatly depending on the assumptions of how inflation/deflation expectations are generated by households.

3.6 Depreciating the exchange rate

A sharp depreciation of the exchange rate raises prices of imported goods and services and may therefore temporarily cause deflation to change to inflation. If households change $\hat{\theta}_G$ or $\hat{\theta}_P$ and consequently π_0 because of this shock on the exchange rate, state (b) may be switched to state (a). Furthermore, if a government or central bank is able to deliberately depreciate the exchange rate sharply (i.e., if it can freely manipulate the exchange rate), it can use this as a tool to switch from state (b) to (a).

This strategy has two problems. First, it is not certain that a shock on the exchange rate will always affect the expectations of θ_G or θ_P, because the exchange rate is irrelevant to θ_G or θ_P directly according to the law of motion for inflation/deflation. If a change in the exchange rate possesses some forward-looking information, $\hat{\theta}_G$ or $\hat{\theta}_P$ may be influenced indirectly by the change in the exchange rate, but that is not a theoretical certainty. Second, and more importantly, it is difficult for a government or central bank to freely manipulate the exchange rate. A change in exchange rates affects international trade and finance; thus, unilateral manipulation of exchange rates is problematic in the international community. At the least, this type of action will be fiercely condemned internationally. As a result, this strategy most likely will not be adoptable, at least not overtly.

Depreciating the exchange rate may be judged as effective in NKPC models (Svensson, 2001; Coenen and Wieland, 2003, 2004). Depreciated exchange rates and the ensuing temporary inflation may change households' inflation/deflation expectations. However, it is not clear why, or how, households change their expectations. In addition, the important problems related to the reaction of the international community remain the same. Therefore, in NKPC models, this strategy also seems unlikely to be used, at least not overtly.

3.7 Waiting for a lucky event

The expectations of θ_G and θ_P will of course be affected also by various exogenous shocks. There may be an exogenous shock that is large enough to make households switch from state (b) to (a). For example, if there is a large upward shock on the prices of imported goods (e.g., due to a hike in oil prices or a sharp exogenous depreciation in the exchange rate), households may think that a switch from state (b) to state (a) is better for them because it may be easier for them to adapt to the shock in inflationary steady state (a) than in deflationary steady state (b). Other examples include a large upward shock on $\hat{\theta}_G$ and a large downward shock on $\hat{\theta}_P$. There are many other possible exogenous shocks that may affect the households' expectations and cause them to switch from state (b) to state (a). Nevertheless, these types of exogenous shock represent luck or randomness (given that exchange rates cannot be manipulated unilaterally by a government or central bank).

3.8 Imposing taxes on money

Several economists have proposed a tax on money to generate a negative rate of real interest (Fukao, 2005; Buiter, 2005). By using these taxes, nominal interest rates again become useful to achieve a target rate of inflation. This measure is therefore different from the previously discussed ones that are intended to influence the households' expectations. Because taxes are imposed by the government and not the central bank, the independence of the central bank is meaningless and the role of central bank becomes ambiguous. That said, once the deflation is reversed to inflation, the independent central bank again takes the initiative in controlling inflation.

This measure has the same problem as raising nominal interest rates. Imposing taxes on money may affect θ_P; in particular, it may increase θ_P. An increased θ_P leads to recession and the acceleration of deflation according to the law of motion for inflation/deflation. Imposing taxes on money therefore is very risky. If θ_P is not affected, this measure would be effective, but it seems likely that θ_P will be affected and any existing recession will be aggravated and deflation will accelerate. Because of these risks, this measure will almost certainly not actually be undertaken.

This measure may be considered to be effective in NKPC models. Because θ_P is usually exogenously given and constant in these models, inflation can be controlled by the monetary taxes without raising θ_P—that is, without the risk of aggravating an existing recession or accelerating deflation. Hence, this measure may be predicted to succeed with a high probability in examinations based NKPC models.

3.9 Difficulty in switching from state (b) to (a)

Examinations in this section have shown that it is difficult for a central bank to make households switch from a deflationary steady state (b) to an inflationary steady state (a). Verbal intervention and QE are basically ineffective. The extraordinary and unconventional monetary policy of renouncing central bank independence will have a large impact on the households' expectations and reverse deflation, but it will be accompanied by a serious negative side effect—that is, high or hyper-inflation. Another extraordinary and unconventional monetary policy—raising nominal interest rates—has a high risk of introducing or worsening a recession without reversing the ongoing deflation. Imposing taxes on money shares the same risks. It seems that the most effective policy to realize a switch from state (b) to (a) is to wait for an exogenous event—that is, to get lucky.

This conclusion suggests that, when inflation shows a sign of changing to deflation in the near future, it is extremely important for the central bank to make households continue to choose inflationary steady state (a). It will be easier for a central bank to make them do so before deflation sets in, for example, by increasing the target rate of inflation and lowering nominal interest rates far more than usual (e.g., Williams, 2009). If a central bank can successfully make households continue to choose state (a), the deflationary state can be averted altogether.

Verbal intervention and use of QE were also considered to be basically ineffective in the NKPC models, but the effects of renouncing central bank independence and raising nominal interest rates are unclear. On the other hand, imposing taxes on money may be effective, but this measure has never actually been tried.

4 CONCLUDING REMARKS

Once deflation takes root, it is not easy to reverse because of the ZLB. If nominal interest rates are stuck at the ZLB, the central bank loses power to manipulate nominal interest rates. The manipulation of expectations is instead regarded as an important alternative tool. In this paper, the feasibility and effectiveness of the manipulation of expectations were examined in the inflation/deflation model shown by Harashima (2004b, 2008, 2015a) as well as in conventional inflation models—particularly new-Keynesian Phillips curve (NKPC) models.

There are only two steady states where both $\theta_P = \bar{r}$ and $\hat{\theta}_G = \hat{\theta}_P$ are satisfied:

inflationary steady state (a), at which $\pi_0 = \pi^*$, and deflationary steady state (b), at which $\pi_0 = -\bar{r}$. Deflationary steady state (b) can be a steady state because the condition $\hat{\theta}_G = \hat{\theta}_P$ can hold owing to the inhibitory effect of the independent central bank, even if nominal interest rates are stuck at ZLB. To switch from state (b) to (a), the households' expectations need to be deliberately changed by the central bank. There are several possible ways to influence households' expectations, but there is no decisive measure to certainly change their expectations. Verbal intervention and QE are not effective. Renouncing the independence of the central bank may be effective but has very negative side effects, and raising nominal interest rates is also very risky. It is uncertain whether depreciating the exchange rate is effective, and it is practically infeasible for international political reasons. Imposing taxes on money is another measure that is very risky and not guaranteed to work. The best way to switch from state (b) to (a) may simply be to wait for a fortuitous exogenous event (i.e., to be lucky). It may therefore be prudent for central banks to act drastically when inflation shows a sign of changing to deflation in the near future, for example, by increasing the target rate of inflation and lowering nominal interest rates far more than usual.

Even at deflationary steady state (b), however, the economy proceeds as normally as it does at inflationary steady state (a). Therefore, it may not be necessary for households to struggle to switch from deflationary steady state (b) to inflationary steady state (a), even though the central bank is very dissatisfied with deflationary steady state (b).

APPENDIX A

A1 The law of motion for inflation/deflation

A1.1 The government

A1.1.1 The government budget constraint

The government budget constraint is

$$\dot{B}_t = B_t i_t + G_t - X_t - \vartheta_t \quad ,$$

where B_t is the nominal obligation of the government to pay for its accumulated bonds, i_t is the nominal interest rate for government bonds, G_t is the nominal government expenditure, X_t is the nominal tax revenue, and ϑ_t is the nominal amount of seigniorage at time t. The tax is assumed to be lump sum, the government bonds are long term, and the returns on the bonds are realized only after the bonds are held during a unit period (e.g., a year). The government bonds are redeemed in a unit period, and the government successively refinances the bonds by issuing new ones at each time t. Let $b_t = \frac{B_t}{P_t}$, $g_t = \frac{G_t}{P_t}$, $x_t = \frac{X_t}{P_t}$, and $\varphi_t = \frac{\vartheta_t}{P_t}$, where P_t is the price level at time t. Let also $\pi_t = \frac{\dot{P}_t}{P_t}$ be the inflation rate at time t. By dividing by P_t, the budget constraint is transformed to $\frac{\dot{B}_t}{P_t} = b_t i_t + g_t - x_t - \varphi_t$, which is equivalent to

$$\dot{b}_t = b_t i_t + g_t - x_t - \varphi_t - b_t \pi_t = b_t (i_t - \pi_t) + g_t - x_t - \varphi_t \quad . \tag{A1}$$

Because the returns on government bonds are realized only after holding the bonds during a unit period, investors buy the bonds if $\bar{i}_t \geq E_t \int_t^{t+1} (\pi_s + r_s) ds$ at time t, where $\bar{i}_t$ is the nominal interest rate for bonds bought at t and r_t is the real interest rate in markets at t. Hence, by arbitrage, $\bar{i}_t = E_t \int_t^{t+1} (\pi_s + r_s) ds$ and if r_t is constant such that $r_t = r$ (i.e., if it is at steady state), then

$$\bar{i}_t = E_t \int_t^{t+1} \pi_s \, ds + r \quad .$$

The nominal interest rate $\bar{i}_t = E_t \int_t^{t+1} \pi_s \, ds + r$ means that, during a sufficiently small period between t and $t + dt$, the government's obligation to pay for the bonds' return in the future increases not by $dt(\pi_t + r)$ but by $dt\left(E_t \int_t^{t+1} \pi_s \, ds + r \right)$. If π_t is constant, then $E_t \int_t^{t+1} \pi_s \, ds = \pi_t$ and $\bar{i}_t = \pi_t + r$, but if π_t is not constant, these equations do not necessarily hold.

Since bonds are redeemed in a unit period and successively refinanced, the bonds the government is holding at t have been issued between t - 1 and t. Hence, under perfect foresight, the average nominal interest rate for all government bonds at time t is the weighted sum of $\bar{i}_t$ such that

$$i_t = \int_{t-1}^{t} \bar{i}_s \left(\frac{\overline{B}_{s,t}}{\int_{t-1}^{t} \overline{B}_{v,t} \, dv} \right) ds = \int_{t-1}^{t} \int_{s}^{s+1} \pi_v \, dv \left(\frac{\overline{B}_{s,t}}{\int_{t-1}^{t} \overline{B}_{v,t} \, dv} \right) ds + r \quad ,$$

where $\overline{B}_{s,t}$ is the nominal value of bonds at time t that were issued at time s. If the weights $\frac{\overline{B}_{s,t}}{\int_{t-1}^{t} \overline{B}_{v,t} \, dv}$ between t - 1 and t are not so different from each other, then approximately $i_t = \int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \, ds + r$. To be precise, if the absolute values of π_s for $t-1 < s \leq t+1$ are sufficiently smaller than unity, the differences among the weights are negligible and then approximately

$$i_t = \int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \, ds + r \tag{A2}$$

(see Harashima, 2008). The average nominal interest rate for the total government bonds, therefore, develops by $i_t = \int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \, ds + r$. If π_t is constant, then $\int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \, ds = \pi_t$; thus, $i_t = \pi_t + r$. If π_t is not constant, however, the equations $\int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \, ds = \pi_t$ and $i_t = \pi_t + r$ do not necessarily hold.

A1.1.2 An economically Leviathan government

Under a proportional representation system, the government represents the median household whereas the representative household from an economic perspective represents the mean household.[2] Because of this difference, they usually have different preferences. To account for this essential difference, a Leviathan government is assumed in the model.[3] There are two extremely different views regarding government's behavior in the literature on political economy: the Leviathan view and the benevolent view (e.g., Downs 1957; Brennan and Buchanan 1980; Alesina and Cukierman 1990). From an economic point of view, a benevolent government maximizes the expected economic utility of the representative household, but a Leviathan government does not. Whereas the expenditure of a benevolent government is a tool used to maximize the economic utility of the representative household, the expenditure of a Leviathan government is a tool used to achieve the government's own policy objectives.[4] For example, if a Leviathan government considers national security to be the most important political issue, defense spending will increase greatly, but if improving social welfare is the top political priority, spending on social welfare will increase dramatically, even though the increased expenditures may not necessarily increase the economic utility of the representative household.

Is it possible, however, for such a Leviathan government to hold office for a long period? Yes, because a government is generally chosen by the median of households under a

[2] See the literature on the median voter theorem (e.g., Downs 1957). Also see the literature on the delay in reforms (e.g., Alesina and Drazen 1991; Cukierman et al. 1992).
[3] The most prominent reference to Leviathan governments is Brennan and Buchanan (1980).
[4] The government behavior assumed in the fiscal theory of the price level reflects an aspect of a Leviathan government. Christiano and Fitzgerald (2000) argue that non-Ricardian policies correspond to the type of policies in which governments are viewed as selecting policies and committing themselves to those policies in advance of prices being determined in markets.

proportional representation system (e.g., Downs 1957), whereas the representative household usually presumed in the economics literature is the mean household. The economically representative household is not usually identical to the politically representative household, and a majority of people could support a Leviathan government even if they know that the government does not necessarily pursue only the economic objectives of the economically representative household. In other words, the Leviathan government argued here is an economically Leviathan government that maximizes the political utility of people, whereas the conventional economically benevolent government maximizes the economic utility of people. In addition, because the politically and economically representative households are different (the median and mean households, respectively), the preferences of future governments will also be similarly different from those of the mean representative household. In this sense, the current and future governments presented in the model can be seen as a combined government that goes on indefinitely; that is, the economically Leviathan government always represents the median representative household.

The Leviathan view generally requires the explicit inclusion of government expenditure, tax revenue, or related activities in the government's political utility function (e.g., Edwards and Keen 1996). Because an economically Leviathan government derives political utility from expenditure for its political purposes, the larger the expenditure is, the happier the Leviathan government will be. But raising tax rates will provoke people's antipathy, which increases the probability of being replaced by the opposing party that also nearly represents the median household. Thus, the economically Leviathan government regards taxes as necessary costs to obtain freedom of expenditure for its own purposes. The government therefore will derive utility from expenditure and disutility from taxes. Expenditure and taxes in the political utility function of the government are analogous to consumption and labor hours in the economic utility function of the representative household. Consumption and labor hours are both control variables, and as such, the government's expenditure and tax revenue are also control variables. As a whole, the political utility function of economically Leviathan government can be expressed as $u_G(g_t, x_t)$.[5] In addition, it can be assumed on the basis of previously mentioned arguments that $\frac{\partial u_G}{\partial g_t} > 0$ and $\frac{\partial^2 u_G}{\partial g_t^{\,2}} < 0$, and therefore that $\frac{\partial u_G}{\partial x_t} < 0$ and $\frac{\partial^2 u_G}{\partial x_t^{\,2}} > 0$.[6] An economically Leviathan government therefore maximizes the expected sum of these utilities discounted by its time preference rate under the constraint of deficit financing.

A1.1.3 The optimization problem

The optimization problem of an economically Leviathan government is

$$Max\ E\int_0^\infty u_G(g_t, x_t)\exp(-\theta_G\, t)dt$$

[5] It is possible to assume that governments are partially benevolent. In this case, the utility function of a government can be assumed to be $u_G(g_t, x_t, c_t, l_t)$, where c_t is real consumption and l_t is the leisure hours of the representative household. However, if a lump-sum tax is imposed, the government's policies do not affect steady-state consumption and leisure hours. In this case, the utility function can be assumed to be $u_G(g_t, x_t)$.

[6] Some may argue that it is more likely that $\frac{\partial u_G}{\partial x_t} > 0$ and $\frac{\partial^2 u_G}{\partial x_t^{\,2}} < 0$. However, the assumption used is not an important issue here because $-x_t\left[\frac{\partial u_G(g_t, x_t)}{\partial x_t}\right]^{-1}\frac{\partial^2 u_G(g_t, x_t)}{\partial x_t^{\,2}}\frac{\dot{x}_t}{x_t} = 0$ at steady state. Thus, the results are not affected by which assumption is used.

subject to the budget constraint

$$\dot{b}_t = b_t(i_t - \pi_t) + g_t - x_t - \varphi_t \quad , \tag{A3}$$

where u_G is the constant relative risk aversion utility function of the government, θ_G is the government's rate of time preference, and E is the expectation operator. All variables are expressed in per capita terms, and population is assumed to be constant. The government maximizes its expected political utility considering the behavior of the economically representative household that is reflected in i_t in its budget constraint.

A1.2 Households

The economically representative household maximizes its expected economic utility. Sidrauski (1967)'s well-known money in the utility function model is used for the optimization problem. The representative household maximizes its expected utility

$$E\int_0^\infty u_P(c_t, m_t)\exp(-\theta_P t)dt$$

subject to the budget constraint

$$\dot{a}_t = (r_t a_t + w_t + \sigma_t) - [c_t + (\pi_t + r_t)m_t] - g_t \quad ,$$

where u_P and θ_P are the utility function and the time preference rate of the representative household, c_t is real consumption, w_t is real wage, σ_t is lump-sum real government transfers, m_t is real money, $a_t = k_t + m_t$, and k_t is real capital. It is assumed that $r_t = f'(k_t)$, $w_t = f(k_t) - k_t f'(k_t)$, $u_P' > 0$, $u_P'' < 0$, $\frac{\partial u_P(c_t, m_t)}{\partial m_t} > 0$, and $\frac{\partial^2 u_P(c_t, m_t)}{\partial m_t^2} < 0$, where $f(\cdot)$ is the production function. Government expenditure (g_t) is an exogenous variable for the representative household because it is an economically Leviathan government. It is also assumed that, although all households receive transfers from a government in equilibrium, when making decisions, each household takes the amount it receives as given, independent of its money holdings. Thus, the budget constraint means that the real output $f(k_t)$ at any time is demanded for the real consumption c_t, the real investment $\dot{k}_t$, and the real government expenditure g_t such that $f(k_t) = c_t + \dot{k}_t + g_t$. The representative household maximizes its expected economic utility considering the behavior of government reflected in g_t in the budget constraint. In this discussion, a central bank is not assumed to be independent of the government; thus, the functions of the government and the central bank are not separated. This assumption can be relaxed, and the roles of the government and the central bank are explicitly separated in Section A2.

Note that the time preference rate of government (θ_G) is not necessarily identical to that of the representative household (θ_P) because the government and the representative household represent different households (i.e., the median and mean households, respectively). In addition, the preferences will differ because (1) even though people want to choose a government that has the same time preference rate as the representative household, the rates may differ owing to errors in expectations (e.g., Alesina and Cukierman 1990); and (2) current voters cannot bind the choices of future voters and, if current voters are aware of this possibility, they may vote more myopically as compared with their own rates of impatience in private economic activities (e.g., Tabellini and Alesina 1990). Hence, it is highly likely that the time preference rates of a government and the representative household are heterogeneous. It should be also noted, however, that even though the rates of time preference are heterogeneous, an economically Leviathan government behaves based only on its own time preference rate, without hesitation.

A1.3 The simultaneous optimization

First, I examine the optimization problem of the representative household. Let Hamiltonian H_P be $H_P = u_P(c_t, m_t)\exp(-\theta_P t) + \lambda_{P,t}[r_t a_t + w_t + \sigma_t - c_t - (\pi_t + r_t)m_t - g_t]$, where $\lambda_{P,t}$ is a costate variable, c_t and m_t are control variables, and a_t is a state variable. The optimality conditions for the representative household are;

$$\frac{\partial u_P(c_t, m_t)}{\partial c_t}\exp(-\theta_P t) = \lambda_{P,t} \quad , \tag{A4}$$

$$\frac{\partial u_P(c_t, m_t)}{\partial m_t}\exp(-\theta_P t) = \lambda_{P,t}(\pi_t + r_t) \quad , \tag{A5}$$

$$\dot{\lambda}_{P,t} = -\lambda_{P,t}\, r_t \quad , \tag{A6}$$

$$\dot{a}_t = (r a_t + w_t + \sigma_t) - [c_t + (\pi_t + r_t)m_t - g_t] \quad , \tag{A7}$$

$$\lim_{t\to\infty} \lambda_{P,t}\, a_t = 0 \quad . \tag{A8}$$

By conditions (A4) and (A5), $\left[\frac{\partial u_P(c_t, m_t)}{\partial c_t}\right]^{-1} \frac{\partial u_P(c_t, m_t)}{\partial m_t} = \pi_t + r_t$, and by conditions (A4) and (A6),

$$-c_t\left[\frac{\partial u_P(c_t, m_t)}{\partial c_t}\right]^{-1} \frac{\partial^2 u_P(c_t, m_t)}{\partial c_t^2}\frac{\dot{c}_t}{c_t} + \theta_P = r_t \quad . \tag{A9}$$

Hence,

$$\theta_P = r_t = r \tag{A10}$$

at steady state such that $\dot{c}_t = 0$ and $\dot{k}_t = 0$.

Next, I examine the optimization problem of the economically Leviathan government. Let Hamiltonian H_G be $H_G = u_G(g_t, x_t)\exp(-\theta_G t) + \lambda_{G,t}[b_t(i_t - \pi_t) + g_t - x_t - \varphi_t]$, where $\lambda_{G,t}$ is a costate variable. The optimality conditions for the government are;

$$\frac{\partial u_G(g_t, x_t)}{\partial g_t}\exp(-\theta_G t) = -\lambda_{G,t} \quad , \tag{A11}$$

$$\frac{\partial u_G(g_t, x_t)}{\partial x_t}\exp(-\theta_G t) = \lambda_{G,t} \quad , \tag{A12}$$

$$\dot{\lambda}_{G,t} = -\lambda_{G,t}(i_t - \pi_t) \quad , \tag{A13}$$

$$\dot{b}_t = b_t(i_t - \pi_t) + g_t - x_t - \varphi_t \quad , \tag{A14}$$

$$\lim_{t \to \infty} \lambda_{G,t}\, b_t = 0 \quad . \tag{A15}$$

Combining conditions (A11), (A12), and (A13) and equation (A2) yields the following equations:

$$-g_t \left[\frac{\partial u_G(g_t, x_t)}{\partial g_t} \right]^{-1} \frac{\partial^2 u_G(g_t, x_t)}{\partial g_t^2} \frac{\dot{g}_t}{g_t} + \theta_G = i_t - \pi_t = r_t + \int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \;\; ds - \pi_t \tag{A16}$$

and

$$-\; x_t \left[\frac{\partial u_G(g_t, x_t)}{\partial x_t} \right]^{-1} \frac{\partial^2 u_G(g_t, x_t)}{\partial {x_t}^2} \frac{\dot{x}_t}{x_t} + \theta_G = i_t - \pi_t = r_t + \int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \;\; ds - \pi_t \quad . \tag{A17}$$

Here, $g_t \left[\frac{\partial u_G(g_t, x_t)}{\partial g_t} \right]^{-1} \frac{\partial^2 u_G(g_t, x_t)}{\partial g_t^2} \frac{\dot{g}_t}{g_t} = 0$ and $x_t \left[\frac{\partial u_G(g_t, x_t)}{\partial x_t} \right]^{-1} \frac{\partial^2 u_G(g_t, x_t)}{\partial {x_t}^2} \frac{\dot{x}_t}{x_t} = 0$ at steady state such that $\dot{g}_t = 0$ and $\dot{x}_t = 0$; thus,

$$\theta_G = r_t + \int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \, ds - \pi_t \quad . \tag{A18}$$

Hence, by equation (A10),

$$\int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \, ds = \pi_t + \theta_G - \theta_P \tag{A19}$$

at steady state such that $\dot{g}_t = 0$, $\dot{x}_t = 0$, $\dot{c}_t = 0$, and $\dot{k}_t = 0$.[7]

Equation (A19) is a natural consequence of simultaneous optimization by the economically Leviathan government and the representative household. If the rates of time preference are heterogeneous between them, then

$$i_t - r = \int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \, ds \neq \pi_t \quad .$$

This result might seem surprising because it has been naturally conjectured that $i_t = \pi_t + r$. However, this is a simple misunderstanding because π_t indicates the instantaneous rate of

[7] If and only if $\theta_G = -\frac{g_t - x_t - \varphi_t}{b_t}$ at steady state, then the transversality condition (A1-15) $\lim_{t \to \infty} \lambda_{G,t}\, b_t = 0$ holds. The proof is shown in Harashima (2008).

inflation at a point such that $\pi_t = \frac{\dot{P}_t}{P_t}$, whereas $\int_{t-1}^{t}\int_{s}^{s+1}\pi_\upsilon \, d\upsilon \, ds$ roughly indicates the average inflation rate in a period. Equation (A19) indicates that π_t develops according to the integral equation $\pi_t = \int_{t-1}^{t}\int_{s}^{s+1}\pi_\upsilon \, d\upsilon \, ds - \theta_G + \theta_P$. If π_t is constant, the equations $i_t = \pi_t + r$ and $\int_{t-1}^{t}\int_{s}^{s+1}\pi_\upsilon \, d\upsilon \, ds = \pi_t$ are true. However, if π_t is not constant, the equations do not necessarily hold. Equation (A19) indicates that the equations $i_t = \pi_t + r$ and $\int_{t-1}^{t}\int_{s}^{s+1}\pi_\upsilon \, d\upsilon \, ds = \pi_t$ hold only in the case where $\theta_G = \theta_P$ (i.e., a homogeneous rate of time preference). It has been previously thought that a homogeneous rate of time preference naturally prevails; thus, the equation $i_t = \pi_t + r$ has not been questioned. As argued previously, however, a homogeneous rate of time preference is not usually guaranteed.

A1.4 The law of motion for trend inflation

Equation (A19) indicates that inflation accelerates or decelerates as a result of the government and the representative household reconciling the contradiction in heterogeneous rates of time preference. If π_t is constant, the equation $\pi_t = \int_{t-1}^{t}\int_{s}^{s+1}\pi_\upsilon \, d\upsilon \, ds$ holds; conversely, if $\pi_t \neq \int_{t-1}^{t}\int_{s}^{s+1}\pi_\upsilon \, d\upsilon \; ds$, then π_t is not constant. Without the acceleration or deceleration of inflation, therefore, equation (A19) cannot hold in an economy in which $\theta_G \neq \theta_P$. In other words, it is not until $\theta_G \neq \theta_P$ that inflation can accelerate or decelerate. Heterogeneous time preferences ($\theta_G \neq \theta_P$) bend the path of inflation and enables inflation to accelerate or decelerate. The difference of time preference rates ($\theta_G - \theta_P$) at each time needs to be transformed to the accelerated or decelerated inflation rate π_t at each time.

Equation (A19) implies that inflation accelerates or decelerates nonlinearly in the case in which $\theta_G \neq \theta_P$. For a sufficiently small period dt, π_{t+1+dt} is determined with π_s $(t-1 < s \leq t+1)$ that satisfies $\int_{t-1}^{t}\int_{s}^{s+1}\pi_\upsilon \, d\upsilon \; ds - \pi_t = \theta_G - \theta_P$, so as to hold the equation $\int_{t}^{t+dt}\int_{s}^{s+1}\pi_\upsilon \, d\upsilon \, ds = \int_{t-1}^{t-1+dt}\int_{s}^{s+1}\pi_\upsilon \, d\upsilon \, ds + \pi_{t+dt} - \pi_t$. A solution of the integral equation (A19) for given θ_G and θ_P is

$$\pi_t = \pi_0 + 6(\theta_G - \theta_P)t^2 \quad . \tag{A20}$$

Generally, the path of inflation that satisfies equation (A19) for $0 \leq t$ is expressed as

$$\pi_t = \pi_0 + 6(\theta_G - \theta_P)\exp[z_t \ln(t)] \quad ,$$

where z_t is a time dependent variable. The stream of z_t is various depending on the boundary condition, i.e., the past and present inflation during $-1 < t \leq 0$ and the path of inflation during $0 < t \leq 1$ that is set to make π_0 satisfy equation (A19). However, z_t has the following important property. If π_t satisfies equation (A19) for $0 \leq t$, and $-\infty < \pi_t < \infty$ for $-1 < t \leq 1$, then

$$\lim_{t \to \infty} z_t = 2 \quad .$$

Proof is shown in Harashima (2008). Any inflation path that satisfies equation (A19) for $0 \leq t$ therefore asymptotically approaches the path of equation (A20). The mechanism behind the law of motion for inflation (equation [A20]) is examined more in detail in Harashima (2008).

A2 The central bank

A central bank manipulates the nominal interest rate according to the following Taylor-type instrument rule in the model;

$$i_t = \bar{\gamma} + \gamma_\pi \left(\pi_t - \pi^*\right) + \gamma_x x_t \quad , \tag{A21}$$

where π^* is the target rate of inflation and $\bar{\gamma}$, γ_π, and γ_x are constant coefficients. $\bar{\gamma} = \pi^* + r$ as is usually assumed.

In Section A1, central banks are not explicitly considered because they are not assumed to be independent of governments. However, in actuality, central banks are independent organizations in most countries even though some of them are not sufficiently independent. Furthermore, in the conventional inflation model, it is the central banks that control inflation and governments have no role in controlling inflation. Conventional inflation models show that the rate of inflation basically converges at the target rate of inflation set by a central bank. The target rate of inflation therefore is the key exogenous variable that determines the path of inflation in these models.

Both the government and the central bank can probably affect the development of inflation, but they would do so in different manners, as equation (A20) and conventional inflation models indicate. However, the objectives of the government and the central bank may not be the same. For example, if trend inflation is added to conventional models by replacing their aggregate supply equations with equation (A20), inflation cannot necessarily converge at the target rate of inflation because another key exogenous variable (θ_G) is included in the models. A government makes inflation develop consistently with the equation (A20), which implies that inflation will not necessarily converge at the target rate of inflation. Conversely, a central bank makes inflation converge at the target rate of inflation, which implies that inflation will not necessarily develop consistently with equation (A20). That is, unless either θ_G is adjusted to be consistent with the target rate of inflation or the target rate of inflation is adjusted to be consistent with θ_G, the path of inflation cannot necessarily be determined. Either θ_G or the target rate of inflation need be an endogenous variable. If a central bank dominates, the target rate of inflation remains as the key exogenous variable and θ_G should then be an endogenous variable. The reverse is also true.

A central bank will be regarded as truly independent if θ_G is forced to be adjusted to the one that is consistent with the target rate of inflation set by the central bank. For example, suppose that $\theta_G > \theta_P$ and a truly independent central bank manipulates the nominal interest rate according to the Taylor-type instrument rule (equation [A21]). Here,

$$i_t = \int_{t-1}^{t} \int_{s}^{s+1} \pi_\upsilon \, d\upsilon \, ds + r = \theta_G + \pi_t \tag{A22}$$

at steady state such that $\dot{g}_t = 0$, $\dot{x}_t = 0$, $\dot{c}_t = 0$, and $\dot{k}_t = 0$ by equations (A2), (A10), and (A19). If the accelerating inflation rate is higher than the target rate of inflation, the central bank can raise the nominal interest rate from $i_t = \theta_G + \pi_t$ (equation [A22]) to

$$i_t = \theta_G + \pi_t + \psi$$

by positive ψ by intervening in financial markets to lower the accelerating rate of inflation. In this case, the central bank keeps the initial target rate of inflation because it is truly independent. The government thus faces a rate of increase of real obligation that is higher than θ_G by the extra rate ψ.[8] If the government lowers θ_G so that $\theta_G < \theta_P$ and inflation stops accelerating, the central bank will accordingly reduce the extra rate ψ. If, however, the government does not accommodate θ_G to the target rate of inflation, the extra rate ψ will increase as time passes because of the gap between the accelerating inflation rate and the target rate of inflation widens by equation (A20) and γ_x in Taylor-type instrument rules is usually larger than unity, say 1.5. Because of the extra rate ψ, the government has no other way to achieve optimization unless it lowers θ_G to one that is consistent with the target rate of inflation. Once the government recognizes that the central bank is firmly determined to be independent and it is in vain to try to intervene in the central bank's decision makings, the government would not dare to attempt to raise θ_G again anymore.

Equation (A20) implies that a government allows inflation to accelerate because it acts to maximize its expected utility based only on its own preferences. A government is hardly the only entity that cannot easily control its own preferences even when these preferences may result in unfavorable consequences. It may not even be possible to manipulate one's own preferences at will. Thus, even though a government is fully rational and is not weak, foolish, or untruthful, it is difficult for it to self-regulate its preferences. Hence, an independent neutral organization is needed to help control θ_G. Delegating the authority to set and keep the target rate of inflation to an independent central bank is a way to control θ_G. The delegated independent central bank will control θ_G because it is not the central bank's preference to stabilize the price level—it is simply a duty delegated to it. An independent central bank is not the only possible choice. For example, pegging the local currency with a foreign currency can be seen as a kind of delegation to an independent neutral organization. In addition, the gold standard that prevailed before World War II can be also seen as a type of such delegation.

Note also that the delegation may not be viewed as bad from the Leviathan government's point of view because only its rate of time preference is changed, and the government can still pursue its political objectives. One criticism of the argument that central banks should be independent (e.g., Blinder 1998) is that, since the time-inconsistency problem argued in Kydland and Prescott (1977) or Barro and Gordon (1983) is more acute with fiscal policy, why is it not also necessary to delegate fiscal policies? An economically Leviathan government, however, will never allow fiscal policies to be delegated to an independent neutral organization because the Leviathan government would then not be able to pursue its political objectives, which in a sense would mean the death of the Leviathan government. The median household that backs the Leviathan government, but at the same time dislikes high inflation, will therefore support the delegation of authority but only if it concerns monetary policy. The independent central bank will then be given the authority to control θ_G and oblige the government to change θ_G in order to meet the target rate of inflation.

Without such a delegation of authority, it is likely that generally $\theta_G > \theta_P$ because θ_G represents the median household whereas θ_P represents the mean household. Empirical studies indicate that the rate of time preference negatively correlates with permanent income (e.g., Lawrance 1991), and the permanent income of the median household is usually lower than that of the mean household. If generally $\theta_G > \theta_P$, that suggests that inflation will tend to accelerate unless a central bank is independent. The independence of the central bank is therefore very important in keeping the path of inflation stable.

Note also that the forced adjustments of θ_G by an independent central bank are exogenous

---

[8] The extra rate ψ affects not only the behavior of government but also that of the representative household, in which the conventional inflation theory is particularly interested. In this sense, the central bank's instrument rule that concerns and simultaneously affects both behaviors of the government and the representative household is particularly important for price stability.

shocks to both the government and the representative household because they are planned solely by the central bank. When a shock on the expected θ_G is given, the government and the representative household must recalculate their optimal paths including the path of inflation by resetting θ_G, π_t, and φ.

APPENDIX B

B1 Preference vs. rationality

The law of motion for inflation/deflation discussed in Section 2.1 indicates that, if the government behaves on the basis of its intrinsic RTP, inflation will accelerate. On the other hand, if people strongly dislike inflation acceleration, a government has to behave so as to not accelerate inflation; however, this conflicts with its own intrinsic preference.

B1.1 The conflict between preference and rationality

Behaving on the basis of its own intrinsic preferences does not mean that a government acts in a stupid, foolish, or irrational manner; rather, it behaves quite normally by naturally adhering to its intrinsic preferences. A fundamental question arises, however: Even if the government is acting quite normally, is this behavior rational? In economics, rationality usually means that, given the available information, optimal decisions are made to achieve an objective, and rational behavior is generally assumed. However, can rational behavior still prevail when a government cannot optimize its behavior to achieve its objective? This special situation emerges if the central bank is perfectly independent and is firmly determined to stabilize inflation and if, at the same time, the intrinsic time preference rate of government is unchangeable. In this situation, the economy will become severely destabilized because it is impossible to satisfy equation (1). Therefore, the government cannot achieve its objective (i.e., cannot maximize its expected utility) and can behave only irrationally. Conversely, if the government wants to optimize its objective and behave rationally, it must change its time preference. Clearly, trade-offs between rationality and time preference exist in some situations, and either rationality or time preference must be endogenized.

Nevertheless, it is highly unlikely that people will not optimize their behavior to meet their objectives (i.e., maximize utility) if they have complete knowledge of the optimal path. Hence, rationality should prevail over preferences, and time preference will be endogenized when a clash between rationality and time preference occurs. If time preference is endogenized, rational decisions become possible. Even though rationality should eventually prevail over preferences, governments will not easily change their own preferences. They will resist endogenizing them and search for options to escape from doing so—it is this stubborn nature that drives a government to deviate from the path specified by its central bank. Even though unfavorable consequences are expected if no change is made, it can be very difficult to change one's own preferences alone. Controlling preferences therefore usually requires the help of other people or institutions; this is one of the reasons why independent central banks were established to stabilize inflation.

If a central bank is not sufficiently independent, the government must change its RTP on its own so as to not accelerate inflation. A government must then rein in its preferences on its own. The RTP of government, therefore, is determined through the struggle between preference and rationality inside the government. If rationality prevails, inflation does not accelerate, but if preference prevails, inflation will accelerate.

B1.2 Two environments

Models are simplified representations of reality. Therefore, models can be classified by how far the chosen model simplifies reality. In particular, models are classified by whether they are based on the assumption that all agents are homogeneous (i.e., a homogeneous environment) or on the assumption that agents are heterogeneous (i.e., a heterogeneous environment).

In models based on a homogeneous environment, it is usually assumed that rationality always prevails over preference, because it has generally been regarded that there is no conflict between preference and rationality in a homogeneous environment. In general, the dominance of rationality in a homogeneous environment has been undoubted (i.e., the rational expectation hypothesis has been accepted).

On the other hand, dominance of rationality in a heterogeneous environment is not necessarily guaranteed because, unlike in a homogeneous environment, serious contradictions between preference and rationality arise in a heterogeneous environment. For example, Becker (1980) showed that, if the RTPs of households are heterogeneous, the most patient household will eventually own all the capital in an economy and the other households cannot achieve optimality. That is, all households except the most patient household cannot behave rationally in the sense that rational households behave in such a way to achieve optimality, if they adhere to their own intrinsic RTPs. Harashima (2004b, 2008, 2015a) showed another case. If a government adheres to its own intrinsic RTP that is higher than the RTP of the representative household, inflation accelerates. If people dislike inflation acceleration and thereby the government has to behave under the condition that it does not accelerate inflation, there is no path that satisfies all optimality conditions for the government as long as it adheres to its own intrinsic RTP. In a heterogeneous environment, therefore, conflicts between preference and rationality can occur.

B1.3 Necessary intelligence

The struggle between preference and rationality is dealt with in the human brain. To resolve conflicts, humans need particular powers or functions—that is, different types of intelligence.

B1.3.1 Sustainability in a union or society

Properly dealing with the struggle between preference and rationality is essential for humans because humans do not live alone—they are social and live in groups. However, the struggle has the potential to destroy a society. In a heterogeneous environment, if preference prevails over rationality, there is no guarantee that a political union or society is sustainable because some members of society cannot achieve optimality. In theory, this problem does not exist in a homogeneous environment, because the conflict basically does not exist and competitive equilibria are optimal for all people. On the other hand, in a heterogeneous environment, competitive equilibria are not necessarily optimal for all people because people have heterogeneous preferences, as discussed in Section B1.2. Many of the people who cannot achieve optimality will strongly oppose the government or other people, and it is likely that the political union or society will collapse, possibly violently.

A political union or society is formed and maintained because it provides benefits to its members. Behaviors that support a union or society are important for humans to survive. The type of potential vulnerability in heterogeneous environments that is discussed above indicates that various types of intelligence are essential to properly manage the struggle between preference and rationality.

B1.3.1.1 Calculations

In a heterogeneous environment, relationships among people are far more complicated than in a homogeneous environment because people do not all behave in the same way in a heterogeneous environment. Humans must possess the intelligence to cope with these complicated relationships. They need to be able to calculate the outcomes of various activities in a heterogeneous group of people, evaluate the outcomes, and select the best action to take among many options in their brains.

The number of calculations required to reach an optimal solution is far larger in a heterogeneous environment than in a homogeneous one because the number of types of people that must be considered and the number of interconnections among heterogeneous people are far greater in a heterogeneous environment. If each person's brain can cope with this extremely large number of calculations, people can behave rationally (i.e., always take the best actions that are calculated to be optimal, that is, the ones that are consistent with the model) even in a heterogeneous environment. If this does not occur, rationality may not prevail over preference.

B1.3.1.2 Evaluation

After a variety of potential outcomes are calculated, many options are evaluated on the basis of the results of calculations to select the optimal option. Therefore, people must have the intelligence to evaluate options. The optimal future path is more complicated in a heterogeneous environment than in a homogeneous environment, because households act differently. The intelligence needed for evaluation allows people to accurately identify the optimal future path by comparing and evaluating various aspects of many different complicated paths.

B1.3.1.3 Self-control

In addition, another type of intelligence is required—that which allows people to align their preferences so as to follow the optimal option. Even if an optimal option is appropriately calculated and evaluated, the optimal option cannot be implemented if people's preferences are not properly controlled. That is, people must exercise self-control. This type of intelligence applies to other activities as well—for example, when a person is on a diet. Children often have difficulty exercising self-control because this type of intelligence is not yet fully developed in childhood. In addition, it seems highly likely that it is also not necessarily sufficiently developed in many adults, and even adults will often lose the battle when forced to choose an option that is against their own preferences.

B1.3.2 Intelligence needed when the three types of subordinate intelligence are deficient

It remains unclear whether humans are sufficiently equipped with the necessary types of intelligence to deal with the calculation, evaluation, and self-control aspects of decision-making in a heterogeneous environment. For example, the capacity of a human's brain may be insufficient to process the extremely large number of calculations necessary in a heterogeneous environment. If this first type of intelligence is insufficient, it will be even more difficult to evaluate which option is appropriate to prevent disrupting the political union or society. Furthermore, even if the intelligence needed for calculations is sufficient, actions taken will not be optimal if the evaluation process is biased or poor.

If any part of the three subordinate intelligences is deficient, however, humans still have alternative methods to employ. For example, they can use approximations. The number of calculations needed will be significantly reduced if an appropriate approximation method is used. The intelligence needed for approximation is likely basically different from the three types of subordinate intelligence, although there may be partial overlap between them. For appropriate approximations, the concept of "fluid intelligence" will be particularly important.

B1.3.3 Fluid intelligence

In psychology and psychometrics, many types of intelligence have been considered, including fluid intelligence, crystallized intelligence, short-term memory, long-term storage and retrieval, reading and writing ability, and visual processing. Among these, the importance of the difference between fluid intelligence and crystallized intelligence has been particularly emphasized. According to Cattell (1963, 1971), fluid intelligence is the ability to solve novel problems by thinking logically without only depending on knowledge previously acquired. This type of intelligence signifies the ability to deal with new situations without relying on knowledge gained at school or through experience. With the help of fluid intelligence, people can flexibly adapt their thinking to new kinds of problems or situations. By contrast, crystallized intelligence is the capacity to acquire and use previously obtained knowledge.

Fluid intelligence is essential when people make approximate calculations and need to judge which approximation is the best among many choices. These judgments are very difficult because we do not know the true values. Therefore, judgments must be made after comprehensive consideration of various choices. Such judgments represent “something new” in the sense that they will not necessarily be judged as best in future periods and under different circumstances. People need to make new judgments in any future period. That is, we must solve an “unknown problem” on each occasion to make the best approximation. Thus, these judgments are innovations that are made by using a person’s fluid intelligence. Fluid intelligence is therefore essential in a heterogeneous environment.

These types of judgments are similar to decisions made in politics. Political conditions change from moment to moment. Yesterday’s optimal political decision may be a non-optimal political decision today. Furthermore, nobody knows for certain whether today’s political decision is truly optimal. Historians examine whether past political decisions were optimal, but there are many political decisions over which even historians cannot reach consensus about their optimality.

B1.4 The degree of rationality in a heterogeneous environment

B1.4.1 The item response theory

Fluid intelligence can be modeled on the basis of the item response theory, which is used widely in psychometric studies (e.g., Lord and Novick, 1968; van der Linden and Hambleton, 1997). In particular, the item response function is used to describe the relationship between abilities and item responses.

A typical item response function is

$$\tilde{p}(\tilde{\mu}) = \tilde{c} + \frac{1-\tilde{c}}{1+e^{-\tilde{a}(\tilde{\mu}-\tilde{b})}} \quad ,$$

where $\tilde{p}$ is the probability of a correct response (e.g., answer) to an item (e.g., test or question), $\tilde{\mu}$ ($\infty > \tilde{\mu} > -\infty$) is a parameter that indicates an individual’s ability, $\tilde{a}$ (> 0) is a parameter that characterizes the slope of the function, $\tilde{b}$ ($\infty \geq \tilde{b} \geq -\infty$) is a parameter that represents the difficulty of an item, and $\tilde{c}$ ($1 \geq \tilde{c} \geq 0$) is a parameter that indicates the probability that an item can be answered correctly by chance.

B1.4.2 The probability of dominance of rationality

How frequently rationality prevails over preference can be modeled with an item response function. Let *FI* be the degree of fluid intelligence in a person. Larger values of *FI* indicate stronger fluid intelligence in the sense that a person more correctly grasps (approximates) a situation by using fluid intelligence. Let also p_{HE} be the probability that rationality prevails over preference in a heterogeneous environment. On the basis of the item response theory, p_{HE} can be modeled as a function of *FI* such that

$$p_{HE}(FI) = \hat{c} + \frac{1-\hat{c}}{1+e^{-\hat{a}(FI-\hat{b})}} \quad , \tag{B1}$$

where $\hat{a}$ (> 0) is a parameter that characterizes the slope of the function, $\hat{b}$ ($\infty \geq \hat{b} \geq -\infty$) is a parameter that represents the difficulty and complexity of a situation, and $\hat{c}$ ($1 \geq \hat{c} \geq 0$) is a parameter that indicates the probability that rationality prevails over preference by exogenous factors. If *FI* is sufficiently large, rationality almost always prevails over preference in a heterogeneous environment, but if it is very small, preference almost always prevails over rationality.

An important implication of equation (B1) is that the rational expectation hypothesis is not necessarily acceptable in a heterogeneous environment. If *FI* is small (i.e., fluid intelligence is weak), preference will often prevail over rationality and thus the rational expectation hypothesis cannot be unconditionally accepted.

B1.5 Fluid intelligence of government

According to the median voter theorem (e.g., Downs, 1957), a government behaves just as the median voter prefers in a one-person one-vote democratic political system. This theorem suggests that the fluid intelligence of government is equal to that of the median voter. On the other hand, the top-level positions in government are usually occupied by the best and brightest in a country, and they will almost certainly have stronger fluid intelligence than the median voter. However, does that mean these officials will make decisions that are different from those of the median voter? If they do so, they will be forced to step down in the next election according to the median voter theorem. Only politicians who make the same decisions as the median voter will be able to occupy top-level positions. Hence, it is likely that the fluid intelligence of government is practically equal to that of the median voter when dealing with issues in which preference and rationality conflict.

B1.6 The nature of $\hat{c}$

The value of $\hat{c}$ is affected by exogenous factors. For example, if the central bank is sufficiently independent and capable, $\hat{c}$ becomes unity—that is, the central bank makes rationality always prevail over preference with regard to the RTP of government. The government is always forced to change its RTP as the central bank orders. It is likely that many institutions or mechanisms work to raise the value of $\hat{c}$. For example, constitutions, laws, treaties, and many government and international organizations will raise the value of $\hat{c}$ by urging governments to maintain rationality. Such institutions and mechanisms have probably been adopted in many societies, because experience has taught us that they help ensure that rationality prevails over preference in a heterogeneous environment. As new institutions or mechanisms were invented and adopted, the probability that rationality prevails over preference may have gradually increased (by increasing the value of $\hat{c}$) through time. Therefore, it is likely that, as civilization has progressed, $\hat{c}$ has increased, and rationality more frequently prevails over preference in a heterogeneous environment.

B2 A model of government RTP

B2.1 Determinants of θ_G

The value of θ_G will usually be equal to the RTP of the median voter, as discussed in Section B1.5. However, in some cases, other elements will also affect the value of θ_G. The determinants of θ_G will be basically classified into the following two elements.

B2.1.1 Preference element

In this paper, I call the determinant that is equal to that of the median voter's RTP the "preference element." This element usually determines the main body of θ_G. Let $\theta_{G,pre}$ be the preference element component of θ_G, and $\theta_{P,med}$ be the intrinsic RTP of the median voter. As discussed in Section B1.5, the intrinsic $\theta_{G,pre}$ is basically equal to the intrinsic $\theta_{P,med}$ in a one-person one-vote democratic political system. Therefore, in the following sections, I assume that $\theta_{G,pre} = \theta_{P,med}$.

B2.1.2 Political element

The determinant that is peculiar to the government's RTP is the "political element." Let $\theta_{G,pol}$ be the political element component of θ_G. If a political system is maintained and stable forever, the political element will be nil, and θ_G will be determined only by the preference element. However, if a political system is unstable, the political element component is not zero, and it increases as the political system becomes more unstable. Although rare, it is possible for a political system to collapse. There are many historical examples of the collapse of a political system. These have been often observed, for example, after a defeat in a large-scale war or after a revolution. The political element is of great significance when a political system is on the brink of collapse. Faced with an impending collapse of the system, the incumbent government will do anything possible to survive the crisis. From the government's perspective, the far future is meaningless—survival is the primary objective. It imposes taxes and increases expenditures so as to avoid immediate collapse. As a result, its actions become increasingly myopic and impatient in the sense that it does not concern itself with future economic conditions. This behavior indicates an increase in $\theta_{G,pol}$.

For most democratic countries, the probability of an imminent collapse of the political system will be negligible, and we may assume that $\theta_{G,pol}$ is zero in those countries, but the political element is very important in politically unstable countries.

B2.2 The model

Section B1 indicates that p_{HE} needs to be expected to generate an expected $\theta_{G,pre}$. Let $p_{HE,G}$ be the p_{HE} of the government and $p_{HE,P}$ be the p_{HE} of the median voter. Because basically $\theta_{G,pre} = \theta_{P,med}$ as discussed in Section B2.1, $p_{HE,G} = p_{HE,P}$ generally, and thereby it is reasonable to assume that $p_{HE,G} = p_{HE,P}$. Therefore, in a one-person one-vote political system,

$$p_{HE,G} = p_{HE,P} = \hat{c} + \frac{1-\hat{c}}{1+e^{-\hat{a}\left(FI_{P,med}-\hat{b}\right)}} \quad , \tag{B2}$$

where $FI_{P,med}$ is the FI of the median voter. Equation (B2) indicates that the smaller $FI_{P,med}$ is, the smaller $p_{HE,G}$ is and the higher the probability of inflation acceleration.

Suppose that the central bank is not independent of the government. Thereby, the government has to control its RTP by itself, that is, without being forced to so by the central bank. (The case for an independent central bank is discussed in Section B2.4.) Suppose also for simplicity that the probability that a political system is on the brink of collapse is p_{inst} and $\theta_{G,pol}$ takes a unique positive value, and the probability of a stable political system is then $1-p_{inst}$ and $\theta_{G,pol} = 0$. The model of the government's RTP that is used to generate the expected RTP of government is therefore

$$\begin{aligned}\theta_G &= p_{HE,G}\theta_P + \left(1-p_{HE,G}\right)\left(p_{inst}\theta_{G,pol} + \theta_{G,pre}\right) \\ &= \left(\hat{c} + \frac{1-\hat{c}}{1+e^{-\hat{a}\left(FI_{P,med}-\hat{b}\right)}}\right)\theta_P + \left[1-\left(\hat{c} + \frac{1-\hat{c}}{1+e^{-\hat{a}\left(FI_{P,med}-\hat{b}\right)}}\right)\right]\left(p_{inst}\theta_{G,pol} + \theta_{G,pre}\right) \quad .\end{aligned} \tag{B3}$$

Equation (B3) indicates that the RTP of government is equal to θ_P when rationality prevails over preference with the probability $p_{HE,G}$. When preference prevails over rationality with the probability $1 - p_{HE,G}$, the RTP of government is equal to the intrinsic RTP of government. The intrinsic RTP of government consists of $\theta_{G,pol}$ with the probability p_{inst} and $\theta_{G,pre}$.

Because $\theta_{G,pre} = \theta_{P,med}$ (as assumed in Section B2.1.1), then by equation (B3),

$$\theta_G = p_{HE,G}\theta_P + \left(1-p_{HE,G}\right)\left(p_{inst}\theta_{G,pol} + \theta_{P,med}\right).$$

$$= \left(\hat{c} + \frac{1-\hat{c}}{1+e^{-\hat{a}\left(FI_{P,med}-\hat{b}\right)}} \right)\theta_P + \left[1 - \left(\hat{c} + \frac{1-\hat{c}}{1+e^{-\hat{a}\left(FI_{P,med}-\hat{b}\right)}} \right) \right]\left(p_{inst}\theta_{G,pol} + \theta_{P,med} \right) . \qquad \text{(B4)}$$

In most democratic countries, the probability of the occurrence of extreme political instability is very low. For those countries, therefore, the model is reduced to a more simple form by assuming $p_{inst} = 0$ such that

$$\theta_G = p_{HE,G}\theta_P + \left(1 - p_{HE,G}\right)\theta_{P,med}$$
$$= \left(\hat{c} + \frac{1-\hat{c}}{1+e^{-\hat{a}\left(FI_{P,med}-\hat{b}\right)}} \right)\theta_P + \left[1 - \left(\hat{c} + \frac{1-\hat{c}}{1+e^{-\tilde{a}\left(FI_{P,med}-\hat{b}\right)}} \right) \right]\theta_{P,med} . \qquad \text{(B5)}$$

Note that if the people and the government have sufficiently strong fluid intelligences, an independent central bank may not be necessary. However, p_{HE} will not be unity even in a country whose people have the highest p_{HE} in the world. Therefore, it is possible for $\theta_G > \theta_P$ in some period in any country; thus, an independent central bank is still important for all countries.

B2.3 Generating an expected θ_G by using heuristics

B2.3.1 Difficulty in expecting θ_G

Specifying the functional form of the structural model of θ_G is only half of the problem of generating an expected θ_G. Although we have the functional form of the model, as shown in equation (B5), we still cannot generate an expected θ_G unless we specify appropriate values of the parameters $\hat{a}$, $\hat{b}$ and $\hat{c}$. Furthermore, to generate the expected θ_G, we must also know the expected values of θ_P, $\theta_{P,med}$, and $FI_{P,med}$.

We may roughly specify the parameter values of $\hat{a}$, $\hat{b}$ and $\hat{c}$ through the results of some type of social experiment, or we may use the estimates derived from other kinds of model concerning fluid intelligence. By substituting these values for the parameter values in the structural model of θ_G, the model could be calibrated. However, expectations based on these estimates will most likely be rather inaccurate and therefore problematic in terms of decision-making on future actions.

A far more serious problem is obtaining the expected future values of θ_P, $\theta_{P,med}$, and $FI_{P,med}$. It is not certain whether the values of θ_P and $\theta_{P,med}$ are constant across time; in fact, many researchers have posited that it is much more likely that they are temporally variable (e.g., Uzawa, 1968; Epstein and Hynes, 1983; Lucas and Stokey, 1984; Parkin, 1988; Obstfeld, 1990; Becker and Mulligan, 1997). Therefore, there is no guarantee that the future values of θ_P and $\theta_{P,med}$ will equal past ones, so the past values cannot be used as substitutes for the expected future values of θ_P and $\theta_{P,med}$. Hence, to generate the expected future values of θ_P and $\theta_{P,med}$, we have to calculate them on the basis of structural models of θ_P and $\theta_{P,med}$. Even if we knew the functional forms of these structural models, we would still need to determine the parameter values for the models. To determine them, however, we would need to obtain a sufficiently large amount of data on the past values of θ_P and $\theta_{P,med}$—that is, the intrinsic RTPs of the representative household and the median voter. Although a household knows its own RTP, it cannot directly observe θ_P and $\theta_{P,med}$ in the same way that it can obtain data on aggregate consumption, investment, production, inflation, trade, and other indicators at relatively low cost. Without data on the past values of θ_P and $\theta_{P,med}$, we cannot estimate the parameter values. Therefore, we cannot generate the expected future values of θ_P and $\theta_{P,med}$ on the basis of their structural models.

Past data on the real interest rate may be used as a substitute for past θ_P because θ_P is basically equal to the real interest rate at steady state (Fisher, 1930). However, during a transition period after θ_P changes, θ_P is not equal to the rate of real interest. Therefore, unless θ_P is constant across time, this substitution does not seem to be sufficiently useful. In addition, if $\theta_{P,med}$ is constant across time, we may approximate the value of $\theta_{P,med}$ on the basis of historical economic and political (election) data. However, as stated in the previous paragraph, it is not known whether θ_P and $\theta_{P,med}$ are constant across time.

Note that, if we assume that RTP is identical for all households, an expected θ_P and $\theta_{P,med}$ are no longer necessary because the RTP of any household is equal to both θ_P and $\theta_{P,med}$. This assumption is very problematic, however, because it is not merely expedient for the sake of simplicity. It is also a critical requirement to eliminate the need for generating an expected θ_P and $\theta_{P,med}$. Therefore, any rationale for assuming identical RTPs should be validated; that is, it should be demonstrated that identical RTPs do exist and are universally observed. In any case, RTP is unquestionably not identical among households. Therefore, households must generate the expected values of θ_P and $\theta_{P,med}$.

B2.3.2 Expectations based on beliefs

Faced with the difficulty of generating expected values of θ_P and $\theta_{P,med}$ and knowing the parameter values in the model of θ_G, households may have to use the concept of bounded rationality to make decisions. One of a few alternatives available for a household to use is its "beliefs" in θ_P and $\theta_{P,med}$ as well as in $\hat{a}$, $\hat{b}$, $\hat{c}$, and $FI_{P,med}$. The use of beliefs does not mean that households deviate from rationality; rather, it is the most rational option in an environment where insufficient information is available.

Belief is merely that, however—belief. There is no guarantee that the value a household believes to be true is actually the correct value. Therefore, it may often change, but it will be changed only if forward-looking information becomes available. In some cases, a household will change its belief when new data are obtained, but in other cases the household will not, depending on how it interprets the new information. This is particularly true when the household believes that it has extracted forward-looking information about θ_P and $\theta_{P,med}$ from the newly obtained data.

B2.3.3 Heuristics

When households interpret the information extracted from new data, they may use heuristic methods such as a simplified linear reduced form model of θ_G. Studies of the use of heuristics and bounded rationality in this context would be useful for better understanding the interpretation mechanism. Heuristic methods will be implemented through the use of fluid intelligence. Hence, the value of $FI_{P,med}$ will also be important in improving the accuracy of expectations generated on the basis of heuristics.

There may be many possible simplified linear reduced form models of θ_G that could be used as heuristic methods, although most of them may be *ad hoc*. Even though such reduced form models are far less credible than a structural model, they may be utilized as a heuristic method of interpretation. Although simplified linear reduced form models may often result in misleading conclusions, they may sometimes provide useful information.

B2.4 Independent central bank

B2.4.1 Generating expected $\theta_G - \theta_P$ through the actions of a central bank

A heuristic way of generating an expected θ_G is to use information about $\theta_G - \theta_P$. The model of inflation acceleration presented in Section 2 indicates that inflation acceleration and deceleration are governed by the value of $\theta_G - \theta_P$. Therefore, what people really need to know is not the expected θ_G but the expected $\theta_G - \theta_P$. If the central bank is sufficiently independent, θ_G is determined by the central bank. In this case, people do not need to know the RTP of the government, but rather the responses of the central bank to $\theta_G - \theta_P$. If an easy method exists to know the response of the central bank to $\theta_G - \theta_P$, households will not have to generate expected θ_G; they need only observe the decisions of the central bank.

Of course, people cannot directly observe the value of $\theta_G - \theta_P$, but they can observe the response of the central bank to $\theta_G - \theta_P$. An independent central bank will raise interest rates if it judges that $\theta_G - \theta_P > 0$. Households can then adjust their expectations accordingly.

B2.4.2 Guaranteed $\theta_G = \theta_P$

If the central bank is sufficiently independent and capable, and successfully controls θ_G, then it is not even necessary for households to generate an expected value for $\theta_G - \theta_P$ because, in this case, θ_G will also equal θ_P. As discussed in Section B1.6, if the central bank is sufficiently independent and capable, then $\hat{c} = 1$ in equation (B2) and thereby, by equations (B4) and (B5), $\theta_G = \theta_P$. The central bank ensures that rationality always prevails over preference with regard to the RTP of government. If the independence of the central bank is very credible, households will always expect that $\theta_G = \theta_P$ at all times in the future.

APPENDIX C

C1 The representative household in dynamic models

C1.1 The assumption of the representative household

The concept of the representative household is a necessity in macroeconomic studies. It is used as a matter of course, but its theoretical foundation is fragile. The representative household has been used given the assumption that all households are identical or that there exists one specific individual household, the actions of which are always average among households (I call such a household "the average household" in Appendix C). The assumption that all households are identical seems to be too strict; therefore, it is usually assumed explicitly or implicitly that the representative household is the average household. However, the average household can exist only under very strict conditions. Antonelli (1886) showed that the existence of an average household requires that all households have homothetic and homogeneous utility functions. This type of utility function is not usually assumed in macroeconomic studies because it is very restrictive and unrealistic. If more general utility functions are assumed, however, the assumption of the representative household as the average household is inconsistent with the assumptions underlying the utility functions.

Nevertheless, the assumption of the representative household has been widely used, probably because it has been believed that the representative household can be interpreted as an approximation of the average household. Particularly in static models, the representative household can be seen to approximate the average household. However, in dynamic models, it is hard to accept the representative household as an approximation of the average household because, if RTPs of households are heterogeneous, there is no steady state where all of the optimality conditions of the heterogeneous households are satisfied (Becker, 1980). Therefore, macroeconomic studies using dynamic models are fallacious if the representative household is assumed to approximate the average household.

C1.2 The representative household in static models

Static models are usually used to analyze comparative statics. If the average household is represented by one specific unique household for any static state, there will be no problem in assuming the representative household as an approximation of the average household. Even though the average household is not always represented by one specific unique household in some states, if the average household is always represented by a household in a set of households that are very similar in preferences and other features, then the representative household assumption can be used to approximate the average household.

Suppose, for simplicity, that households are heterogeneous such that they are identical except for a particular preference. Because of the heterogeneous preference, household consumption varies. However, levels of consumption will not be distributed randomly because the distribution of consumption will correspond to the distribution of the preference. The consumption of a household that has a very different preference from the average will be very different from the average household consumption. Conversely, it is likely that the consumption of a household that has the average preference will nearly have the average consumption. In addition, the order of the degree of consumption will be almost unchanged for any static state because the order of the degree of the preference does not change for the given state.

If the order of consumption is unchanged for any given static state, it is likely that the household with consumption that is closest to the average consumption will also always be a household belonging to a group of households that have very similar preferences. Hence, it is possible to argue that, approximately, one specific unique household's consumption is always average for any static state. Of course, it is possible to show evidence that is counter to this argument, particularly in some special situations, but it is likely that this conjecture is usually true in normal situations, and the assumption that the representative household approximates the average household is acceptable in static models.

C1.3 The representative household in dynamic models

In dynamic models, however, the story is more complicated. In particular, heterogeneous RTPs pose a serious problem. This problem is easily understood in a dynamic model with exogenous technology (i.e., a Ramsey growth model). Suppose that households are heterogeneous in RTP, degree of risk aversion (ε), and productivity of the labor they provide. Suppose also for simplicity that there are many "economies" in a country, and an economy consists of a household and a firm. The household provides labor to the firm in the particular economy, and the firm's level of technology (A) varies depending on the productivity of labor that the household in its economy provides. Economies trade with each other: that is, the entire economy of a country consists of many individual small economies that trade with each other.

A household maximizes its expected utility, $E\int_0^\infty u(c_t)\exp(-\theta t)dt$, subject to $\dot{k}_t = f(k_t) - c_t$, where $u(\bullet)$ is the utility function; $f(\bullet)$ is the production function; θ is RTP; E is the expectation operator; $y_t = \frac{Y_t}{L_t}$, $k_t = \frac{K_t}{L_t}$, and $c_t = \frac{C_t}{L_t}$; Y_t (≥ 0) is output, K_t (≥ 0) is capital input, L_t (≥ 0) is labor input, and C_t (≥ 0) is consumption in period t. The optimal consumption path of this Ramsey-type growth model is

$$\frac{\dot{c}_t}{c_t} = \varepsilon^{-1}\left(\frac{\partial y_t}{\partial k_t} - \theta\right) ,$$

and at steady state,

$$\frac{\partial y_t}{\partial k_t} = \theta \quad . \tag{C1}$$

Therefore, at steady state, the heterogeneity in the degree of risk aversion (ε) is irrelevant, and the heterogeneity in productivity does not result in permanent trade imbalances among economies because $\frac{\partial y_t}{\partial k_t}$ in all economies is kept equal by market arbitrage. Hence, heterogeneity in the degree of risk aversion and productivity does not matter at steady state. Therefore, the same logic as that used for static models can be applied. Approximately, one specific unique household's consumption is always average for any time in dynamic models, even if the degree of risk aversion and the productivity are heterogeneous. Thus, the assumption of the representative household is also acceptable in dynamic models even if the degree of risk aversion and the productivity are heterogeneous.

However, equation (C1) clearly indicates that heterogeneity in RTP is problematic. As Becker (1980) shows, if RTP is heterogeneous, the household that has the lowest RTP will eventually possess all capital. With heterogeneous RTPs, there is no steady state where all households achieve all of their optimality conditions. In addition, the household with consumption that is average at present has a very different RTP from the household with consumption that is average in the distant future. The consumption of a household that has the average RTP will initially be almost average, but in the future the household with the lowest RTP will be the one with consumption that is almost average. That is, the consumption path of the household that presently has average consumption is notably different from that of the household with average consumption in the future. Therefore, any individual household cannot be almost average in any period and thus cannot even approximate the average household. As a result, even if the representative household is assumed in a dynamic model, its discounted expected utility $E\int_0^\infty u(c_t)\exp(-\theta t)dt$ is meaningless, and analyses based on it are fallacious.

If we assume that RTP is identical for all households, the above problem is solved. However, this solution is still problematic because that assumption is not merely expedient for the sake of simplicity; rather, it is a critical requirement to allow for an assumed representative household. Therefore, the rationale for identical RTPs should be validated; that is, it should be demonstrated that identical RTPs are actually and universally observed. RTP is, however, unquestionably not identical among households. Hence, it is difficult to accept the representative household assumption in dynamic models based on the assumption of identical RTP.

The conclusion that the representative household assumption in dynamic models is meaningless and leads to fallacious results is very important, because a huge number of studies have used the representative household assumption in dynamic models. To solve this severe problem, an alternative interpretation or definition of the representative household is needed.

Note that in an endogenous growth model the situation is even more complicated. Because a heterogeneous degree of risk aversion also matters, the assumption of the representative household is more difficult to accept, so an alternative interpretation or definition is even more important when endogenous growth models are used.

C2 Sustainable heterogeneity

C2.1 The model

Suppose that two heterogeneous economies—economy 1 and economy 2—are identical except for their RTPs. Households within each economy are assumed to be identical for simplicity. The population growth rate is zero. The economies are fully open to each other, and goods, services, and capital are freely transacted between them, but labor is immobilized in each economy.

Each economy can be interpreted as representing either a country (the international interpretation) or a group of identical households in a country (the national interpretation). Because the economies are fully open, they are integrated through trade and form a combined economy. The combined economy is the world economy in the international interpretation and the national economy in the national interpretation. In the following discussion, a model based on the international interpretation is called an international model and that based on the national interpretation is called a national model. Usually, the concept of the balance of payments is used only for the international transactions. However, because both national and international interpretations are possible, this concept and terminology are also used for the national models in Appendix C.

RTP of household in economy 1 is θ_1 and that in economy 2 is θ_2, and $\theta_1 < \theta_2$. The production function in economy 1 is $y_{1,t} = A^\alpha f(k_{1,t})$ and that in economy 2 is $y_{2,t} = A^\alpha f(k_{2,t})$, where $y_{i,t}$ and $k_{i,t}$ are, respectively, output and capital per capita in economy i in period t for $i = 1, 2$; A is technology; and α $(0 < \alpha < 1)$ is a constant. The population of each economy is $\frac{L}{2}$; thus, the total for both is L, which is sufficiently large. Firms operate in both economies. The current account balance in economy 1 is τ_t and that in economy 2 is $-\tau_t$. The production functions are specified as

$$y_{i,t} = A^\alpha k_{i,t}^{1-\alpha} \quad ;$$

thus, $Y_{i,t} = K_{i,t}^{1-\alpha}(AL)^{\alpha}$ $(i = 1,2)$. Because A is given exogenously, this model is an exogenous technology model (Ramsey growth model). The examination of sustainable heterogeneity based on an endogenous growth model is shown in Harashima (2014a).

Because both economies are fully open, returns on investments in each economy are kept equal through arbitration, such that

$$\frac{\partial y_{1,t}}{\partial k_{1,t}} = \frac{\partial y_{2,t}}{\partial k_{2,t}} \quad . \tag{C2}$$

Because equation (C2) always holds through arbitration, equations $k_{1,t} = k_{2,t}$, $\dot{k}_{1,t} = \dot{k}_{2,t}$, $y_{1,t} = y_{2,t}$, and $\dot{y}_{1,t} = \dot{y}_{2,t}$ also hold.

The accumulated current account balance $\int_0^t \tau_s ds$ mirrors capital flows between the two economies. The economy with current account surpluses invests them in the other economy. Because $\frac{\partial y_{1,t}}{\partial k_{1,t}}\left(=\frac{\partial y_{2,t}}{\partial k_{2,t}}\right)$ are returns on investments, $\frac{\partial y_{1,t}}{\partial k_{1,t}}\int_0^t \tau_s ds$ and $\frac{\partial y_{2,t}}{\partial k_{2,t}}\int_0^t \tau_s ds$ represent income receipts or payments on the assets that an economy owns in the other economy. Hence,

$$\tau_t - \frac{\partial y_{2,t}}{\partial k_{2,t}}\int_0^t \tau_s ds$$

is the balance on goods and services of economy 1, and

$$\frac{\partial y_{1,t}}{\partial k_{1,t}}\int_0^t \tau_s ds - \tau_t$$

is that of economy 2. Because the current account balance mirrors capital flows between the economies, the balance is a function of capital in both economies, such that

$$\tau_t = \kappa(k_{1,t}, k_{2,t}) \quad .$$

The government (or an international supranational organization) intervenes in the activities of economies 1 and 2 by transferring money from economy 1 to economy 2. The amount of transfer in period t is g_t, and it is assumed that g_t depends on capital inputs, such that

$$g_t = \bar{g}k_{1,t} \quad ,$$

where $\bar{g}$ is a constant. Because $k_{1,t} = k_{2,t}$ and $\dot{k}_{1,t} = \dot{k}_{2,t}$,

$$g_t = \bar{g}k_{1,t} = \bar{g}k_{2,t} \quad .$$

Each household in economy 1 therefore maximizes its expected utility

$$E\int_0^{\infty} u_1(c_{1,t})\exp(-\theta_1 t)dt \quad ,$$

subject to

$$\dot{k}_{1,t} = A^{\alpha} k_{1,t}^{1-\alpha} - c_{1,t} + (1-\alpha) A^{\alpha} k_{1,t}^{-\alpha} \int_0^t \tau_s ds - \tau_t - \bar{g} k_{1,t} \quad , \tag{C3}$$

and each household in economy 2 maximizes its expected utility

$$E \int_0^{\infty} u_2(c_{2,t}) \exp(-\theta_2 t) dt \quad ,$$

subject to

$$\dot{k}_{2,t} = A^{\alpha} k_{2,t}^{1-\alpha} - c_{2,t} - (1-\alpha) A^{\alpha} k_{2,t}^{-\alpha} \int_0^t \tau_s ds + \tau_t + \bar{g} k_{2,t} \quad , \tag{C4}$$

where $u_{i,t}$ and $c_{i,t}$, respectively, are the utility function and per capita consumption in economy i in period t for i = 1, 2; and E is the expectation operator. Equations (C3) and (C4) implicitly assume that each economy does not have foreign assets or debt in period t = 0.

C2.2 Sustainable heterogeneity without government intervention

Heterogeneity is defined as being sustainable if all of the optimality conditions of all heterogeneous households are satisfied indefinitely. First, the natures of the model when the government does not intervene (i.e., $\bar{g} = 0$) are examined. The growth rate of consumption in economy 1 is

$$\frac{\dot{c}_{1,t}}{c_{1,t}} = \varepsilon^{-1} \left\{ (1-\alpha) A^{\alpha} k_{1,t}^{-\alpha} + (1-\alpha) A^{\alpha} k_{1,t}^{-\alpha} \frac{\partial \int_0^t \tau_s d\, s}{\partial k_{1,t}} - \alpha (1-\alpha) A^{\alpha} k_{1,t}^{-\alpha-1} \int_0^t \tau_s d\, s - \frac{\partial \tau_t}{\partial k_{1,t}} - \theta_1 \right\} \quad .$$

Hence,

$$\lim_{t \to \infty} \frac{\dot{c}_{1,t}}{c_{1,t}} = \varepsilon^{-1} \lim_{t \to \infty} \left\{ (1-\alpha) A^{\alpha} k_{1,t}^{-\alpha} + (1-\alpha) A^{\alpha} k_{1,t}^{-\alpha} \frac{\partial \int_0^t \tau_s ds}{\partial k_{1,t}} - \alpha (1-\alpha) A^{\alpha} k_{1,t}^{-\alpha-1} \int_0^t \tau_s ds - \frac{\partial \tau_t}{\partial k_{1,t}} - \theta_1 \right\} = 0$$

and thereby

$$\lim_{t \to \infty} (1-\alpha) A^{\alpha} k_{1,t}^{-\alpha} \left[1 + (1-\alpha) \Psi \right] - \Xi - \theta_1 = 0 \quad ,$$

where $\Xi = \lim_{t \to \infty} \frac{\tau_t}{k_{1,t}} = \lim_{t \to \infty} \frac{\tau_t}{k_{2,t}}$ and $\Psi = \lim_{t \to \infty} \frac{\int_0^t \tau_s ds}{k_{1,t}} = \lim_{t \to \infty} \frac{\int_0^t \tau_s ds}{k_{2,t}}$. $\lim_{t \to \infty} \frac{\dot{y}_{1,t}}{y_{1,t}} = \lim_{t \to \infty} \frac{\dot{c}_{1,t}}{c_{1,t}} =$ $\lim_{t \to \infty} \frac{\dot{k}_{1,t}}{k_{1,t}} = \lim_{t \to \infty} \frac{\dot{\tau}_t}{\tau_t} = 0$, and Ψ is constant at steady state because $k_{1,t}$ and τ_t are constant; thus, $\Xi = \lim_{t \to \infty} \frac{\tau_t}{k_{1,t}}$ is constant at steady state. For Ψ to be constant at steady state, it is necessary that $\lim_{t \to \infty} \tau_t = 0$ and thus $\Xi = 0$. Therefore,

$$\lim_{t\to\infty}(1-\alpha)A^{\alpha}k_{1,t}^{-\alpha}\left[1+(1-\alpha)\Psi\right]-\theta_1=0 \quad , \tag{C5}$$

and

$$\lim_{t\to\infty}(1-\alpha)A^{\alpha}k_{2,t}^{-\alpha}\left[1-(1-\alpha)\Psi\right]-\theta_2=0$$

because

$$\lim_{t\to\infty}\frac{\dot{c}_{2,t}}{c_{2,t}}=\varepsilon^{-1}\lim_{t\to\infty}\left\{(1-\alpha)A^{\alpha}k_{2,t}^{-\alpha}-(1-\alpha)A^{\alpha}k_{2,t}^{-\alpha}\frac{\partial\int_0^t\tau_s ds}{\partial k_{2,t}}+\alpha(1-\alpha)A^{\alpha}k_{2,t}^{-\alpha-1}\int_0^t\tau_s ds+\frac{\partial\tau_t}{\partial k_{2,t}}-\theta_2\right\}=0$$

Because $\lim_{t\to\infty}(1-\alpha)A^{\alpha}k_{1,t}^{-\alpha}\left[1+(1-\alpha)\Psi\right]=\theta_1$, $\lim_{t\to\infty}(1-\alpha)A^{\alpha}k_{2,t}^{-\alpha}\left[1-(1-\alpha)\Psi\right]=\theta_2$, and $\frac{\partial y_{1,t}}{\partial k_{1.t}}=\frac{\partial y_{2,t}}{\partial k_{2.t}}=A^{\alpha}k_{1,t}^{-\alpha}=A^{\alpha}k_{2,t}^{-\alpha}$, then

$$\Psi=\frac{\theta_1-\theta_2}{2(1-\alpha)\lim_{t\to\infty}\frac{\partial y_{1,t}}{\partial k_{1.t}}} \quad . \tag{C6}$$

By equations (C5) and (C6),

$$\lim_{t\to\infty}\frac{\partial y_{1,t}}{\partial k_{1,t}}+\lim_{t\to\infty}\frac{\partial y_{1,t}}{\partial k_{1,t}}(1-\alpha)\Psi=\theta_1 \quad ;$$

thus,

$$\lim_{t\to\infty}\frac{\partial y_{1,t}}{\partial k_{1.t}}=\frac{\theta_1+\theta_2}{2}=\lim_{t\to\infty}\frac{\partial y_{2,t}}{\partial k_{2.t}} \quad . \tag{C7}$$

If equation (C7) holds, all of the optimality conditions of both economies are indefinitely satisfied. The state indicated by equation (C7) is called the “multilateral steady state” or “multilateral state” in the following discussion. By procedures similar to those used for the endogenous growth model in Harashima (2014a), the condition of the multilateral steady state for *H* economies that are identical except for their RTPs is shown as

$$\lim_{t\to\infty}\frac{\partial y_{i,t}}{\partial k_{i.t}}=\frac{\sum_{q=1}^{H}\theta_q}{H} \tag{C8}$$

for any i, where $i = 1, 2, \dots , H$.

Because

$$\Psi = \frac{\theta_1 - \theta_2}{2(1-\alpha)\lim_{t\to\infty}\frac{\partial y_{1,t}}{\partial k_{1,t}}} = \frac{\theta_1 - \theta_2}{(1-\alpha)(\theta_1 + \theta_2)} < 0$$

by equation (C7), then by $\lim_{t\to\infty}\frac{\int_0^t \tau_s ds}{k_{1,t}} = \Psi < 0$,

$$\lim_{t\to\infty}\int_0^t \tau_s ds < 0 \quad ;$$

that is, economy 1 possesses accumulated debts owed to economy 2 at steady state, and economy 1 has to export goods and services to economy 2 by

$$\left|(1-\alpha)A^{\alpha}k_{1,t}^{-\alpha}\int_0^t \tau_s ds\right|$$

in every period to pay the debts. Nevertheless, because $\lim_{t\to\infty}\tau_t = 0$ and $\Xi = 0$, the debts do not explode but stabilize at steady state. Because of the debts, the consumption of economy 1 is smaller than that of economy 2 at steady state under the condition of sustainable heterogeneity.

Note that many empirical studies conclude that RTP is negatively correlated with income (e.g., Lawrance, 1991; Samwick, 1998; Ventura, 2003). Suppose that, in addition to the heterogeneity in RTP ($\theta_1 < \theta_2$), the productivity of economy 1 is higher than that of economy 2. At steady state, the consumption of economy 1 would be larger than that of economy 2 as a result of the heterogeneity in productivity. However, as a result of the heterogeneity in RTP, the consumption of economy 1 is smaller than that of economy 2 at steady state under sustainable heterogeneity. Which effect prevails will depend on differences in the degrees of heterogeneity. For example, if the difference in productivity is relatively large whereas that in RTP is relatively small, the effect of the productivity difference will prevail and the consumption of economy 1 will be larger than that of economy 2 at steady state under sustainable heterogeneity.

C3 An alternative definition of the representative household

C3.1 The definition

Section C2 indicates that, when sustainable heterogeneity is achieved, all heterogeneous households are connected (in the sense that all households behave by considering other households' optimality) and appear to be behaving collectively as a combined supra-household that unites all households, as equations (C7) and (C8) indicate. The supra-household is unique and its behavior is time-consistent. Its actions always and consistently represent those of all households. Considering these natures of households under sustainable heterogeneity, I present the following alternative definition of the representative household: "the behavior of the representative household is defined as the collective behavior of all households under sustainable heterogeneity."

Even if households are heterogeneous, they can be represented by a representative household as defined above. Unlike the representative household defined as the average household, the collective representative household reaches a steady state where all households satisfy all of their optimality conditions in dynamic models. In addition, this representative household has a RTP that is equal to the average RTP as shown in equations (C7) and (C8).[9]

[9] If sustainable heterogeneity is achieved with the help of the government's intervention, the time preference rate of

Hence, we can assume not only a representative household but also that its RTP is the average rate of all households.

C3.2 Universality of sustainable heterogeneity

An important point, however, is that this alternatively defined representative household can be used in dynamic models only if sustainable heterogeneity is achieved, but this condition is not necessarily always naturally satisfied. Sustainable heterogeneity is achieved only if households with lower RTPs behave multilaterally or the government appropriately intervenes. Therefore, the representative household assumption is not necessarily naturally acceptable in dynamic models unless it is confirmed that sustainable heterogeneity is usually achieved in an economy.

Notwithstanding this flaw, the representative household assumption has been widely used in many macroeconomic studies that use dynamic models. Furthermore, these studies have been little criticized for using the inappropriate representative household assumption. In addition, in most economies, the dire state that Becker (1980) predicts has not been observed even though RTPs of households are unquestionably heterogeneous. These facts conversely indicate that sustainable heterogeneity—probably with government interventions—has been usually and universally achieved across economies and time periods. In a sense, these facts are indirect evidence that sustainable heterogeneity usually prevails in economies.

Note that because the representative household's behavior in dynamic models is represented by the collective behavior of all households under sustainable heterogeneity, RH's RTP is not intrinsically known to households, but they do need to have an expected rate. Each household intrinsically knows its own preferences, but it does not intrinsically know the collective preference of all households. Therefore, in dynamic models, it must be assumed that all households do not *ex ante* know RH's RTP, but households estimate it from information on the behaviors of other households and the government.

C4 NEED FOR AN EXPECTED RTP RH

C4.1 The behavior of household

Achieving sustainable heterogeneity affects the behavior of the individual household because sustainable heterogeneity indicates that each household must consider the other households' optimality (as well as the behavior of the government, if necessary). This feature does not mean that households behave cooperatively with other households. Each household behaves non-cooperatively based on its own RTP, but at the same time, it behaves considering whether the other households' optimality conditions are achieved or not. This consideration affects the actions a household takes in that it affects the choice of a household's initial consumption.

Sustainable heterogeneity indicates that a household's future path of consumption has to be consistent with the future path of sustainable heterogeneity. Thereby, a household sets its initial consumption such that it will proceed on the path that is consistent with the path of sustainable heterogeneity and eventually reach a steady state.

C4.2 Deviation from sustainable heterogeneity

C4.2.1 Political elements

What happens if a household deviates from sustainable heterogeneity? A deviation means that a household sets its initial consumption at a level that is not consistent with sustainable heterogeneity. For less advantaged households (i.e., households with higher RTPs), the only way to satisfy all of their optimality conditions is to set their initial consumption consistent with sustainable heterogeneity. Therefore, they will not take the initiative to deviate. In contrast, the most advantaged households (i.e., those with the lowest RTP) can satisfy all of their optimality

---

the representative household will not be exactly equal to the average rate of time preference.

conditions even if they set initial consumption independent of sustainable heterogeneity. The incentive for the most advantaged household to select a multilateral path will be weak because the growth rate of the most advantaged household on the multilateral path is lower than that on the unilateral path.

When economy 1 selects the unilateral path, does economy 2 quietly accept the unfavorable consequences shown in Becker (1980)? From an economic perspective, the optimal response of economy 2 is the one shown in Harashima (2010): economy 2 should behave as a follower and accept the unfavorable consequences. However, if other factors—particularly political ones—are taken into account, the response of economy 2 will be different. Faced with a situation in which all the optimality conditions cannot be satisfied, it is highly likely that economy 2 would politically protest and resist economy 1. It should be emphasized economy 2 is not responsible for its own non-optimality, which is a result of economy 1's unilateral behavior in a heterogeneous population. Economy 2 may overlook the non-optimality if it is temporary, but it will not if it is permanent. As shown in Harashima (2010), the non-optimality is permanent, it is quite likely that economy 2 will seriously resist economy 1 politically.

If economy 1 could achieve its optimality only on the unilateral path, economy 1 would counter the resistance of economy 2, but this is not the case. Because of this, economy 2's demand does not necessarily appear to be unreasonable or selfish. Faced with the protest and resistance by economy 2, economy 1 may compromise or cooperate with economy 2 and select the multilateral path.

C4.2.2 Resistance

The main objective of economy 2 is to force economy 1 to select the multilateral path and to establish sustainable heterogeneity. This objective may be achieved through cooperative measures, non-violent civil disobedience (e.g., trade restrictions), or other more violent means.

Restricting or abolishing trade between the two economies will cost economy 1 because it necessitates a restructuring of the division of labor, and the restructuring will not be confined to a small scale. Large-scale adjustments will develop that involve all levels of divided labor, because they are all correlated with each other. For example, if an important industry had previously existed only in one economy, owing to a division of labor, and trade between the two economies was no longer permitted, the other economy would have to establish this industry while also maintaining other industries. As a result, economy 1 would incur non-negligible costs. More developed economies have more complicated and sophisticated divisions of labor, and restructuring costs from the disruption of trade will be much higher in developed economies. In addition, more resources will need to be allocated to the generation of technology because technology will also no longer be traded. Finally, all of the conventional benefits of trade will be lost. Trade is beneficial because of the heterogeneous endowment of resources, as the Heckscher-Ohlin theorem shows. Because goods and services are assumed to be uniform in the models presented in Appendix C, the benefits of trade are implicit in the models. However, in the real word, resources such as oil and other raw materials are unevenly distributed, so a disruption or restriction of trade will substantially damage economic activities on both national and international levels.

The damage done by trade restrictions has an upper limit, however, because the restructuring of the division of labor, additional resource allocation to innovation, and loss of trade benefits are all finite. Therefore, in some cases, particularly if economies are not sufficiently developed and division of labor is not complex, the damage caused will be relatively small. Hence, a disruption of trade (non-violent civil disobedience in the national models) may not be sufficiently effective as a means of resistance under some these conditions.

In some cases, harassment, sabotage, intimidation, and violence may be used, whether legal or illegal. In extreme cases, war or revolution could ensue. In such cases, economy 1 will be substantially damaged in many ways and be unable to achieve optimality. The resistance and resulting damages will continue until sustainability is established.

In any case, the objective of economy 2's resistance conversely implies that establishing sustainability eliminates the risk and cost of political and social instability. The resistance of economy 2 will lower the desire of economy 1 to select the unilateral path.

C4.2.3 United economies

An important countermeasure to the fragility of sustainable heterogeneity for less advantaged economies is the formation of a union of economies. If economies other than economy 1 are united by commonly selecting the multilateral path within them, their power to resist economy 1 will be substantially enhanced. Consider the multi-economy model shown in Harashima (2010). If the economies do not form a union, the power to resist the unilateral actions of economy 1 is divided and limited to the power of each individual economy. However, if the economies are united, the power to resist economy 1 increases. If a sufficient number of economies unite, the multilateral path will almost certainly be selected by economy 1.

To maintain the union, any economy in the union should have the explicit and resolved intention of selecting the multilateral path within the union, even if it is relatively more advantaged within the union. To demand that relatively more advantaged economies select the multilateral path, less advantaged economies themselves must also select the multilateral path in any case. Otherwise, less advantaged economies will be divided and ruled by more advantaged economies. For all heterogeneous people to happily coexist, all of them should behave multilaterally. At the same time, Harashima (2010) indicates that the more advantaged an economy is, the more modestly it should behave, i.e., the more it should restrain itself from accumulating extra capitals.

In general, therefore, the most advantaged (the lowest RTP) household will be forced to set its initial consumption consistent with sustainable heterogeneity.

C4.3 Need for an expected RTP RH

Because all households need to set their initial consumption consistent with sustainable heterogeneity to achieve it, households must calculate the path of sustainable heterogeneity before setting their initial consumption levels. To calculate this level, each household first must know the value of RTP RH. However, although a household naturally knows the value of its own RTP, it does not intrinsically know the value of RTP RH. To know this, a household would have to know the values of all of the other households' RTPs. Hence, the expected value of RTP RH must somehow be generated utilizing all other relevant available information. The necessity of an expected RTP RH is critically important because RTP plays a crucial role as the discount factor in dynamic models.

Note that, if we assume that RTP is identical for all households, an expected RTP RH is no longer needed because any household's own RTP is equal to the RTP RH. This solution is still problematic, however, because the assumption is not merely expedient for the sake of simplicity; rather, it is a critical requirement to eliminate the need for an expected RTP RH. Therefore, any rationale for assuming identical RTPs should be validated; that is, it should be demonstrated that identical RTPs do exist and are universally observed. However, RTP is unquestionably not identical among households. Therefore, households must use expected values of RTP RH.

C5 The RTP model

C5.1 Need to know the structural model

If RTP RH is a constant parameter, as has been long and widely assumed, the need for an expected RTP RH would not be a serious problem. The historical mean of an unchanging RTP RH could be estimated relatively precisely based on long-term data of various economic indicators even if the structural model remained unknown. The RTP RH could be specified as the RTP that is most consistent with long-term trends of the indicators.

Although RTP has been treated as a constant parameter in many studies, this feature has not been demonstrated either empirically or theoretically. Rather, the assumption is merely expedient for the sake of simplicity. There is another practical reason for this treatment: models with a permanently constant RTP exhibit excellent tractability (see Samuelson, 1937). However, some have argued that it is natural to view RTP as temporally variable, and the concept of a temporally varying RTP has a long history (e.g., Böhm-Bawerk, 1889; Fisher, 1930). More recently, Lawrance (1991) and Becker and Mulligan (1997) showed that people do not inherit permanently constant RTPs by nature and that economic and social factors affect the formation of RTPs. Their arguments indicate that many incidents can affect and change RTP. Models of endogenous RTP have been presented, the most familiar of which is Uzawa's (1968) model.

If the RTP RH is temporally variable, its future stream must be expected by households, and a rational expectation is a model-consistent expectation. To generate rational expectations of RTP RH, therefore, the structural model of the RTP RH (i.e., equations that fundamentally describe how it is endogenously formed) needs to be known.

C5.2 Endogenous RTP models

C5.2.1 Uzawa's (1968) model

The most well-known endogenous RTP model is that of Uzawa (1968). It has been applied in many analyses (e.g., Epstein and Hynes, 1983; Lucas and Stokey, 1984; Epstein, 1987; Obstfeld, 1990). However, Uzawa's model has not necessarily been regarded as a realistic expression of the endogeneity of RTP because it has a serious drawback in that impatience increases as income, consumption, and utility increase. The basic structure of Uzawa's model is

$$\theta_t = \theta^*[u(c_t)] \ ,$$

$$0 < \frac{d\theta_t}{du(c_t)} \ ,$$

in which RTP in period t (θ_t) is temporally variable and an increasing function of present utility $u(c_t)$ where c_t is consumption in period t. The condition $0 < \frac{d\theta_t}{du(c_t)}$ is necessary for the model to be stable. This property is quite controversial and difficult to accept *a priori* because many empirical studies have indicated that RTP is negatively correlated with permanent income (e.g., Lawrance, 1991); thus, many economists are critical of Uzawa's model. Epstein (1987), however, discussed the plausibility of increasing impatience and offered some counter-arguments. However, his view is in the minority, and most economists support arguments in favor of a decreasing RTP, such that $\frac{d\theta_t}{du(c_t)} < 0$. Hence, although Uzawa's model attracted some attention, the analysis of the endogeneity of RTP has progressed very little. Although Uzawa's model may be flawed, it does not mean that the conjecture that RTP is influenced by future income, consumption, and utility is fallacious. Rather, it means that an appropriate model in which RTP is negatively correlated with income, consumption, and utility has not been presented.

C5.2.2 Size effect on impatience

The problem of $0 < \frac{d\theta_t}{du(c_t)}$ in Uzawa's model arises because distant future levels of consumption have little influence on factors that form RTP; that is, RTP is formed only with the information on present consumption, and it must be revised every period in accordance with consumption growth. However, there is no *a priori* reason why information on distant future activities should be far less important than the information on the present and near future activities. Fisher (1930) argued that

> [O]ur first step, then, is to show how a person's impatience depends on the *size* of his income, assuming the other three conditions to remain constant; for, evidently, it is possible that two incomes may have the same time shape, composition and risk, and yet differ in size, one being, say, twice the other in every period of time.
>
> In general, it may be said that, other things being equal, the smaller the income, the higher the preference for the present over the future income. It is true of course that a permanently small income implies a keen appreciation of wants as well as of immediate wants. … But it increases the want for immediate income *even more* than it increases the want for future income. (p. 72)

According to Fisher's (1930) view, a force that influences RTP is a psychological response derived from the perception of the "size of the entire income or utility stream." This view indicates that it is necessary to probe how people perceive the size of the entire income or utility stream.

Little effort has been directed toward probing the nature of the size of the utility or income stream on RTP, although numerous psychological experiments have been performed with regard to the anomalies of the expected utility model with a constant RTP (e.g., Frederick et al., 2002). Analyses using endogenous RTP models so far have merely introduced the *a priori* assumption of endogeneity of RTP without explaining the reasoning for doing so in detail. Hence, even now, Fisher′s (1930) insights are very useful for the examination of the size effect. An important point in Fisher′s quote is that the size of the infinite utility stream is perceived as "permanently" high or low. The size difference among the utility streams may be perceived as a permanently continuing difference of utilities among different utility streams. Anticipation of a permanently higher utility may enhance an emotional sense of well-being because people feel they are in a long-lasting secure situation, which will generate a positive psychological response and make people more patient. If that is true, distant future utilities should be taken into account equally with present utility. Otherwise, it is impossible to distinguish whether the difference of utilities will continue permanently.

From this point of view, the specification that only the present utility influences the formation of RTP, as is the case of Uzawa′s model, is inadequate. Instead, a simple measure of the size where present and future utilities are summed with equal weight will be a more appropriate measure of the size of a utility stream.[10]

C5.3 Model of RTP[11]

C5.3.1 The model

The representative household solves the maximization problem as shown in Section C1.3. Taking the arguments in Section C5.2 into account, the "size" of the infinite utility stream can be defined as follows.

Definition 1: The size of the utility stream W for a given technology A is

[10] Das (2003) showed another stable endogenous time preference model with decreasing impatience. Her model is stable, although the rate of time preference is decreasing because endogenous impatience is almost constant. In this sense, the situation her model describes is very special.

[11] The idea of this type of endogenous time preference model was originally presented in Harashima (2004a).

$$W = \lim_{T \to \infty} E \int_0^T \rho(t) u(c_t) dt \quad ,$$

where E is the expectation operator, and

$$\rho(t) = \frac{1}{T} \quad \text{if} \quad 0 \le t \le T$$
$$\rho(t) = 0 \qquad \text{otherwise.}$$

$\rho(t)$ indicates weights and has the same value in any period. Thus, the weights for the evaluation of future utilities are distributed evenly over time, as discussed in Section C5.2.

To this point, technology A has been assumed to be constant. If A is temporally variable (A_t) and grows at a constant rate and the economy is on a balanced growth path such that A_t, y_t, k_t, and c_t grow at the same rate, then the definition of W needs to be modified because any stream of c_t and $u(c_t)$ grows to infinity. It is then impossible to distinguish the sizes of the utility stream by simply summing up c_t as $T \to \infty$ as shown in Definition 1. Because balanced growth is possible only when technological progress is Harrod neutral, I assume a Harrod neutral production function such that

$$y_t = \omega A_t^{\varpi} k_t^{1-\varpi} \quad ,$$

where $\varpi(0 < \varpi < 1)$ and $\omega(0 < \omega)$ are constants. To distinguish the sizes of utility stream, the following value is set as the standard stream of utility,

$$u(\tilde{c}\, e^{\psi t}) \quad ,$$

where $\tilde{c}\,(0 < \tilde{c})$ is a constant and $\psi(0 < \psi)$ is a constant rate of growth. Streams of utility can be compared with this standard stream. If a constant relative risk aversion utility function is assumed, a stream of utility can be compared with the standard stream of utility as follows:

$$\frac{u(c_t)}{u(\tilde{c} e^{\psi t})} = \frac{c_t^{1-\gamma}}{(\tilde{c} e^{\psi t})^{1-\gamma}} = \frac{1-\gamma}{\tilde{c}^{1-\gamma}} u\left(\frac{c_t}{e^{\psi t}}\right) \quad .$$

By using this ratio, a given stream of utility can be distinguished from the standard stream of utility. That is, the size of a utility stream W for a given stream of technology A_t that grows at the same rate ψ as y_t, k_t, and c_t can be alternatively defined as

$$W = \lim_{T \to \infty} E \int_0^T \rho(t) u\left(\frac{c_t}{e^{\psi t}}\right) dt \quad .$$

Clearly, if $\psi = 0$, then the size (W) degenerates into the one shown in Definition 1.

If there is a steady state such that

$$\lim_{t \to \infty} E\,[u(c_t)] = E[u(c^*)] \quad ,$$

or for the case of expected balanced growth,

$$\lim_{t \to \infty} E\left[u\left(\frac{c_t}{e^{\psi t}} \right) \right] = E\left[u\left(c^* \right) \right] \quad ,$$

where c^* is a constant and indicates steady-state consumption, then

$$W = E\left[u\left(c^* \right) \right]$$

for the following reason. Because $\lim_{t \to \infty} E[u(c_t)] = E\left[u\left(c^* \right) \right]$ (or $\lim_{t \to \infty} E\left[u\left(\frac{c_t}{e^{\psi t}} \right) \right] = E\left[u\left(c^* \right) \right]$), then

$$\lim_{T \to \infty} \int_0^T \rho(t) \left\{ E\left[u\left(c^* \right) \right] - E[u(c_t)] \right\} dt = E\left[u\left(c^* \right) \right] - W$$

$$\left(\text{or } \lim_{T \to \infty} \int_0^T \rho(t) \left\{ E\left[u\left(c^* \right) \right] - E\left[u\left(\frac{c_t}{e^{\psi t}} \right) \right] \right\} dt = E\left[u\left(c^* \right) \right] - W \right) .$$

In addition,

$$\lim_{T \to \infty} \int_0^T \rho(t) \left\{ E\left[u\left(c^* \right) \right] - E[u(c_t)] \right\} dt = 0$$

$$\left(\text{or } \lim_{T \to \infty} \int_0^T \rho(t) \left\{ E\left[u\left(c^* \right) \right] - E\left[u\left(\frac{c_t}{e^{\psi t}} \right) \right] \right\} dt = 0 \right) .$$

Hence, $W = E\left[u\left(c^* \right) \right]$; that is, RTP is determined by steady-state consumption (c^*).

The RTP model presented in Appendix C is constructed on the basis of this measure of W. An essential property that must be incorporated into the model is that RTP is sensitive to, and a function of, W such that

$$\theta = \theta^{**}(W) \quad ,$$

where $\theta^{**}(W)$ is monotonically continuous and continuously differentiable. Because W is a sum of utilities, this property simply reflects the core idea of an endogenous RTP. However, this property is new in the sense that RTP is sensitive not only to the present utility but also to the entire stream of utility, that is, the size of the utility stream represented by the utility of steady-state consumption. This property is intuitively acceptable because it is likely that people set their principles or parameters for their behaviors considering the final consequences of their behavior (i.e., the steady state; see, e.g., Barsky and Sims, 2012).

Another essential property that must be incorporated into the model is

$$\frac{d\theta}{dW} < 0 \quad .$$

Because $W = E\left[u\left(c^* \right) \right]$ and $0 < \frac{du(c_t)}{dc_t}$, RTP is inversely proportionate to c^*. This property is consistent with the findings in many empirical studies, which have shown that RTP is negatively correlated with permanent income (e.g., Lawrance, 1991).

In summary, the basic structure of the model is:

$$\theta = \theta^{**}(W) = \theta^{**}\{E[u(c^*)]\} \ ,$$
$$\frac{d\theta}{dW} = \frac{d\theta}{dE[u(c^*)]} < 0 \ . \tag{C9}$$

This model is deceptively similar to Uzawa's endogenous RTP model and simply replaces c_t with c^* and $0 < \frac{d\theta_t}{du(c_t)}$ with $\frac{d\theta}{dE[u(c^*)]} < 0$. However, the two models are completely different because of the opposite characteristics of $0 < \frac{d\theta_t}{du(c_t)}$ and $\frac{d\theta}{dE[u(c^*)]} < 0$.

C5.3.2 Nature of the model

The model can be regarded as successful only if it exhibits stability. In Uzawa's model, the economy becomes unstable if $0 < \frac{d\theta_t}{du(c_t)}$ is replaced with $\frac{d\theta_t}{du(c_t)} < 0$. In this section, I examine the stability of the model.

C5.3.2.1 Equilibrium RTP

In Ramsey-type models, such as shown in Section C1.3, if a constant RTP is given, the value of the marginal product of capital (i.e., the value of the real interest rate) converges to that of the given RTP as the economy approaches the steady state. Hence, when a RTP is specified at a certain value, the corresponding expected steady-state consumption is uniquely determined. Given fixed values of other exogenous parameters, any predetermined RTP has unique values of expected consumption and utility at steady state. There is a one-to-one correspondence between the expected utilities at steady state and the RTPs; therefore, the expected utility at steady state can be expressed as a function of RTP. Let c_x^* be a set of steady-state consumption levels, given a set of RTPs (θ_X) and other fixed exogenous parameters. The concept of $\theta \rightarrow W$ discussed above can be described as

$$g(\theta) = E[u(c^*)](=W) \ , \tag{C10}$$

where $c^* \in c_x^*$ and $\theta \in \theta_x$. On the other hand, RTP is a continuous function of steady-state consumption as shown in equation (C9) such that $\theta = \theta^{**}(W) = \theta^{**}\{E[u(c^*)]\}$. The reverse function is

$$h(\theta) = E[u(c^*)](=W) \ . \tag{C11}$$

The equilibrium RTP is determined by the point of intersection of the two functions, $g(\theta)$ and $h(\theta)$, as shown in Figure C1. Figure C2 shows the special but conventionally assumed case for $h(\theta)$ in which θ is not sensitive to W, and RTP is constant. There exists a point of intersection because both $g(\theta)$ and $h(\theta)$ are monotonically continuous for $\theta > 0$. $h(\theta)$ is monotonically continuous because $\theta^{**}(W)$ is monotonically continuous. $g(\theta)$ is monotonically continuous because, as a result of utility maximization, $c^* = f(k^*)$ and

$\theta = \dfrac{df\left(k^*\right)}{dk^*}$, where k^* is capital input per capita at steady state such that $k^* = \lim_{t\to\infty}\left(k_t\right)$. Because $f\left(k^*\right)$ and $\dfrac{df\left(k^*\right)}{dk^*}$ are monotonically continuous for $k^* > 0$, c^* is a monotonically continuous function of θ for $\theta > 0$. Here, because u is monotonically continuous, then $E\left[u\left(c^*\right)\right] = g(\theta)$ is also monotonically continuous for $\theta > 0$.

The function $g(\theta) = E\left[u\left(c^*\right)\right] = W$ is a decreasing function of θ because higher RTP results in lower steady state consumption. The function $h(\theta) = E\left[u\left(c^*\right)\right] = W$ is also a decreasing function of θ because $\dfrac{d\theta}{dW} < 0$. Thus, both $g(\theta)$ and $h(\theta)$ are decreasing, but the slope of $h(\theta)$ is steeper than that of $g(\theta)$ as shown in Figure C1. This is true because $g(\theta) = W$ is the consequence of a Ramsey-type model as shown in Section C1.3; thus, if $\theta \to \infty$, then $g(\theta) = W \to 0$ because $\theta = i_t \to \infty$ and $k_t \to 0$, and if $\theta \to 0$, then $g(\theta) = W \to \infty$ because $\theta = i_t \to 0$ and $k_t \to \infty$. The function $h(\theta) = W$ indicates the endogeneity of RTP, and because RTP is usually neither zero nor infinity, then even if $h(\theta) = W \to 0$, $\theta < \infty$, and $h(\theta) = W \to \infty$, $0 < \theta$. Hence, the locus $h(\theta) = W$ cuts the locus $g(\theta) = W$ downward from the top, as shown in Figure C1. Hence, the locus $h(\theta) = W$ is more vertical than $g(\theta) = W$, and thereby a permanently constant RTP, as shown in Figure C2, has probably been used as an approximation of the locus $h(\theta) = W$ for simplicity.

C5.3.2.2 Stability of the model

RTP is constant unless a shock that changes the expected c^* occurs because W does not depend on t but on the expected c^*. Thus, the same RTP and steady state continue until such a shock hits the economy. Therefore, the endogeneity of RTP only matters when a shock occurs. This constancy is the key for the stability of the model. Once the RTP corresponding to the intersection (Fig. 1) is determined, it is constant and the economy converges at a unique steady state unless a shock that changes the expected c^* occurs. The shock is exogenous to the model, and the economy does not explode endogenously but stabilizes at the steady state. Hence, the property $\dfrac{d\theta}{dW} < 0$ in the model, which is consistent with empirical findings, does not cause instability.

The model is therefore acceptable as a model of endogenous RTP. Furthermore, because RTP is endogenously determined, the assumption of irrationality is not necessary for the determination of RTP. Nevertheless, a shock on RTP can be initiated by a shock on the expected c^*; thus, even if the so-called animal spirits are directly irrelevant to determination of RTP, they may be relevant in the generation of shocks on the expected c^*.

C6 FREQUENT RTP SHOCKS

C6.1 Difficulty in knowing RTP RH

To estimate the parameter values of equation (C11) in the structural model of RTP RH, it is necessary to obtain a sufficiently large amount of data on the value of RTP RH. To obtain these data, a household must know the RTPs of all the other households. Although a household knows its own RTP, it has almost no information about the RTPs of all the other households much less time-series data on each household's RTP. Because of the lack of available data, a household cannot estimate the parameter values in equation (C11) in the structural model of RTP RH even if it knows the functional forms of equations in the structural model.

We can easily generate data on aggregate consumption, investment, production, inflation, trade, and other factors at a relatively low cost, but we cannot directly observe the value of RTP RH. Nonetheless, many estimates of RTP have been reported, but they are not based on a structural model of RTP. Most are the results of experimental studies or indirect estimates based on other models (e.g., Ramsey growth models) on the assumption that RTP is constant. Experiments can give us some information on the RTPs of test subjects, but we should not naively use these estimates as the RTP RH in the calculation of the future path of economy because they vary widely according to the experimental environments. Furthermore, most of the indirect estimates were calculated on the assumption that RTP is constant, which as discussed previously, is most likely not the case. The basic problem is that no credible estimation method of RTP RH has been established.

C6.2 Expectations based on beliefs

The lack of observable data on RTP RH will significantly hinder households from generating rational expectations of the future path of economy. How do households rationally expect their future streams of consumption and production and calculate their optimal paths without information on RTP RH, which is indispensable as the discount factor? The historical mean of RTP RH estimated by long-term data is not consistent with a rational expectation of the future stream because RTP is not constant. Without a reliable method for estimating the parameters of the structural model, it is impossible for households to generate rational expectations of the future path of the economy.

An alternative way of estimating expected values of RTP RH is needed, but even if an alternative method is utilized, households still have to behave as rationally as possible even in an environment of significantly incomplete information. In this situation, household may have to use the concept of bounded rationality to make decisions. It is possible that the only alternative for a household is to use its "belief" about the RTP RH. The use of a belief does not mean that households deviate from rationality; rather, it is the most rational behavior they can use in an environment where insufficient information is available.

Such a belief is defined in Appendix C as the range of values of RTP RH within which a household believes that the true RTP RH exists. Households utilize the belief in place of equation (C11). More specifically, suppose that household i ($i \in N$) believes that the RTP RH in the future is situated in the range λ_i, where the subjective probability density at any point on λ_i is identical (i.e., its distribution shape is uniform). Because households have no information about the shape of the distribution, they assume that it is uniform. This supposition means that household i believes that λ_i is stationary. Let $\bar{\lambda}_i$ be the mean of λ_i. Suppose that household i calculates its optimal future path on the belief that the mean of future values of RTP RH is $\bar{\lambda}_i$. By equation (C10), W can be calculated based on $\bar{\lambda}_i$, and the expected future path of economy can be calculated.

Households can equally access all relevant information. Therefore, if the belief of a household is very different from those of the majority, the household will soon perceive that its belief is different, through observing the behavior of majority. The household will change its belief to the almost same as those of the majority because otherwise it cannot achieve optimality as expected on the assumption that sustainable heterogeneity is achieved. Hence, it is likely that households' beliefs become similar, and thereby, it is assumed for simplicity that households' beliefs are identical.

Note that households do not cooperatively and collectively expect the future path of economy (i.e., the representative household's future path), but each household independently and individually generates its own expectations based on its belief in RTP RH. The household thereby creates its own expected future path considering the expected representative household's future path. The aggregates are the sum of all household's independent and individual activities, but if sustainable heterogeneity is achieved, the aggregates appear to be the same as the results of the representative household's activities.

C6.3 Refining beliefs

A household knows that its expectation is based on its beliefs and not the structural model. Therefore, it will always want to refine the belief, that is, raise the probability that the belief is the correct value, by exploiting all currently available relevant information. Let a set of currently available economic indicators be $\boldsymbol{I}_t$ (e.g., the observed data on consumption, production, inventory, etc.). These data may provide some useful information on the past RTP RH, and a household may refine its belief based on this information. These data and equation (C10) can be used to generate estimates of past values of RTP RH. However, $\boldsymbol{I}_t$ includes noise, and data in $\boldsymbol{I}_t$ will usually be somewhat inconsistent between the elements of $\boldsymbol{I}_t$. In addition, because equation (C10) indicates the steady state values that are achieved after a long-period transition, the short-term past data included in $\boldsymbol{I}_t$ are basically insufficient to obtain a credible estimate. Therefore, the estimate of the past values of RTP RH based on $\boldsymbol{I}_t$ and equation (C10) will usually have a large confidence interval. Let $\bar{\mu}_I$ be the estimated past RTP RH and μ_I be its confidence interval of, for example, 95%. Because households can equally access all relevant information, assume for simplicity that μ_I and $\bar{\mu}_I$ are identical for all households.

Although a household knows that $\bar{\mu}_I$ is not a credible estimate, has a large confidence interval, and is merely an estimate (usually a point estimate) of a past value, it will strive to utilize the information derived from $\bar{\mu}_I$ to refine its beliefs in the future value of RTP RH. Usually $\bar{\mu}_I$ will not be equal to $\bar{\lambda}_i$, but the ranges of λ_i and μ_I may partly overlap. Household i may utilize the information from this partial overlap to refine its belief (i.e., information of how λ_i is different from μ_I). $\bar{\mu}_I \neq \bar{\lambda}_i$ indicates that the belief $\bar{\lambda}_i$ is wrong, $\bar{\mu}_I$ is wrong, both are wrong, or both are right if the true past RTP RH is $\bar{\mu}_I$ but the true future RTP RH is $\bar{\lambda}_i$. The belief $\bar{\lambda}_i$ may be wrong because the RTP RH will change in the near future, and $\bar{\mu}_I$ may be wrong because the RTP RH changed during the period in which the data were obtained. In addition, a household knows that μ_I is the result of all households' activities based on their beliefs, not on the true value of RTP RH. These uncertainties arise because households cannot know the parameters of the structural model. Without using the structural model, household i cannot judge whether $\bar{\lambda}_i$ is wrong, $\bar{\mu}_I$ is wrong, both are wrong, or both are right. As a result, household i will not easily adjust its belief from $\bar{\lambda}_i$ to $\bar{\mu}_I$.

However, it is still likely that information about the difference between λ_i and μ_I can be used to refine the belief. To extract the useful information, the following rules may be used:

Rule 1: if $\bar{\mu}_I$ is included in λ_i, the belief is not adjusted; otherwise, the belief is adjusted from $\bar{\lambda}_i$ to $\bar{\mu}_I$.

Rule 2: if $\bar{\lambda}_i$ is included in μ_I, the belief is not adjusted; otherwise, the belief is adjusted from $\bar{\lambda}_i$ to $\bar{\mu}_I$.

Rule 3: if λ_i and μ_I overlap at or above a specified ratio, the belief is not adjusted; otherwise, the belief is adjusted from $\bar{\lambda}_i$ to $\bar{\mu}_I$.

The above rules may be seen as a type of adaptive expectation because μ_I indicates the past RTP RH. However, in the situation where the parameters of the structural model of the RTP RH are unknown, it may be seen as rational to utilize the information contained in μ_I by adopting one of these rules.

C6.4 Changing beliefs

However, it does not seem likely that a household will refine its belief following one of the rules shown above because the rules are basically backward looking and will not be adopted as a tool for refining the belief if a household is convinced that the RTP RH is temporally variable. The belief will only be changed if forward-looking information is available, that is, when a household becomes aware of information about the future RTP RH in μ_I. For example, the difference between λ_i and μ_I may reflect an unexpected and large positive technology shock that occurred after the formation of belief λ_i. Because the effects of the technology shock will persist for long periods in the future, household i will most likely change its belief. In this case, a household will not simply refine its belief from $\bar{\lambda}_i$ to $\bar{\mu}_I$; it will change to another value that is formed as an entirely new belief.

Whether a household changes its belief or not, therefore, will depend not simply on μ_I but on the information the household can extract from μ_I about the future path of the economy. Hence, in some cases, a household will change its belief when new values of μ_I are obtained, but in other cases, it will not, depending on how the household interprets the information contained in μ_I.

C6.5 Heuristics

When a household interprets μ_I, it may also use heuristic methods, for example, a simplified linear reduced form model of RTP RH. Studies of the use of heuristics and bounded rationality in this context would be useful for better understanding the interpretation mechanism of μ_I. There are many possible simplified linear reduced form models of RH's RTP that could be used as heuristic methods although most of them may be ad hoc. Even though such reduced form models are far less credible than a structural model, they may be utilized as a heuristic method of interpreting μ_I by households. Although these types of models may often result in misleading conclusions, they may sometimes provide useful information. For example, if a linear correlation between RTP RH and a financial indicator exists, even if it is weak or temporary, changes in the financial indicator may contain useful information about changes in the RTP RH. Therefore, if a household believes that this correlation exists, it will use this information to interpret μ_I.

C6.6 Frequent RTP shocks

Households must have expected values of RTP RH for sustainable heterogeneity, but as previously discussed, the expectations are not based on the structural model but rather on a belief that is not guaranteed to generate the correct value. In addition, the belief can be influenced by heuristic considerations. These features indicate that the expected values of RTP RH will fluctuate more frequently than the intrinsic RTP RH.

Households' expectations of RTP RH will change when the intrinsic RTP RH shifts, for example, when new information about shocks on the factors that determine equation (C10) becomes available. For a given θ, $E[u(c^*)]$ changes if the expectation of future productivity changes. Productivity at the macro level will be influenced by scientific technology, financial technology, social infrastructure, and other factors. If expectations about these factors in the future changes, the expected future productivity and $E[u(c^*)]$ will also change. In addition, even if intrinsic RTP RH does not change, the expected RTP RH will change if a household's belief is altered because of new information contained in μ_I. Hence, the expected RTP RH can change independently of intrinsic changes in RTP RH. Therefore, even if intrinsic changes in RTP RH occur infrequently, changes in the expected RTP RH may occur more frequently.

A household's expected RTP RH can potentially change every time new information on μ_I becomes available if it contains the information that makes beliefs change. Information concerning factors that affect the expected RTP RH will become available frequently, and at least some of the information may be both very important and unexpected. In addition, there will be many disturbances in the fundamental factors that affect equation (C10), and many of these disturbances will also cause μ_I to change. As discussed previously, a household may interpret these changes in μ_I as a change in the true RTP RH. Therefore, it is likely that households' expected RTP RH change more frequently than the intrinsic RTP RH, and thereby, that time preference shocks also occur more frequently than previously thought.

Even a small piece of additional information about the belief can significantly change the path of the economy. For example, if many households believe a rumor (whether it is true or not) related to information about the interpretation of μ_I and respond similarly to it, their expectations will be changed in the same direction by the rumor. If all households respond similarly to an untrue rumor and change their expectations equally to an untrue value, the economy will proceed based on the incorrect expectation of RTP RH. The $\overline{\mu}_I$ that is observed a few periods later will follow these wrongly expected values of RTP RH. Upon obtaining new data of $\overline{\mu}_I$ that are consistent with these wrongly expected values, households will judge that their (incorrect) changes were in fact correct. As a result, the incorrect expectations become self-fulfilling. This spurious situation may reach an impasse at some point in the future because the expectations are based not on a structural model but on the (incorrect) beliefs. Households will not anticipate the impasse until the economy reaches it because they believe that the wrongly expected RTP RH (i.e., the currently held belief) is true.

References

Alesina, Alberto and Alex Cukierman. (1990) "The Politics of Ambiguity," *The Quarterly Journal of Economics*, Vol. 105, No. 4, pp. 829–850.

Alesina, Alberto and Allen Drazen. (1991) "Why Are Stabilizations Delayed?" *International Economic Review,* Vol. 81, No. 5, pp. 1170–88.

Antonelli, Giovanni Battista (1886) *Sulla Teoria Matematica Della Economia Politica*, Pisa, Tipografia del Falchetto.

Barro, Robert J. and David B. Gordon. (1983) "A Positive Theory of Monetary Policy in a Natural Rate Model," *The Journal of Political Economy*, Vol. 91, No. 4. pp. 589–610.

Barsky, Robert B. and Eric R. Sims (2012) "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence," *American Economic Review*, Vol. 102, No. 4. pp. 1343-77.

Becker, Gary S. and Casey Mulligan (1997) "The Endogenous Determination of Time Preference," *Quarterly Journal of Economics*, Vol. 112, No. 3, pp. 729–758.

Becker, Robert A. (1980) "On the Long-run Steady State in a Simple Dynamic Model of Equilibrium with Heterogeneous Households," *The Quarterly Journal of Economics*, Vol. 95, No. 2, pp. 375–382.

Bernanke, B. S. and Reinhart, V.R. (2004) "Conducting Monetary Policy at Very Low Short-Term Interest Rates," *American Economic Review*, Vol. 94, pp. 85-90.

Bernanke, Ben, Vincent R. Reinhart, and Brian P. Sack (2004) "Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment," *Brookings Papers on Economic Activity*, No. 2, pp. 1-100.

Blinder, Alan S. (1998) *Central Banking in Theory and Practice*, The MIT Press, Cambridge.

Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J. and Jansen, D. (2008) "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence," *Journal of Economic Literature*, Vol. 46, pp. 910-45.

Böhm-Bawerk, Eugen von (1889) *Capital and Interest*, Reprinted by Libertarian Press, South Holland, IL, 1970.

Brennan, Geoffrey and James M. Buchanan. (1980) *The power to tax: analytical foundations of a fiscal constitution*, Cambridge University Press, Cambridge MA.

Buiter, Willem H. (2005) "Overcoming the Zero Bound on Nominal Interest Rates: Gesell's Currency Carry Tax vs. Eisler's Parallel Virtual Currency," *Hitotsubashi University Research Unit for Statistical Analysis in Social Sciences, Discussion Paper Series*, No.96.

Cattell, Raymond Bernard (1963) "Theory of Fluid and Crystallized Intelligence: A Critical Experiment," *Journal of Educational Psychology*, Vo. 54, pp. 1-22.

Cattell, Raymond Bernard (1971) *Abilities: Their structure, growth, and action*, Houghton Mifflin, Boston.

Christiano, Lawrence and Terry J. Fitzgerald. (2000) "Understanding the Fiscal Theory of the Price Level," *NBER Working Paper* No. 7668.

Cukierman, Alex, Steven B. Webb, and Bilin Neyapti. (1992) "Measuring the Independence of Central Banks and Its Effect on Policy Outcomes," *World Bank Economic Review*, Vol.6, No.3, pp. 353-398.

Coenen, G. and V. Wieland (2003) "The Zero-interest-rate Bound and the Role of the Exchange Rate for Monetary Policy in Japan," *Journal of Monetary Economics*, Vol. 50, pp. 1071-1101.

Coenen, G. and V. Wieland (2004) "Exchange Rate Policy and the Zero Bound on Nominal Interest Rates," *American Economic Review: Papers and Proceeding*, Vol. 94, pp. 80-84.

Das, Mausumi. (2003) "Optimal Growth with Decreasing Marginal Impatience," *Journal of Economic Dynamics & Control*, Vol. 27, No. 10, pp. 1881–1898.

Downs, Anthony. (1957) *An Economic Theory of Democracy*, Harper, New York.

Edwards, Jeremy and Michael Keen. (1996) "Tax Competition and Leviathan," *European Economic Review*, Vol. 40, No. 1, pp. 113–134.
Eggertsson, G. B. and Woodford, M. (2003) "The Zero Bound on Interest Rates and Optimal Monetary Policy," Brookings Papers on Economic Activity, No. 1, pp. 193-233.
Epstein, Larry G. and J. Allan Hynes (1983) "The Rate of Time Preference and Dynamic Economic Analysis," *Journal of Political Economy*, Vol. 91, No. 4, pp. 611-635.
Fisher, Irving (1930) *The Theory of Interest*, Macmillan, New York
Frederick, Shane, George Loewenstein and Ted O′ Donoghue (2002) ‟Time Discounting and Time Preference: A Critical Review,″ *Journal of Economic Literature*, Vol. 40, No. 2, pp. 351-401.
Fukao, Mitsuhiro (2005) "The Effects of 'Gesell' (Currency) Taxes in Promoting Japan's Economic Recovery," *Hitotsubashi University Research Unit for Statistical Analysis in Social Sciences, Discussion Paper Series*, No.94.
Galí, Jordi, Mark Gertler, and David López-Salido. (2005) "Robustness of the estimates of the hybrid New Keynesian Phillips curve," *Journal of Monetary Economics*, Vol. 52, No. 6, pp. 1107-1118.
Groves, P. M. and R. F. Thompson (1970) "Habituation: A Dual-process Theory," *Psychological Review*, Vol. 77, pp. 419–450.
Harashima, Taiji. (2004a) "A More Realistic Endogenous Time Preference Model and the Slump in Japan," *EconWPA Working Papers*, ewp-mac0402015.
Harashima, Taiji. (2004b) "The Ultimate Source of Inflation: A Microfoundation of the Fiscal Theory of the Price Level," *EconWPA Working Papers*, ewp-mac/ 0409018.
Harashima, Taiji. (2008) "A Microfounded Mechanism of Observed Substantial Inflation Persistence," *MPRA* (*The Munich Personal RePEc Archive*) *Paper*, No. 10668.
Harashima, Taiji. (2010) "Sustainable Heterogeneity: Inequality, Growth, and Social Welfare in a Heterogeneous Population," *MPRA* (*The Munich Personal RePEc Archive*) *Paper*, No. 24233.
Harashima, Taiji (2014a) "The Representative Household Assumption Requires Sustainable Heterogeneity in Dynamic Models," *Theoretical and Practical Research in Economic Fields*, Vol. 5, No. 1.
Harashima, Taiji (2014b) ‟Time Preference Shocks,″ *MPRA* (*The Munich Personal RePEc Archive*) *Paper*, No. 60205.
Harashima, Taiji. (2015a) "A Way Out of the Euro Crisis: Fiscal Transfers Are Indispensable for Sustainability in a Union with Heterogeneous Members," *MPRA* (*The Munich Personal RePEc Archive*) *Paper*, No. 63025.
Harashima, Taiji. (2015b) "The Rate of Time Preference of Government," *MPRA* (*The Munich Personal RePEc Archive*) *Paper*, No. 65387.
Kydland, Finn E. and Edward C. Prescott. (1977) "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, Vol. 85, No. 3, pp. 473–491.
Lawrance, Emily C. (1991) "Poverty and the Rate of Time Preference: Evidence from Panel Data," *Journal of Political Economy*, Vol. 99, No. 1, pp. 54–77.
Lord, Frederic M. and Melvin R. Novick. (1968) *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, MA.
Lucas, Robert E., Jr. and Nancy L. Stokey (1984) "Optimal Growth with Many Consumers," *Journal of Economic Theory*, Vol. 32, pp. 139-171.
Obstfeld, Maurice (1990) "Intertemporal Dependence, Impatience, and Dynamics," *Journal of Monetary Economics*, Vol. 26, No. 1, pp. 45-75.
Parkin, Michael (1988) "A Method for Determining Whether Parameters in Aggregative Models Are Structural," *Carnegie-Rochester Conference Series on Public Policy*, Vol. 29, No. 1, pp. 215–252.

Rankin, Catharine H., Thomas Abrams, Robert J. Barry, Seema Bhatnagar, David F. Clayton, John Colombo, Gianluca Coppola, Mark A. Geyer, David L. Glanzman, Stephen Marsland j, Frances K. McSweeney, Donald A. Wilson, Chun-Fang Wum, and Richard F. Thompson (2009) "Habituation Revisited: An Updated and Revised Description of the Behavioral Characteristics of Habituation," *Neurobiology of Learning and Memory*, Vol. 92, No. 2, pp. 135-138.

Samuelson, Paul (1937) "A Note on Measurement of Utility," *Review of Economic Studies*, Vol. 4, No. 2, pp. 155–161.

Samwick, Andrew A. (1998) "Discount Rate Heterogeneity and Social Security Reform," *Journal of Development Economics*, Vol. 57, No. 1, pp. 117–146.

Sidrauski, Miguel. (1967) "Rational Choice and Patterns of Growth in a Monetary Economy," *American Economic Review*, Vol. 57, No. 2, pp. 387–393.

Svensson, L. O. E. (2001) "The Zero Bound in an Open Economy: A Foolproof Way of Escaping from a Liquidity Trap," *Monetary and Economic Studies*, Vol. 19(S-1), pp. 277-312.

Tabellini, Guido and Alberto Alesina. (1990) "Voting on the Budget Deficit," *American Economic Review*, Vol. 80, No. 1, pp. 37–49.

Thompson, R. F. and W. A. Spencer (1966) "Habituation: A Model Phenomenon for the Study of Neuronal Substrates of Behavior," *Psychological Review*, Vol. 73, pp. 16–43.

Ugai, H (2007), "Effects of the Quantitative Easing Policy: a Survey of Empirical analyses", *Bank of Japan Monetary and Economic Studies*, 25, 1-47.

Uzawa, Hirofumi. (1968) "Time Preference, the Consumption Function, and Optimal Asset Holdings," in *Value, Capital, and Growth: Papers in Honour of Sir John Hicks* (J. N. Wolfe ed.), University of Edinburgh Press, Edinburgh, Scotland.

van der Linden, Wim J. and Ronald K. Hambleton (Eds.) (1997) *Handbook of Modern Item Response Theory*, Springer-Verlag, New York.

Ventura, Luigi. (2003) "Direct Measure of Time-preference," *Economic and Social Review*, Vol. 34, No. 3, pp. 293–310.

Williams, John (2009) "Heeding Daedalus: Optimal Inflation and the Zero Lower Bound," *Brookings Papers on Economic Activity*, No. 2, pp.1–37.

Woodford, Michael (2012) "Methods of Policy Accommodation at the Interest-rate Lower Bound," in Jackson Hole Symposium, August, Federal Reserve Bank of Kansas City.

Figure C1: Endogenous time preference

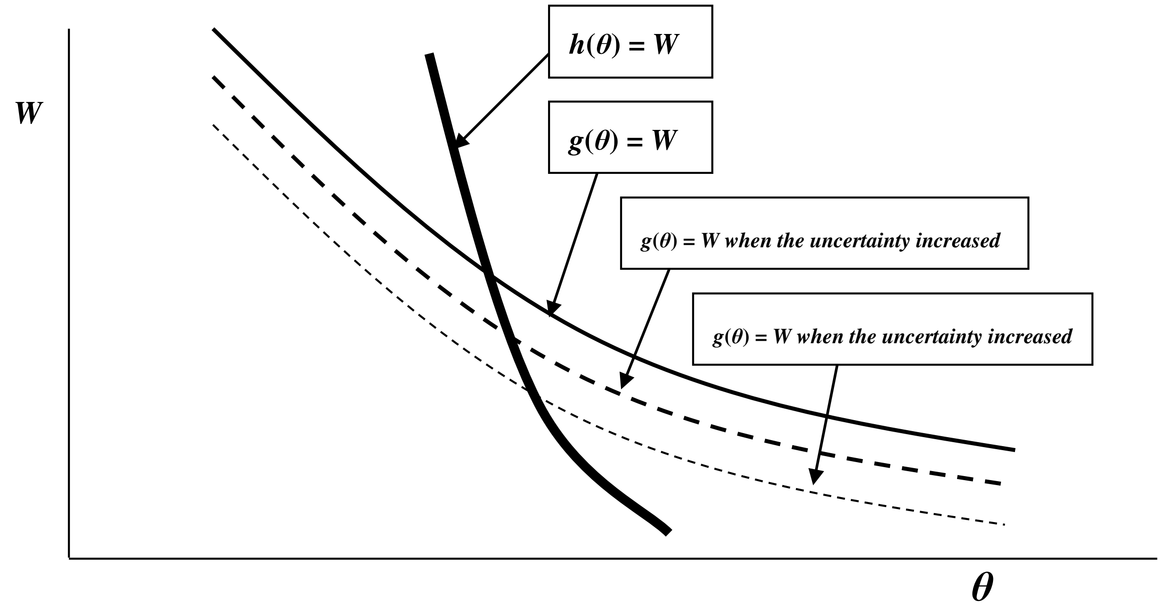

Figure C2: Permanently constant time preference