The Expectations Hypothesis of the Term Structure When Interest Rates Are Close to Zero*

Francisco J. Ruge-Murcia
Département de sciences économiques and CIREQ
Université de Montréal

Forthcoming in the *Journal of Monetary Economics*

Abstract

In an economy where cash can be stored costlessly in nominal terms, the nominal interest rate is bounded below by zero. This paper derives the implications of this non-negativity constraint for the term structure and shows that it induces a nonlinear and convex relation between short- and long-term interest rates. The long-term rate responds asymmetrically to changes in the short-term rate, and by less than is predicted by the benchmark linear model. In particular, a decrease in the short-term rate produces a smaller response in the long-term rate than an increase of the same magnitude. The empirical predictions of the model are examined using data from Japan.

JEL Classification: E43
Key Words: Limited-dependent rational-expectations models, nonlinear forecasting, monetary policy, Japan.

*This project was started while I was visiting scholar at the Research Department of the International Monetary Fund. I am grateful to the IMF for its hospitality and to its staff for their help collecting the data. I have benefited from comments by Rui Castro, Robert King, Benoit Perron, and an anonymous referee. Financial support from the Social Sciences and Humanities Research Council and the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche is gratefully acknowledged. Previous versions of this paper were circulated under the title "Some Implications of the Zero Lower Bound on Interest Rates for the Term Structure and Monetary Policy." *Correspondence:* Département de sciences économiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec) H3C 3J7, Canada. E-mail: francisco.ruge-murcia@umontreal.ca. Telephone: (514) 343 2396. Fax: (514) 343 5831.

1 Introduction

This paper studies the implications of the zero lower bound for the term structure of interest rates. The zero bound arises because agents would rather store the currency themselves, rather than lend it at a loss (Fisher, 1896). Since the values that the nominal interest rate can take are limited to the interval $[0, \infty)$, the bound is imposed here by modeling the one-period interest rate as a limited-dependent variable. The limited-dependent variable model is a device that forces agents to explicitly consider the zero lower bound when constructing their forecasts, even if all observations of the nominal interest rate to date are strictly positive. Closed-form analytical results are obtained for the case of a two-period bond and normally distributed disturbances. Numerical results, using a frequency simulator to compute the forecasts of the nonlinear model, are obtained for longer maturities under more general distributional assumptions.

The main implications of the non-negativity constraint are the following. First, the zero lower bound induces a nonlinear and convex relation between the long-term interest rate and the level and standard deviation of the short-term interest rate. Second, the response of long-term interest rates to changes in the short-term rate is asymmetric. A decrease in the short-term rate produces a smaller response (in absolute value) in the long-term rate than an increase of the same magnitude. Third, the response of long-term rates to changes in the short-term rate (whether an increase or a decrease) is smaller in the neighborhood of the zero lower bound, specially for longer maturities. Finally, the yield curve becomes steeper as the short-term interest rate approaches zero. The reason is that when the short-term interest rate is close to zero, agents understand that the range of its possible future realization is larger above than below the current rate. Hence, their forecasts of future short-term interest rates are above the current rate and, by the Pure Expectations Hypothesis (PEH), the spread between current long- and short-term rates increases. All these results, coupled with the observation that when the short-term interest rate is low the scope for further interest rate cuts is limited by the zero lower bound, imply that the power of monetary policy to affect long-term interest rates through the term structure is considerably reduced at low interest rates. Iwata and Wu (2005) construct a nonlinear Vector Autoregression (VAR) where the nominal interest rate is a limited-dependent variable and show that the effect of monetary policy on output is also smaller in the neighborhood of the zero lower bound.

The magnitude of these effects diminishes as the short-term interest rate rises above zero and it is negligible when the short-term rate is at a safe distance from the non-negativity constraint. However, in some OECD countries today, nominal interest rates are at historically low levels. For example, *The Economist*, 28 June 2003 (p. 97) reports that the three-month money market rates in Japan, Switzerland, and the United States are (as of 25 June) 0.01, 0.28 and 0.88 per cent per year, respectively. Thus, the implications of the zero lower bound outlined above may be empirically relevant for these countries. This paper examines the implications of the model using Japanese data because its recent experience provides us with sufficiently long, high frequency series of very low nominal interest rates. Results indicate that the nonlinear model, which takes into account the effect of the non-negativity constraint on expectations, delivers smaller forecast errors than a benchmark linear model. This is basically due to the fact that the nonlinear model more accurately predicts the slope of the Japanese yield curve at low interest rates. Impulse-response analysis indicates that adjustments to the short-term interest rate lead to asymmetric responses by the long-term rates. This asymmetry is proportionally more important for longer maturities. Finally, Ordinary Least Squares (OLS) results are roughly in line with these predictions and deliver quantitative implications that are numerically closer to the ones of the nonlinear model compared to those of the linear model.

The paper is organized as follows. Section 2 introduces the process for the one-period rate and derives the implications of the nonlinear model for the two-period rate when shocks are normally distributed. Section 3 derives the conditional expectations for longer horizons and under general distributional assumptions. Section 4 examines the empirical predictions of the model using data from Japan. Section 5 concludes.

2 The Two-Period Bond

This section *i*) presents a limited-dependent variable model for the one-period nominal interest rate that captures the idea that nominal interest rates are bounded below by zero, and *ii*) derives the implications of this non-negativity constraint for a long-term bond with maturity equal to two periods. Considering this special case first is instructive because, under certain conditions, it is possible to write a closed-form expression linking the short- and long-term interest rates and derive analytically the implications of the zero lower bound.

2.1 The Model for the Short-Term Rate

The model for the one-period nominal interest rate is based on Black (1995). Black restates Fisher's argument for non-negative nominal interest rates by interpreting currency and interest rates as options. Currency is an option in the sense that were the bond return negative, agents could hold currency instead.[1] In addition, Black proposes a characterization of the

[1]Keynes (1936, p. 202) contains a similar idea whereby the lower the nominal interest rate, the smaller the "earnings from illiquidity" that compensate the risk of "loss on capital account." Keynes maintains

short-term interest rate that distinguishes between the observed and the "shadow" nominal interest rates. The observed rate is a call option on the shadow rate, where the latter is what the interest rate would be in the absence of the currency option. The payoff of this option is

$$r_t = \begin{cases} r_t^*, & \text{if} \quad r_t^* > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

where r_t and r_t^* are the one-period observed and shadow nominal interest rates, respectively.[2]

Under (1), the nominal interest rate may also be interpreted as a limited-dependent variable censored at zero, with r_t^* as the associated latent variable. A model of this form was first studied in a regression context by Tobin (1958) and is a standard tool in cross-section econometrics. The use of limited dependent variable models in a time-series framework is usually more complicated because economic models frequently predict that an expectation is one of the explanatory variables. Limited-dependent models with rational expectations have been employed by Shonkwiler and Maddala (1985) and Holt and Johnson (1989) to study the determination of commodity prices, by Baxter (1990) to study adjustable-peg exchange rate regimes, and by Pesaran and Samiei (1992, 1995) and Pesaran and Ruge-Murcia (1999) to study exchange rates subject to two-sided limits.

By the Pure Expectations Hypothesis (PEH) of the term structure of interest rates, the nominal return of a two-period zero-coupon bond must equal the average expected return of the sequence of two one-period bonds held over its lifetime

$$R_t = (1/2)(r_t + E(r_{t+1}|I_t)) + \theta_t, \tag{2}$$

where R_t is the return of the two-period bond, I_t is the non-decreasing set of information available to market participants at time t and it is assumed to include observations of the variables up to and including period t, $E(r_{t+1}|I_t)$ is the conditional expectation of the nominal return of the one-period bond acquired at time t+1, and θ_t is a serially uncorrelated stochastic term that includes a liquidity premium and has variance σ_θ^2.

In order to give empirical content to the theory, assume that r_t^* is generated by

$$r_t^* = \alpha + \psi(L)r_t + \boldsymbol{\beta}\mathbf{x}_t + \epsilon_t, \tag{3}$$

that this is "the chief obstacle to a fall in the rate of interest to a very low level." He does not specify whether this low level is zero, but indicates (p. 207) that once the interest rate has fallen to a "certain level", liquidity preference is essentially absolute and agents prefer "cash to holding a debt which yields so low a rate of interest." This is Keynes' liquidity trap where the monetary authority loses control over the interest rate and monetary policy becomes ineffectual.

[2] In the continuous-time literature in finance, Cox, Ingersoll, and Ross (1985) propose a model where the volatility of the short-term interest rate is proportional to the square root of its level. With additional parameter restrictions, this model can also insure the non-negativity of interest rates. Kariya and Kamizono (1997) examine the performance of this model using Japanese data and report very limited empirical success.

where α is a non-negative intercept, L is the lag operator, $\psi(L)$ stands for the polynomial $\sum_{j=1}^{p} \psi_j L^j$, $\boldsymbol{\beta}$ is a $1 \times m$ vector of parameters, $\mathbf{x}_t$ is an $m \times 1$ vector of explanatory variables, and ϵ_t is a disturbance term with mean zero and variance σ_ϵ^2, serially uncorrelated, and uncorrelated with θ_t. The unconditional variance of ϵ_t is constant, but its conditional variance could be time-varying. For example, the conditional variance of ϵ_t could be described by an ARCH-type model in order to capture the observed volatility changes in short-term interest rates. Wolman (1999) considers a deterministic version of the system (1)-(3) where r_t^* is the central bank's desired short-term nominal interest rate and results from a Taylor-type policy rule.

The explanatory variables in $\mathbf{x}_t$ could be generated by the stochastic process

$$\mathbf{x}_t = \mathbf{A}\mathbf{H}_{t-1} + \mathbf{w}_{1,t}, \tag{4}$$

where $\mathbf{A}$ is an $m \times b$ matrix of coefficients, $\mathbf{H}_t$ is a $b \times 1$ vector of predetermined variables possibly including past values of $\mathbf{x}_t$, and $\mathbf{w}_{1,t}$ is an $m \times 1$ vector of random disturbances assumed to be independently and identically distributed $(i.i.d.)\,(0, \boldsymbol{\Omega}^{1/2})$ and uncorrelated with θ_t and ϵ_t.

The construction of the conditional expectation $E(r_{t+1}|I_t)$ is not trivial because the process of the short-term nominal interest rate is nonlinear. However, it is possible to find a closed-form expression for this expectation in the case where shocks are normally distributed. In addition, it is easiest to see the relation between the short- and long-term interest rates if one assumes that r_t^* depends only on the first lag of r_t

$$r_t^* = \alpha + \psi r_{t-1} + \epsilon_t, \tag{5}$$

where $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ and $\psi > 0$. The assumption $\psi > 0$ means that the nominal interest rate is positively autocorrelated, as in the data. The closed-form expression for $E(r_{t+1}|I_t)$ is given in the following proposition:

Proposition 1. *Assume that the short-term interest rate follows the limited-dependent process (1) where r_t^* is determined according to (5). Define the variable $c_{t+1} = -E(r_{t+1}^*|I_t)/\sigma_\epsilon = -(\alpha + \psi r_t)/\sigma_\epsilon$. Then, the conditional expectation of the short-term nominal interest rate at time $t+1$ constructed at time t is given by*

$$E(r_{t+1}|I_t) = (\alpha + \psi r_t)(1 - \Phi(c_{t+1})) + \sigma_\epsilon \phi(c_{t+1}), \tag{6}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the cumulative and density functions of a standard normal variable, respectively.

Proof. See Appendix A.

2.2 Implications

The conditional expectation (6) may be substituted in equation (2) to obtain

$$R_t = r_t/2 + ((\alpha + \psi r_t)(1 - \Phi(c_{t+1}) + \sigma_\epsilon \phi(c_{t+1}))/2 + \theta_t. \tag{7}$$

Equation (7) is used to derive the implications of the zero lower bound for the relation between the long- and the short-term interest rates. First, R_t is increasing and convex in r_t. This can be seen from the first- and second-order derivatives $\partial R_t/\partial r_t = 1/2 + (\psi/2)(1 - \Phi(c_{t+1})) > 0$ and $\partial^2 R_t/\partial r_t^2 = (\psi^2/2\sigma_\epsilon)\phi(c_{t+1}) > 0$. The nonlinearity arises from the effect of the zero lower bound on the current expectation about the future short-term interest rate. This effect is similar to the honeymoon effect in continuous- and discrete-time exchange-rate target-zone models (see Krugman, 1991, and Pesaran and Samiei, 1995, respectively). As the nominal interest rate rises above zero, c_{t+1} decreases monotonically, $\Phi(c_{t+1})$ and $\phi(c_{t+1})$ tend to zero, and the model approaches a linear forecasting model (see below).

Second, even if the current short-term interest rate is zero, the long-term rate is strictly positive. To see this, note that when $r_t = 0$, the future short-term rate can only take values equal to or larger than the current rate. Thus, the conditional expectation of r_{t+1} must be strictly positive. It follows that R_t must be strictly positive as well.

Third, given the current short-term interest rate, R_t is increasing and convex in the conditional standard deviation of r_t. As before, this can be seen from the derivatives $\partial R_t/\partial\sigma_\epsilon = \phi(c_{t+1})/2 > 0$ and $\partial^2 R_t/\partial\sigma_\epsilon^2 = c_{t+1}^2\phi(c_{t+1})/2\sigma_\epsilon > 0$. Under Black's interpretation of interest rates as options, this is precisely the result that option pricing theory would predict.

Finally, a decrease in the short-term rate produces a smaller response (in absolute value) in the long-term rate than an increase of the same magnitude. That is, $|R_t(r_t - \Delta r_t) - R_t(r_t)| < |R_t(r_t + \Delta r_t) - R_t(r_t)|$, where Δr_t is the change in the short-term interest rate. In order to verify this claim, use $\partial R_t/\partial r_t > 0$ and the definition of absolute value to write $R_t(r_t + \Delta r_t) - R_t(r_t) > -(R_t(r_t - \Delta r_t) - R_t(r_t))$. Rearranging this expression yields $(R_t(r_t + \Delta r_t) + R_t(r_t - \Delta r_t))/2 > R_t(r_t)$, that is satisfied because the function $R_t(r_t)$ is convex in its argument.

It is useful to compare the predictions of the nonlinear model with the ones of a benchmark linear forecasting model. The counterpart of the process (1)-(5) when one ignores the zero lower bound on interest rates is $r_t = \alpha + \psi r_{t-1} + \epsilon_t$. In this case, $E(r_{t+1}|I_t) = \alpha + \psi r_t$, and

$$R_t = r_t/2 + (\alpha + \psi r_t)/2 + \theta_t. \tag{8}$$

Note that R_t is linear in r_t. Because $\partial R_t/\partial r_t = (1+\psi)/2$ is constant and independent of the short-term rate, changes in r_t produce changes in R_t that are symmetric, proportional, and history-independent. Also, the long-term rate is independent of the conditional standard deviation of the short-term rate. Finally, since $(1+\psi)/2 > 1/2 + (\psi/2)(1-\Phi(c_{t+1}))$, where the left-hand-side (right-hand-side) is $\partial R_t/\partial r_t$ under the linear (nonlinear) model, the marginal effect on R_t of a change in r_t is smaller in the neighborhood of the zero lower bound.

Up to the extent that monetary policy affects long-term interest rates through the term structure, these results imply that its power is considerably reduced at low interest rates. First, when the short-term interest rate is low, the scope for further interest rate cuts is limited by the zero lower bound. Second, decreases in short-term interest rates have smaller effects than increases on long-term rates (in absolute value). Finally, adjustments to the short-term interest rate (whether decreases or increases) have a smaller effect on long-term rates. Section 4 uses stochastic simulation to show that these conclusions hold in the general case and are empirically relevant for the case of Japan.

3 The Multi-Period Bond

Consider now the more general case where the long-term bond has maturity of n periods. The PEH predicts that the return of the n-period, pure-discount bond must equal the average expected return of the sequence of n one-period bonds held over its lifetime. That is,

$$R_t = (1/n)(r_t + E(r_{t+1}|I_t) + \cdots + E(r_{t+n-1}|I_t)) + \theta_t, \tag{9}$$

where R_t is now the nominal return of the n-period bond. The remaining notation is as previously defined.

Proposition 2 derives the conditional expectations of the short-term interest rate when r_t is subject to the non-negativity constraint. This proposition generalizes Proposition 1 in that it makes no assumptions regarding the distribution of the shocks, allows a multivariate process for r_t^*, and considers horizons $s \geq 1$.

Proposition 2. *Assume that the short-term interest rate follows the limited-dependent process (1) where r_t^* is determined according to (3). Assume that the explanatory variables $\mathbf{x}_t$ follow the process (4). Define the composite error term*

$$u_{s,t+s} = \epsilon_{t+s} + \sum_{k=1}^{s-1} \psi_{s-k}\mu_{k,t+k} + \boldsymbol{\beta}\mathbf{w}_{s,t+s}, \tag{10}$$

where $\mu_{k,t+k} = r_{t+k} - E(r_{t+k}|I_t)$ and $\mathbf{w}_{s,t+s} = \mathbf{x}_{t+s} - E(\mathbf{x}_{t+s}|I_t)$, with cumulative distribution and density functions denoted by $F_s(\cdot)$ and $f_s(\cdot)$, respectively. Define the variable

$$c_{t+s} = -E(r_{t+s}^*|I_t) = -\left(\alpha + \sum_{k=1}^{\min\{p,s-1\}} \psi_k E(r_{t+s-k}|I_t) + \sum_{j=s}^{p} \psi_j L^j r_{t+s} + \boldsymbol{\beta} E(\mathbf{x}_{t+s}|I_t)\right). \tag{11}$$

Then, the conditional expectation of the short-term nominal interest rate at time $t+s$ constructed at time t is given by

$$E(r_{t+s}|I_t) = (E(r^*_{t+s}|I_t) + E(u_{s,t+s}|I_t, u_{s,t+s} > c_{t+s}))(1 - F_s(c_{t+s})). \tag{12}$$

Proof. See Appendix B.

Note that for horizons $s > 1$, the conditional expectations $E(r_{t+s}|I_t)$ cannot be expressed in closed-form. The reason is that $u_{s,t+s}$ includes interest-rate forecast errors that, due to the limited-dependent nature of r_t, do not follow a standard distribution. In particular, the density of the forecast errors is non-normal and asymmetric, even if shocks are normally distributed. The asymmetry is due to the fact that forecast errors are bounded below by $-E(r_{t+s}|I_t)$. This asymmetry disappears as the current short-term rate rises away from zero and the effect of the non-negativity constraint disappears. Since $E(r_{t+s}|I_t)$ cannot be computed analytically, this paper develops a Monte-Carlo procedure to compute this expectation numerically. The procedure extends the frequency simulators by Lerman and Manski (1981) and McFadden (1989) to dynamic nonlinear rational-expectations models. It constructs estimates of $F_s(c_{t+s})$ and $E(u_{s,t+s}|I_t, u_{s,t+s} > c_{t+s})$ by simulating paths of the short-term nominal interest rate subject to the non-negativity constraint and computing the relative frequency with which the zero lower bound is hit and the sample average of the realizations of the nominal rate above the bound. Appendix C contains a more detailed description of this procedure.

Due to the lack of closed-form solution to describe $E(r_{t+s}|I_t)$, it is not possible to derive general analytical results. However, it is clear that $E(r_{t+s}|I_t)$ is a nonlinear function of current and past realizations of the short-term interest rate. This follows from the observation that $E(u_{s,t+s}|I_t, u_{s,t+s} > c_{t+s})$ and $F_s(c_{t+s})$ are nonlinear functions of c_{t+s}, that, in turn, is a function of current and past realizations of r_t. Moreover, since the long-term interest rate depends on the average forecast of future short-term rates, R_t is also a nonlinear function of current and lagged short-term interest rates. The next section examines empirically the implications of this nonlinear relationship using Japanese data and shows that the results derived in Section 2.2 hold more generally.

4 Interest Rates in Japan

This section examines empirically the model of the term structure using Japanese data. Taking as given the estimated process of the short-term interest rate, this section derives the implied long-term interest rate under the PEH, computes the response of the long rate to a change in the short rate, and compares the forecasting power of the linear and nonlinear models.

The dataset consists of 304 weekly (Thursday) observations of the one-, three-, six-, and twelve-month nominal interest rates on zero-coupon Treasury Bills for Japan. The data source is *Datastream* and the sample period is 6 July 1995 to 26 April 2001. The sample starts in 6 July 1995 for two reasons. First, the sequential test by Bai and Perron (1998) identifies a structural break in the one-month nominal interest rate on 29 June 1995. Andrew's (1993) sup$F$ statistic is 18.01, which is well above the 5 per cent critical level of 8.58. The 95 per cent confidence interval for the break date is 22 June to 6 July 1995. Second, contemporaneous accounts indicate that a more expansionist monetary policy was undertaken by the Bank of Japan, starting in the third quarter of 1995 (see, for example, *The Economist*, 16 September 1995, p. 86). The sample ends on 26 April 2001 because this was the latest available observation when the data was collected.

During this period, Japanese interest rates were remarkably close to the zero lower bound. The data set features observations as low as 0.03 for the one-month interest rate on 20 May 1999 and 0.06 for the twelve-month interest rate on 29 October 1998. The medians for the one-, three-, six-, and twelve-month interest rates are only 0.38, 0.44, 0.44, and 0.47, respectively.

4.1 The Short-Term Rate

For the analysis that follows, the short-term rate is the one-month interest rate. The process of r_t^* is described in terms of past realizations of r_t with lag length determined using sequential Likelihood Ratio (LR) tests.[3] The conditional variance of ϵ_t is modeled as a GARCH(1,1). That is, $\epsilon_t = \sqrt{h_t} v_t$, where v_t is $i.i.d.N(0,1)$ and $h_t = \zeta + \delta\epsilon_{t-1}^2 + \varrho h_{t-1}$. The estimated process is

$$r_t = \begin{cases} r_t^*, & \text{if} \quad r_t^* > 0, \\ 0, & \text{otherwise,} \end{cases}$$

with

$$r_t^* = \underset{(0.0106)}{0.0185} + \underset{(0.079)}{0.598 r_{t-1}} + \underset{(0.077)}{0.127 r_{t-2}} + \underset{(0.055)}{0.214 r_{t-3}} + \epsilon_t,$$

and

$$h_t = \underset{(0.0008)}{0.004} + \underset{(0.111)}{0.539 \epsilon_{t-1}^2} + \underset{(0.079)}{0.223 h_{t-1}^2}.$$

[3] In principle, the model can accommodate a multivariate representation of r_t^*. However, many series, like money and output growth, are unavailable on a weekly basis. Using a univariate process for the short-term interest rate makes the model more parsimonious and allows one to concentrate on a well-defined statistical object, namely the bivariate process (r_t, R_t), for the purpose of econometric inference.

A standard misspecification test for ARCH-type models is the Ljung-Box Q-statistic applied to the squared standardized residuals. If the ARCH model is correctly specified, then the squared residuals corrected for heteroskedasticity should be serially uncorrelated. Under the null hypothesis of no autocorrelation, the Q-statistic is chi-square distributed with degrees of freedom equal to the number of autocorrelations. The Q-statistics for up to five autocorrelations are 0.004, 0.62, 0.63, 0.72, and 0.74, respectively. Since all statistics are below the 5 per cent critical value of their appropriate distribution, the null hypothesis cannot be rejected. Hence, it would appear that a GARCH(1,1) model adequately captures the conditional heteroskedasticity in the Japanese one-month interest rate.

4.2 Predicted Long-Term Rates

This section examines the conditional expectations of the short rate and the predicted long rates under the linear and nonlinear models. First, the top three panels in Figure 1 plot the conditional expectations of the future short-term rate three-, six-, and twelve-months ahead for different values of the current short rate. In all panels, the linear (nonlinear) model is represented by the dotted (continuous) line. The conditional expectations of r_t are computed using the frequency simulator in Appendix C and the conditional variance of ϵ_t is set to its unconditional mean.

For the nonlinear model that takes into account the effect of the zero lower bound, the conditional expectations "bend" upward at low nominal interest rates and, consequently, they are nonlinear and convex in the current short-term interest rate. This effect is similar to the honeymoon effect in exchange-rate target-zone models. At low nominal interest rates, the range of possible future realizations of r_t is larger above than below the current rate as a result of the non-negativity constraint. Hence, the conditional forecast of future short-term rates is above the current rate. This effect increases as r_t approaches the zero lower bound and it is more pronounced as the horizon increases. For the linear model, the conditional expectations are linear in the current short-term interest rate. For interest rates well above the non-negativity constraint, the forecasts from both models coincide, but for rates close to zero, the forecasts from the nonlinear model are higher than those from the linear model at all horizons.

Second, the bottom three panels of Figure 1 plot the three-, six-, and twelve-month interest rates predicted by both models. Since under the PEH long-term interest rates are an average of forecasted future short-term rates, the properties the long-term rates follow from those of the conditional expectations of the short-term rates. For the nonlinear model, long-term interest rates are nonlinear and convex in the current short-term rate. For the linear model, the long-term rates are linear functions of the current short-term rate. When the short-term rate is well above the non-negativity constraint, the long-term rates predicted by both models are identical, but when r_t is close to zero, the nonlinear model predicts substantially higher long-term interest rates than the linear model. This means that the nonlinear model predicts a steeper yield curve than the linear model at low short-term interest rates. We will see below that this prediction accords well with Japanese data and explains the superior forecasting performance of the nonlinear model.[4]

4.3 Impulse-Response Analysis

This section examines the response of the long-term interest rate associated with a change in the short-term interest rate. Figure 2 plots the responses of the three-, six-, and twelve-month interest rates to an increase and to a decrease of 25 basis points in the short-term rate under the linear (dotted line) and nonlinear (continuous line) models. In constructing these responses, the level and the conditional variance of the one-month interest rate are fixed to their unconditional means.

Note that under the linear forecasting model, the response of the long-term rate to an increase in the short-term rate is exactly the mirror image of the response to a decrease. In general, for linear models of the term structure, an innovation to the short-term rate yields movements in the long-term rate that are symmetric, proportional, and history-independent. That is, the impulse-response associated with a shock of size 1 (standard deviation) would be the mirror image of the response to a shock of size -1, one-half the response of shock size 2, and independent of the moment the shock is assumed to take place.

Under the nonlinear model, the response of the long-term rate to an innovation in the short-term rate is asymmetric. The change in the long-term rate when the one-month rate increases by 25 basis points is larger (in absolute value) than its change when the one-month rate decreases by 25 basis points. When the one-month rate increases (decreases) by 25 basis points, the three-, six-, and twelve-month rates increase (decrease) by 12, 9, and 5.6 (11, 7.8, and 4.7) basis points, respectively. Note that this asymmetric effect is proportionally larger for longer maturities. This result reflects the more general proposition that in nonlinear systems, impulse responses can vary with the size and sign of the shock and the initial conditions (see Koop, Pesaran, and Potter, 1996). This result is due to the fact that, when interest rates are close to zero, agents understand that future values of the short-term rate

[4] The working paper version of this article (Ruge-Murcia, 2002) examines the relation between the level of the long-term rate and the conditional standard deviation of the short-term rate. Results indicate that the relation is nonlinear and convex as predicted by the theory, but that the empirical magnitude of the volatility effects described in Black (1995) are negligible.

are limited below, but not above, the current rate. Thus, the effect of a decrease in the short-term rate on the long-term rate is smaller than that of an increase.

In both models, the effect of a change in the short-term rate on the long-term rate decreases with the horizon, but it is more severe for the nonlinear model. For example, an interest rate cut by 25 basis points in the one-month rate produces an immediate reduction in the three-, six, and twelve-month interest rates of 12.6, 10 and 6.9 points, respectively, under the linear model, but of 11, 7.8, and 4.7, respectively, under the nonlinear model. This means that changes in the short-term rate induce smaller changes in the long-term rate in the neighborhood of the zero lower bound. Consequently, the power of monetary policy to affect long-term rates through the term structure is smaller at low nominal interest rates.

A simple way to examine the model predictions is to perform an OLS regression of the change in the long-term interest rate on a constant, and positive and negative changes (separately) of the short-term interest rate. The coefficients of the explanatory variables are reported in Panel A of Table 1. These coefficients can be thought of as reduced-form estimates of $\partial R_t / \partial r_t$ under the PEH of the term structure. Notice that both increases and decreases of the short-term rate produce statistically significant changes in the three- and six-month interest rates. However, while increases of the short-term rate produce a statistically significant change in the twelve-month interest rate, decreases produce a quantitatively small and statistically insignificant response. This is in line with the prediction of the nonlinear model that the asymmetric response is more important for longer maturities.[5]

Note that the coefficients in Panel A measure the effect of a change by 1 basis point in the short-term rate on the long-term rate. Hence, in order to compare the magnitude of the responses predicted by OLS and by the impulse-response analysis, it is necessary to multiply the former by 25; these are the numbers reported in Panel B of Table 1. The initial effects predicted by the impulse-response analysis for both the linear and nonlinear models are included in the 95 per cent confidence interval of the effects predicted by OLS, but those of the nonlinear model are quantitatively closer.

The results from the impulse-response analysis imply that the real effects of easing monetary policy may be smaller when interest rates are close to zero. VAR evidence consistent with this prediction is reported by Iwata and Wu (2005) for the case of Japan. In particular, these authors find that the zero bound reduces the effect of monetary policy shocks on output and limits the central bank's ability to purse counter-cyclical policy.

[5]This prediction of the model is for a given short-term rate. However, when one includes the current short-term rate as a control in the regression, results are essentially the same as reported.

4.4 Comparing the Linear and Nonlinear Forecasting Models

One strategy to compare empirically the relative merits of the linear and nonlinear models is to compute forecast error statistics. To that effect, the in-sample and the one-step-ahead out-of-sample Root Mean Squared Error (RMSE) for both models are constructed. The out-of-sample measures are computed for the last 50 observations in the sample by recursively estimating the model and constructing the forecasts. The linear model constructs the long-term interest rate using linear forecasts of the one-month interest rate. The nonlinear models take into account the effect of the zero lower bound on expectations, but differ on their treatment of the conditional standard deviation of ϵ_t. The nonlinear model *I* computes the forecasts of the one-month interest rates using the GARCH(1,1) estimates of the conditional standard deviation of ϵ_t. The nonlinear model *II* fixes the conditional standard deviation of ϵ_t to its unconditional mean. All statistics are reported in Table 2.

In some cases, the linear model has smaller in-sample RMSE's than the nonlinear models for the three- and six-month rates. However, the nonlinear models have smaller in-sample RMSE's for the twelve-month rate, and superior out-of-sample performances for all maturities in all cases. The reason for this result is the following. The nonlinear model recognizes that at very low short-term interest rates, future short-term rates are more likely to increase than to decrease further. This is a direct consequence of the zero lower bound on interest rates. As seen in Figure 1, this effect is larger as the horizon increases. Hence, the nonlinear model predicts a steep yield curve. In contrast, the linear model underpredicts the long-term rate and implies a flatter yield curve than found in the data. Finally note that, in all cases, the nonlinear model *I* is inferior to model *II*. Hence, the nonlinear model which is the more parsimonious and which abstains from incorporating the volatility effect described by Black (1995) appears to better predict the Japanese long-term interest rates.

The gain of using the nonlinear models is quantitatively small but might be economically important when pricing bonds and other debt instruments. In order to examine whether it might be profitable to make trades on the basis of the nonlinear model, consider a fictional weekly auction of Japanese Treasury Bills. Sealed bids are made by agents for contracts that deliver 1 Yen at maturity. There are three risk-neutral agents in the auction. The first agent uses the PEH and a linear forecasting model to construct her bid. The second agent uses the PEH and a nonlinear forecasting model that takes into account the zero-lower bound. In order to make this an out-of-sample exercise, these two agents are assumed to only have access to past observations of the variables, with current variables being determined at the end of the auction. These agents use an AR(3) process with GARCH(1,1) disturbances to model the one-month interest rate and incorporate new observations to their data set as they become available. The third agent has perfect foresight and submits bids that correspond exactly to what the interest rate will be at the end of the auction. Contracts are allocated from the highest bid down, but only price bids higher than the one made by the fictional perfect-foresight agent are allowed. Otherwise, the implied interest rate would differ from the one observed in the data.

Table 3 reports the average annual return of portfolios composed entirely of bonds of one maturity. Notice that, in all cases, the nonlinear forecasting model delivers higher returns than the linear model. The reason is that the linear model underpredicts the nominal interest rate or, equivalently, overpredicts the contract price. The agent using the linear forecasting model bids a higher price, buys less contracts (for a given wealth), and obtains a lower return than the other two market participants. In order to put the difference between forecasting models in perspective, assume that the wealth invested by each agent is 10000 Yen. Then, Table 3 means that, on average, and at the maturity date, the agent using the nonlinear forecasting model has 0.06, 0.9, 3.9, and 12.4 more Yen in her pocket than the agent using the linear model, depending on whether she invests in one-, three-, six-, or twelve-month Treasury Bills. Since the nonlinear model dominates the linear one for all maturities, a portfolio that consists of nonzero holdings of more than one bond type would also have a higher return when constructed using a forecasting model that takes into account the zero lower bound.

5 Conclusions

The recent Japanese experience provides us with a natural experiment to study a number of important issues in macroeconomics. In particular, given concerns about deflation in a number of central banks, Japan allows us to study the implications of the zero lower bound for the term structure and monetary policy. To that end, this paper constructs a dynamic, limited-dependent variable model that imposes this non-negativity constraint. The limited-dependent variable model is a device that forces agents to explicitly incorporate the zero lower bound in their forecasts, even if all observations of the nominal interest rate to date are strictly positive. The term-structure model is general in that it nests linear models as the short-term interest rate rises above zero. However, in the neighborhood of the zero lower bound, the predictions of the nonlinear model differ sharply from those of a benchmark linear model. These predictions imply that the power of monetary policy to affect long-term interest rates through the term structure is reduced at low interest rates. The empirical analysis of recent Japanese data indicates that although the nonlinear and asymmetric effects predicted by the model are quantitatively small, the nonlinear model provides a better fit of the Japanese term structure than a linear model that ignores the effect of the zero lower bound on expectations.

A Proof of Proposition 1

Define the standardized normal variable $\xi_t = \epsilon_t/\sigma_\epsilon$ and use the definition of c_{t+1} to write the process of the short-term interest rate at time $t+1$ as

$$r_{t+1} = \begin{cases} E(r^*_{t+1}|I_t) + \epsilon_{t+1}, & \text{if} \quad \xi_{t+1} > c_{t+1}, \\ 0, & \text{otherwise.} \end{cases}$$

Note that c_{t+1} is known at time t. The conditional expectation of r_{t+1} is the weighted average

$$\begin{aligned} E(r_{t+1}|I_t) = \; & E(r_{t+1}|I_t, \xi_{t+1} > c_{t+1}) \Pr(\xi_{t+1} > c_{t+1}) \\ & + E(r_{t+1}|I_t, \xi_{t+1} \leq c_{t+1}) \Pr(\xi_{t+1} \leq c_{t+1}). \end{aligned}$$

Since the forecast $E(r^*_{t+1}|I_t) = \alpha + \psi r_t$ is also known at time t,

$$E(r_{t+1}|I_t, \xi_{t+1} > c_{t+1}) = \alpha + \psi r_t + E(\epsilon_{t+1}|I_t, \xi_{t+1} > c_{t+1}).$$

Write $E(\epsilon_{t+1}|I_t, \xi_{t+1} > c_{t+1}) = \sigma_\epsilon \phi(c_{t+1})/(1-\Phi(c_{t+1}))$, where $1-\Phi(c_{t+1})$ stands for $\Pr(\xi_{t+1} > c_{t+1})$. Note that $E(r_{t+1}|I_t, \xi_{t+1} \leq c_{t+1}) = 0$. With these intermediate results

$$E(r_{t+1}|I_t) = (\alpha + \psi r_t)(1 - \Phi(c_{t+1})) + \sigma_\epsilon \phi(c_{t+1}),$$

as claimed.¶

B Proof of Proposition 2

Use the definitions of $u_{s,t+s}$ and c_{t+s} to write the process of the short-term interest rate at time $t+s$ as

$$r_{t+s} = \begin{cases} E(r^*_{t+s}|I_t) + u_{s,t+s}, & \text{if} \quad u_{s,t+s} > c_{t+s}, \\ 0, & \text{otherwise.} \end{cases}$$

The conditional expectation of r_{t+s} is the weighted average

$$\begin{aligned} E(r_{t+s}|I_t) = \; & E(r_{t+s}|I_t, u_{s,t+s} > c_{t+s}) \Pr(u_{s,t+s} > c_{t+s}), \\ & + E(r_{t+s}|I_t, u_{s,t+s} \leq c_{t+s}) \Pr(u_{s,t+s} \leq c_{t+s}). \end{aligned}$$

Since the forecast $E(r^*_{t+s}|I_t)$ is known at time t,

$$E(r_{t+s}|I_t, u_{s,t+s} > c_{t+s}) = E(r^*_{t+s}|I_t) + E(u_{s,t+s}|I_t, u_{s,t+s} > c_{t+s}).$$

Note that $E(r_{t+s}|I_t, u_{s,t+s} \leq c_{t+s}) = 0$. With these intermediate results and using $\Pr(u_{s,t+s} > c_{t+s}) = 1 - F_s(c_{t+s})$, the conditional expectation of the short-term interest rate at time $t+s$ is

$$E(r_{t+s}|I_t) = (E(r^*_{t+s}|I_t) + E(u_{s,t+s}|I_t, u_{s,t+s} > c_{t+s}))(1 - F_s(c_{t+s})),$$

as claimed.¶

C Computation of the Conditional Expectations

The procedure involves the following steps:
Step 1: Having found analytically or numerically, the one-step-ahead conditional expectation of the nominal interest rate, $E(r_{t+1}|I_t)$, use the definition of (11) for $s = 2$ to obtain c_{t+2}.
Step 2: Simulate M observations of the short-term interest rate at time $t+1$ using (1) and (3). The non-negativity constraint may be numerically enforced by substituting negative realization of r_{t+1} with zeroes. Compute the M realizations of the forecast error, $\mu_{1,t+1} = r_{t+1} - E(r_{t+1}|I_t)$.
Step 3: Draw M realizations of ϵ_{t+2} and $\mathbf{w}_{t+2}$, and combine them with the μ's according to (10) to obtain M realizations of $u_{2,t+2}$.
Step 4: Construct an estimate of $F_2(c_{t+2})$ as the proportion of observations of $u_{2,t+2}$ that are larger than c_{t+2} :

$$F_s(c_{t+s}) = (1/M) \sum_{j=1}^{M} \Im(u_{2,t+2} > c_{t+2}),$$

where $\Im(\cdot)$ is an indicator function which takes the value 1 when its argument is true and 0 otherwise. Construct an estimate of $E(u_{2,t+2}|I_t, u_{2,t+2} > c_{t+2})$ by taking the arithmetic average of observations of $u_{2,t+2}$ that fall above c_{t+2} :

$$E(u_{s,t+s}|I_t, u_{s,t+s} > c_{t+s}) = (1/M) \sum_{j=1}^{M} u_{s,t+s} \Im(u_{2,t+2} > c_{t+2}).$$

Step 5: Applying relation (12) for $s = 2$ delivers $E(r_{t+2}|I_t)$. Using $E(r_{t+2}|I_t)$, the procedure can then be repeated recursively for $s = 3, 4, \ldots, n-1$.

Notice that subsequent iterations make use of the conditional expectations computed previously. When computing $E(r_{t+3}|I_t)$, one requires $E(r_{t+1}|I_t)$ and $E(r_{t+2}|I_t)$ to construct c_{t+3}, and so on. The one-step-ahead forecast, $E(r_{t+1}|I_t)$, required to start the recursion, can be found analytically in the special case where ϵ_t is normally distributed, or numerically using the same procedure above. (In the latter case, the recursion would start with r_t, rather than with $E(r_{t+1}|I_t)$.) In the case where ϵ_t is conditionally heteroskedastic, the draw of future realizations of ϵ_t needs to take into account the fact that its conditional variance changes over time. See Baillie and Bollerslev (1992) for the construction of forecasts of the conditional variance of many common parametric models.

Table 1. Results of OLS Regression

Variable	Maturity		
	3 Months	6 Months	12 Months
	A. OLS Coefficients		
Intercept	0.001	−0.005	−0.005
	(0.006)	(0.005)	(0.006)
Positive	0.342*	0.338*	0.201*
	(0.099)	(0.152)	(0.092)
Negative	−0.420*	−0.258*	−0.141
	(0.056)	(0.078)	(0.092)
	B. Predicted Response to a Change of 25 Basis Points in the Short-Term Rate		
Positive	8.55	8.45	5.03
	(2.48)	(3.80)	(2.30)
Negative	−10.5	−6.45	−3.53
	(1.40)	(1.95)	(2.30)

Notes: This table reports the results (in basis points) of the projection of ΔR_t on a constant, $\Im(\Delta r_t > 0)\Delta r_t$, and $\Im(\Delta r_t < 0)\Delta r_t$, where $\Im(\cdot)$ is an indicator function which takes the value 1 when its argument is true and 0 otherwise. Standard errors are robust to serial correlation and heteroskedasticity.

Table 2. Root Mean Squared Error of Linear and Nonlinear Forecasting Models

Maturity	Model		
	Linear	Nonlinear *I*	Nonlinear *II*
		A. In-Sample	
3 Months	12.848	12.948	12.729
6 Months	15.267	15.307	15.111
12 Months	26.066	24.947	24.857
		B. Out-of-Sample	
3 Months	13.289	13.208	12.986
6 Months	14.234	13.604	13.172
12 Months	15.949	13.853	13.457

Notes: This table reports the RMSE in basis points. The linear model forecasts the one-month interest rate linearly. The nonlinear model *I* takes into account the effect of the zero lower bound and the conditional standard deviation of ϵ_t. The nonlinear model *II* takes into account the effect of the zero lower bound but fixes the conditional standard deviation of ϵ_t to its unconditional mean.

Table 3. Average Annualized Return on Bond Portfolio from 18 May 2000 to 26 April 2001

Composition of Bond Portfolio	Model		
	Perfect Foresight	Linear	Nonlinear *II*
1-Month Bonds Only	0.881	0.683	0.690
3-Month Bonds Only	0.887	0.644	0.679
6-Month Bonds Only	0.888	0.596	0.674
12-Month Bonds Only	0.984	0.590	0.714

Note: This table reports returns in percentage points.

References

[1] Andrews, D., 1993, Tests for Parameter Instability and Structural Change with Unknown Change Point, Econometrica 61, 821-856.

[2] Bai, J. and P. Perron, 1998, Estimating and Testing Linear Models with Multiple Structural Changes, Econometrica 66, 47-78.

[3] Baillie, R. T. and T. Bollerslev, 1992, Prediction in Dynamic Models with Time-Dependent Conditional Variances, Journal of Econometrics 52, 91-113.

[4] Black, F., 1995, Interest Rates as Options, Journal of Finance 50, 1371-1376.

[5] Baxter, M., 1990, Estimating Rational Expectations Models with Censored Variables: Mexico's Adjustable Peg Regime of 1973-1982, University of Rochester, Mimeo.

[6] Cox, J. C., J. E. Ingersoll and S. A. Ross, 1985, A Theory of the Term Structure of Interest Rates, Econometrica 53, 385-407.

[7] Fisher, I., 1896, Appreciation and Interest (August M. Kelley Bookseller, New York).

[8] Holt, M. T. and S. R. Johnson, 1989, Bounded Price Variation and Rational Expectations in an Endogenous Switching Model of the US Corn Market, Review of Economics and Statistics 71, 605-613.

[9] Iwata, S. and S. Wu, 2005, Estimating Monetary Policy Effects When Interest Rates are Bounded at Zero, University of Kansas, Mimeo.

[10] Kariya, T. and K. Kamizono, 1997, Testing the Validity of the Cox-Ingersoll-Ross Type Model for Japanese Interest Rates, Economic Review 48, 195-206.

[11] Keynes, J. M., 1936, The General Theory of Employment, Interest and Money (Macmillan Press, London).

[12] Koop, G., M. H. Pesaran and S. M. Potter, 1996, Impulse Response Analysis in Nonlinear Multivariate Models, Journal of Econometrics 74, 119-147.

[13] Krugman, P., 1991, Target Zones and Exchange Rate Dynamics, Quarterly Journal of Economics 106, 669-682.

[14] Lerman, S. and C. Manski, 1981, On the Use of Simulated Frequencies to Approximate Choice Probabilities, in: C. Manski and D. McFadden, eds., Structural Analysis of Discrete Data with Econometric Applications (MIT Press, Cambridge).

[15] McFadden, D., 1989, A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration, Econometrica 57, 995-1026.

[16] Pesaran, M. H. and H. Samiei, 1992, Estimating Limited-Dependent Rational Expectations Models: With an Application to Exchange Rate Determination in a Target Zone, Journal of Econometrics 53, 141-163.

[17] Pesaran, M. H. and H. Samiei, 1995, Limited-Dependent Rational Expectations Models with Future Expectations, Journal of Economic Dynamics and Control 19, 1325-1353.

[18] Pesaran, M. H. and F. J. Ruge-Murcia, 1999, Analysis of Exchange Rate Target Zones Using a Limited-Dependent Rational Expectations Model with Jumps, Journal of Business and Economic Statistics 17, 50-66.

[19] Ruge-Murcia, F. J., 2002, Some Implications of the Zero Lower Bound on Interest Rates for the Term Structure and Monetary Policy, University of Montreal, CRDE Working Paper 06-2002.

[20] Shonkwiler, J. and G. Maddala, 1985, Modeling Expectations of Bounded Prices: An Application to the Market for Corn, Review of Economics and Statistics 38, 634-641.

[21] Tobin, J., 1958, Estimation of Relationships for Limited Dependent Variables, Econometrica 26, 24-36.

[22] Wolman, A. L., 1999, Staggered Price Setting and Zero Bound on Nominal Interest Rates, Federal Reserve Bank of Richmond Quarterly Review 84, 1-24.

Fig. 1: Predicted Long Rates

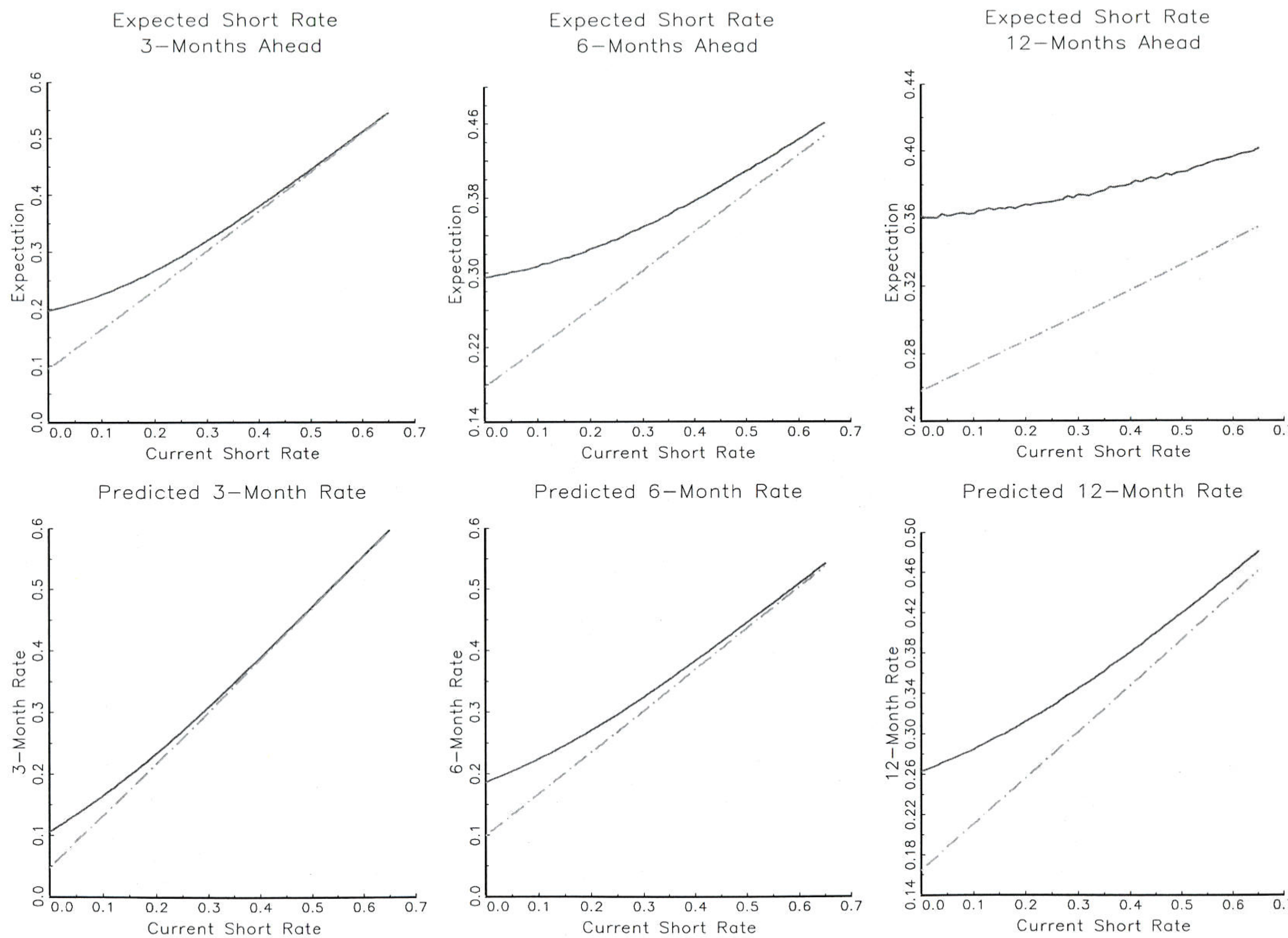

Fig. 2: Response of the Long Rate to a Change of 25 Basis Points in the Short Rate