Estimating Diffusion Models of Interest Rates at the Zero Lower Bound: From the Great Depression to the Great Recession and Beyond

Lealand Morin[*]
Department of Economics, University of Central Florida

October 13, 2022

Abstract

The time series of the federal funds rate has recently been extended back to 1928, now including several episodes during which interest rates remained near the lower bound of zero. I analyze this series, using the method of indirect inference, by applying recent research on bounded time series to estimate a set of bounded parametric diffusion models. This combination uncouples the specification of the bounds from the law of motion. Although Louis Bachelier was the first to use arithmetic Brownian motion to model financial time series, he has often been criticized for this proposal, since the process can take on negative values. Most researchers favor processes such as geometric Brownian motion, which remains positive. Under my framework, Bachelier's proposal remains valid when specified with bounds and is shown to compare favorably when modeling the federal funds rate.

Keywords: federal funds target rate, interest rate, zero lower bound, diffusion processes, regulated Brownian motion, bounded processes.

1 Introduction

Recent research has uncovered the historical series of the federal funds rate, back to 1928—the earliest date that this series was recorded in newspapers. The extension

[*]I thank Harry J. Paarsch for helpful comments on the manuscript. I am also grateful for helpful feedback on an earlier version of this manuscript from Morten Nielsen, James MacKinnon, Allan Gregory, and Prosper Dovonan, as well as from seminar participants at Queen's University and annual meetings of the Canadian Economics Association. I also gratefully acknowledge SSHRC for their generous support of this research during my doctoral studies at Queen's University. All errors are my own.

of this data series affords the opportunity to study interest rates during another chapter of history in which the economy dipped into a severe recession. Both the Great Depression and, more recently, the Great Recession feature extended periods of time during which the federal funds rate remained near the zero lower bound. The current state of the economy during the coronavirus pandemic has again brought interest rates to the lower bound. This suggests that, although these events are rare, the benchmark interest rate can remain near zero for extended periods of time. Properly accounting for these episodes is especially important when discounting cash flows over a period of decades, for example, to evaluate securities or long-lived investment projects.

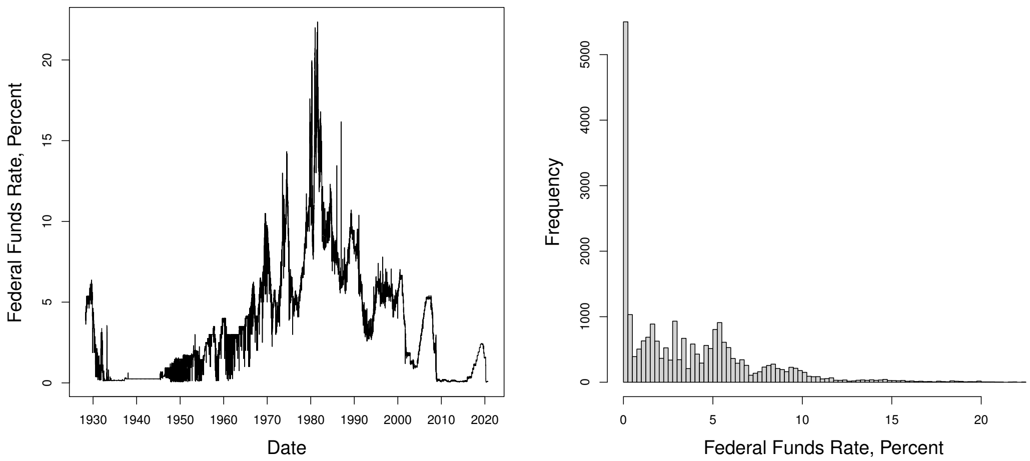
This research follows a long history of models for interest rates using a variety of diffusion models. The tendency of these series to remain near the zero lower bound suggests that a model should incorporate this characteristic. When used for this purpose, diffusion models typically impose a specification of variance that vanishes as the series approaches zero—to enforce the boundary—such as with geometric Brownian motion. In contrast, the historical record shows non-zero variance that characterizes the changes that, by constraint, are skewed upward, when the interest rate is near zero. In this paper, I combine research on time series in the presence of bounds with estimation methods for diffusion models for the federal funds rate. This combination produces a model that remains above the bound of zero, without placing constraints directly on the specification of the drift and diffusion terms, in a way that that adequately characterizes the movement in the federal funds rate.

The remainder of the paper proceeds as follows. In the next section, I describe the historical series of the federal funds rate. Then, I outline a set of diffusion models appropriate for the features of the series. I also augment this set of models with the explicit specification of the zero lower bound, without imposing constraints on the parameters in the law of motion of the series. In the next section, I describe the empirical methodology of indirect inference, a method of simulated minimum distance, applied to this estimation problem. I present the empirical results in the following section and then draw conclusions.

2 Federal Reserve Data

The sample of 24,121 daily observations of the federal funds rate spans the period 4 April 1928 to 15 September 2020. The data were obtained from two sources. For the period beginning 1 July 1954, the sample was drawn from a single series in the FRED database at the Federal Reserve Bank of St. Louis. Over this period, the series was collected by the Federal Reserve Bank of New York. Until recently, the earlier values of this series were not listed in any digital form. Through the extensive data collection efforts of Anbil, Carlson, Hanes, and Wheelock [2020], the series has been extended back to 1928 by transcribing the rates published in printed copies of the *New York Herald Times* and the *Wall Street Journal*. From these reports, four series are available from the FRED database, specifically, the high and low reported federal funds rates from each of the two newspapers. These series were combined to extend the history of the federal funds rate to cover the nearly century-long period 1928 to 2020. The process of combining and cleaning the data is described in Appendix A and is further described in an appendix to Anbil et al. [2020], which is the documentation for the data source.

The entire series of the Federal Reserve rate is depicted in panel (a) of Figure 1. During the Great Depression and also the Great Recession, the prolonged sojourns at the lower bound of zero are clearly visible. In the histogram shown in panel (b), it is clear that the rate has been near the zero lower bound for several thousand days throughout the sample, which represents a considerable fraction of the observations.



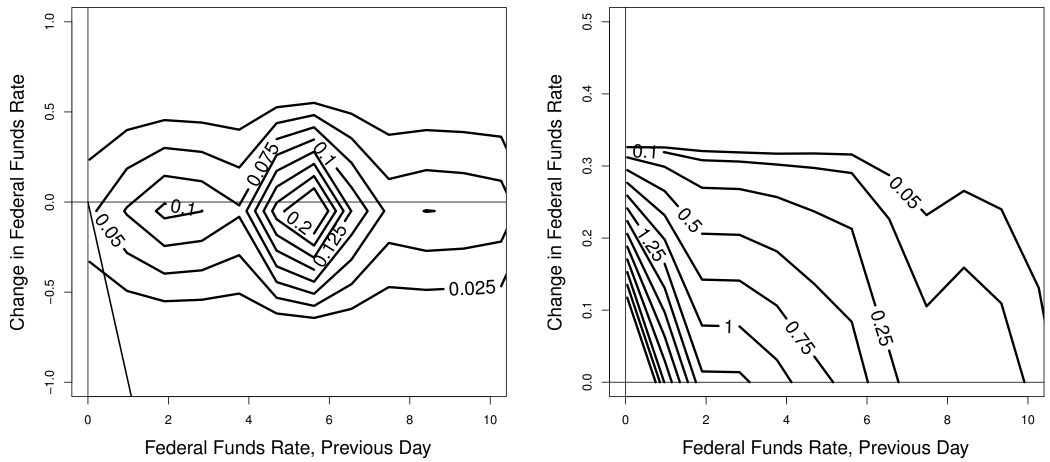
(a) Time Series Plot (b) Histogram

Fig. 1: Daily Series of the Federal Funds Rate, 1928–2020

Periods in which the federal funds rate was near zero are clearly visible in panel (a) after the Great Depression and the Great Recession and, more recently, during the coronavirus pandemic. These near-zero observations represent the mode of the distribution in the histogram in panel (b).

The distribution is skewed to the right, with a short visit to double-digit interest rates during the period of high inflation in the late 1970s and the early 1980s.

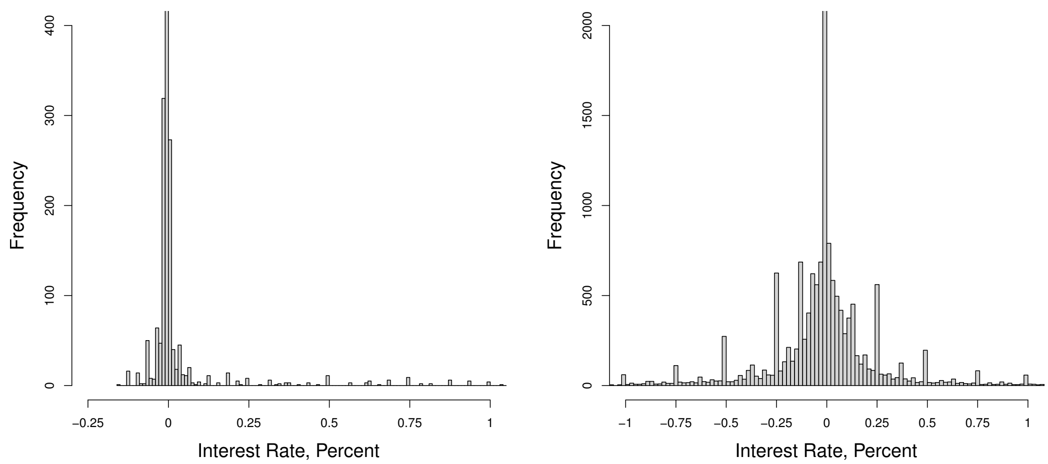
Figure 2 depicts level curves of kernel-smoothed density plots of the daily changes in the federal funds rate against the levels of the federal funds rate. The mode of the joint distribution appears at an interest rate just above five percent, with daily changes lower than ten basis points, centered just below zero. To illustrate the variability of changes in the series, panel (b) of Figure 2 depicts the kernel-smoothed density of absolute changes in the federal funds rate. When the federal funds rate is lower, the variability tends to be concentrated around smaller changes. The changes in the series are smaller near the zero lower bound, with the mode of the joint distribution near the origin. Nevertheless, the series exhibits daily variation of up to ten basis points when the rate is near zero, which is a large percentage change of the interest rate.



(a) Changes in Federal Funds Rate (b) Absolute Changes in Federal Funds Rate

Fig. 2: Density Plots of Daily Changes in the Federal Funds Rate, 1928–2020
Panel (a) depicts the level curves of the kernel-smoothed joint density of the daily changes in the federal funds rate against the level of the federal funds rate the previous day. Panel (b) depicts the level curves of the joint density of the absolute value of the daily changes against the level of the federal funds rate the previous day.

Figure 3 depicts two histograms of the changes in the federal funds rate divided into two subsamples. Panel (a) of Figure 3 depicts the histogram from the sample that includes only the days in which the federal funds rate was 25 basis points or less the previous day. The histogram in panel (b) covers the remainder of the sample. An important feature of the series, with regards to the specification of an econometric model, is that the series is highly skewed to the right when the series is near zero. With the constraint of the zero lower bound there is little else that can happen. In contrast, in the histogram in panel (b), the distribution appears symmetric. The distribution also appears leptokurtic, with long tails and a sharp peak at zero. This is partly an artifact of the large rate changes that occurred when interest rates were very high, during the 1970s and 1980s. These findings suggest that the variability of changes in the series should be modeled as an increasing function of the current rate. In addition, the econometric model should be specified with a non-negligible degree of variability at low interest rates as well.



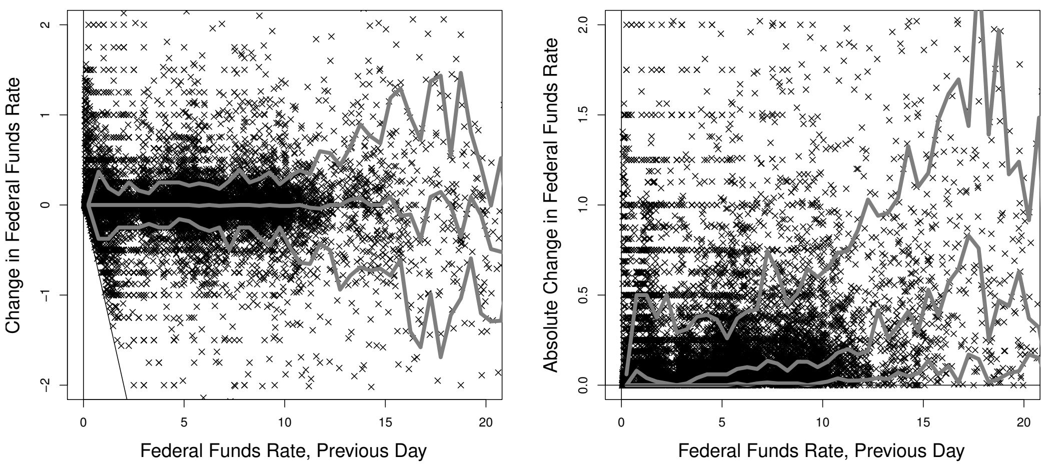
(a) Near the Zero Lower Bound

(b) Away from the Zero Lower Bound

Fig. 3: Histograms of Daily Changes in the Federal Funds Rate, 1928–2020
The sample was divided into two parts based on whether the federal funds rate was above or below 25 basis points the day before each observation. On the days after the federal funds rate was greater than 25 basis points, in panel (b), the distribution appears symmetric, with some point masses on changes that are multiples of 25 basis points. In the rest of the sample, in panel (a), when the series was within 25 basis points of the zero lower bound, the series is skewed to the right but displays a similar degree of variability.

Figure 4 presents a rudimentary attempt at modeling the changes in the interest rate as a function of the level. The daily changes in the rate are shown in a pair of scatter graphs, with the changes in the interest rate in panel (a) and the absolute changes in panel (b), both of which are plotted against the level of the federal funds rate the previous day. The dashed lines represent the ten-, fifty- and ninety-percent quantiles of the changes in the interest rate. In both panels, the dashed lines of quantiles were drawn with observations grouped into intervals fifty basis points wide.

Much of the variability remains within 25 basis points for interest rates as high as ten percent, with the variability gradually increasing in this range. The variability appears much higher in the sparsely-populated region in which the federal funds rate is above ten percent. For interest rates arbitrarily close to the lower bound of zero, the variability appears to vanish. The variance increases sharply, however, as soon as the federal funds rate is above the zero lower bound.



(a) Changes in Federal Funds Rate (b) Absolute Changes in Federal Funds Rate

Fig. 4: Quantiles of Daily Changes in the Federal Funds Rate, 1928–2020
Panel (a) depicts a scatter graph of the daily changes in the federal funds rate plotted against the federal funds rate on the previous day. Panel (b) depicts a plot of the absolute value of the daily changes against the federal funds rate on the previous day. The ten-, fifty-, and ninety-percent quantiles are superimposed over these plots, indicated by the grey lines. These figures indicate a steep increase in variability of the series for small values of the federal funds rate, followed by a curve that remains fairly flat up to a federal funds rate of seven percent. A gradual rise in variability appears as the federal funds rate rises above ten percent.

In panel (a), the restriction to the zero lower bound is represented by a diagonal line with a slope of negative one. The restriction to the zero lower bound is clearly visible, with all observed changes restricted above the line where the maximum change equals the negative of the current interest rate. In an empirical specification for this series, the variability of the process should increase sharply for low values of the interest rate. Above an interest rate of two percent, the rate of variability should increase gradually. These characteristics will guide the specification of the set of diffusion models described in the next section.

3 Diffusion Models of the Interest Rate

The history of diffusion models extends back as far as the history of the Federal Reserve. Louis Bachelier [1900] was the first to apply the stochastic process today referred to as Brownian motion—using the process as a model of prices in financial markets. Soon after, Albert Einstein [1905] made a pioneering effort to characterize Brownian motion formally—using the botanist Robert Brown's observation of the motion of particles suspended in a liquid as a metaphor, and expanded on the topic in Einstein [1926]. Norbert Wiener [1923] built a rigorous theoretical foundation, including a statement of the conditions for the existence of Brownian motion, which was named a Wiener process in his honour in some parts of the literature.

A few decades later, Kiyoshi Itô [1944] produced what would later be referred to as the Itô Formula, which greatly expanded the set of processes that could be applied to phenomena measured in continuous time. This modeling tool eventually led to an explosion of continuous-time models in finance, especially after Black and Scholes [1973] assumed geometric Brownian motion (GBM) for the law of motion of securities and portfolios that were valued with their celebrated Black–Scholes option pricing formula. This body of theoretical work was consolidated into many references that are well-known for applications in economics and finance—including Duffie [2001], Cochrane [2005], and Shreve [2004]. More generally, Karlin and Taylor [1981] as well as Karatzas and Shreve [1991] all presented modern introductions to diffusion models.

Although the models have become commonly used to evaluate securities, it is an altogether different matter to estimate the parameters of these models. The technique proposed by Äit Sahalia [2002] is a popular method for an approximate maximum-likelihood estimator (MLE). Phillips and Yu [2009] have also provided a survey of various approaches to likelihood-based estimation of diffusion models. Contributing to this literature, Äit Sahalia and Park [2012] have developed specification tests for diffusion models. More recently, Choi et al. [2014] have developed a framework for model selection among different models of diffusion. Kim and Park [2017] also considered the special case of recurrent diffusions, with an application to high-frequency regression. Using a nonparametric approach, Äit Sahalia and Park [2016] considered an application to nonstationary continuous-time models.

A similar alternative to the estimation of diffusion models involves using a time-series model with time-varying coefficients, such as in Park and Hahn [1999] as well as Cai et al. [2009]. Ever since Engle [1982], much attention has been paid to models of time-varying and nonlinear heteroskedasticity. Notable examples include Han and Park [2008] as well as Chung and Park [2007], and Park [2002].

Models with time-varying coefficients have a long history in the macroeconomic literature, especially in term-structure models of the interest rate. This line of research dates back to Cox et al. [1981], and the expectations hypothesis of the term structure of interest rates. This was followed by Campbell [1986] as well as Campbell and Shiller [1991], who provided an early prescription such that default-risk-free zero-coupon bonds could be valued based on expected returns with term premiums that are constant through time. Within this framework, affine term-structure models have been the focus of attention because of their analytical tractability. Duffie [2001] has presented an authoritative treatment of asset pricing and term-structure modeling.

Early examples of interest rate models included the foundational single-factor models pioneered by Vasicek [1977] and Cox et al. [1985]. Vasicek [1977] proposed an Ornstein–Uhlenbeck process, also called a mean-reverting process, which is a model having the following form:

$$dr_t = \kappa(\rho - r_t)dt + \sigma dW_t, \tag{1}$$

where ρ is the unconditional mean and κ is known as the speed of mean reversion. Cox et al. [1985] proposed a modification to this model, in the form of a state-dependent rate of diffusion

$$dr_t = \kappa(\rho - r_t)dt + \sigma\sqrt{r_t}dW_t. \tag{2}$$

This formulation is sometimes known as a square-root process. The mean-reverting characteristic of these models make for an interesting comparison, when analyzed in combination with a series constrained by bounds. These models fit within a broader set of models estimated in this paper.

More generally, a diffusion process X_t is a continuous-time process defined for each $t \in \mathbb{R}_+$. The instantaneous change in X_t satisfies the stochastic differential equation

$$dX_t = \mu(X_t; \boldsymbol{\alpha})dt + \sigma(X_t; \boldsymbol{\beta})dW_t, \tag{3}$$

where W_t is a standard Brownian motion, $\mu(X_t; \boldsymbol{\alpha})$ is the drift function with parameter $\boldsymbol{\alpha}$, and $\sigma(X_t; \boldsymbol{\beta})$ is the diffusion function with parameter $\boldsymbol{\beta}$. Collect the parameters into one vector $\boldsymbol{\theta} = [\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top]^\top \in \boldsymbol{\Theta}$ and define $\mathcal{D}$ as the domain of the diffusion process. This definition includes the following examples, each of which has parameter values that lie within a convex parameter space Θ.

Example 1: Some examples of diffusion processes include

- Arithmetic Brownian Motion (ABM):

  $\mu(X_t;\boldsymbol{\alpha}) = \alpha_I$ and $\sigma(X_t;\boldsymbol{\beta}) = \beta_I$ with $\beta_I > 0$ and $\mathcal{D} = \mathbb{R}$.

- Geometric Brownian Motion (GBM):

  $\mu(X_t;\boldsymbol{\alpha}) = \alpha_X X_t$ and $\sigma(X_t;\boldsymbol{\beta}) = \beta_X X_t$ with $\beta_X > 0$ and $\mathcal{D} = \mathbb{R}_+$.

- Ornstein–Uhlenbeck Process (OU):

  $\mu(X_t;\boldsymbol{\alpha}) = \alpha_I + \alpha_X X_t$ and $\sigma(X_t;\boldsymbol{\beta}) = \beta_I$ with $\alpha_X < 0, \beta_I > 0$ and $\mathcal{D} = \mathbb{R}$.

- Square-Root Process (SQR):

  $\mu(X_t;\boldsymbol{\alpha}) = \alpha_I + \alpha_X X_t$ and $\sigma(X_t;\boldsymbol{\beta}) = \beta_X\sqrt{X_t}$ with $\alpha_X < 0, \beta_X > 0$ and $\mathcal{D} = \mathbb{R}_+$.

In the above notation, parameters are assigned subscripts according to the term multiplied by the parameter: α_I is an intercept, β_X is a slope coefficient and β_E appears in the exponent of X_t.

The set of models presented above includes the models proposed by Vasicek [1977] and Cox et al. [1985]. In particular, the mean-reverting drift function $\kappa(\rho - r_t)$ is equivalent to $\mu(X_t;\boldsymbol{\alpha}) = \alpha_I + \alpha_X X_t$, in the OU and SQR models shown above. The parameter α_I is the negative of the speed of mean reversion κ. The unconditional mean ρ is captured by the ratio $\alpha_I/\kappa = -\alpha_I/\alpha_X$.

The observed interest rate, denoted by r_t in the literature, takes the place of X_t in the notation above and is the variable of interest in this paper. In the usual framework, the process X_t is free to move throughout $\mathcal{D}$, the domain of the diffusion process. In this sense, the process is already bounded by its domain. In this paper, however, the focus is on the interest rate r_t, which is bounded from below by zero. Several models listed above already impose this constraint: with GBM and the SQR model, the specification of the diffusion function dictates that the rate of diffusion vanishes as the process approaches zero, keeping the process above the lower bound of zero. The OU process and ABM do not restrict the process to positive values, which is a common criticism of Bachelier [1900] for modeling stock prices. I augment the menu of stochastic processes in a way that decouples the specification of the process from the restriction to the region above the lower bound of zero, leaving open the possibility of exploring the full set of processes to model interest rates.

Researchers have developed a budding literature concerned with statistical methodologies for estimating bounded processes, with Cavaliere [2005] as well as Cavaliere and Xu [2014] having highlighted key examples in econometric theory. This literature built on the foundational treatment of Harrison [1985], setting the stage for the theoretical analysis of stochastic processes within bounds, with the primary example referred to as regulated Brownian motion. Specifically, a realization of regulated Brownian motion X_t is constrained to remain within the interval $[\underline{b}, \overline{b}]$ at all times t. In the case of interest rates studied in this paper, only the lower bound applies with $\underline{b} = 0$ and the upper bound is not binding. The theory presented in Morin [2017] built on this body of knowledge by merging this line of research with the maximum-likelihood methodology developed in Jeong [2008] as well as Jeong and Park [2013] to estimate parametric diffusion models. In this paper, I take a different approach—implementing a variant of indirect inference, following Gouriéroux et al. [1993], except with bounds imposed.

This combination opens up the set of models available to the applied researcher studying variables that lie within bounds. In the absence of an explicit specification of the bounds, modelers are often constrained to choose a model specification that enforces the bound. As mentioned above, the models GBM and SQR satisfy this criteria, since the rate of diffusion vanishes as the process nears zero. This specification might be in conflict with the reality of a non-zero rate of diffusion near the bound, as was documented above in the analysis of the federal funds rate. Using the flexibility offered by the explicit modeling of the bounds, however, there is no need to be constrained to such a specification. In particular, it is possible to include a positive intercept in the diffusion function, specifying a positive rate of diffusion near the lower bound of zero, matching the conditions observed in the historical series of interest rates. Thus, the processes listed above are augmented in this paper to allow for a nonzero constant at the lower bound of zero by appending the following set of models.

Example 2: The augmented versions of selected diffusion processes are as follows:

- Augmented Geometric Brownian Motion (GBM-I):
  $\mu(X_t; \boldsymbol{\alpha}) = \alpha_X X_t$ and $\sigma(X_t; \boldsymbol{\beta}) = \beta_X X_t + \beta_I$ with $\beta_I > 0, \beta_X > 0$ and $\mathcal{D} = \mathbb{R}$.

- Augmented Ornstein–Uhlenbeck Process (OU-X):
  $\mu(X_t; \boldsymbol{\alpha}) = \alpha_I + \alpha_X X_t$ and $\sigma(X_t; \boldsymbol{\beta}) = \beta_I + \beta_X X_t$ with $\alpha_2 < 0, \beta_I > 0, \beta_X > 0$ and $\mathcal{D} = \mathbb{R}$.

- Augmented Square-Root Process (SQR-I):
  $\mu(X_t; \boldsymbol{\alpha}) = \alpha_I + \alpha_X X_t$ and $\sigma(X_t; \boldsymbol{\beta}) = \beta_X \sqrt{X_t} + \beta_I$ with $\alpha_X < 0, \beta_I > 0, \beta_X > 0$ and $\mathcal{D} = \mathbb{R}$.

In the naming convention for the augmented models, the hyphenated name has a suffix that indicates the subscript of the parameter appended to the model. Notice that the domain of these processes are no longer restricted to the positive real line, since the rate of diffusion no longer vanishes toward zero. These processes only remain within bounds when regulated to remain above zero, so these processes are now candidates for modeling the interest rate in the bounded specification presented in this paper.

This list of models represents a first step toward a model of the term structure that incorporates the constraints imposed by the zero lower bound. This approach does sacrifice accuracy by using a single-factor model, since it is well known that up to three factors will improve prediction, as shown by Littermann and Scheinkman [1991]. These factors correspond to the level, the slope, and the curvature of the yield curve. In this paper, I restrict attention to the level, which is of first-order importance, especially considering the proximity to the zero lower bound. The hypothesis tested here is that a one-factor model that accounts for the zero lower bound will be better-suited than a one-factor model without, for modeling interest rates after the Great Recession and the Great Recession. A question left for further research regards the comparison of performance between a multifactor model with the zero lower bound, versus one that ignores the bounds. The results presented below support this approach as a promising avenue of research. This comparison is made possible using the econometric methodology presented next.

4 Indirect Inference

Although the exact likelihood functions are known for the unbounded ABM, GBM, and OU processes, those for the others are not. As a substitute, the likelihood function of the corresponding discrete-time process is used to approximate the likelihood function of each continuous-time model. Under such an approach, however, the MLE has an additional source of bias, often referred to as the convexity effect, which obtains because under the approximate approach the drift and the diffusion functions are assumed constant over the intervals of time between observations. Because the exact discretizations of the processes with bounds are unknown, for all stochastic processes that I considered (except for the ABM process), an alternative estimation strategy was required.

Gouriéroux et al. [1993] estimated the GBM, OU, and, by extension, the ABM processes, as applications of indirect inference. Indirect inference was first introduced by Smith [1990], and in Smith [1993], before it was generalized by Gouriéroux et al. [1993], and later refined by Gallant and Tauchen [2006] as well as Gouriéroux et al. [2010] and, more recently, by Bruins et al. [2018]. Hall and Rust [2021] employed the method in a time-series model and introduced the term simulated minimum distance (SMD) estimation to describe this procedure. As Guvenen and Smith [2014] have noted, indirect inference is very useful when estimating models for which the likelihood function or the criterion function to be optimized is analytically intractable, or too computationally burdensome to evaluate. This method is admissible here because all the processes that I seek to estimate can be simulated—both with and without the zero lower bound imposed.

The essence of indirect inference is to use an auxiliary model, which is easy to compute, to capture features of the data and then to form a criterion function on which to base the estimation. The mapping from the structural parameters of interest to the auxiliary parameters that optimize this auxiliary model define a *binding function* through which one can indirectly draw inference on the structural parameters in the model of interest. In this case, following the approach of Gouriéroux et al. [1993], I used the logarithm of the likelihood function for the corresponding discrete-time process as the objective function for the auxiliary model. In the notation from above, this objective function is

$$\begin{aligned} Q_T(\boldsymbol{\alpha},\boldsymbol{\beta};\{X_t\}_{t=1}^T) =& \frac{1}{T}\sum_{t=2}^{T} -\log[\sigma(\boldsymbol{\beta};X_{t-1})] - \\ & \frac{1}{2T}\sum_{t=2}^{T}\frac{[X_t - X_{t-1} - \mu(\boldsymbol{\alpha};X_{t-1})]^2}{\sigma(\boldsymbol{\beta};X_{t-1})^2}. \end{aligned} \tag{4}$$

This objective function is a suitable choice for an auxiliary model because there exists a clear correspondence between the structural parameters $\boldsymbol{\theta}^\top = [\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top]$ that generate the simulated series $\{X_t^*\}_{t=1}^T$ and the auxiliary parameters $\tilde{\boldsymbol{\theta}}^{*\top} = [\tilde{\boldsymbol{\alpha}}^{*\top}, \tilde{\boldsymbol{\beta}}^{*\top}]$ that optimize the objective function $Q_T(\boldsymbol{\theta}, \{X_t^*\}_{t=1}^T)$. This optimization defines a binding function $\boldsymbol{b} : \boldsymbol{\Theta} \to \boldsymbol{\Theta}$ defined as $\boldsymbol{b}(\boldsymbol{\theta}_0) = (\tilde{\boldsymbol{\theta}}_0)$, where $\tilde{\boldsymbol{\theta}}_0$ is the vector of the auxiliary parameters that optimizes the following limiting objective function:

$$Q_\infty(\boldsymbol{\theta}_0; \{X_t\}_{t=1}^T) = \lim_{T\to\infty} Q_T(\boldsymbol{\theta}_0; \{X_t\}_{t=1}^T).$$

A necessary condition for estimation is that $\boldsymbol{b}(\cdot)$ is one-to-one, so the matrix of partial derivatives $\partial \boldsymbol{b}(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$, evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, must have full column rank; see Gouriéroux et al. [1993].

The estimated structural parameters in $\boldsymbol{\theta}$ is the vector $\hat{\boldsymbol{\theta}}$ that generates the series $\{X_t^*\}_{t=1}^T$ with auxiliary parameter estimates $\tilde{\boldsymbol{\theta}}^*$ that closely match the auxiliary parameters $\tilde{\boldsymbol{\theta}}$, which are estimated from the observed interest-rate series. The formulation in this paper is just identified because a one-to-one mapping exists between the structural parameters and the auxiliary parameters. Here, the estimated structural parameters $\hat{\boldsymbol{\theta}}$ produce a series $\{X_t^*\}_{t=1}^T$ with estimates of auxiliary parameters $\tilde{\boldsymbol{\theta}}^*$ that exactly match the auxiliary parameters $\tilde{\boldsymbol{\theta}}$ obtained from the observed interest-rate series. The estimation procedure is summarized in the following steps:

1. Calculate the auxiliary parameters from the data to obtain $\tilde{\boldsymbol{\theta}}$, which is a vector of the parameter estimates from optimizing the objective function for the auxiliary model equation (4) evaluated using the observed interest-rate series $\{X_t\}_{t=1}^T$.

2. Conduct a numerical optimization to search over candidate values of $\hat{\boldsymbol{\theta}}$ by iterating over the following steps:

   (a) draw a realization of the series $\{X_t^*\}_{t=1}^T$, given the candidate parameter $\hat{\boldsymbol{\theta}}$, following the law of motion specified by the diffusion model, and conforming to the zero lower bound, if imposed;

(b) calculate the auxiliary parameters $\tilde{\boldsymbol{\theta}}^*$ from the simulated series $\{X_t^*\}_{t=1}^T$, following the same procedure as in step 1 to optimize the objective function for the auxiliary model equation (4);

(c) calculate the weighting matrix $\boldsymbol{\Omega}$, described below, to normalize the auxiliary parameters;

(d) calculate the weighted distance between the auxiliary parameters from the observed interest-rate series and the simulated data

$$(\tilde{\boldsymbol{\theta}}^* - \tilde{\boldsymbol{\theta}})^\top \boldsymbol{\Omega}^{-1} (\tilde{\boldsymbol{\theta}}^* - \tilde{\boldsymbol{\theta}}). \tag{5}$$

3. Minimize the weighted distance by iterating over steps 2 (a)–2 (d), to find the parameter $\hat{\boldsymbol{\theta}}$ that minimizes the distance criterion in equation (5).

Given a discrete series $\{X_t\}_{t=1}^T$, the objective function of equation (4) in steps 1 and 2 (b) can be optimized using standard numerical methods, such as quasi-Newton methods. To ease the computational burden, I concentrate out the parameter $\boldsymbol{\alpha}$ in $\mu(\boldsymbol{\alpha}; X_t)$ with a weighted least-squares approach, with weights $\sqrt{\sigma(\boldsymbol{\beta}; X_t)}$ to obtain the optimal value of the auxiliary parameter $\tilde{\boldsymbol{\alpha}}$ for a candidate value of $\tilde{\boldsymbol{\beta}}$. Then I conduct the outer numerical optimization over the space of the auxiliary parameter $\boldsymbol{\beta}$ to obtain the estimate $\tilde{\boldsymbol{\beta}}$.

To generate the simulated series $\{X_t^*\}_{t=1}^T$ in step 2 (a), I generated a realization of the discretized process with a small time step: the complete series has two million observations and I dropped 99 observations between each consecutive pair of the remaining twenty thousand observations, roughly matching the 24,121 observations in the sample.

The models I estimated are pure time-series models; no exogenous variables were included. Gouriéroux et al. [1993] showed that the optimal weighting matrix $\mathbf{\Omega}^*$ has the form $\mathbf{J}_0\mathbf{V}_0^{-1}\mathbf{J}_0$. The matrix $\mathbf{J}_0$ is consistently estimated by

$$-\frac{\partial^2 Q_T}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}(\tilde{\boldsymbol{\theta}};\{X_t\}_{t=1}^T), \tag{6}$$

which is calculated by taking numerical derivatives of the objective function with the series $\{X_t\}_{t=1}^T$ generated at the parameter estimates $\hat{\boldsymbol{\theta}}$. Since the objective function Q_T can be written as a sum of contributions to the objective function $\sum_{t=2}^T q_t(\boldsymbol{\theta}; X_t, X_{t-1})$,

$$\mathbf{V}_0 = \lim_{T\to\infty} \mathbb{V}\left[\frac{1}{\sqrt{T}}\sum_{t=2}^T \frac{\partial q_t}{\partial\tilde{\boldsymbol{\theta}}}(\boldsymbol{\theta}; X_t, X_{t-1})\right], \tag{7}$$

which I estimated using the approach of Newey and West [1987], as outlined by Gouriéroux et al. [1993]. The estimates and standard errors obtained from this procedure are documented in the next section.

5 Empirical Results

For the first step in the estimation procedure, I estimated the auxiliary parameters from the observed series of the federal funds rate. These estimates are shown in Table 1.[1] For all models, the drift functions are negatively sloped, with a daily reduction on the order of half of a basis point for each percent interest rate. In terms of the diffusion rate, the estimates range between a constant rate of 35 basis points per day, for the OU and ABM models, to a constant slope of 64 basis points for each percentage point of interest, for GBM. For the models with non-zero intercepts,

[1] One would normally expect that standard errors be shown with such estimates, however, the estimates from the auxiliary model are known to be biased, which is the primary motivation for using indirect inference, in this case, and for the applications presented in Gouriéroux et al. [1993].

the intercept ranges from thirteen to twenty basis points, with a slope coefficient of eleven down to three basis points per day times an exponent of the interest rate.

	α_I	α_X	β_I	β_X	β_E
ABM	–0.000159	0.0	0.3500	0.0	0.0
GBM	0.0	–0.000043	0.0	0.6493	1.0
OU	0.017460	–0.004725	0.3495	0.0	0.0
SQR	0.021664	–0.005451	0.0	0.2748	0.5
GBM-C	0.0	–0.001737	0.1998	0.0334	1.0
OU-X	0.016417	–0.004445	0.1990	0.0336	1.0
SQR-C	0.018110	–0.004781	0.1284	0.1174	0.5
FULL	0.018039	–0.004769	0.1331	0.1116	0.5192

Tab. 1: Estimated Auxiliary Parameters

The auxiliary parameters were estimated by optimizing the auxiliary model using the entire series of the federal funds rate from 1928 to 2000. In the models in which some parameter values are implied by the specification of the model, those values are shown in the table alongide the estimates from the other models, with minimal significant digits.

Table 2 collect the estimates from the models with series that are regulated to remain above zero. Under this approach, I used these series to calculate auxiliary parameters and solved for the values of the structural parameters so that the auxiliary parameters from the simulated series matched those in Table 1.

The ABM process is the simplest model, which was estimated to have a constant downward drift of 6 basis points per day, however, this number is not statistically significant, since the objective function is very flat in this region, with many drift values mapped to similar auxiliary parameters. Aside from the SQR and GBM processes, the remaining processes also have negative drift functions over the observed range of the interest rate, however, nearly all of those coefficients in those drift functions are statistically insignificant. The SQR process differs in that the drift has the mean-reverting characteristic, with a negative drift above around 3.2 percent. For the GBM, a positive drift was predicted.

The most important differences were found in the specification of the diffusion

Model	α_I	α_X	β_I	β_X	β_E	$\Pr(\chi^2_{5-k} > d)$
ABM	–0.0611	0.0	0.3801	0.0	0.0	0.9348
	(0.5843)		(0.0030)			
GBM	0.0	0.1853	0.0	0.5328	1.0	0.0000
		(0.5150)		(0.0905)		
OU	–0.0215	0.0005	0.3593	0.0	0.0	0.9373
	(0.0145)	(0.0011)	(0.0027)			
SQR	0.0225	–0.0070	0.0	0.5050	0.5	0.1009
	(0.0008)	(0.2774)		(0.0044)		
GBM-C	0.0	–0.0021	0.2249	0.0315	1.0	0.9849
		(0.0012)	(0.0034)	(0.0009)		
OU-X	–0.0006	–0.0020	0.2113	0.0309	1.0	0.8845
	(0.0131)	(0.0012)	(0.0033)	(0.0009)		
SQR-C	–0.0109	–0.0007	0.2094	0.0863	0.5	0.8748
	(0.0844)	(0.0035)	(0.0153)	(0.0064)		

Tab. 2: Estimated Structural Parameters (Zero Lower Bound Imposed)
The estimates of the structural parameters solve for equality between the auxiliary parameters from the observed federal funds rate and those estimates from the series generated from each model, with the realization regulated to remain above zero. The column labeled $\Pr(\chi^2_{5-k} > d)$, lists the p-value for a test of the restrictions to the model in each row, each of which is a special case of the full model. The degrees of freedom for this test is $5 - k$, where k is the number of parameters in each model, compared to five parameters in the full model.

function. The ABM and OU processes, each with a constant drift function, have a value near the estimated auxiliary parameter values in the neighborhood of 35 basis points per day. The GBM-C and OU-X models, each of which have a linear diffusion function, have an estimated intercept near twenty basis points per day, and a slope coefficient that increases three basis points per day for each percentage point of the federal funds rate. The GBM and SQR models from Table 3, both of which are missing the intercept in the diffusion function, have much higher values of the slope coefficient β_X of fifty basis points per day, for each percentage point of the interest rate. These steeper slopes are required to fit the noticeable variation in the interest rate when it is close to zero, however, the slope overestimates the variation in the rate of diffusion when the interest rate is higher. The SQR-C model accommodates both the non-zero intercept and the declining rate of increase in the diffusion rate for higher values of the interest rate. In the SQR-C model, the intercept of the diffusion rate is similar to that of the other models augmented with an intercept, with a value of twenty basis points per day at an interest rate of zero. As the interest rate rises, the rate of diffusion rises at a rate of eight basis points times the square root of the federal funds rate, measured in percentage points.

To test for differences between the goodness of fit of the models, I calculated the weighted distance of the auxiliary parameters from those from the full model. All the models are nested in the full model that has all five of the parameters, with drift function $\mu(\alpha; X_t) = \alpha_I + \alpha_X X_t$ and diffusion rate $\sigma(\beta; X_t) = \beta_I + \beta_X X_t^{\beta_E}$, in which the exponent can be estimated. To facilitate a comparison between the models, I calculate the minimized distance between the auxiliary parameters for each of the models and those from the full model. The results appear in the column labeled $\Pr(\chi^2_{5-k} > d)$, which is the p-value of a test of the restrictions to the model in each row, each of which is a special case of the full model. The degrees of freedom for this test is $5 - k$, where k is the number of parameters in each model, compared to five parameters in the full model.

The results are split into a several-way tie. With the exception of the GBM model, the p-values are greater than ten percent, indicating that the models have similar predictions. By definition, the p-values are decreasing in the distance from the auxiliary parameters from the full model. The two models with the farthest distance from the full model—GBM and SQR—are the two models with a vanishing rate of diffusion at the zero lower bound. The remaining models have non-zero intercepts in the diffusion function and the performance is similar whether or not the diffusion rate is increasing in the level of the interest rate. With non-zero rates of diffusion at zero, all of these other models have the potential to cross below the zero lower bound. This suggests that the most important model specification decision is to choose a regulated process with a strictly positive rate of diffusion. This matches the findings in Figure 3 and Figure 4, in which the federal funds rate was shown to exhibit a substantial degree of variation near the zero lower bound, as well as relatively flat quantiles of absolute changes in the range from zero to seven percent. Furthermore, the simplest model, the ABM, which is a regulated Brownian motion when combined with bounds, appears to fit the data well enough without the added complexity of the other alternatives.

Model	α_I	α_X	β_I	β_X	β_E	$\Pr(\chi^2_{5-k} > d)$
ABM	–0.0002	0.0	0.3492	0.0	0.0	0.7721
	(0.0034)		(0.0024)			
OU	0.0209	–0.0057	0.3498	0.0	0.0	0.8223
	(0.0052)	(0.0009)	(0.0024)			
GBM-C	0.0	–0.0022	0.2146	0.0338	1.0	0.9185
		(0.0020)	(0.0030)	(0.0008)		
OU-X	0.0176	–0.0050	0.2112	0.0319	1.0	0.7665
	(0.0079)	(0.0017)	(0.0034)	(0.0009)		
SQR-C	0.0208	–0.0058	0.1859	0.0993	0.5	0.8620
	(0.0401)	(0.0116)	(0.0118)	(0.0059)		

Tab. 3: Estimated Structural Parameters (Zero Lower Bound Ignored)
The estimates of the structural parameters solve for equality between the auxiliary parameters from the observed federal funds rate and those estimates from the series generated from each model, with no restriction of the series. The column labeled $\Pr(\chi^2_{5-k} > d)$, lists the p-value for a test of the restrictions to the model in each row, each of which is a special case of the full model. The degrees of freedom for this test is $5 - k$, where k is the number of parameters in each model, compared to five parameters in the full model.

With the exception of GBM and SQR, the remaining models will produce series that can move below the zero lower bound, if not regulated to remain above the bound. Table 3 collects the estimates ignoring the fact that the interest rate series must remain above zero. Under this approach, the series $\{X_t^*\}_{t=1}^T$ is calculated by strictly following the law of motion with no adjustments. Strictly speaking, all of these models are misspecified, since they can produce series with negative values. I still investigated these models, however, to compare the coefficients with those from the corresponding bounded processes.

The diffusion functions all exhibit a similar change with this change in specification. The slope coefficients β_X are higher for all three models when bounds are ignored. Further, the intercept terms β_I are lower for all models. This suggests a lower rate of diffusion for lower interest rates, which appears to be a symptom of the censoring from the zero lower bound.

The drift functions changed predictably as well. The drift slope coefficients α_X moved further into negative territory. The intercepts α_I switched to positive values for all but the ABM. The constant drift of the ABM process changed from a large negative value to a negative value near zero. Overall, the pattern suggests a switch between zero-drift or mean-reverting processes, without bounds, and processes with negative drift pushing toward a lower bound.

In terms of the apparent accuracy of the estimates, the standard errors are smaller, for all but α_X, when the bounds are ignored. Most parameters show a small change in the magnitude of the standard errors, however, the change is particularly large for the intercept of the drift function. The presence of the zero lower bound, when it is ignored in the estimation, gives the illusion that the process is better behaved, with less variability in the estimates of the coefficients in the models.

6 Conclusion

In this paper, I have followed a long history of applications modeling interest rates with diffusion models. This history also includes prolonged periods in which the interest rate remains near the zero lower bound. Thanks to an exhaustive data-collection effort by Anbil et al. [2020], it is now possible to combine these to produce a model that matches this characteristic of the interest rate over long periods of time. I applied recently-developed econometric techniques for bounded processes to estimate simple models of the federal funds rate in the presence of the zero lower bound. I estimated a set of diffusion models and demonstrated the difference in interpretation between the cases with and without the bounds explicitly incorporated into the estimation technique.

When the bounds are taken into account, the estimated drift has a greater negative slope and is higher near the zero lower bound, and the estimated rate of diffusion is also higher near the bound. When ignored, the zero lower bound created the false impression of a mean-reverting process, compared to a process with negative drift moving toward the lower bound, when the bound was taken into account. Ignoring the bound also created the illusion that the coefficients were more precisely estimated. The series appeared more variable when the lower bound was acknowledged, which avoided mistaking a process stuck near bounds for one that is less variable. Although the differences in the estimated drift functions were not statistically significant, the differences in the diffusion functions were more important. The imposition of the lower bound allows greater flexibility for model specification, particularly in achieving a non-zero diffusion rate at the zero lower bound, which matches the behavior observed in the historical series of the federal funds rate.

In this paper, I have addressed a single issue in a way that has not appeared in the literature. Other specifications exist as well: considering them would be fruitful avenues for future research. In particular, one could extend this analysis to include discontinuities in the series, since changes of this type tend to occur after FOMC meetings. The fit of the model could be considerably improved by introducing a more detailed specification on those dates. A multivariate model that takes the information available at the time of the meeting is an option that also lies outside of the univariate framework presented here. Clearly, the members of the Federal Reserve Board respond to factors that measure the state of the economy, in addition to subjective conditions, such as investor sentiment. The analysis of these rate changes is certainly more complicated than a univariate model would permit. Still, the approach taken here factors in the zero lower bound, which is shown to have a notable effect on the estimation results. This work is one step toward a multivariate framework that can fruitfully answer such questions. Morin and Shang [2020] have taken another step in this direction.

In hindsight, the current economic crisis during the coronavirus pandemic is not an unprecedented event. The experience during the Great Recession has shown that the economy can return to normal after a visit to the zero lower bound. With the benefit of further hindsight, the experience during the Great Depression has shown that these episodes can be prolonged for many years and may be an expected part of the experience to be considered when evaluating investments with a lifespan of decades. The approach taken in this paper provides an empirical model of interest rates that is designed to endure for the decades to come.

Another problem that has endured in the literature—for more than a century—is that whenever the pioneering work of Bachelier [1900] is mentioned, the next statement is an indictment of the lack of suitability of arithmetic Brownian motion for modeling non-negative financial time series. Often, geometric Brownian motion is discussed next for its conformity to positive values. The trade-off, however, is that the GBM process specifies a vanishing rate of diffusion as the process nears zero, which was shown to fit the data poorly. In contrast, Bachelier's framework does specify a non-zero rate of diffusion near the bound, demonstrating performance comparable to the other models, and conforms to the characteristics of the historical series of interest rates. At long last, his pioneering model has been given a fair treatment and shown to be respectable with the proper guidance.

References

Yacine Äit Sahalia. Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Econometrica*, 70(1):223–262, January 2002.

Yacine Äit Sahalia and Joon Y. Park. Stationarity-based specification tests for diffusions when the process is nonstationary. *Journal of Econometrics*, 169:279–292, 2012.

Yacine Äit Sahalia and Joon Y. Park. Bandwidth selection and asymptotic properties of local nonparametric estimators in possibly nonstationary continuous-time models. *Journal of Econometrics*, 192:119–138, 2016.

Sriya Anbil, Mark A. Carlson, Christopher Hanes, and David C. Wheelock. A new daily federal funds rate series and history of the federal funds market, 1928–1954. Federal Reserve Bank of St. Louis Working Paper Series 2020-016B, Federal Reserve Bank of St. Louis, July 2020.

Louis Bachelier. Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, 3(17):21–86, 1900.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

Marianne Bruins, James A. Duffy, Michael P. Keane, and Anthony A. Smith, Jr. Generalized indirect inference for discrete choice models. *Journal of Econometrics*, 205(1):177–203, July 2018.

Zongwu Cai, Qi Li, and Joon Y. Park. Functional-coefficient models for nonstationary time series data. *Journal of Econometrics*, 148(2):101–113, February 2009.

John Y. Campbell. A defense of traditional hypotheses about the term structure of interest rates. *Journal of Finance*, 41(1):183–193, 1986.

John Y. Campbell and Robert J. Shiller. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies*, 58(3):495–514, 1991.

Giuseppe Cavaliere. Limited time series with a unit root. *Econometric Theory*, 21 (05):907–945, October 2005.

Giuseppe Cavaliere and Fang Xu. Testing for unit roots in bounded time series. *Journal of Econometrics*, 178(2):259–272, 2014.

Hwan-sik Choi, Minsoo Jeong, and Joon Y. Park. An asymptotic analysis of likelihood-based diffusion model selection using high frequency data. *Journal of Econometrics*, 178:539–557, 2014.

Heetaik Chung and Joon Y. Park. Nonstationary nonlinear heteroskedasticity in regression. *Journal of Econometrics*, 137(1):230–259, March 2007.

John H. Cochrane. *Asset Pricing*. Princeton University Press, Princeton, NJ, 2005.

John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross. A re-examination of traditional hypotheses about the term structure of interest rates. *Journal of Finance*, 36(4):769–799, 1981.

John C. Cox, Jonathan E. Ingersoll, and Stephen A Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

Darrell Duffie. *Dynamic Asset Pricing Theory*, volume 4. Princeton University Press, Princeton, third edition, 2001.

Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 17:549–560, 1905.

Albert Einstein. *Investigations on the Theory of Brownian Movement*. Dover, New York, NY, 1926.

Robert F. Engle, III. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

A. Ronald Gallant and George Tauchen. Which moment to match? *Econometric Theory*, 12(4):657–681, 2006.

Christian Gouriéroux, Alain Monfort, and Eric Renault. Indirect inference. *Journal of Applied Economics*, 8:S85–S118, 1993.

Christian Gouriéroux, Peter C. B. Phillips, and Jun Yu. Indirect inference for dynamic panel models. *Journal of Econometrics*, 157(1):68–77, 2010.

Fatih Guvenen and Anthony A. Smith, Jr. Inferring labor income risk and partial insurance from economic choices. *Econometrica*, 82(6):2085–129, 2014.

George Hall and John Rust. Estimation of endogenously sampled time series: The case of commodity price speculation in the steel market. *Journal of Econometrics*, 222(1, Part A):219–243, 2021.

Heejoon Han and Joon Y. Park. Time series properties of ARCH processes with persistent covariates. *Journal of Econometrics*, 146(2):275–292, October 2008.

J. Michael Harrison. *Brownian Motion and Stochastic Flow Systems*. John Wiley and Sons, Inc., New York, NY, 1985.

Kiyoshi Itô. Stochastic integral. *Proc. Imperial Acad. Tokyo*, 20:519–524, 1944.

Minsoo Jeong. *Asymptotics for the Maximum Likelihood Estimation of Diffusion Models*. PhD thesis, Texas A & M University, College Station, TX, 2008.

Minsoo Jeong and Joon Y. Park. Asymptotic theory of maximum likelihood estimation of diffusion models. Technical report, 2013.

Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, NY, 1991.

Samuel Karlin and Howard M. Taylor. *A Second Course in Stochastic Processes*. Academic Press, New York, NY, 1981.

Jihyun Kim and Joon Y. Park. Asymptotics for recurrent diffusions with application to high frequency regression. *Journal of Econometrics*, 196(1):37–54, January 2017.

Robert Littermann and José Scheinkman. Common factors affecting bond returns. *Journal of Fixed Income*, 1(1):54–61, 1991.

Lealand Morin. *Keeping Diffusion Processes within Bounds: Using Information between Observations*. PhD thesis, Queen's University, Kingston, ON, 2017.

Lealand Morin and Ying Shang. Federal Reserve policy after the zero lower bound: an indirect inference approach. *Empirical Economics*, 60:2105–2124, 2020.

Whitney K. Newey and Kenneth D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987.

Joon Y. Park. Nonstationary nonlinear heteroskedasticity. *Journal of Econometrics*, 110(2):383–415, October 2002.

Joon Y. Park and Sang B. Hahn. Cointegrating regressions with time varying coefficients. *Econometric Theory*, 15(5):664–703, October 1999.

Peter C. B. Phillips and J. Yu. Maximum likelihood and Gaussian estimation of continuous time models in finance. In T.G. Andersen, editor, *Handbook of Financial Time Series*, Mathematics and Statistics. Springer London, Limited, 2009. ISBN 9783540712978.

Steven E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance. Springer Science + Business Media, LLC, New York, NY, 2004.

Anthony A. Smith, Jr. *Three Essays on the Solution and Estimation of Dynamic Macroeconomic Models*. PhD thesis, Duke University, Durham, NC, 1990.

Anthony A. Smith, Jr. Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics*, 8:S63–S84, 1993.

Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188, 1977.

Norbert Wiener. Differential space. *Journal of Mathematical Physics*, 2:131–174, 1923.

Appendix A.1: Data Processing

The data were joined from two main sources, both of which are available from the FRED database at the Federal Reserve Bank of St. Louis. For the period from 1 July 1954 to 15 September 2020, the series comprises daily observations of a single series, the *Effective Federal Funds Rate*, which is labeled *DFF* in the FRED database. This series covers the period that begins when the Federal Reserve Bank of New York began to record the daily figure. The data before 1954 were collected by Anbil, Carlson, Hanesi, and Wheelock [2020], who transcribed four series from microfiche records of daily newspapers. The following list describes the series name in the FRED database, the series label, in upper case letters, along with the dates defining the period in which the series were observed. All series were downloaded on 3 July 2020.

- *High Value of the Federal Funds Rate for the Indicated Date Published in The New York Herald-Tribune (FFHTHIGH)* from 1928-04-04 to 1937-08-12.
- *High Value of the Federal Funds Rate for the Indicated Date Published in The Wall Street Journal (FFWSJHIGH)* from 1932-06-01 to 1954-06-28.
- *Low Value of the Federal Funds Rate for the Indicated Date Published in The New York Herald-Tribune (FFHTLOW)* from 1928-04-04 to 1938-03-01.
- *Low Value of the Federal Funds Rate for the Indicated Date Published in The Wall Street Journal (FFWSJLOW)* from 1932-06-01 to 1954-06-30.

These series were then aggregated into a single series using the following procedure, as outlined by Anbil et al. [2020]. Each day, the aggregation method depends on the availability of the series because the dates of publication do not completely overlap and all four series were not reported every day.

1. For days when both newspapers reported a high and a low rate, the daily value was calculated using observations from both newspapers, using the simple average of the midpoints of the high and low rates reported in the *Herald Tribune* and *Wall Street Journal*. The high and low were not necessarily bid and offered rates; on dates when the *Wall Street Journal* provided an offer range, the midpoint of the bid-offer range was used to calculate the average value between the two newspapers.

2. For days when one newspaper reported only a single rate, that rate was taken as that newspaper's rate for the market. The average was then calculated with this value combined with the midpoint of the high and low (or bid and ask) provided by the other newspaper.

3. If only one newspaper provided data, the daily value was recorded as the midpoint of rates or the single rate provided by that newspaper.

To extend this series to the current decade, the digital series of the *Effective Federal Funds Rate*, from its first date of availability in the FRED database, was appended to the historical series from 4 April 1928 to 30 June 1954. After constructing these series, Anbil et al. [2020] verified that connecting these series was reasonable, as the behavior near the end of the newspaper series was comparable to that of the beginning of the official series published by the Federal Reserve. For this paper, weekends and holidays were excluded from the series, to restrict analysis to business days and to avoid dates with partial coverage in the historical series from newspapers.