Estimation of Cox-Ingersoll-Ross (CIR) Parameters: OLS and Maximum Likelihood Estimation (MLE) Methods

Andrea Buffoli, Davide Rolfi

June 23, 2025

1 Introduction

The Cox Ingersoll Ross (CIR) model, introduced in 1985 by Cox, Ingersoll and Ross (1), marked a key development in the modeling of short-term interest rates. While not explicitly framed as a correction to the earlier Vasicek model (2) introduced in 1977, the CIR model addressed one of its most significant limitations, namely the possibility of negative interest rates. The Vasicek model, relying on a Gaussian Ornstein Uhlenbeck process, allows the short rate to become negative due to the normal distribution of its innovations, an outcome that was considered implausible at the time of its development.

The CIR model resolves this issue by incorporating a square root diffusion term, which ensures the short rate remains non-negative under specific conditions. This feature made the model particularly appealing in the academic and financial communities of the 1980s, when negative rates were not observed and were generally considered unrealistic.

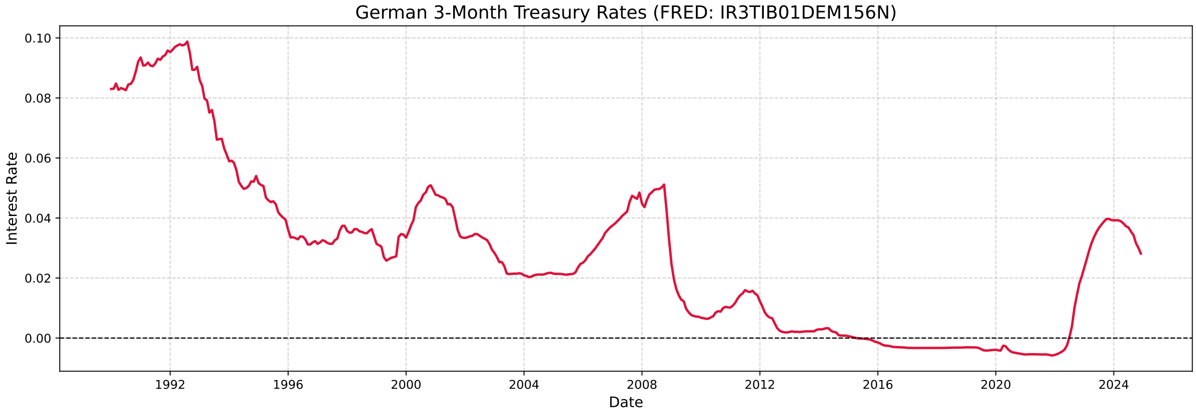

Figure 1: Historical 3-month German Treasury Bill rates (FRED series: IR3TIB01DEM156N) from 1990 to 2024. The plot illustrates prolonged periods of near-zero and negative short-term interest rates, particularly from 2015 to 2022. This empirical evidence challenges the applicability of models like CIR, which enforce non-negativity by construction, and highlights the practical relevance of alternative models that accommodate negative rates.

Although negative interest rates have since become an empirical reality in several markets, including Germany and Switzerland, Figure 1, the CIR model remains a foundational framework in the term structure literature. Its ability to capture mean reverting behavior while preserving non-negativity laid the groundwork for further theoretical developments and practical applications in fixed income modeling.

2 Mathematical Formulation of the CIR model

The CIR process is defined by the following stochastic differential equation:

$$dr(t) = a(b - r(t))\, dt + \sigma\sqrt{r(t)}\, dW(t) \tag{1}$$

where $r(t)$ denotes the short-term interest rate at time t, $a > 0$ is the speed of mean reversion, $b > 0$ is the long-term mean level, $\sigma > 0$ is the volatility parameter, and $W(t)$ is a standard Brownian motion. The CIR model differs from the Vasicek model in its diffusion term: the presence of $\sqrt{r(t)}$ ensures that the interest rate remains non-negative by making volatility state-dependent.

The Feller condition,

$$2ab \geq \sigma^2,$$

ensures that the origin is an inaccessible boundary for the CIR process. When this condition is satisfied, the process remains strictly positive almost surely. If the condition is violated, the process may reach zero with positive probability, although it stays non-negative due to the square-root diffusion term.

Given the interest rate dynamics and the fact that the CIR model belongs to the class of affine term structure models[1], the price of a zero-coupon bond at time t, maturing at time T, can be written as:

$$P(t,T) = A(\tau) \cdot e^{-B(\tau)r(t)}, \qquad \tau = T - t \tag{2}$$

where $A(\tau)$ and $B(\tau)$ are deterministic functions of the time to maturity. These functions can be derived by applying Itô's lemma to the expression for $P(t,T)$ and matching the resulting dynamics with the pricing equation under the risk-neutral measure.

These function are given by

$$B\left(\tau\right) = \frac{2\left(e^{\gamma\tau} - 1\right)}{\left(a+\gamma\right)\left(e^{\gamma\tau} - 1\right) + 2\gamma} \tag{3}$$

$$A\left(\tau\right) = \left(\frac{2\gamma e^{\frac{(a+\gamma)\tau}{2}}}{\left(a+\gamma\right)\left(e^{\gamma\tau} - 1\right) + 2\gamma}\right)^{\frac{2ab}{\sigma^2}} \tag{4}$$

where $\gamma = \sqrt{a^2 + 2\sigma^2}$. To give an idea of the role of this two functions, $B\left(\tau\right)$ accounts for the sensitivity of the bond to the current interest rate, while $A\left(\tau\right)$ accounts for the expected evolution of rates and risk-neutral discounting.

[1]For an overview of affine term structure models, see the following reference: Affine Term Structure Model on Wikipedia.

3 Parameter Estimation via Ordinary Least Squares (OLS)

One approach to estimate the parameters of the CIR model is to apply *Ordinary Least Squares (OLS)* to a transformed version of the original stochastic differential equation. Starting from the discretized form of the CIR process:

$$\Delta r(t) = a(b - r(t))\Delta t + \sigma\sqrt{r(t)}\Delta W(t),$$

we can divide both sides by $\sqrt{r(t)}$, under the condition that $r(t) \neq 0$, to obtain:

$$\frac{\Delta r(t)}{\sqrt{r(t)}} = ab\Delta t \cdot \frac{1}{\sqrt{r(t)}} - a\Delta t \cdot \sqrt{r(t)} + \sigma\sqrt{\Delta t} \cdot \varepsilon_t,$$

where $\varepsilon_t \sim \mathcal{N}(0, 1)$. This transformation linearizes the model in terms of the regressors $\frac{1}{\sqrt{r(t)}}$ and $\sqrt{r(t)}$, making it possible to estimate the parameters a, b, and σ using linear regression. Specifically, the regression takes the form:

$$y_t = \beta_1 \cdot x_{1,t} + \beta_2 \cdot x_{2,t} + \text{error},$$

where:

$$\begin{aligned} y_t &= \frac{r(t + \Delta t) - r(t)}{\sqrt{r(t)}}, \\ x_{1,t} &= \frac{1}{\sqrt{r(t)}}, \\ x_{2,t} &= \sqrt{r(t)}. \end{aligned}$$

The parameters are then recovered as follows:

$$\begin{aligned} \hat{a} &= -\frac{\beta_2}{\Delta t}, \\ \hat{b} &= \frac{\beta_1}{\beta_2}, \\ \hat{\sigma} &= \frac{\text{std(residuals)}}{\sqrt{\Delta t}}. \end{aligned}$$

The estimated parameters are presented in Table 1. The estimated speed of mean reversion, $a \approx 1.50$, suggests a relatively fast adjustment of the short-term interest rate toward its long-term mean. The long-term level $b \approx 1.35\%$ reflects the persistently low-rate environment observed during the sample period. The volatility parameter $\sigma \approx 0.23$ indicates moderate fluctuations around the mean, consistent with the model's square-root diffusion structure. Taken together, these estimates describe a process that tends to revert quickly toward a low equilibrium, with volatility increasing with the level of the interest rate.

Moreover, Figure 2 illustrates the comparison between the observed 3-month Treasury Bill rates and the simulated paths generated using the CIR model with parameters estimated via OLS. The median simulation trajectory captures the overall declining trend in the interest rates over the considered time period, while the 5th–95th percentile band highlights the degree of uncertainty implied by the model. Despite some discrepancies in periods of sharp rate movements, the CIR model provides a reasonably good fit to the historical data given its parsimony and mean-reverting structure.

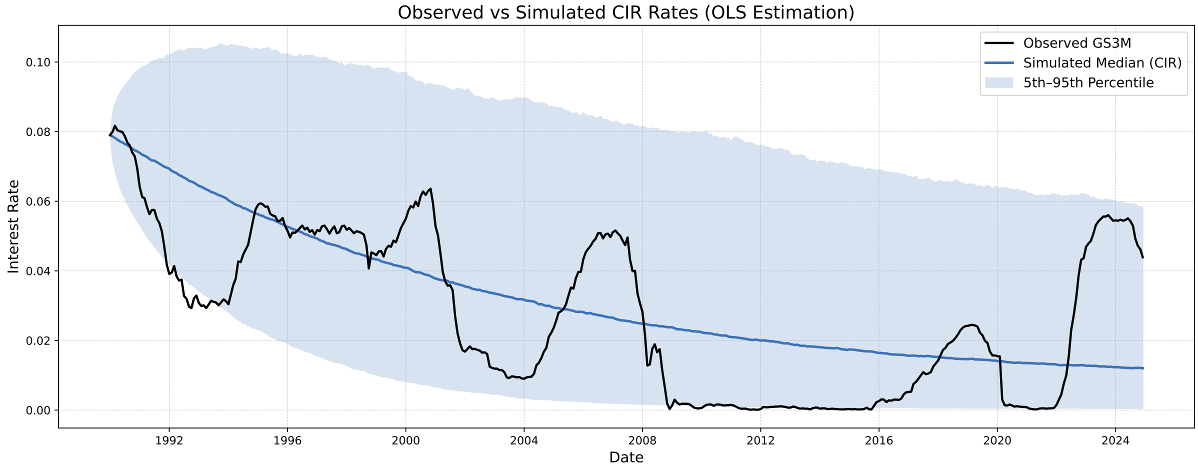

Figure 2: Comparison between observed 3-month Treasury Bill rates (GS3M) and the CIR model simulation based on OLS estimates. The solid black line represents the actual observed interest rates, while the blue line shows the median path of 10,000 CIR-simulated trajectories. The shaded area denotes the 5th to 95th percentile range, capturing the model's uncertainty. Parameters used in the simulation are obtained from the linearized OLS estimation of the CIR process.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import statsmodels.api as sm
from pandas_datareader import data as pdr
import datetime

# --- Load GS3M data from FRED
start = datetime.datetime(1990, 1, 1)
end = datetime.datetime(2024, 12, 31)
df = pdr.DataReader("GS3M", "fred", start, end).dropna()
r = df['GS3M'].values / 100  # Convert to decimal
dates = df.index
dt = 1 / 250  # Daily step

# --- OLS estimation (linearized CIR)
r_t = r[:-1]
r_tp1 = r[1:]
sqrt_r_t = np.sqrt(r_t)
y = (r_tp1 - r_t) / sqrt_r_t
x1 = 1 / sqrt_r_t
x2 = sqrt_r_t
X = np.column_stack([x1, -x2])  # no intercept
model = sm.OLS(y, X)
results = model.fit()
beta1, beta2 = results.params

# Parameter recovery
a = beta2 / dt
```

```python
b = beta1 / beta2
sigma = np.std(results.resid) / np.sqrt(dt)

print("OLS estimation (CIR linearized form):")
print(f"a     = {a:.6f}")
print(f"b     = {b:.6f}")
print(f"sigma = {sigma:.6f}")
print(f"R    = {results.rsquared:.6f}")

# --- Residuals plot + save
plt.figure(figsize=(14, 5))
plt.scatter(range(len(results.resid)), results.resid, color="crimson", s=10,
    alpha=0.7)
plt.axhline(0, color="black", linestyle="--")
plt.title("Regression Residuals (CIR OLS)", fontsize=15)
plt.xlabel("Time", fontsize=12)
plt.ylabel("Residual", fontsize=12)
plt.grid(True, linestyle="--", alpha=0.6)
plt.tight_layout()

# Salva e mostra
filename_resid = "cir_ols_residuals_plot.pdf"
plt.savefig(filename_resid, dpi=600, bbox_inches='tight')
plt.show()

# Download (Colab)
files.download(filename_resid)

# --- CIR Simulation
N = len(r)
n_simulations = 10000
r0 = r[0]
paths = np.zeros((N, n_simulations))
paths[0] = r0
np.random.seed(42)

for t in range(1, N):
    z = np.random.normal(size=n_simulations)
    rt = paths[t - 1]
    dr = a * (b - rt) * dt + sigma * np.sqrt(np.maximum(rt, 0)) * np.sqrt(dt) * z
    paths[t] = rt + dr

# --- Percentile bands
median = np.percentile(paths, 50, axis=1)
p5 = np.percentile(paths, 5, axis=1)
p95 = np.percentile(paths, 95, axis=1)

# --- Plot simulated vs observed
fig, ax = plt.subplots(figsize=(15, 6))
main_color = "#3B73B9"
ax.plot(dates, r, color="black", label="Observed GS3M", linewidth=2, zorder=3)
ax.plot(dates, median, color=main_color, label="Simulated Median (CIR)",
    linewidth=2.2, zorder=2)
ax.fill_between(dates, p5, p95, color=main_color, alpha=0.2, label="5th95th
    Percentile", zorder=1)

ax.set_title("Observed vs Simulated CIR Rates (OLS Estimation)", fontsize=16)
ax.set_xlabel("Date", fontsize=13)
```

```python
ax.set_ylabel("Interest Rate", fontsize=13)
ax.grid(True, linestyle='--', linewidth=0.5, alpha=0.7)
ax.legend(fontsize=12)
plt.tight_layout()

# --- Save and download
filename = "cir_simulation_plot_ols.pdf"
plt.savefig(filename, dpi=600, bbox_inches='tight')
plt.show()

# --- Download in Colab
files.download(filename)
```

Table 1: OLS estimation for CIR model parameters (3-Month Treasury Bill Rates)

Parameter	Estimate	Description
a	1.502519	Speed of mean reversion
b	0.013517	Long-term mean level
σ	0.231797	Volatility coefficient
R^2	0.018455	R-squared

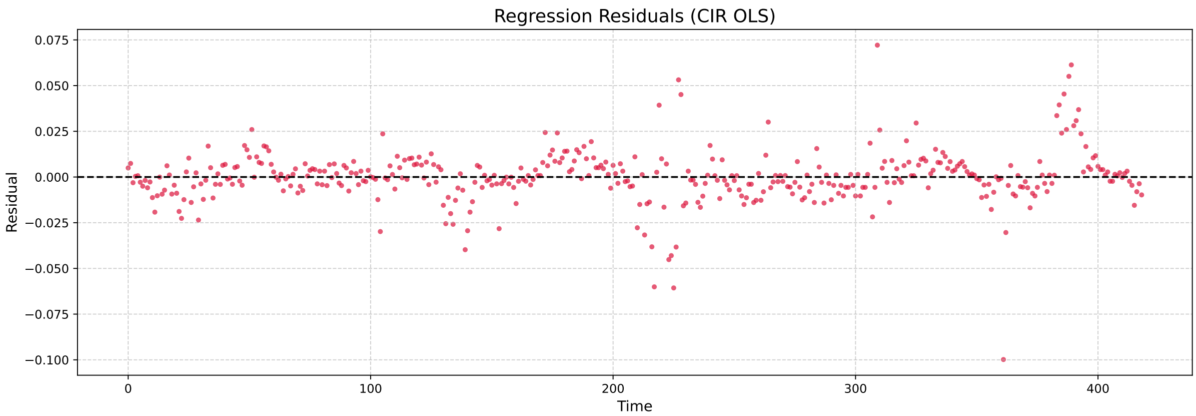

Figure 3: Regression residuals from the linearized CIR model estimated via OLS. The residuals, computed as the difference between observed and predicted values of the transformed interest rate process, appear centered around zero with some dispersion, suggesting potential heteroskedasticity.

The residual plot in Figure 3 provides insight into the goodness-of-fit of the linearized OLS regression. Although the residuals appear to be centered around zero, indicating the absence of strong bias, there is clear heteroskedasticity, as the variance of the residuals varies across time. This is a known limitation of OLS applied to models with state-dependent volatility such as the CIR process. The low R^2 value reported in Table 1 is expected due to the inherently stochastic nature of the CIR process, whose dynamics are predominantly driven by the volatility term rather than the deterministic drift. Consequently, the linearized OLS regression, focusing exclusively on the drift component, captures only a limited fraction of the total variability present in the interest rate data. Nonetheless, the estimated parameters obtained via OLS remain valuable as initial approximations for more precise estimation methods such as MLE.

4 Parameter Estimation via Maximum Likelihood Estimation (MLE)

While the OLS approach provides a fast and interpretable estimation of the CIR model parameters, it is based on a linearized and discretized version of the original continuous-time process. To overcome the limitations of this approximation, we also estimate the parameters using Maximum Likelihood Estimation (MLE) based on the exact transition density of the CIR process.

One important property of the CIR process is that it admits a closed-form expression for the transition probability density function (pdf). Specifically, the distribution of $r(t + \Delta t)$ given $r(t)$ is a scaled non-central chi-squared distribution:

$$r(t + \Delta t) \mid r(t) \sim c \cdot \chi'^2_d(\lambda), \tag{5}$$

where the parameters of the distribution are given by:

$$\begin{aligned} c &= \frac{\sigma^2(1 - e^{-a\Delta t})}{4a}, \\ d &= \frac{4ab}{\sigma^2}, \\ \lambda &= \frac{4ae^{-a\Delta t}}{\sigma^2(1 - e^{-a\Delta t})} r(t). \end{aligned}$$

This result can be derived by solving the Fokker–Planck equation (also known as the forward Kolmogorov equation) for the CIR process. Due to the affine structure of the model, the transition density can be shown to coincide with that of a scaled non-central chi-squared distribution. Specifically, by applying an appropriate transformation to the CIR process, it is possible to express $r(t)$ in terms of a process whose dynamics match those of a non-central chi-squared variable with time-dependent parameters.

This connection to the non-central chi-squared distribution is crucial for constructing the likelihood function. It allows us to write down an explicit expression for the conditional density $p(r(t + 1) \mid r(t); a, b, \sigma)$, and consequently, build the full likelihood as a product of these conditional densities over the observed time series. The parameter estimates obtained via Maximum Likelihood Estimation (MLE) are reported in Table 2. It is important to note that the estimates obtained through Ordinary Least Squares (OLS) were used as initial values for the MLE procedure.

```
import numpy as np
import pandas as pd
import datetime
from scipy.optimize import minimize
from scipy.stats import ncx2
import matplotlib.pyplot as plt
from pandas_datareader import data as pdr
from google.colab import files
import warnings
warnings.filterwarnings("ignore")
```

```python
# Load GS3M data from FRED
start = datetime.datetime(1990, 1, 1)
end = datetime.datetime(2024, 12, 31)
df = pdr.DataReader("GS3M", "fred", start, end).dropna()
r = df['GS3M'].values / 100
dates = df.index
dt = 1 / 250

# CIR negative log-likelihood using non-central chi-squared density
def cir_ncx2_neg_log_likelihood(params, r, dt):
    alpha, b, sigma = params
    if alpha <= 0 or b <= 0 or sigma <= 0:
        return 1e10

    eps = 1e-8
    ll = 0.0

    for i in range(1, len(r)):
        rt = max(r[i - 1], eps)
        rt1 = max(r[i], eps)

        c = (sigma**2 * (1 - np.exp(-alpha * dt))) / (4 * alpha)
        d = 4 * alpha * b / sigma**2
        lambd = (4 * alpha * np.exp(-alpha * dt) / (sigma**2 * (1 - np.exp(-alpha *
    dt)))) * rt

        if c <= 0 or d <= 0 or lambd < 0:
            return 1e10

        x = rt1 / c
        try:
            density = ncx2.pdf(x, d, lambd) / c
            if density <= 0 or np.isnan(density):
                return 1e10
            ll += np.log(density)
        except:
            return 1e10

    return -ll

# MLE optimization
initial_guess = [1.502519, 0.013517, 0.231797]
result = minimize(cir_ncx2_neg_log_likelihood, initial_guess, args=(r, dt),
                  method='L-BFGS-B', bounds=[(1e-5, None), (1e-5, None), (1e-5,
    None)])
alpha_hat, b_hat, sigma_hat = result.x
loglik = -result.fun

print("Estimated Parameters (CIR MLE via nc-chi2):")
print(f"alpha = {alpha_hat:.5f}")
print(f"b     = {b_hat:.5f}")
print(f"sigma = {sigma_hat:.5f}")
print(f"Log-likelihood = {loglik:.5f}")

# CIR simulation
N = len(r)
n_simulations = 10000
r0 = r[0]
```

```
paths = np.zeros((N, n_simulations))
paths[0] = r0
np.random.seed(42)

for t in range(1, N):
    z = np.random.normal(size=n_simulations)
    rt = paths[t - 1]
    dr = alpha_hat * (b_hat - rt) * dt + sigma_hat * np.sqrt(np.maximum(rt, 0)) *
    np.sqrt(dt) * z
    paths[t] = rt + dr

# Percentiles
median = np.percentile(paths, 50, axis=1)
p5 = np.percentile(paths, 5, axis=1)
p95 = np.percentile(paths, 95, axis=1)

# Plot
fig, ax = plt.subplots(figsize=(15, 6))
main_color = "#3B73B9"
ax.plot(dates, r, color="black", label="Observed GS3M", linewidth=2, zorder=3)
ax.plot(dates, median, color=main_color, label="Simulated Median (CIR  MLE)",
    linewidth=2.2, zorder=2)
ax.fill_between(dates, p5, p95, color=main_color, alpha=0.2, label="5th95th
    Percentile", zorder=1)

ax.set_title("Observed vs Simulated CIR Rates (MLE via nc-chi2)", fontsize=16)
ax.set_xlabel("Date", fontsize=13)
ax.set_ylabel("Interest Rate", fontsize=13)
ax.grid(True, linestyle='--', linewidth=0.5, alpha=0.7)
ax.legend(fontsize=12)
plt.tight_layout()

# Save and download
filename = "mle_simulations.pdf"
plt.savefig(filename, dpi=600, bbox_inches='tight')
plt.show()
files.download(filename)
```

Table 2: MLE estimation for CIR model parameters (3-Month Treasury Bill Rates)

Parameter	Estimate	Description
a	1.90007	Speed of mean reversion
b	0.01639	Long-term mean level
σ	0.24153	Volatility coefficient

It can be observed that the estimates are similar to those obtained via OLS, although the speed of mean reversion and the long-term mean level are higher when estimated using MLE. The estimate $\hat{a} \approx 1.90$ implies an even faster adjustment toward the long-run level compared to the OLS estimate, indicating a strong pullback mechanism in the rate dynamics. The long-term mean level increases slightly to $\hat{b} \approx 1.64\%$, suggesting a marginally more optimistic view of the equilibrium interest rate. The volatility parameter $\hat{\sigma} \approx 0.2415$ is also higher, which enhances the model's ability to capture dispersion in observed rates. Overall, these estimates provide a more flexible fit to the data, benefiting from the likelihood-based approach and the exact transition density of the CIR process.

To assess the internal consistency of the model, we can verify the Feller condition: $2ab \geq \sigma^2$. Substituting the MLE estimates gives $2 \cdot 1.90007 \cdot 0.01639 \approx 0.0623$ and $\sigma^2 \approx 0.0583$, so the condition is satisfied. This ensures that the origin is an inaccessible boundary and the short rate remains strictly positive with probability one under the estimated parameters. Figure 4 displays the comparison between the actual historical dynamics of interest rates and the 5th and 95th percentiles of simulations generated with the CIR model calibrated using the MLE-estimated parameters. The simulated median closely follows the general downward trend of the observed interest rates, particularly in the earlier years. However, the model struggles to capture abrupt shifts and prolonged periods of near-zero rates, as seen post-2008. The wide confidence band reflects the high uncertainty in long-run simulations due to the stochastic nature of the CIR process. This highlights the trade-off between analytical tractability and empirical flexibility when using affine term structure models.

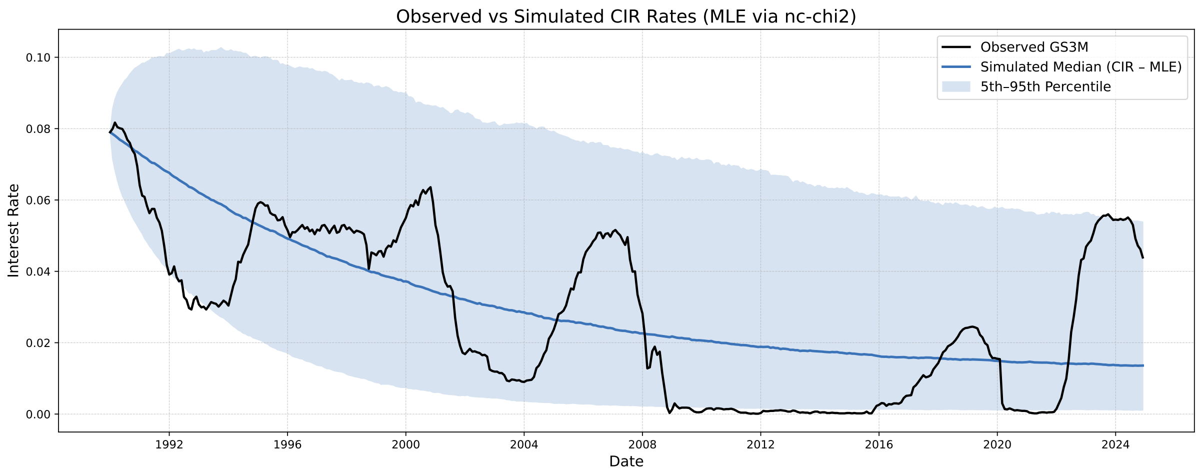

Figure 4: Comparison between observed and simulated 3-Month Treasury Bill rates using the CIR model estimated via Maximum Likelihood. The black line represents the historical GS3M rates, while the blue line shows the median of 10,000 Monte Carlo simulations based on CIR dynamics with parameters estimated via maximum likelihood using the non-central chi-squared transition density. The shaded area corresponds to the 5th and 95th percentiles of the simulated paths, illustrating the model's stochastic dispersion.

5 Conclusion

The Cox-Ingersoll-Ross (CIR) model provides a mathematically elegant and economically grounded framework for modeling the evolution of short-term interest rates. Its most notable advantage lies in the non-negativity constraint on interest rates, ensured by the square-root diffusion term. This feature was especially appealing at the time of its development, when negative rates were considered a theoretical impossibility.

In this work, we applied both an Ordinary Least Squares (OLS) and a Maximum Likelihood Estimation (MLE) approach to estimate the parameters of the CIR model using U.S. 3-month Treasury Bill data. The OLS method enabled fast and interpretable point estimates, while the MLE approach—based on the exact transition density of the

CIR process modeled as a non-central chi-squared distribution—provided a more rigorous statistical foundation for inference. Both methods yielded consistent results, suggesting that the CIR model adequately captures the key dynamics of interest rate mean reversion and volatility.

However, a limitation emerges when applying the CIR model in the context of modern interest rate environments. As shown in Figure 1, short-term interest rates in countries such as Germany have remained negative for extended periods, particularly from 2015 to 2022. Since the CIR model enforces strictly non-negative rates by construction, it becomes unsuitable in these empirical settings. Alternative models such as Vasicek or Hull-White, which allow for negative interest rates due to their Gaussian or additive noise structure, are often more appropriate under such monetary regimes.

Nevertheless, the CIR model remains highly relevant in contexts where negative rates are implausible or undesirable, including insurance applications, credit risk modeling, and regulatory scenarios where preserving positivity is essential. It is also widely used as a building block in multi-factor term structure models.

In summary, while the CIR model may not be universally applicable—especially in low-rate or negative-rate environments—it continues to serve as a foundational reference in interest rate theory. Its combination of analytical tractability and economic plausibility ensures its enduring utility in financial modeling, particularly when the non-negativity constraint is desired or required.

References

[1] 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.

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