# Testing for Omitted Variables in the Conditional Variance Function in a Heteroscedastic Regression Model

Fuchun Li [1] Jiajing Sun [2] and Yuanlin Wang[3]

[1]Payments Canada

[2]University of Chinese Academy of Science

[3]Dongbei University of Economics and Finance

June 3, 2026

# Overview

- 1. Motivation
- 2. Contributions
- 3. The model, hypothesis and test statistic
- 4. The limiting distribution of the test statistic
- 5. Monte Carlo studies
- 6. An empirical application
- 7. Conclusions

# 1. Motivation

- The homoscedasticity assumption in a regression model can make the statistical inference substantially simplified
  - Homoscedasticity: the variance of the error term in a regression model is constant.
    - The OLS estimator is unbiased and efficient
    - unbiased estimation of the standard error.
- The heteroskedasticity is a major concern in regression analysis.
  - heteroskedasticity: the variance of the error term in a regression model is not constant.
    - The OLS estimator is inefficient
    - Inference based on the assumption of homoskedasticity is misleading

# 1. Motivation

- Tremendous work exists in this literature on testing for heteroskedasticity.
    - The early tests for heteroscedasticity are based on either the Lagrange multiplier principle or the least squares residuals
        - Examples of these early tests include the tests by Breusch and Pagan (1979), White (1980), Koenker and Bassett (1982), and Newey and Powell (1987), among others.
    - A problem associated with these early tests is that many of these tests are not consistent (Hong, 1993).
- Nonparametric consistent tests for heteroscedasticity are proposed, including the nonparametric tests by Hong (1993), Dette and Munk (1998), Hsiao and Li (2001), Su and Ullah (2013), and Xu and Cao (2021), among others.

# 1. Motivation

- When the homoscedasticity assumption is not met, efficient statistical inference for the regression model requires that heteroscedasticity be taken into account
  - A serious problem with a heteroscedastic regression model is the misspecification of the conditional variance function of the error term
  - The misspecification of the conditional variance function could lead to misleading conclusions in inference and hypothesis test
  - It is therefore important to develop reliable tests for the functional form of the conditional variance function of the error term
- Much work has been devoted to using a nonparametric method to testing a parametric form of the conditional variance function in a heteroscedastic regression model

# 1. Motivation

- Examples include those tests developed by Hu, Li, and Tan (2024), Wang and Zhou (2007), and Samarakoon and Song (2012), just to mention a few.
- Relatively, the issue of testing the null hypothesis of a nonparametric form of the conditional variance function against a nonparametric alternative has attracted less attention
  - A leading case where such a situation naturally arises is to test the significance of certain conditional variables in a nonparametric conditional variance function
  - Namely whether certain conditional variables can be omitted from the nonparametric conditional variance function.

# 1. Motivation

- A simplified conditional variance function derived from omitted a subset of the conditional variables could be analytically tractable and easy to implement the inference analysis.
  - However, such a simplification specification of the conditional variance function should be justified quantitatively
- This motivates us to develop a test for the significance of a subset of conditional variables in the conditional variance function of the error term in a general regression model.

# 2.Contribution

- Propose a consistent test for the significance of a subset of conditional variables in the conditional variance function of the error term in a general regression model
  - The proposed test is shown to have an asymptotic normal distribution under the null hypothesis that a subset of conditional variables can be omitted from the conditional variance function
  - The test statistic diverges to positive infinity against fixed alternatives and has nontrivial power against a class of local alternatives.
- An online supplement provides an empirical illustration using Canadian interest-rate data.

# 2. Contribution

- The supplement uses daily 3-month Canadian Treasury bill rates and 10-year Government of Canada benchmark bond yields as short- and long-rate proxies.
- The empirical test examines whether the long-term rate contains additional information for the conditional variance of short-rate innovations.
- The corrected studentized statistic rejects the omitted-variable null; the long-term rate should not be omitted from the variance specification.

# 3. The model, hypotheses and test statistic

- Following Su and Ullah (2010) and Hsiao and Li (2001), consider a time-series regression model:

$$Y_t = g(Z_t, \theta_0) + U_t$$

  - $g(\cdot, \cdot)$ is a known function with $Z_t$ being a $\kappa \times 1$ vector of regressors, $\theta_0$ being a $l \times 1$ vector of unknown parameters, and $U_t$ being a scalar error term such that $E[U_t|Z_t] = 0$. We also assume $E(U_t|X_t) = 0$.
  - $X_t = (W_t, V_t)$ is a vector $d \times 1$ of conditional variables. $X_t$ may be different from regressors $Z_t$, where $W_t \in \mathbb{R}^p, V_t \in \mathbb{R}^q$, and $p + q = d$.
  - Under $E(U_t|X_t) = 0$, the variance of $U_t$ conditional on $X_t = x = (w, v)$ is

$$\sigma^2(x) = E(U_t^2|X_t = x)$$

and the reduced conditional variance is

$$\sigma_1^2(w) = E(U_t^2|W_t = w)$$

# 3. The model, hypotheses and test statistic

- 3.1. The model and hypotheses
  - We are interested in whether $V_t$ can be omitted from the conditional variance function $\sigma^2(W_t, V_t)$,
    namely the null hypothesis,

$$H_0 : \sigma^2(w, v) = \sigma_1^2(w) \text{ almost everywhere,}$$

    the alternative hypothesis is that the two conditional variance functions differ on a compact set with positive measure, that is,

$$H_1 : \sigma^2(w, v) \neq \sigma_1^2(w) \text{ on a compact set } S \subset \mathbb{R}^d \text{ with positive measure.}$$

# 3. The model, hypotheses and test statistic

- 3.2. Test statistic
  - Our test statistic is based on the squared difference between $\sigma^2(w,v)$ and $\sigma_1^2(w)$:

$$J = \int (\sigma^2(w,v) - \sigma_1^2(w))^2 a(w,v) dF(w,v)$$

    - $a(w,v)$ is a bounded nonnegative weighting function with compact support $S$. It localizes the comparison to the region where the two variance functions are being compared.
    - The squared-distance interpretation is preserved when $a(w,v)$ is nonnegative.
    - $F(w,v)$ is the cumulative distribution function of $(W_t, V_t)$.

# 3. The model, hypotheses and test statistic

- To get a feasible test statistic, we need to nonparametrically estimate the unknown functions, $\sigma^2(x), \sigma_1^2(w)$, and $F(w,v)$. $\sigma^2(x) = E[U_t^2|X_t = x]$ can be estimated nonparametrically

$$\hat{\sigma}^2(w,v) = \frac{\sum_{t=1}^{n} z_t K(\frac{W_t - w}{h}, \frac{V_t - v}{h})}{\sum_{t=1}^{n} K(\frac{W_t - w}{h}, \frac{V_t - v}{h})},$$

  $z_t = (Y_t - g(Z_t, \hat{\theta}))^2$ is the estimation of $U_t^2 = (Y_t - g(Z_t, \theta))^2$.

  $\sigma_1^2(w) = E[U_t^2|W_t = w]$ can be estimated nonparametrically,

$$\hat{\sigma}_1^2(w) = \frac{\sum_{t=1}^{n} z_t K(\frac{W_t - w}{h_0})}{\sum_{t=1}^{n} K(\frac{W_t - w}{h_0})},$$

# 3. The model, hypotheses and test statistic

- $h \equiv h_n$ is a sequence of smoothing parameters for kernel estimation of $\sigma^2(w, v)$
- $h_0 \equiv h_{0,n}$ is a sequence of smoothing parameters for kernel estimation of $\sigma_1^2(w)$,
- $K(\cdot)$ is a kernel function, where we take the multivariate kernel function to be a product of the univariate kernel functions $k(\cdot)$.

# 3. The model, hypotheses and test statistic

- Inserting $\sigma^2(x)$ and $\sigma_1^2(w)$ into $J$ and replacing $F(w,v)$ with its empirical distribution function, $F_n(w,v)$, we have,

$$\begin{aligned} J_n &= \int (\hat{\sigma}^2(w,v) - \hat{\sigma}_1^2(w))^2 a(w,v) dF_n(w,v) \\ &= \frac{1}{n}\sum_{t=1}^{n}[\hat{\sigma}^2(W_t,V_t) - \hat{\sigma}_1^2(W_t)]^2 a(W_t,V_t). \end{aligned}$$

- The test statistic for $H_0$ versus $H_1$ is an appropriately centered and scaled version of $J_n$,

$$\hat{J}_n \equiv nh^{d/2}\hat{\sigma}_0^{-1}(J_n - n^{-1}h^{-d}\hat{\delta}_1 - n^{-1}h^{-p}\hat{\delta}_2 - n^{-1}h_0^{-p}\hat{\delta}_3),$$

where $\hat{\sigma}_0, \hat{\delta}_1, \hat{\delta}_2$, and $\hat{\delta}_3$ are the estimators of $\sigma_0, \delta_1, \delta_2$, and $\delta_3$.

# 3. The model, hypotheses, and test statistic

- $\sigma_0, \delta_1, \delta_2$, and $\delta_3$ are defined as follows,

$$\begin{aligned}
\sigma_0^2 &= 2\int \left[ \int \left( \int K(u)K(u+v)du \right)^2 dv \right] \sigma_{U^2}^4(x)a^2(x)dx, \\
\delta_1 &= \int K^2(u)du \int \sigma_{U^2}^2(x)a(x)dx, \\
\delta_2 &= -2K(0) \int\int \sigma_{U^2}^2(w,v)(f(w,v)/f(w))a(w,v)dwdv, \\
\delta_3 &= \int K^2(w)dw \int\int \sigma_{U^2}^2(w)(a(w,v)f(w,v)/f(w))dwdv,
\end{aligned}$$

where $\sigma_{U^2}^2(x) = E[(U_t^2 - \sigma^2(X_t))^2 | X_t = x], \sigma_{U^2}^2(w) = E[(U_t^2 - \sigma_1^2(W_t))^2 | W_t = w]$, and $f(w)$ is the density function of $W_t$.

# 3. The model, hypotheses, and test statistic

- $\hat{\sigma}_0, \hat{\delta}_1, \hat{\delta}_2$, and $\hat{\delta}_3$ are defined as follows,

$$\begin{aligned}
\hat{\sigma}_0 &= \sqrt{\frac{2\int(\int K(u)K(u+v)du)^2 dv}{n}\sum_{t=1}^{n}\frac{\hat{\sigma}^4_{U^2}(X_t)a^2(X_t)}{\hat{f}(X_t)}}, \\
\hat{\delta}_1 &= \frac{\int K^2(u)du}{n}\sum_{t=1}^{n}\frac{\hat{\sigma}^2_{U^2}(X_t)a(X_t)}{\hat{f}(X_t)}, \\
\hat{\delta}_2 &= -\frac{2K(0)}{n}\sum_{t=1}^{n}\frac{\hat{\sigma}^2_{U^2}(X_t)a(X_t)}{\hat{f}(W_t)}, \\
\hat{\delta}_3 &= \frac{\int K^2(w)dw}{n}\sum_{t=1}^{n}\frac{\hat{\sigma}^2_{U^2}(W_t)a(X_t)}{\hat{f}(W_t)},
\end{aligned}$$

# 3. The model, hypotheses, and test statistic

- where

$$\hat{\sigma}^2_{U^2}(x) = \frac{\sum_{t=1}^n \hat{U}_t^4 K(\frac{X_t - x}{h})}{\sum_{t=1}^n K(\frac{X_t - x}{h})} - \left(\frac{\sum_{t=1}^n \hat{U}_t^2 K(\frac{X_t - x}{h})}{\sum_{t=1}^n K(\frac{X_t - x}{h})}\right)^2,$$

$$\hat{\sigma}^2_{U^2}(w) = \frac{\sum_{t=1}^n \hat{U}_t^4 K(\frac{W_t - w}{h_0})}{\sum_{t=1}^n K(\frac{W_t - w}{h_0})} - \left(\frac{\sum_{t=1}^n \hat{U}_t^2 K(\frac{W_t - w}{h_0})}{\sum_{t=1}^n K(\frac{W_t - w}{h_0})}\right)^2,$$

$\hat{U}_t = Y_t - g(Z_t, \hat{\theta})$, and $\hat{f}(x)$ and $\hat{f}(w)$ are the nonparametric density estimators of $f(x)$ and $f(w)$,

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{t=1}^n K(\frac{X_t - x}{h}),$$

and

$$\hat{f}(w) = \frac{1}{nh_0^p} \sum_{t=1}^n K(\frac{W_t - w}{h_0}).$$

# 4. The limiting distribution of the test statistic

- To derive the asymptotic distribution of our test statistic, we need the following assumptions
  - Assumption 1. $\{(U_t, Z_t', X_t')'\}_{t=1}^n$ is a strictly stationary absolutely regular process with the mixing coefficient $\beta_\tau$ that satisfies $\beta_\tau = O(\rho^\tau)$ for some $0 < \rho < 1$.
  - Assumption 2. The weight function $a(w, v)$ is bounded, nonnegative, and compactly supported on $S$, with $\Pr\{a(X_t) > 0\} > 0$; the density $f(w, v)$ is bounded away from zero on $S$, namely $\inf_{(w,v)\in S} f(w, v) > 0$.
  - Assumption 3. The functions $\sigma^2(x)$, $\sigma_1^2(w)$, $\mu_4(x)$, and $\eta_4(w)$ are $r$-times continuously differentiable, or belong to a Hölder class of order $r$, on the relevant compact supports.
  - Assumption 4. The kernel function $K(\cdot)$ is a product of the univariate kernel function $k(\cdot)$ that is a bounded function on $\mathbb{R}$, symmetric about 0, with $\int |k(u)|du < \infty$, $\int k(u)du = 1$, $\int u^j k(u)du = 0$ for $1 \leq j < r$, where $r > 3d/4$.

# 4. The limiting distribution of the test statistic

- Assumption 5. The smoothing parameters satisfy $h \asymp n^{-1/\gamma}$ and $h_0 \asymp n^{-1/\lambda}$, with $2d < \gamma < 2r + d/2$, $p < \lambda \leq 2r + p$, and $\gamma p/d < \lambda < \gamma$.
- Assumption 6. (i) The parameter space $\Theta$ of $\theta$ is a compact subset of $\mathbb{R}^l$. $E[Y_t - g(Z_t, \theta_0)]^2$ is uniquely minimized at $\theta_0$ in $\Theta$. (ii) The regression function $g(z, \theta)$ is twice continuously differentiable in $\theta$. Let $\nabla g(z, \theta) \equiv \partial g(z, \theta)/\partial\theta$ and $\nabla^2 g(z, \theta) \equiv \partial^2 g(z, \theta)/\partial\theta\theta'$ be continuous in $z$ and dominated by a function $M_g(z)$ with $E[M_g^2(Z_t)|X_t = x]$ being a continuous function on $S$. (iii) $E[\nabla g(z, \theta)\nabla' g(z, \theta)]$ is nonsingular for $\theta$ in a neighborhood of $\theta_0$. (iv) $E[\nabla' g(Z_t, \theta)U_t|X_t] = 0$.

# 4. The limiting distribution of the test statistic

- Briefly comment on the above assumptions.
  - Assumption 1 is needed to apply the Central Limit Theorem (CLT) for second order degenerate U-statistics of strictly stationary and absolutely regular processes (Fan and Li, 1999)
  - Assumption 2 requires that $a(x)$ be bounded with compact support. As a result, we can only detect deviation between $\sigma(x)$ and $\sigma_1(w)$ on $S$.
  - Assumption 3 imposes moment conditions on $(U_t^2 - \sigma^2(X_t))$ under both the alternative and null hypotheses.
  - Assumption 4 is a standard assumption on the kernel function with order $r$.
  - Under Assumption 5, we have $nh^d \to \infty, nh_0^p \to \infty$, $nh^{d/2+2r} \to 0$, $h_0/h \to 0$ and $h^d/h_0^p \to 0$.
  - Assumption 6 (i)-(iii) together with Assumption 1 ensure that $\theta_n - \theta_0 = O_p(n^{-1/2})$.

# 4. The limiting distribution of the test statistic

- Asymptotic distributions
  - Assumption 1 is needed to apply the Central Limit Theorem (CLT) for second order degenerate U-statistics of strictly stationary and absolutely regular processes (Fan and Li, 1999)
  - Assumption 2 requires that $a(x)$ be bounded with compact support. As a result, we can only detect deviation between $\sigma(x)$ and $\sigma_1(w)$ on $S$.
  - Assumption 3 imposes moment conditions on $(U_t^2 - \sigma^2(X_t))$ under both the alternative and null hypotheses.
  - Assumption 4 is a standard assumption on the kernel function with order $r$.
  - Under Assumption 5, we have $nh^d \to \infty, nh_0^p \to \infty$, $nh^{d/2+2r} \to 0$, $h_0/h \to 0$ and $h^d/h_0^p \to 0$.
  - Assumption 6 (i)-(iii) together with Assumption 1 ensure that $\theta_n - \theta_0 = O_p(n^{-1/2})$.

# 4. The limiting distribution of the test statistic

- **Theorem 1.** Assume that Assumptions 1-6 hold. Then we have, (i) Under $H_0$, $\hat{J}_n \Rightarrow N(0,1)$ in distribution.
(ii)Under $H_1$, $\Pr[\hat{J}_n > B_n] \to 1$ for any non-stochastic sequence $B_n = o(nh^{d/2})$.
    - In practice, to test the null hypothesis $H_0$ against $H_1$ at the level $\alpha$, we need to compare $\hat{J}_n$ to the critical value $z_\alpha$ from the $N(0,1)$ distribution,
        - for example, $z_{0.01} = 2.33, z_{0.05} = 1.64$, and $z_{0.10} = 1.28$ because the test statistic $\hat{J}_n$ is one-sided.
        - The null hypothesis is rejected if $\hat{J}_n > z_\alpha$.

# 4. The limiting distribution of the test statistic

- To examine the asymptotic local power of $\hat{J}_n$, we consider the following sequence of Pitman local alternatives,

$$H_1^{\gamma_n}: \quad \sigma^2(x) = \sigma_1^2(w) + \gamma_n \Delta(x),$$

  where $\Delta(x)$ is a continuous function, $E[\Delta(W_t, V_t) \mid W_t = w] = 0$ a.s., $\int |\Delta(x)| dx < \infty$, $\int \Delta^2(x) dx < \infty$, $\Pr\{\Delta(X_t) \neq 0, X_t \in S\} > 0$, and $\gamma_n \to 0$ as $n \to \infty$.

- The following theorem establishes the local power property of our test statistic.

- **Theorem 2.** Assume that Assumptions 1-6 hold. Let $\mu_\Delta = \sigma_0^{-1} \int [2f^2(x)/f^2(w)] \Delta^2(x) a(x) dF(x)$. If $H_1^{\gamma_n}$ holds and $\gamma_n = n^{-1/2} h^{-d/4}$, then $\hat{J}_n \Rightarrow N(\mu_\Delta, 1)$ and $\Pr[\hat{J}_n > z_\alpha] \to 1 - \Phi(z_\alpha - \mu_\Delta)$.

# 4. The limiting distribution of the test statistic

- Theorem 2 indicates that $\hat{J}_n$ can detect local alternatives in $H_1^{\gamma_n}$ that differ from the null hypothesis at the rate of $n^{-1/2}h^{-d/4}$.
- The slower $h$ converges to zero, the greater is the asymptotic power of $\hat{J}_n$ in the sense that it can detect local alternatives in $H_1^{\gamma_n}$ that are closer to the null hypothesis.

# 5. Monte Carlo studies

- Monte Carlo simulations are conducted to assess the finite sample performance of our test statistic.
- The following data generating processes (DGPs), DGP1 and DGP2, to study the size performance of $\hat{J}_n$.
  - DGP1: $Y_t = 1 + Z_t + \eta_t \epsilon_t$, $Z_t = 0.5Z_{t-1} + \varepsilon_t$, $W_t = Z_t$, $V_t \sim N(0,1)$, and $\eta_t^2 = \exp(0.2 + 0.8W_t)$. Under DGP1, $V_t$ is irrelevant for the conditional variance.
  - DGP2: $Y_t = 0.5Y_{t-1} + \eta_t \epsilon_t$, $W_t = Y_{t-1}$, $V_t \sim N(0,1)$, and $\eta_t^2 = \exp\{0.2 + 0.4g_s(Y_{t-1})\}$, where $g_s(y) = 0.75 \tanh(y/0.75)$. Under DGP2, $V_t$ is irrelevant for the conditional variance.

# 5. Monte Carlo studies

- The following data generating processes (DGPs), DGP3 and DGP4, to study the power performance of $\hat{J}_n$.
  - DGP3: $Y_t = 1 + Z_t + \zeta_t \epsilon_t$, where $Z_t = 0.5 Z_{t-1} + \varepsilon_t$, and $\varepsilon_t$ and $\epsilon_t$ are independent processes, both of which follow i.i.d. $N(0,1)$. Let $V_t$ be i.i.d. $N(0,1)$.
  $\zeta_t = \zeta(X_t) = \sqrt{exp(0.2 + 0.3 Z_t + 0.5 V_t)}$.
  - DGP4: $Y_t = 0.5 Y_{t-1} + \xi_t \epsilon_t$, $W_t = Y_{t-1}$, $V_t \sim N(0,1)$, and $\xi_t^2 = \exp\{0.2 + 0.3 g_s(Y_{t-1}) + 0.5 V_t\}$, where $g_s(y) = 0.75 \tanh(y/0.75)$.
- Note that DGP1 is tested as $H_0$ against DGP3 as $H_1$, and DGP2 is tested as $H_0$ against DGP4 as $H_1$.

# Monte Carlo studies: size

- The table reports estimated rejection frequencies at the 5% nominal level for the current DGP1–DGP2 designs.

| | | DGP1 | | | DGP2 | |
|---|---|---|---|---|---|---|
| $n$ | $c = 1.0$ | $c = 1.5$ | $c = 2.0$ | $c = 1.0$ | $c = 1.5$ | $c = 2.0$ |
| 250 | 0.026 | 0.017 | 0.021 | 0.044 | 0.019 | 0.021 |
| 500 | 0.024 | 0.024 | 0.037 | 0.049 | 0.025 | 0.033 |
| 1000 | 0.031 | 0.025 | 0.042 | 0.050 | 0.042 | 0.050 |
| 2500 | 0.035 | 0.033 | 0.049 | 0.064 | 0.045 | 0.046 |

The estimated sizes generally move toward the nominal level as $n$ increases, with some finite-sample sensitivity to bandwidth choices.

# Monte Carlo studies: power

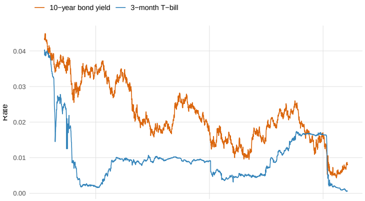
- The table reports estimated powers at the 5% nominal level for the current DGP3–DGP4 alternatives.

| | | DGP3 | | | DGP4 | |
|---|---|---|---|---|---|---|
| $n$ | $c = 1.0$ | $c = 1.5$ | $c = 2.0$ | $c = 1.0$ | $c = 1.5$ | $c = 2.0$ |
| 250 | 0.521 | 0.552 | 0.528 | 0.527 | 0.578 | 0.586 |
| 500 | 0.817 | 0.896 | 0.909 | 0.838 | 0.910 | 0.923 |
| 1000 | 0.983 | 0.996 | 0.997 | 0.990 | 0.998 | 0.999 |
| 2500 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Power increases with sample size. Within the reported bandwidth range, larger bandwidth multipliers tend to give higher finite-sample power.

# 6. Empirical illustration

- The online supplement applies the test to Canadian interest-rate data.




- Question: should the current long-term rate be included when modeling the conditional variance of short-rate innovations?
- Conditioning variables: $W_t = r_t^1$ is the 3-month Canadian Treasury bill rate, and $V_t = r_t^2$ is the 10-year Government of Canada benchmark bond yield.

# 6. Empirical illustration

- The response variable is

$$Y_t = (r^1_{t+1} - r^1_t)/\sqrt{\Delta}, \qquad \Delta = 1/250.$$

- The mean equation is $Y_t = \alpha + \beta r^1_t + U_t$. The full variance estimator conditions on $(r^1_t, r^2_t)$ and the reduced estimator conditions only on $r^1_t$.
- The empirical weight is the indicator of the central 95% rectangle of $(r^1_t, r^2_t)$, consistent with the compact-support condition.
- A standardized residual bootstrap is used as a finite-sample calibration check.

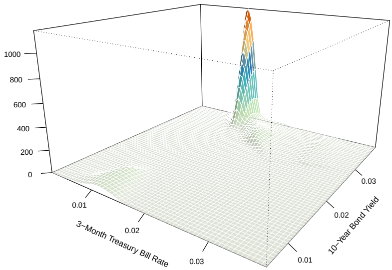
# 6. Empirical illustration

| Specification | $n$ | $\hat{J}_n$ | Asym. $p$ | Bootstrap $B$ | Boot. $p$ |
|---|---|---|---|---|---|
| Full sample, central 95% support | 3316 | 3.177 | 0.0007 | 499 | 0.002 |
| Full sample, all observations | 3316 | 3.177 | 0.0007 | – | – |
| Early-sample robustness, central 95% support | 3083 | 3.317 | 0.0005 | 499 | 0.004 |
| Early-sample robustness, all observations | 3083 | 3.317 | 0.0005 | – | – |

- The test rejects conditioning only on the current short rate $r_t^1$.
- The current long-term rate $r_t^2$ should be included in the conditioning vector for the conditional variance specification.

# 6. Empirical illustration

- The estimated full conditional variance surface changes in the long-rate direction.
- This supports including $r_t^2$ in the conditioning set, rather than using only $r_t^1$.



# 7. Conclusions

- We propose a nonparametric specification test for omitted variables in the conditional variance function of a heteroscedastic regression model.
  - The statistic is asymptotically standard normal under the omitted-variable null.
  - The statistic diverges under fixed alternatives and has nontrivial power under local alternatives.
- Monte Carlo simulations show that estimated sizes generally move toward nominal levels as the sample size increases, with some bandwidth sensitivity, and that the test is powerful against the alternatives considered.
- The online supplement provides an empirical illustration using Canadian interest-rate data.

# 7. Conclusions

- Using daily 3-month Canadian Treasury bill rates and 10-year Government of Canada benchmark bond yields, the empirical illustration rejects conditioning only on the current short rate.
- The empirical evidence indicates that the current long-term rate should be included in the conditioning set for the conditional variance of short-rate innovations.
- The method can be used more broadly to assess which variables are necessary in multivariate conditional-volatility and conditional second-moment specifications.

# Key references


- Aït-Sahalia (1996), testing continuous-time spot-rate models, *Review of Financial Studies*.
- Fan and Yao (1998), conditional variance estimation, *Biometrika*.
- Fan and Li (1999), degenerate U-statistics and specification testing, *Journal of Nonparametric Statistics*.
- Zheng (1996), nonparametric functional-form testing, *Journal of Econometrics*.
- Dette, Neumeyer, and Van Keilegom (2007), testing variance functions, *JRSS-B*.

# Thank you

Any feedback or suggestions are very welcome.