Econometric Theory, **0**, 2026, 1–32.
doi:10.1017/S0266466626100395

UNLOCKING THE REGRESSION SPACE

Liudas Giraitis
Queen Mary University of London

George Kapetanios
King's College London

Yufei Li
King's College London

Alexia Ventouri
King's College London

This article introduces and analyzes a framework that accommodates general heterogeneity in regression modeling. It demonstrates that regression models with fixed or time-varying parameters can be estimated using the ordinary least squares (OLS) and time-varying OLS methods, respectively, across a broad class of regressors and noise processes not covered by existing theory. The proposed setting facilitates the development of asymptotic theory and the estimation of robust standard errors. The robust confidence interval estimators accommodate substantial heterogeneity in both regressors and noise. The resulting robust standard error estimates coincide with White's (1980, Econometrica *48*, 817–838) heteroskedasticity-consistent estimator but are applicable to a broader range of conditions, including models with missing data. They are computationally simple and perform well in Monte Carlo simulations. Their robustness, generality, and ease of implementation make them highly suitable for empirical applications. Finally, the article provides a brief empirical illustration.

1. INTRODUCTION

Regression analysis is the cornerstone of statistical theory and practice. Ordinary least squares (OLS) has been applied, within various regression contexts, to build an extensive toolkit, for the exploration of economic and financial datasets. The basic theory underlying OLS estimation and inference in regression models has been largely established for over half of a century (see, e.g., Lai and Wei, 1982). The problem of robust estimation has long been a focus of empirical work in economics, beginning with the seminal work by White (1980). Its importance is well understood in applied econometrics. At the same time, several important concerns

The authors thank the Editor Peter C. B. Phillips and two anonymous referees for insightful comments and suggestions that significantly improved the article. Address correspondence to Liudas Giraitis, School of Economics and Finance, Queen Mary University of London, United Kingdom, e-mail: l.giraitis@qmul.ac.uk.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (https://creativecommons.org/licenses/by-nc-nd/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use and/or adaptation of the article.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

have been raised by applied researchers. Angrist and Pischke (2010) noted that "Leamer (1983) diagnosed his contemporaries' empirical work as suffering from a distressing lack of robustness to changes in key assumptions," and Leamer (2010) later reflected that "sooner or later, someone articulates the concerns that gnaw away in each of us and asks if the Assumptions are valid." Similarly, Karmakar, Richter, and Wu (2022) observed that the assumption of parameter constancy, or "stationarity is often an oversimplified assumption that ignores systematic deviations of parameters from constancy." Clearly, this concern extends beyond parameter stability to encompass the stability of regressors, regression noise, and the underlying modeling assumptions.

In this article, we focus on the inherent capacity of regression modeling to accommodate the effects of structural change in settings with both fixed and time-varying parameters. Many such structural changes influence not only the model parameters but also the regression space itself. This space comprises both the regressors and regression noise, and improper treatment of these components may result in incorrect inferences, misinterpretations, and forecasting distortions. We therefore examine which specifications of the regression space can flexibly account for structural change while still enabling estimation of both fixed and time-varying regression parameters, the construction of confidence intervals, and the computation of standard errors.

Among recent developments, Wu (2005), Hall, Han, and Boldea (2012), and others have proposed advanced theoretical methods for the estimation of the fixed parameters, while Cattaneo, Jansson, and Newey (2018) and Jochmans (2019) developed procedures to estimate both fixed parameters and standard errors in regression models with an increasing number of covariates and heteroskedasticity. Meanwhile, Li, Phillips, and Gao (2020), Sun et al. (2021), and Linton and Xiao (2019) introduced new modeling frameworks that explicitly account for structural change. A common response to concerns about heteroskedasticity in the recent literature is the use of the heteroskedasticity-robust variance and standard error estimators for linear regression models (see Eicker, 1963; White, 1980; MacKinnon, 2012; Cattaneo et al., 2018; among others).

There is also a sizeable and growing literature on the estimation of time-varying coefficient regression models, including the works of Fan and Zhang (1999) and Vogt (2012), among others. This literature further explores tests for different types of parameter variation (see, e.g., Bai and Perron, 1998; Zhang and Wu, 2012; Zhang and Wu, 2015; Hu, Kasparis, and Wang, 2025). In addition, specification tests and tests for parameter instability have received significant attention, with important contributions by Hansen (2000), Georgiev et al. (2018), Hidalgo, Lee, and Seo (2019), Boldea, Cornea-Madeira, and Hall (2019), Fu et al. (2023), and others.

The modeling of deterministic, smooth parameter evolution has a long history in statistics. Early examples include linear processes with time-varying spectral densities, introduced by Priestley (1965). This framework is essentially nonparametric and it has been further developed by Robinson (1989), Robinson (1991),

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

Dahlhaus (1997), Dahlhaus, Richter, and Wu (2019), and Dahlhaus and Richter (2023), some of whom refer to these processes as *locally stationary*. The estimation of time-varying parameters, as well as fixed parameters under heteroskedasticity in time-series models, has been studied in Dahlhaus and Giraitis (1998), Xu and Phillips (2008), and Dalla, Giraitis, and Robinson (2020), among others. Nonlinear time-varying time-series models have also been developed by Doukhan and Wintenberger (2008), Bardet and Wintenberger (2009), Vogt (2012), and Karmakar et al. (2022). Despite these advances, such approaches have not been widely adopted in applied economics, where random coefficient models remain more prevalent.

Various methods have been proposed over the years to identify and handle structural change. Early contributions assumed that changes were deterministic, rare, and abrupt. Testing for parameter breaks dates back to the pioneering work by Chow (1960), with further contributions by Brown, Durbin, and Evans (1975) and Ploberger and Krämer (1992), among others. More recent approaches allow for random evolution of parameters, where changes may be discrete, as in Markov switching models by Hamilton (1989) or threshold models by Tong (1990), or continuous as in smooth transition models by Terasvirta (1998), or those driven by unobservable shocks, as in random coefficient models by Nyblom (1989). For example, Cogley and Sargent (2005) use random coefficient models to study stochastic volatility, while Primiceri (2005) examines whether changes in parameters or in the variance of shocks—policy-induced or otherwise—contributed to the period of macroeconomic calmness known as the "Great Moderation" after 1985. In these frameworks, parameters typically evolve as random walks or autoregressive processes.

Building on this literature, Giraitis, Kapetanios, and Yates (2014), Giraitis, Taniguchi, and Taqqu (2018), Dendramis, Giraitis, and Kapetanios (2021), and others have developed a theoretical time-series framework for random coefficient models and their estimation using kernel-based methods, which perform well in finite samples. These methods are computationally simple and straightforward to implement in applied research. For example, Chronopoulos, Giraitis, and Kapetanios (2022) demonstrated the empirical prevalence of persistent volatility, suggesting that GARCH-type volatility structures may be less common than previously thought. Nevertheless, a full treatment of estimation and inference within a general regression framework has, surprisingly, not yet been provided.

In this article, we provide a rigorous validation of the asymptotic normality of the feasible t-statistic for the estimation of both fixed and time-varying parameters in linear regression models within an extended regression space of regressors and regression disturbances. Our main objective is to describe, in transparent terms, the extended regression space under which such normality is preserved. The class of admissible regressors and regression noises is broad. Regressors are obtained by rescaling and shifting stationary short-memory (SM) sequences, while regression errors are generated by arbitrary rescaling of a stationary martingale difference

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

(m.d.) sequence. The restrictions imposed on the scale factors and mean processes are weak, allowing these to be either deterministic or stochastic, and to vary over time, possibly abruptly or through non-stationary (e.g., unit-root) dynamics. Some assumptions on the scale factors are necessary and resemble the Lindeberg condition in the classical Lindeberg–Feller central limit theorem. Importantly, the robust feasible t-statistic retains the same form and limiting distribution as in the standard setting. The estimator of feasible robust standard errors coincides with the heteroskedasticity-consistent standard error estimator of White (1980). Our assumptions do not rely on mixing or near-epoch dependence conditions, which commonly prevail throughout the existing literature. Given the generality of the regression space, these assumptions typically require no additional empirical verification.

The estimation framework for fixed regression parameters is developed in Section 2, which introduces the extended regression space, the underlying assumptions, and the main theoretical results. Section 3 establishes the estimation theory for time-varying regression parameters within the same framework. The proofs highlight how the results for the fixed-parameter case naturally extend to the time-varying settings, with only negligible additional terms.

Our results are complementary to existing frameworks. The novelty lies in providing a methodological foundation that confirms the validity of robust regression estimation in an extended regression space. The fundamental theory in this area traces back to Lai and Wei (1982) who studied regression models with heteroskedastic m.d. noise under eigenvalue-based assumptions. Alternative methods, such as bootstrap procedures, see Hall et al. (2012) and Boldea et al. (2019), are widely used in regression analysis but may not be directly applicable to such general class of regressors and regression noises. In contrast, we demonstrate that White-type standard errors remain applicable and computationally straightforward.

All theoretical results are supported by detailed, rigorous proofs. Monte Carlo simulations confirm that the proposed robust regression estimators perform well in finite samples. Overall, the framework developed in this article is particularly suited to modeling economic and financial data, where heterogeneity, structural change, and dependence are inherent features.

The remainder of the article is organized as follows. Section 2 presents the regression setting with the extended regression space, accommodating heterogeneity and dependence, and outlines the theoretical results for infeasible and feasible t-statistics in the case of fixed parameters. Section 3 extends the analysis to time-varying regression parameters. Section 4 addresses regression modeling with missing data patterns. Section 5 illustrates the flexibility of our robust estimation method through its application to the estimation of an AR(p) model generated by a stationary m.d. noise. Section 6 presents Monte Carlo simulation results. In Section 7, we provide an empirical application of the robust regression framework to modeling asset returns. Finally, Section 8 concludes. Proofs and additional simulations are provided in the Supplementary Material.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

2. OLS ESTIMATION IN GENERAL REGRESSION SPACE

In this section, we focus on OLS estimation in an environment that permits general heterogeneity in regression modeling. We analyze the model

$$y_t = \beta' z_t + u_t, \tag{1}$$

where β is a p-dimensional parameter vector, $z_t = (z_{1t}, \ldots, z_{pt})'$ is a stochastic regressor, and u_t is an uncorrelated noise term. To include an intercept, the first component can be set as $z_{1t} = 1$. We refer to the collection of $\{z_t, u_t\}$ jointly as "the regression space."

An applied researcher may want to work within a regression space that accommodates a wide range of regressors and regression noises, without being hindered by restrictive technical assumptions. Ideally, such a setting should permit regressors to exhibit non-stationarity and undefined generic structural change, while enabling estimation and inference under weak theoretical constraints that do not require empirical verification.

Our goal is to extend the OLS estimation procedure to a broad regression framework defined by baseline assumptions aligned with empirical research practice. These assumptions cover a wide variety of types of potentially non-stationary regression variables encountered in applied work. The framework achieves a level of generality comparable to that in Giraitis, Li, and Phillips (2024), which addresses testing for the absence of correlation and cross-correlation under general heterogeneity.

We begin with specifying the structure of an uncorrelated regression noise u_t. Suppose that

$$u_t = h_t \varepsilon_t, \tag{2}$$

where $\{\varepsilon_t\}$ is a zero-mean stationary uncorrelated m.d. noise, and $\{h_t\}$ is a deterministic or stochastic scale factor independent of $\{\varepsilon_t\}$. The following assumption formalizes these conditions.

Assumption 2.1. $\{\varepsilon_t\}$ is a stationary m.d. noise with respect to some σ-field filtration $\mathcal{F}_t$, such that

$$\mathbb{E}[\varepsilon_t | \mathcal{F}_{t-1}] = 0, \quad \mathbb{E}\varepsilon_t^8 < \infty, \quad \mathbb{E}\varepsilon_t^2 = 1.$$

The sequence $\{\varepsilon_t\}$ is independent of $\{h_t\}$. Moreover, variable ε_1 has a probability density function $f(x)$ satisfying $f(x) \le c < \infty$ for all $|x| \le x_0$, for some $x_0 > 0$ and $c > 0$.

The information set $\mathcal{F}_t$ is generated by the past history $\mathcal{F}_t = \sigma(\varepsilon_s, s \le t)$ and possibly other variables.

A typical example of m.d. noise in applied work is provided by the ARCH/GARCH family and the class of stochastic volatility processes. The specification (2) therefore allows for conditional heteroskedasticity in u_t.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

We next specify the regressors $z_t = (z_{1t}, \dots, z_{pt})'$ which form the key structural component of our regression space. For $k = 1, \dots, p$, the regressors can be written as

$$z_{kt} = \mu_{kt} + g_{kt}\eta_{kt}, \quad t = 1, \dots, n, \tag{3}$$

where $\eta_t = (\eta_{1t}, \dots, \eta_{pt})'$ is a stationary sequence, $g_t = (g_{1t}, \dots, g_{pt})'$ are deterministic or stochastic scale factors, and $\mu_t = (\mu_{1t}, \dots, \mu_{pt})'$ is a vector of deterministic or stochastic means. We assume that $\{\mu_t, g_t, h_t\}$ are independent of $\{\varepsilon_t, \eta_t\}$. To include an intercept in model (1), we set

$$z_{1t} \equiv 1 = \mu_{1t} + g_{1t}\eta_{1t}, \quad \mu_{1t} = 0, \quad g_{1t} = \eta_{1t} = 1. \tag{4}$$

We further suppose that in (3) $E\eta_{kt} = 0$ except for the intercept (4), where $\eta_{1t} = 1$.

In summary, the admissible regressors $\{z_t\}$ in our setting are obtained by shifting and rescaling an SM stationary process $\{\eta_t\}$ by the mean process μ_t and the scale factor g_t:

$$z_t = \mu_t + I_{gt}\eta_t, \quad I_{gt} = \mathrm{diag}(g_{1t}, \dots, g_{pt}).$$

The underlying stationary sequence $\{\eta_t\}$ is the fundamental component structuring regressors z_t. Estimation of the regression parameter β requires only mild assumptions on $\{\mu_t, g_t\}$, and SM dependence assumption on η_t, satisfied by ARMA and related stationary time-series models. This framework eliminates the need for additional empirical validation.

Definition 2.1. A (univariate) covariance stationary sequence $\{\xi_t\}$ has short-memory (SM) if $\sum_{h=-\infty}^{\infty} |\mathrm{cov}(\xi_h, \xi_0)| < \infty$.

Assumption 2.2. $\eta_t = (\eta_{1t}, \dots, \eta_{pt})'$ is an $\mathcal{F}_{t-1}$-measurable sequence with $E[\eta_{kt}^2] = 1$ and $E[\eta_{kt}^8] < \infty$.

(i) For $k, j = 1, \dots, p$, the sequences $\{\eta_{kt}\}$ and $\{\eta_{jt}\eta_{kt}\}$ are covariance stationary and have SM.

(ii) The matrix $E[\eta_1\eta_1']$ is positive definite.

The novelty of this regression framework lies in the structural specification (3), which accommodates regressors $z_t = (z_{1t}, \dots, z_{pt})'$ that may be deterministic or stochastic, and stationary or non-stationary. This flexibility arises from allowing a broad class of scale factors and mean processes $\{h_t, g_t, \mu_t\}$ which brings the OLS estimation closer to empirical practice.

The estimation framework also accommodates triangular arrays of means and scale factors: $\big(\mu_t, g_t, h_t,\ t = 1, \dots, n\big) = \big(\mu_{nt}, g_{nt}, h_{nt},\ t = 1, \dots, n\big)$. Throughout the article, we assume that these quantities may depend on the sample size n. For brevity of notation, the subscript n is omitted.

The underlying stationary noise component η_t in the regressors z_t in (3) is weakly exogenous with respect to the stationary m.d. noise ε_t in $u_t = h_t\varepsilon_t$. The mean and scale factors $\{\mu_t, g_t\}$ are independent of $\{\varepsilon_t\}$, though they may be dependent on $\{h_t\}$.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

Overall, $\{\mu_t, g_t, h_t\}$ are mutually independent of $\{\eta_t, \varepsilon_t\}$, while potential dependence among $\{\mu_t\}$, $\{g_t\}$, and $\{h_t\}$, and their components, is unrestricted.

The processes μ_{kt} and g_{kt} can be interpreted as conditional mean and variance, $\mu_{kt} = E[z_{kt}|\mathcal{F}_n^*]$, and $g_{kt}^2 = \text{var}(z_{kt}|\mathcal{F}_n^*)$ of z_{kt}, where $\mathcal{F}_n^* = \sigma\big(\mu_t, g_t, h_t, t = 1, \ldots, n\big)$ denotes the information set generated by the means and scale factors.

Denote for $k = 1, \ldots, p$,

$$v_k^2 = \sum_{t=1}^n g_{kt}^2 h_t^2, \quad v_{gk}^2 = \sum_{t=1}^n g_{kt}^2, \tag{5}$$

$$D = \text{diag}(v_1, \ldots, v_p), \quad D_g = \text{diag}(v_{g1}, \ldots, v_{gp}).$$

We write $a_n \asymp_p b_n$ if $a_n = O_p(b_n)$ and $b_n = O_p(a_n)$.

Assumption 2.3. The scale factors $h_t \geq 0$ and $g_t \geq 0$ are deterministic or stochastic non-negative variables such that, for $k = 1, \ldots, p$,

$$\frac{\max_{1\leq t\leq n} g_{kt}^2}{v_{gk}^2} = o_p(1), \quad \frac{\max_{1\leq t\leq n} \mu_{kt}^2}{v_{gk}^2} = o_p(1), \tag{6}$$

$$\frac{\sum_{t=1}^n \mu_{kt}^2}{v_{gk}^2} = O_p(1), \quad \frac{\sum_{t=1}^n \mu_{kt}^2 h_t^2}{v_k^2} = O_p(1), \quad v_k^2 \asymp_p v_{gk}^2, \quad v_k \to_p \infty. \tag{7}$$

Assumptions (6) and (7) impose only mild restrictions on the means μ_t and scale factors g_t. In particular, condition (6) resembles the Lindeberg condition in the classical Lindeberg–Feller central limit theorem, as it excludes the possibility that the OLS estimation is dominated by a single extreme observation of z_t.

The first restriction on g_{kt} in (6) is necessary. For example, consider the regressor $z_t = g_t\eta_t, t = 1, \ldots, n$, with scale factors $g_1 = 1$ and $g_2 = g_3 = \cdots = g_n = 0$, so that $z_2 = z_3 = \cdots = z_n = 0$. In this case, the OLS estimator of β is inconsistent, and such a scale factor g_t does not satisfy (6).

The second condition (7) ensures that OLS estimation is driven by the stochastic component $g_t\eta_t$ of the regressor z_t, rather than by a deterministic or stochastic drift in μ_t. In the presence of an intercept, condition (7) further implies that $\sum_{t=1}^n h_t^2 \asymp_p n$, since $v_1^2 \asymp_p v_{g1}^2$, $g_{1t} = 1$, $v_{g1}^2 = n$, and $v_1^2 = \sum_{t=1}^n h_t^2$.

To estimate $\beta = (\beta_1, \ldots, \beta_p)'$, we use the standard OLS estimator

$$\widehat{\beta} = \left(\sum_{j=1}^n z_j z_j'\right)^{-1} \left(\sum_{j=1}^n z_j y_j\right) \tag{8}$$

computed from the sample $y_j, z_j, j = 1, \ldots, n$.

Consistency: We first establish the consistency of the OLS estimator $\widehat{\beta}$.

THEOREM 2.1. *Suppose that $(y_1, \ldots, y_n)$ is a sample from the regression model (1) and Assumptions 2.1–2.3 are satisfied. Then, the OLS estimator $\widehat{\beta}$ is consistent, that is,*

$$D(\widehat{\beta} - \beta) = \big(v_1(\widehat{\beta}_1 - \beta_1), \ldots, v_p(\widehat{\beta}_p - \beta_p)\big)' = O_p(1). \tag{9}$$

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

This result implies that the k-th component $\widehat{\beta}_k$ of the OLS estimator is v_k-consistent, that is, $\widehat{\beta}_k - \beta_k = O_p(v_k^{-1})$. The convergence rate v_k may deviate from the conventional $\sqrt{n}$ rate and may differ across components. From the definition of v_k and v_{gk}, it follows that

$$\text{if } g_{kt}, h_t \geq c > 0 \text{ for all } t, n, \text{ then } v_k, v_{gk} \geq c\sqrt{n}. \tag{10}$$

Asymptotic Normality: The asymptotic normality of an element $\widehat{\beta}_k$ of the OLS estimator, as well as the computation of its standard errors, requires additional assumptions on the scale factors and the stationary processes $\{\eta_t, \varepsilon_t\}$.

Assumption 2.4. (i) For $k, j = 1, \ldots, p$, the sequences $\{\varepsilon_t^2\}$, $\{\eta_{jt}\varepsilon_t^2\}$, and $\{\eta_{jt}\eta_{kt}\varepsilon_t^2\}$ are covariance stationary and have SM. (ii) For $k = 1, \ldots, p$,

$$\frac{\max_{1\leq t\leq n} g_{kt}^2 h_t^2}{v_k^2} = o_p(1), \quad \frac{\max_{1\leq t\leq n} \mu_{kt}^2 h_t^2}{v_k^2} = o_p(1). \tag{11}$$

Assumption 2.4(i) is not required when ε_t is i.i.d. Together, Assumptions 2.3 and 2.4(ii) exclude cases in which the mean process μ_t or a few extreme observations of z_t or u_t dominate the estimation of the regression parameter. Overall, these assumptions are mild. They accommodate both deterministic and stochastic means μ_t and scale factors h_t, g_t that may change abruptly over time unlike other theoretically rigorous treatments, which restrict structural change to be deterministic and smooth. This flexibility makes the framework particularly suitable for modeling financial data, as it allows for volatility jumps, commonly observed in empirical finance (see, e.g., Eraker, Johannes, and Polson, 2003). In modern macroeconomic VAR models, the scale factor h_t in the uncorrelated noise representation $u_t = h_t\varepsilon_t$ is typically assumed to be stochastic (see, e.g., Chan, Koop, and Xuewen, 2024; Carriero et al., 2024), which our framework naturally encompasses.

Lemma 2.1 below shows that Assumptions 2.3 and 2.4(ii) hold for regressors z_t and noises u_t with bounded $4+\delta$ moments satisfying (10). The following example provides additional sufficient conditions.

Example 2.1. Assumptions 2.3 and 2.4(ii) are satisfied by regressors z_t and noise u_t whose scale factors h_t, g_t and means μ_t satisfy $0 < c \leq h_t, g_{kt} \leq C$, $||\mu_t|| \leq C$, where $0 < c, C < \infty$ do not depend on t, n or $k = 1, \ldots, p$.

When $0 < c \leq h_t \leq C$, $||\mu_t|| \leq C$ for all t, n, Assumptions 2.3 and 2.4(ii) hold for scale factors g_{kt} satisfying

$$\frac{\min_{t=1,\ldots,n} g_{kt}^2}{\sum_{t=1}^n g_{kt}^2} = o_p(1), \quad k = 1, \ldots, p.$$

This condition is, for example, met when g_{kt} follows a unit root process defined by $g_{kt} = \sum_{j=1}^t \xi_j$, where $\{\xi_j\}$ is a sequence of i.i.d. $(0, \sigma^2)$ random variables with finite moments of order $\theta > 2$. The idea of modeling parameters as unit root processes was discussed, for example, in Nyblom (1989).

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

We now describe infeasible standard errors $\sqrt{\omega_{kk}}$ using the notation:

$$S_{zz} = \sum_{t=1}^{n} z_t z_t', \quad S_{zzuu} = \sum_{t=1}^{n} z_t z_t' u_t^2,$$
$$\Omega_n = E[S_{zz}\,|\mathcal{F}_n^*]^{-1} E[S_{zzuu}\,|\mathcal{F}_n^*] E[S_{zz}\,|\mathcal{F}_n^*]^{-1} = (\omega_{jk}), \tag{12}$$

where ω_{jk} denotes the (j,k)-th element of the matrix Ω_n. The infeasible standard error of $\widehat{\beta}_k$ is defined as $\sqrt{\omega_{kk}}$, that is, the square root of the corresponding diagonal element of Ω_n.

The generality of our regression setting limits the multivariate asymptotic theory that can be established for $\widehat{\beta}$. While a full joint distribution of $\widehat{\beta}$ is not available, we can derive asymptotic normality for linear combinations $a'\widehat{\beta}$ and then construct feasible inference for the individual component β_k.

Existing literature typically imposes stronger assumptions on regressors and errors, such as mixing regressors (White, 2014, Thm. 3.78), locally stationary regressors in Zhang and Wu (2012, Eqn. (2.3)), or near-epoch-dependent errors in (Hall et al., 2012, Assumption 8).

THEOREM 2.2. *Suppose that the assumptions of Theorem 2.1 and Assumption 2.4 hold. Then, for any $a = (a_1, \ldots, a_p)' \neq 0$, the OLS estimator $\widehat{\beta}$ satisfies*

$$\frac{a'D(\widehat{\beta} - \beta)}{\sqrt{a'D\Omega_n Da}} \to_d \mathcal{N}(0,1). \tag{13}$$

In particular, for $k = 1, \ldots, p$, the t-statistic for β_k satisfies

$$\frac{\widehat{\beta}_k - \beta_k}{\sqrt{\omega_{kk}}} \to_d \mathcal{N}(0,1). \tag{14}$$

Property (13) is difficult to implement in practice because it requires estimation of the unknown matrices D, Ω_n, except in the special case $a' = (0, \ldots, 1, \ldots, 0)$ with only the k-th element being nonzero. In this case, (13) reduces to (14), and the infeasible standard error $\sqrt{\omega_{kk}}$ can be consistently estimated by

$$\widehat{\Omega}_n = S_{zz}^{-1} S_{zz\widehat{u}\widehat{u}} S_{zz}^{-1} = (\widehat{\omega}_{jk}), \quad \widehat{u}_t = y_t - \widehat{\beta}' z_t. \tag{15}$$

The feasible standard error $\sqrt{\widehat{\omega}_{kk}}$ is the square root of the diagonal element $\widehat{\omega}_{kk}$ of $\widehat{\Omega}_n$.

COROLLARY 2.1. *Under the assumptions of Theorem 2.2, for $k = 1, \ldots, p$, as $n \to \infty$,*

$$\frac{\widehat{\beta}_k - \beta_k}{\sqrt{\widehat{\omega}_{kk}}} \to_d \mathcal{N}(0,1), \quad \frac{\widehat{\omega}_{kk}}{\omega_{kk}} = 1 + o_p(1), \quad \sqrt{\omega_{kk}} \asymp_p v_k^{-1}. \tag{16}$$

This result is the main contribution of Section 2. It enables straightforward computation of standard errors and the construction of confidence intervals for β_k

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

in the extended regression framework. Notably, the estimator $\widehat{\Omega}_n$ coincides with the heteroskedasticity-consistent standard error estimator of White (1980).

Remark 2.1. The consistency rate $v_k = (\sum_{t=1}^n g_{kt}^2 h_t^2)^{1/2}$ for the parameter β_k may take the form $v_k \sim cn^\alpha$ for any $\alpha > 0$, ranging from a super-slow ($0 < \alpha < 1$) to a super-fast ($\alpha > 1$) convergence. To illustrate this, consider the regression model

$$y_t = \beta_1 + \beta_2 z_{2t} + \beta_3 z_{3t} + u_t, \quad u_t = h_t \varepsilon_t \text{ with } h_t = 1,$$
$$z_{kt} = g_{kt}\eta_{kt}, \quad g_{kt} = t^{(\alpha_k - 1)/2} \text{ for } k = 2, 3,$$

where $\alpha_2 > 1$, $0 < \alpha_3 < 1$, and $\{\eta_{2t}\}$, $\{\eta_{3t}\}$, $\{\varepsilon_t\}$ are i.i.d. $\mathcal{N}(0,1)$. Then $v_1 = \sqrt{n}$ and $v_k \sim \alpha_k^{-1/2} n^{\alpha_k/2}$ for $k = 2, 3$, producing different convergence rates across the parameters. Even in this simple case, the usual multivariate asymptotic normality for $\sqrt{n}(\widehat{\beta} - \beta)$ does not hold.

Corollary 2.1 allows us to establish the asymptotic power and consistency of the test for testing the hypothesis

$$H_0 : \beta_k = \beta_k^0, \quad \text{vs.} \quad H_1 : \beta_k \neq \beta_k^0,$$

that is, whether the k-th element of the regression parameter $\beta = (\beta_1, \ldots, \beta_p)'$ is equal to a specific value β_k^0.

COROLLARY 2.2. *Suppose that $\beta_k^0 \neq \beta_k$. Then, under the assumptions of Corollary 2.1,*

$$t = \frac{\widehat{\beta}_k - \beta_k^0}{\sqrt{\widehat{\omega}_{kk}}} \asymp_p v_k \to_p \infty. \tag{17}$$

We conclude this section with a lemma that provides simple sufficient moment-type conditions for the validity of Assumptions 2.3 and 2.4(ii). In particular, condition (10) implies (19).

LEMMA 2.1. *Suppose that for $k = 1, \ldots, p$,*

$$Ez_{kt}^4 \leq c, \quad E|u_t|^{4+\delta} \leq c \text{ for some } \delta > 0, \tag{18}$$

$$n/v_k^2 = O_p(1), \quad n/v_{gk}^2 = O_p(1), \tag{19}$$

where $c < \infty$ does not depend on t, n. Then Assumptions 2.3 and 2.4(ii) hold.

In particular, (19) is satisfied if $\max_{t=1,\ldots,n} h_t^{-1} = O_p(1)$ and $\max_{t=1,\ldots,n} g_{kt}^{-1} = O_p(1)$.

The regular estimator of standard errors in OLS regression estimation is given by

$$\widehat{\Omega}_n^{(st)} = S_{zz}^{-1} \widehat{\sigma}_u^2, \quad \widehat{\sigma}_u^2 = n^{-1} \sum_{j=1}^n \widehat{u}_j^2. \tag{20}$$

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

Unlike the robust standard errors $\sqrt{\widehat{\omega}_{kk}}$, these conventional standard errors may produce coverage distortions, particularly when heteroskedasticity or heterogeneity in g_t, h_t, or μ_t is present (see Section 6). This underscores the robustness and strong empirical relevance of the normal approximation in (16).

This section validates the asymptotic normality of the feasible t-statistics for the components of the OLS estimator in linear regression models with general heterogeneity. The assumptions imposed are mild yet flexible, allowing a wide class of (possibly nonstationary) regressors and noise processes beyond those typically considered in the existing literature. Some conditions on scale factors are analogous to the Lindeberg condition and remain necessary. Our framework complements, rather than replaces, prior approaches; for instance, near-unit-root regressors in Georgiev et al. (2018) require a distinct theoretical treatment. Although bootstrap methods, see, for example, Hall et al. (2012) and Boldea et al. (2019), are widely applied in regression analysis, they may not extend to the heterogeneous structures considered here. By contrast, we demonstrate that the heteroskedasticity-consistent standard errors of White (1980) remain applicable and computationally straightforward.

Our focus in this article is model (1), where the regression noise u_t in (1) is uncorrelated. Extending the asymptotic theory to account for dependence in u_t is a natural next step and is currently under consideration.

Proofs of all results are provided in the Supplementary Material.

3. TIME-VARYING OLS ESTIMATION IN THE EXTENDED REGRESSION SPACE

This section demonstrates further advantages of the theory of regression estimation with a fixed parameter, developed in Section 2. Thanks to the flexible setting, the estimation of time-varying parameters naturally follows from our theory for fixed-parameter regression in the extended space, along with bounding of some negligible terms.

In the previous section, we discussed the estimation of regression model (1), $y_j = \beta' z_j + u_j$, with a fixed parameter β. We now extend the model by allowing the regression parameter to vary over time. Specifically, we consider the model

$$y_j = \beta_j' z_j + u_j, \quad j = 1, \ldots, n, \tag{21}$$

where the regressors z_j and the regression noise u_j, as defined in (3) and (2), remain unchanged. That is, they belong to the same regression space as in Section 2.

The primary objective is to develop a point-wise estimation procedure for the path $\beta_1, \ldots, \beta_n$ of the time-varying parameter β_j in model (21), while preserving the same regression space introduced in Section 2.

The literature on the estimation of time-varying regression parameters β_j is extensive. It primarily focuses on estimation and testing for parameter stability under relatively strong assumptions on the regressors and regression noise. For instance, regressors are assumed to be locally stationary in Vogt (2012, Model

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

(3)), stationary and strongly mixing in Fu et al. (2023, Assumption A.1), and strictly stationary in Hu et al. (2025, Assumption P(d)). It is clear that the class of regressors considered in our setting is broader, and they may be neither mixing nor stationary.

The objective of this section is to describe the extended regression space of regressors z_t and disturbances u_t that ensures the asymptotic normality of the feasible t-statistic for estimating the components of the time-varying parameter β_t. We show that, as long as the regressors and the disturbance follow the structure $z_t = \mu_t + I_{gt}\eta_t$ and $u_t = h_t\varepsilon_t$, the class of admissible means μ_t and scale factors g_t, h_t is very broad and characterized by weak restrictions that may not require empirical verification.

Further extensions of the regression space are possible. For example, the weakly exogenous component η_t of regressors z_t in our article is assumed to be an SM process. In contrast, Hu et al. (2025) demonstrate that estimation of the time-varying parameter β_t also permits weakly exogenous, strictly stationary regressors z_t that exhibit long-memory behavior.

While most assumptions on the regressors z_j and regression noise u_j remain unchanged from Section 2, the estimator requires some modifications. Under an additional smoothness assumption on $\{\beta_j\}$, the time-varying OLS estimator $\widehat{\beta}_t$ of parameter β_t at time t is the standard OLS estimator for a fixed regression parameter, obtained by regressing $\widetilde{y}_j = b_{n,tj}^{1/2} y_j$ on $\widetilde{z}_j = b_{n,tj}^{1/2} z_j$:

$$\widehat{\beta}_t = \left(\sum_{j=1}^{n} \widetilde{z}_j \widetilde{z}_j'\right)^{-1} \left(\sum_{j=1}^{n} \widetilde{z}_j \widetilde{y}_j\right) = \left(\sum_{j=1}^{n} b_{n,tj} z_j z_j'\right)^{-1} \left(\sum_{j=1}^{n} b_{n,tj} z_j y_j\right). \tag{22}$$

The weights $b_{n,tj}$ are generated as follows:

$$b_{n,tj} = K\left(\frac{|t-j|}{H}\right),\ t,j = 1,\dots,n, \tag{23}$$

where $H = H_n$ is a bandwidth parameter such that $H \to \infty$ and $H = o(n)$. The kernel function K is bounded and there exist $a_0, \delta > 0$ and $\theta > 3$ such that

$$\begin{aligned} &K(x) \geq a_0 > 0,\ 0 \leq x \leq \delta, \\ &K(x) \leq Cx^{-\theta},\ x > \delta. \end{aligned} \tag{24}$$

For example, (24) is satisfied by functions $K(x) = I(x \in [0,1])$ and $K(x) = p(x)$, where $p(x)$ is the probability density function of the standard normal distribution.

We impose a smoothness assumption on the time-varying parameter β_j, which may be either deterministic or stochastic, while allowing unrestricted dependence between the parameter $\{\beta_j\}$ and the regressors and regression noise $\{z_j, u_j\}$.

Assumption 3.1. For some $\gamma \in (0,1]$ and for $t,j = 1,\dots,n$,

$$E||\beta_t - \beta_j||^2 \leq c\left(\frac{|t-j|}{n}\right)^{2\gamma}, \tag{25}$$

where $c < \infty$ does not depend on $t, j,$ or n.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

Next, we briefly outline how our asymptotic theory for the time-varying robust estimator builds on the results from Section 2 on fixed-parameter regression estimation and the smoothness assumption (25). To demonstrate this, we introduce the following regression model with a fixed parameter $\beta = \beta_t$:

$$y_j^* = \beta' \widetilde{z}_j + \widetilde{u}_j,\ \widetilde{u}_j = b_{n,tj}^{1/2} u_j, \quad j = 1, \dots, n. \tag{26}$$

Notice that the OLS estimator $\widehat{\beta}$ of the fixed parameter β satisfies:

$$\widehat{\beta} = \left(\sum_{j=1}^{n} \widetilde{z}_j \widetilde{z}_j' \right)^{-1} \left(\sum_{j=1}^{n} \widetilde{z}_j y_j^* \right) = \beta + \left(\sum_{j=1}^{n} \widetilde{z}_j \widetilde{z}_j' \right)^{-1} \left(\sum_{j=1}^{n} \widetilde{z}_j \widetilde{u}_j \right). \tag{27}$$

Since $\widetilde{y}_j = y_j^* + (\beta_j - \beta_t)' \widetilde{z}_j$, the time-varying estimator $\widehat{\beta}_t$ given in (22) satisfies:

$$\begin{aligned} \widehat{\beta}_t - \beta_t &= \left(\sum_{j=1}^{n} \widetilde{z}_j \widetilde{z}_j' \right)^{-1} \left(\sum_{j=1}^{n} \widetilde{z}_j \{ y_j^* + (\beta_j - \beta_t)' \widetilde{z}_j \} \right) - \beta_t \\ &= \widehat{\beta} - \beta + R_t, \quad R_t = \left(\sum_{j=1}^{n} \widetilde{z}_j \widetilde{z}_j' \right)^{-1} \left(\sum_{j=1}^{n} \widetilde{z}_j \widetilde{z}_j' (\beta_j - \beta_t) \right). \end{aligned} \tag{28}$$

Notice that $\widehat{\beta} - \beta$ in (27) and (28) does not depend on β_t. Additionally, the regression space in the estimation of the fixed parameter in Section 2 permits rescaling, so premultiplying by the kernel weights $b_{n,tj}^{1/2}$ does not change the structure of the regressors $\widetilde{z}_j = (\widetilde{z}_{1j}, \dots, \widetilde{z}_{pj})'$ and $\widetilde{u}_j$: they still satisfy the settings (3) and (2). Consequently, the model (26) is covered by the regression model (1) with a fixed parameter, and the asymptotic results for $\widehat{\beta} - \beta$ follow from Section 2. The main technical task in this section is to show that the remainder term R_t in (28) is negligible, which follows from the smoothness assumption (25).

The regressors z_j and regression noise u_j belong to the same regression space as defined in Section 2. While the assumptions on the stationary process $\{\eta_j\}$ and the m.d. noise $\{\varepsilon_j\}$ remain unchanged, for simplicity, we replace the previous conditions on the scale factors g_j, h_j and the means μ_j with simple sufficient assumptions similar to those used in Lemma 2.1. As before, the scale factors $\{h_j, g_j\}$ and $\{\mu_j\}$ can be deterministic or stochastic, may vary with n, and are independent of $\{\eta_j, \varepsilon_j\}$.

Denote

$$v_{kt}^2 = \sum_{j=1}^{n} b_{n,tj}^2 g_{kj}^2 h_j^2, \quad v_{gk,t}^2 = \sum_{j=1}^{n} b_{n,tj}^2 g_{kj}^2, \quad k = 1, \dots, p.$$

Assumption 3.2. z_t and u_t are such that, for $k = 1, \dots, p$,

$$E z_{kt}^4 \le c, \quad E|u_t|^{4+\delta} \le c \text{ for some } \delta > 0, \tag{29}$$

$$H / v_{kt}^2 = O_p(1), \qquad H / v_{gk,t}^2 = O_p(1), \tag{30}$$

where $c < \infty$ does not depend on t, n.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

It is straightforward to show that (30) is valid if $g_{kt}, h_t \geq c > 0$ for all t, n.

To describe the infeasible standard errors $\sqrt{\omega_{kk,t}}$, we use

$$S_{zz,t} = \sum_{j=1}^{n} b_{n,tj} z_j z_j', \quad S_{zzuu,t} = \sum_{j=1}^{n} b_{n,tj}^2 z_j z_j' u_j^2,$$

$$\Omega_{nt} = E[S_{zz,t}|\mathcal{F}_n^*]^{-1} E[S_{zzuu,t}|\mathcal{F}_n^*] E[S_{zz,t}|\mathcal{F}_n^*]^{-1} = (\omega_{jk,t}),$$

where $\omega_{jk,t}$ denotes the (j,k)-th element of the matrix Ω_{nt}. The infeasible standard error $\sqrt{\omega_{kk,t}}$ is defined by the diagonal element $\omega_{kk,t}$ of the matrix Ω_{nt}.

The next theorem establishes the consistency rate and asymptotic normality property for the components of the time-varying OLS estimator $\widehat{\beta}_t = (\widehat{\beta}_{1t}, \ldots, \widehat{\beta}_{pt})'$, and allows for arrays of integers $t = t_n \in [1, \ldots, n]$, which may depend on n.

THEOREM 3.1. *Suppose that* $(y_1, \ldots, y_n)$ *is a sample from the regression model (21). Assume that Assumptions 2.1, 2.2, 2.4(i), 3.1, and 3.2 hold. Then, for* $1 \leq t = t_n \leq n$ *and* $k = 1, \ldots, p$,

$$\widehat{\beta}_{kt} - \beta_{kt} = O_p\big(H^{-1/2} + (H/n)^{\gamma}\big), \tag{31}$$

$$\frac{\widehat{\beta}_{kt} - \beta_{kt}}{\sqrt{\omega_{kk,t}}} \to_d \mathcal{N}(0,1) \quad \text{if } H = o(n^{2\gamma/(2\gamma+1)}), \tag{32}$$

and $\sqrt{\omega_{kk,t}} \asymp_p H^{-1/2}$.

The consistency rate in (31) is determined by the bandwidth parameter H and the smoothness parameter $\gamma \in (0,1)$ in (25). The condition $H = o(n^{2\gamma/(2\gamma+1)})$ ensures that in (32) the bias term remains negligible.

As in the fixed-parameter case, for (z_j, u_j) from the extended regression space, the asymptotic normality can be established in point-wise estimation of each individual component $\widehat{\beta}_{kt}$ of $\widehat{\beta}_t$.

The unknown standard error $\sqrt{\omega_{kk,t}}$ can be consistently estimated by

$$\widehat{\Omega}_{nt} = S_{zz,t}^{-1} S_{zz\widehat{u}\widehat{u},t} S_{zz,t}^{-1} = (\widehat{\omega}_{jk,t}), \quad \widehat{u}_j = y_j - \widehat{\beta}_j' z_j. \tag{33}$$

The feasible standard error $\sqrt{\widehat{\omega}_{kk,t}}$ is defined by the diagonal element $\widehat{\omega}_{kk,t}$ of $\widehat{\Omega}_{nt}$.

COROLLARY 3.1. *Under the assumptions of Theorem 3.1, for* $k = 1, \ldots, p$, *and* $H = o(n^{2\gamma/(2\gamma+1)})$, *it holds*

$$\frac{\widehat{\beta}_{kt} - \beta_{kt}}{\sqrt{\widehat{\omega}_{kk,t}}} \to_d \mathcal{N}(0,1), \quad \frac{\widehat{\omega}_{kk,t}}{\omega_{kk,t}} = 1 + o_p(1). \tag{34}$$

Corollary 3.1 allows us to establish the asymptotic power of the test of the hypothesis

$$H_0 : \beta_{kt} = \beta_{kt}^0, \quad \text{vs.} \quad H_1 : \beta_{kt} \neq \beta_{kt}^0,$$

based on the t-statistics $(\widehat{\beta}_{kt} - \beta_{kt}^0)/\sqrt{\widehat{\omega}_{kk,t}}$.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

COROLLARY 3.2. *Suppose that* $|\beta_{kt}^0 - \beta_{kt}| \geq a > 0$ *for* $t = t_n \in [1, \ldots, n]$ *as* $n \to \infty$. *Then, under the assumption of Corollary 3.1,*

$$\frac{\widehat{\beta}_{kt} - \beta_{kt}^0}{\sqrt{\widehat{\omega}_{kk,t}}} \asymp_p H^{1/2} \to \infty. \tag{35}$$

The estimator $\widehat{\Omega}_{nt}$ used to obtain the robust standard errors in (33) is a time-varying version of the heteroskedasticity-consistent estimator of the standard errors of White (1980). Simulation results confirm that it does not produce coverage distortions in the estimation of β_t under the settings considered in this section.

Finally, we provide examples of smoothly varying deterministic and stochastic parameters β_t that satisfy Assumption 3.1.

Example 3.1. A standard example of a deterministic time-varying parameter β_t which satisfies Assumption 3.1 is $\beta_t = \beta_{t,n} = g(t/n)$, $t = 1, \ldots, n$, where $g(\cdot)$ is a deterministic smooth function that has property $|g(x) - g(y)| \leq C|x - y|$. Such β_t satisfies (25) with $\gamma = 1$.

A standard example of a stochastic smooth parameter β_t is a rescaled random walk $\beta_t = \beta_{t,n} = n^{-1/2} \sum_{j=1}^{t} e_j$, $t = 1, \ldots, n$, where $\{e_j\}$ is an i.i.d. sequence with $E[e_j] = 0$ and $E[e_j^2] < \infty$. It satisfies (25) with $\gamma = 1/2$, that is, for $t > s$,

$$E(\beta_t - \beta_s)^2 = n^{-1} E \left(\sum_{j=s+1}^{t} e_j \right)^2 \leq C(t-s)/n.$$

The above results are equipped with thorough and mathematically rigorous proofs, which can be found in the Supplementary Material.

The key new features in the estimation of the time-varying parameter β_t are similar to those highlighted in the estimation of the fixed parameter in Section 2. Although the computation is straightforward, establishing the validity of the robust standard errors $\sqrt{\widehat{\omega}_{kk,t}}$ in the extended regression space of (z_t, u_t) is challenging because the scale factors h_t, g_t, μ_t in model (21) are unknown and potentially random, and highly general, while the asymptotic behavior of the $\omega_{kk,t}$ may not be well-defined. The asymptotic normality of a single component of the estimator can still be established, even though a full multivariate asymptotic theory is not available. Unlike most existing literature, β_t is permitted to evolve as a smoothly varying stochastic process.

4. REGRESSION WITH MISSING DATA

In the previous sections, we showed that the extended regression space enables the estimation of both fixed and time-varying regression parameters. It offers several theoretical advantages, in particular, the ability to estimate regression models in the presence of missing data. Given the importance in empirical regression

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

analysis in situations where some observations y_t or regressors z_t are missing, see, for example, Enders (2022), we now present new and somewhat unexpected results on regression estimation with missing data. We show that the foundational assumptions underlying the construction of regression space also allow us to accommodate a broad range of missing data patterns.

In this section, we suppose that instead of the full sample $(y_1, z_1), \ldots, (y_n, z_n)$, we observe a subsample

$$(y_{k_1}, z_{k_1}), \ldots, (y_{k_N}, z_{k_N}), \quad N \le n, \tag{36}$$

of the dependent variable y_t and the regressor z_t. Our primary interest is to estimate both fixed and time-varying regression parameters using the subsample (36).

To that end, we represent the observed data as a partially observed sample

$$(\widetilde{y}_j, \widetilde{z}_j) = (\tau_j y_j, \tau_j z_j), \quad j = 1, \ldots, n, \tag{37}$$

where τ_j is the missing-data indicator. In (36), it is defined as

$$\tau_j = \begin{cases} 1 & \text{for} \quad j = k_1, k_2, \ldots, k_N, \quad \text{where} \quad k_1 < k_2 < \cdots < k_N \le n, \\ 0 & \text{otherwise.} \end{cases} \tag{38}$$

We set $\tau_j = 1$ if both y_j and z_j are observed; otherwise, $\tau_j = 0$. Throughout this section, τ_j is treated as a sequence of random or deterministic variables, allowing for regularly missing, block-wise missing, or randomly missing data patterns.

In order for the theoretical results of the previous section to apply, we impose the following assumptions on the missing data indicator τ_t, the regressors $z_{kt} = \mu_{kt} + g_{kt}\eta_{kt}$ in (3), and the regression noise $u_t = h_t\varepsilon_t$ in (2).

Assumption 4.1. The missing-data indicator $\{\tau_t\}$ is assumed to be independent of $\{\varepsilon_t, \eta_t\}$ in (2) and (3).

Assumption 4.2. (i) $Ez_{kt}^4 \le c$ and $E|u_t|^{4+\delta} \le c$ for some $\delta > 0$, where $c > 0$ does not depend on k, t, n.

(ii) $g_{kt} \ge c > 0$ and $h_t \ge c > 0$, where c does not depend on k, t, n.

(iii) ε_t, η_t satisfy Assumptions 2.1, 2.2, and 2.4(i).

Estimation of Fixed Parameters: Suppose that $y_t = \beta' z_t + u_t$ follows the regression model (1) with a fixed parameter β as in Section 2. Our primary interest is to estimate the parameter β using the subsample (36). In view of (1), we can write the partially observed regression model as

$$\begin{aligned} &\widetilde{y}_t = \tau_t y_t = \tau_t(\beta' z_t + u_t), \\ &\widetilde{y}_t = \beta' \widetilde{z}_t + \widetilde{u}_t, \quad \widetilde{u}_t = \tau_t u_t = \{\tau_t h_t\}\varepsilon_t. \end{aligned} \tag{39}$$

In (39), the regressors $\widetilde{z}_t$ and the noise $\widetilde{u}_t$ can be represented as

$$\begin{aligned} &\widetilde{z}_{kt} = \widetilde{\mu}_{kt} + \widetilde{g}_{kt}\eta_{kt}, \quad \widetilde{\mu}_{kt} = \tau_t \mu_{kt}, \quad \widetilde{g}_{kt} = \tau_t g_{kt}, \\ &\widetilde{u}_t = \widetilde{h}_t \varepsilon_t, \quad \widetilde{h}_t = \tau_t h_t. \end{aligned} \tag{40}$$

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

They belong to the regression space described in (2) and (3). Therefore, parameter β and the correspondent standard errors in model (39) can be estimated using the OLS estimator $\widehat{\beta}$ and $\widehat{\omega}_{kk}$:

$$\widehat{\beta} = \left(\sum_{t=1}^{n} \widetilde{z}_t \widetilde{z}_t'\right)^{-1} \left(\sum_{t=1}^{n} \widetilde{z}_t \widetilde{y}_t\right), \quad \widehat{\Omega}_n = S_{\widetilde{z}\widetilde{z}}^{-1} S_{\widetilde{z}\widehat{u}\widehat{u}} S_{\widetilde{z}\widetilde{z}}^{-1} = (\widehat{\omega}_{jk}), \tag{41}$$

$$S_{\widetilde{z}\widetilde{z}} = \sum_{t=1}^{n} \widetilde{z}_t \widetilde{z}_t', \quad S_{\widetilde{z}\widehat{u}\widehat{u}} = \sum_{t=1}^{n} \widetilde{z}_t \widetilde{z}_t' \widehat{u}_t^2, \quad \widehat{u}_t = \widetilde{y}_t - \widehat{\beta}' \widetilde{z}_t.$$

THEOREM 4.1. *The OLS estimator $\widehat{\beta}$ of the parameter β in the regression model (39) with missing data has the following asymptotic properties. If Assumptions 4.1 and 4.2 hold and $n/N = O_p(1)$, then, for $k = 1, \ldots, p$, as $n \to \infty$,*

$$\frac{\widehat{\beta}_k - \beta_k}{\sqrt{\widehat{\omega}_{kk}}} \to_d \mathcal{N}(0,1), \qquad \sqrt{\widehat{\omega}_{kk}} \asymp_p n^{-1/2}. \tag{42}$$

Remark 4.1. Theorem 4.1 shows that ignoring missing data does not affect the estimation of the fixed parameter. That is, the researcher can compute the estimators $\widehat{\beta}$ and $\sqrt{\widehat{\omega}_{kk}}$ directly using subsample $y_{k_j}, z_{k_j}, j = 1, \ldots, N$:

$$\widehat{\beta} = \left(\sum_{j=1}^{N} z_{k_j} z_{k_j}'\right)^{-1} \left(\sum_{j=1}^{N} z_{k_j} y_{k_j}\right), \quad \widehat{\Omega}_n = S_{*,zz}^{-1} S_{*,z\widehat{u}\widehat{u}} S_{*,zz}^{-1} = (\widehat{\omega}_{jk}),$$

$$S_{*,zz} = \sum_{j=1}^{N} z_{k_j} z_{k_j}', \quad S_{*,z\widehat{u}\widehat{u}} = \sum_{j=1}^{N} z_{k_j} z_{k_j}' \widehat{u}_{k_j}^2, \quad \widehat{u}_{k_j} = y_{k_j} - \widehat{\beta}' z_{k_j}.$$

Estimation of Time-Varying Parameters: Assume now that $y_t = \beta_t' z_t + u_t$ follows the regression model (21) with time-varying parameter β_t, where regressors z_t and regression noise u_t are as in (3) and (2). We are interested in estimating the parameter β_t in the presence of missing data using the subsample (36). Similarly to (39), we base the estimation on the partially observed regression model with a time-varying parameter:

$$\widetilde{y}_j = \beta_j' \widetilde{z}_j + \widetilde{u}_j, \quad j = 1, \ldots, n, \tag{43}$$

where regressors $\widetilde{z}_j$ and the noise $\widetilde{u}_j$ are defined as in (40). They belong to the regression space described by (2) and (3) and thus results of Section 3 on the estimation of the time-varying parameter β_j apply.

We show in the following theorem that under Assumptions 4.1 and 4.2, parameter β_t and standard errors can be estimated point-wise at each time $t = 1, \ldots, n$ provided that the missing data pattern satisfies the following condition:

$$H/N_t = O_p(1), \quad \text{where } N_t = \sum_{j=1}^{n} \tau_j b_{n,tj}. \tag{44}$$

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

This condition holds, for example, if $\tau_j = 1$ for $|j-t| \le \epsilon H$ for some $\epsilon > 0$.

The estimator $\widehat{\beta}_t$ and the estimator of the robust standard errors $\widehat{\omega}_{kk,t}$ given in (22) and (33) are defined as

$$\widehat{\beta}_t = \left(\sum_{j=1}^{n} b_{n,tj}\widetilde{z}_j\widetilde{z}_j'\right)^{-1}\left(\sum_{j=1}^{n} b_{n,tj}\widetilde{z}_j\widetilde{y}_j\right), \tag{45}$$

$$\widehat{\Omega}_{nt} = S_{\widetilde{z}\widetilde{z},t}^{-1} S_{\widetilde{z}\widetilde{z}\widehat{u}\widehat{u},t} S_{\widetilde{z}\widetilde{z},t}^{-1} = (\widehat{\omega}_{jk,t}), \quad \widehat{u}_j = \widetilde{y}_j - \widehat{\beta}_j'\widetilde{z}_j.$$

THEOREM 4.2. *The OLS estimator $\widehat{\beta}_t$ of the time-varying parameter β_t in the regression model (43) with missing data has the following properties. Assume that $1 \le t = t_n \le n$, Assumptions 4.1, 3.1, and 4.2 are satisfied and that the condition $H/N_t = O_p(1)$ holds. Then, for $k = 1, \dots, p$, as $n \to \infty$,*

$$\widehat{\beta}_{kt} - \beta_{kt} = O_p\big(H^{-1/2} + (H/n)^{\gamma}\big), \tag{46}$$

$$\frac{\widehat{\beta}_{kt} - \beta_{kt}}{\sqrt{\widehat{\omega}_{kk,t}}} \to_d \mathcal{N}(0,1) \quad \text{if } H = o(n^{2\gamma/(2\gamma+1)}), \tag{47}$$

$$\widehat{\omega}_{kk,t} \asymp_p H^{-1}. \tag{48}$$

5. ESTIMATION OF A STATIONARY AR(*P*) MODEL WITH AN M.D. NOISE

In this section, we focus on another practical application of our regression framework developed in Section 2. We show that it covers the estimation of the parameters of a stationary AR(p) model driven by a stationary m.d. noise ε_t:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \tag{49}$$

where parameters $\phi_0, \dots, \phi_p$ are such that the model (49) has a stationary solution. Xu and Phillips (2008) developed estimation theory for the AR(p) model $y_t = \phi_0 + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + u_t$, when $u_t = h_t\varepsilon_t$ and h_t is a smoothly varying deterministic sequence, and the m.d. sequence ε_t satisfies $E[\varepsilon_t^2|\mathcal{F}_{t-1}] = 1$ a.s. Giraitis, Taniguchi, and Taqqu (2018) were among the first to analyze the distortions of standard errors caused by m.d. noise in the estimation of ARMA models. This article shows that the variance of the parameter vector ϕ converges to a well-defined limit; however, its complex structure complicates the estimation of the limiting variance and the corresponding standard errors in empirical applications. They restricted the estimation of standard errors to AR(1) and MA(1) models. In the case of an AR(p) model, using our method, we are able to estimate standard errors for any p without relying on asymptotic approximations which is the main novelty and contribution of this section. Notice that the model (49) can be written as a special case of the regression model (1):

$$y_t = \beta' z_t + u_t, \quad u_t = \varepsilon_t. \tag{50}$$

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

Here, the parameter $\beta = (\beta_1, \ldots, \beta_{p+1})' = (\phi_0, \ldots, \phi_p)'$ is fixed, and the regressors $z_t = (z_{1t}, z_{2t}, \ldots, z_{p+1,t})' = (1, y_{t-1}, y_{t-2}, \ldots, y_{t-p})'$ are stationary random variables. It is straightforward to verify that the regressors

$$z_{kt} = \mu_{kt} + g_{kt}\eta_{kt}, \quad \mu_{kt} = E[y_{t-k}] = Ey_1, \quad g_{kt} = 1, \quad \eta_{kt} = y_{t-k} - E[y_{t-k}]$$

for $k = 2, \ldots, p+1$ satisfy the regression assumption (3). In the theorem below, we assume that the standard stationarity conditions on parameters of the AR(p) model (49) are satisfied, see for example, Theorem 3.1.1 in Brockwell and Davis (1991), which ensure the existence of a stationary solution

$$y_t = \mu + \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \quad \text{where } \sum_{j=0}^{\infty} |a_j| < \infty, \ \mu = Ey_t. \tag{51}$$

We assume that ε_t satisfies Assumption 2.1 and $\eta_t = (y_{t-1}, y_{t-2}, \ldots, y_{t-p})'$ satisfies Assumptions 2.2 and 2.4(i). These assumptions impose only mild restrictions on the m.d. noise ε_t, and their validity can be verified for typical examples of uncorrelated m.d. noise, such as ARCH-type processes.

The OLS estimator $\widehat{\beta}$ of β in the regression model (50) is defined as in (8) and $\widehat{\omega}_{kk}$ as in (15).

THEOREM 5.1. *Suppose that the AR(p) model (49) with m.d. noise ε_t has a stationary solution as in (51), that $E\varepsilon_t^8 < \infty$ and that (ε_t, η_t) satisfy Assumptions 2.1, 2.2, and 2.4(i). Then the OLS estimator $\widehat{\beta}$ of parameter β in regression model (50) has the following properties: for $k = 1, \ldots, p+1$, as $n \to \infty$,*

$$\frac{\widehat{\beta}_k - \beta_k}{\sqrt{\widehat{\omega}_{kk}}} \to_d \mathcal{N}(0,1), \qquad \sqrt{\widehat{\omega}_{kk}} \asymp_p n^{-1/2}. \tag{52}$$

The Monte Carlo results presented in Section 6.4 demonstrate that the robust OLS (ROLS) estimation produces correct 95% confidence intervals for β_k, whereas the standard OLS method exhibits coverage distortions when the noise ε_t is not i.i.d. This finding indicates that the ROLS estimator has a broader range of applicability than merely addressing heteroskedasticity, and that it can also be effectively used in regression settings not covered by the standard OLS estimation and inference theory.

It is worth noting that the papers by Doukhan and Wintenberger (2008), Bardet and Wintenberger (2009), and Karmakar et al. (2022) provide advanced theoretical results on the modeling and estimation of general nonlinear time-varying time-series models; however, they address the linear AR(p) model (49) only in the trivial case of an i.i.d. noise ε_t.

6. MONTE CARLO SIMULATIONS

In this section, we explore the finite-sample performance of the robust and standard OLS estimation methods in the regression settings, outlined in Sections 2 and 3.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

We examine the impact of time-varying deterministic and stochastic parameters, means, scale factors, and heteroskedasticity of the regression noise on estimation. A comparison of simulation results for standard and robust estimation methods shows that, despite the generality of our regression setting, estimation based on the robust standard errors produces well-sized confidence intervals for fixed and time-varying regression parameters β and β_t, while application of the standard confidence intervals leads to severe distortion of coverage rates.

6.1. Estimation of a Fixed Parameter

We generate arrays of samples of regression models with fixed parameters and an intercept:

$$y_t = \beta_1 + \beta_2 z_{2t} + \beta_3 z_{3t} + u_t, \quad u_t = h_t \varepsilon_t, \quad \beta = (\beta_1, \beta_2, \beta_3)' = (0.5, 0.4, 0.3)'. \tag{53}$$

We set the sample size to $n = 1,500$, conduct $1,000$ replications, and set the nominal coverage probability at 0.95. (Estimation results for $n = 200,800$ are available upon request.) We also include a more complex example in the Supplementary Material.

This model includes three parameters and three regressors. We set $z_{1t} = 1$ and define

$$\begin{aligned} &z_{kt} = \mu_{kt} + g_{kt}\eta_{kt}, \ k = 2, 3, \\ &\mu_{kt} = 0.5\sin(\pi t/n) + 1, \quad \eta_{kt} = 0.5\eta_{k,t-1} + \xi_{kt}, \end{aligned} \tag{54}$$

where $\xi_{2t} = \varepsilon_{t-1}$ and $\xi_{3t} = \varepsilon_{t-2}$. The stationary m.d. noise ε_t in u_t is generated by a GARCH(1, 1) process

$$\varepsilon_t = \sigma_t e_t, \quad \sigma_t^2 = 1 + 0.7\sigma_{t-1}^2 + 0.2\varepsilon_{t-1}^2, \quad e_t \sim i.i.d.\mathcal{N}(0, 1). \tag{55}$$

MODEL 6.1. *y_t follows (53) with deterministic scale factors. We set $h_t = 0.3(t/n)$ and $g_{2t} = g_{3t} = 0.4(t/n)$.*

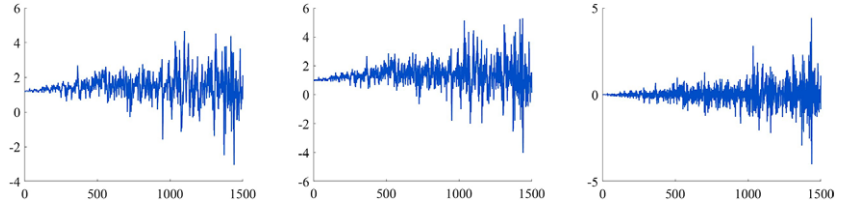
MODEL 6.2. *y_t follows (53) with stochastic scale factors. We set*

$$h_t = \left|\frac{1}{2\sqrt{n}}\sum_{j=1}^{t}\zeta_j\right| + 0.25, \quad g_{2t} = g_{3t} = \left|\frac{1}{2\sqrt{n}}\sum_{j=1}^{t}\nu_{kj}\right| + 0.25.$$

The generating noises $\{\zeta_j, \nu_{2j}, \nu_{3j}\}$ are i.i.d. $\mathcal{N}(0, 1)$ and independent of $\{\varepsilon_t\}$.

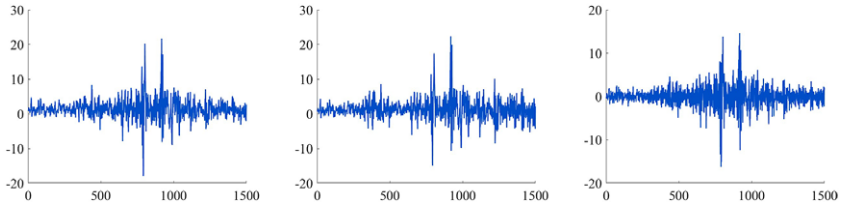
Models 6.1 and 6.2 are regression models with fixed parameters. Examples of plots of the simulated dependent variable, regressors, and regression noise are shown in Figures 1 and 2 (z_{2t} and z_{3t} have similar patterns). To verify the validity of the asymptotic normal approximation of Corollary 2.1 in finite samples, we compute empirical coverage rates (CP) for 95% confidence intervals used in the ROLS estimation, for parameter β. For comparison, we compute the coverage

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press



(a) y_t (b) z_{2t} (c) u_t

FIGURE 1. Plots of y_t, z_{2t}, and u_t in Model 6.1.



(a) y_t (b) z_{2t} (c) u_t

FIGURE 2. Plots of y_t, z_{2t}, and u_t in Model 6.2.

TABLE 1. Robust OLS estimation in Model 6.1

Parameters	Bias	RMSE	CP	CP_{st}	SD
β_1	−0.00570	0.04579	95.0	79.2	0.04544
β_2	0.00206	0.03407	95.4	72.7	0.03401
β_3	0.00204	0.03495	94.0	72.9	0.03489

rates CP_{st} for standard confidence intervals based on the standard errors (20) used in standard OLS estimation. The robust and standard OLS procedures share the same estimator $\widehat{\beta}$, and hence the same bias, root mean square error (RMSE), and standard deviation (SD). Their confidence intervals differ because the variances (and standard errors) in their normal approximations are different.

Table 1 reports estimation results for Model 6.1 which contains deterministic scale factors. It shows that the coverage rate CP for robust confidence intervals is close to the nominal 95%, while the coverage rate CP_{st} of the standard confidence intervals drops below 80%. The Bias, RMSE, and SD are small.

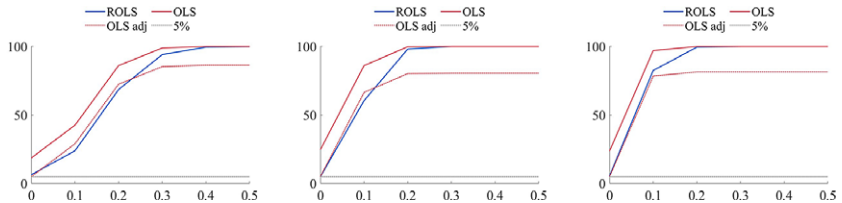
Table 2 presents estimation results for Model 6.2 which includes stochastic scale factors. It shows that the coverage rate CP for robust confidence intervals is close to the nominal 95%, whereas the standard estimation method produces coverage distortions for parameters β_2 and β_3.

To assess power, we vary β_3 in Model 6.1 from 0 to 0.5 and record how often the test rejects $H_0 : \beta_3 = 0$. Figure 3 reports results for ROLS and OLS at sample

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

TABLE 2. Robust OLS estimation in Model 6.2

Parameters	Bias	RMSE	CP	CP_{st}	SD
β_1	−0.00420	0.05117	94.6	92.2	0.05100
β_2	0.00208	0.03205	94.6	87.4	0.03199
β_3	0.00071	0.01542	94.8	85.3	0.01541



(a) $n = 200$ (b) $n = 800$ (c) $n = 1500$

FIGURE 3. Size, power, and adjusted power (%) for the test $H_0 : \beta_3 = 0$ in Model 6.1: $\beta_3 = 0, \cdots, 0.5$, $n = 200, 800, 1,500$.

sizes $n = 200, 800, 1,500$. When $\beta_3 = 0$, ROLS achieves a good size close to the nominal 5%, while the size based on OLS results starts around 20% and remains heavily oversized even as n increases. For $\beta_3 \neq 0$, power rises monotonically with β_3 for both methods. In Figure 3, the blue solid lines represent power based on ROLS, and the red solid lines correspond to standard OLS. Considering that the OLS estimation has large size distortion, we compute its adjusted power, shown by the red dotted lines. With a small sample size $n = 200$, OLS appears more powerful for $\beta_3 \leq 0.2$, whereas ROLS catches up and achieves good power when $\beta_3 \geq 0.3$. For $n = 800$ and $1,500$, both methods already achieve good power around $\beta_3 = 0.2$. Overall, ROLS provides reliable size and competitive power across different sample sizes. Similar results are observed for Model 6.2.

6.2. Estimation of a Time-Varying Parameter

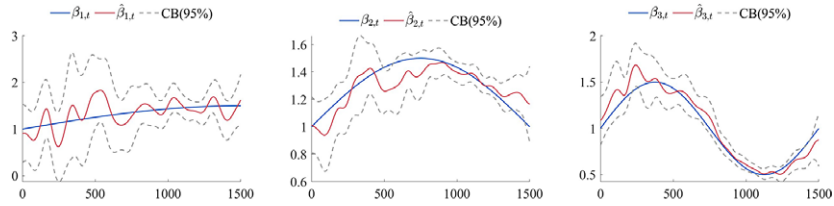
In this section, we examine the validity of the normal approximation for the estimator $\widehat{\beta}_t$, (22), of the time-varying parameter β_t, as established in Corollary 3.1 of Section 3. We replace the fixed regression parameter β in the model (53) by a time-varying parameter $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$:

$$y_t = \beta_{1t} + \beta_{2t} z_{2t} + \beta_{3t} z_{3t} + u_t, \quad u_t = h_t \varepsilon_t, \tag{56}$$

where $z_{1t} = 1$ and z_{2t}, z_{3t} are defined using μ_{2t}, μ_{3t} and η_{2t}, η_{3t} as in (54).

We consider two simulation models. Model 6.3 assumes deterministic parameters and scale factors, while Model 6.4 combines deterministic and stochastic parameters and scale factors.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press



(a) $\beta_{1,t}$ (b) $\beta_{2,t}$ (c) $\beta_{3,t}$

FIGURE 4. Robust 95% confidence intervals for the time-varying parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ in Model 6.3: $n = 1,500$, bandwidth $H = n^{0.5}$. Single replication.

MODEL 6.3. *y_t follows (56) with ε_t as in (55). The scale factors h_t, g_{2t}, g_{3t} and the parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ are deterministic:*

$$h_t = 0.5\sin(2\pi t/n) + 1, \quad g_{2t} = g_{3t} = 0.5\sin(\pi t/n) + 1.$$
$$\beta_{1t} = 0.5\sin(0.5\pi t/n) + 1, \quad \beta_{2t} = 0.5\sin(\pi t/n) + 1, \quad \beta_{3t} = 0.5\sin(2\pi t/n) + 1.$$

MODEL 6.4. *y_t follows (56) with $\varepsilon_t \sim i.i.d.\mathcal{N}(0,1)$ and scale factors:*

$$h_t = 0.5\sin(2\pi t/n) + 1, \quad g_{2t} = \left|n^{-\gamma}\sum_{j=1}^{t}\zeta_j\right| + 0.2, \quad g_{3t} = 0.5\sin(\pi t/n) + 1.$$

Parameters β_{1t}, β_{2t} are the same as in Model 6.3, while β_{3t} is stochastic:

$$\beta_{3t} = \left|n^{-\gamma}\sum_{j=1}^{t}\nu_j\right| + 0.3(t/n),$$

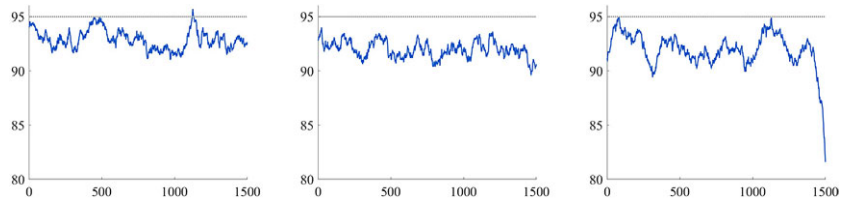
where $\{\zeta_j\}, \{\nu_j\}$ are stationary ARFIMA$(0,d,0)$ processes with memory parameter $d = 0.4$.

We estimate β_t using the estimator $\widehat{\beta}_t$, (22), where the weights $b_{n,tj} = K(|t-j|/H)$ are computed with the Gaussian kernel function $K(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ with bandwidth $H = n^h$, $h = 0.4, 0.5, 0.6, 0.7$.

Figure 4 displays parameter estimation results for a single simulation from Model 6.3. It depicts the estimates $\widehat{\beta}_{k1}, \ldots, \widehat{\beta}_{kn}$ (red line) against the true parameters β_{kt} (blue line), $k = 1,2,3$ obtained with the bandwidth $H = n^{0.5}$, and their point-wise 95% confidence intervals (gray dashed lines), computed using the robust standard errors. The robust time-varying confidence intervals cover the true parameters β_{kt}, $t = 1, \ldots, n$, for most of the time points.

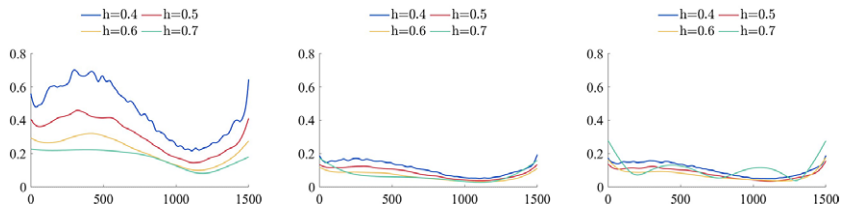
Figure 5 reports the point-wise empirical coverage rates (blue line) in the time-varying robust estimation of parameters β_{kt}, $k = 1,2,3$, which are close to the nominal 95% for most of the time points. Figure 6 shows the RMSEs for different choices of the bandwidth $H = n^h$, $h = 0.4, 0.5, 0.6, 0.7$. As expected, the

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press



(a) $\beta_{1,t}$ (b) $\beta_{2,t}$ (c) $\beta_{3,t}$

FIGURE 5. Coverage rates (in %) of robust confidence intervals for the time-varying parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ in Model 6.3: $n = 1,500$, bandwidth $H = n^{0.5}$.



(a) $\beta_{1,t}$ (b) $\beta_{2,t}$ (c) $\beta_{3,t}$

FIGURE 6. RMSE for the time-varying parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ in Model 6.3: $n = 1,500$, bandwidth $H = n^h$, $h = 0.4, 0.5, 0.6, 0.7$.

RMSE depends on the smoothness of the parameter β_{kt} and is often minimized by moderately large values of H, for example, $H = n^{0.6}$.

Figure 7 reports estimation results for a single simulation from Model 6.4, and Figure 8 displays point-wise empirical coverage rates for robust 95% confidence intervals. For deterministic parameters β_{1t} and β_{2t}, the estimation quality is good and the results are similar to those obtained for Model 6.3. For the stochastic parameter β_{3t}, the robust point-wise confidence intervals cover the path of stochastic parameter β_{3t} for most of the time points (see Figure 7(c)). Figure 8(c) shows that coverage rates of robust time-varying confidence intervals for β_{3t} might be slightly affected by the stochastic variation in the parameter and scale factors. Nevertheless, they are still satisfactory and reasonably close to the nominal 95% coverage.

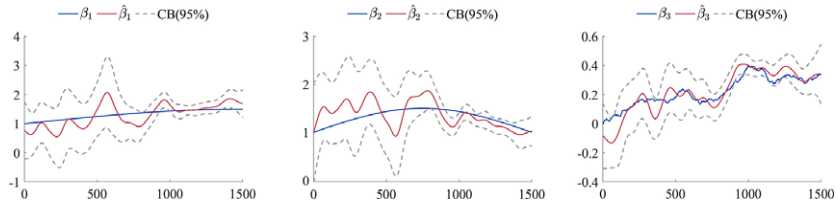
6.3. Estimation of Regression Parameters with Missing Data

To examine the impact of missing data on the robust and standard OLS estimation based on partially observed data $(y_{j_1}, z_{j_1}), (y_{j_2}, z_{j_2}), \ldots, (y_{j_N}, z_{j_N})$, we use two types of missing data patterns over the time period $1, \ldots, 1,500$.

Type 1. The block of data $j \in [650, 850]$ is missing.

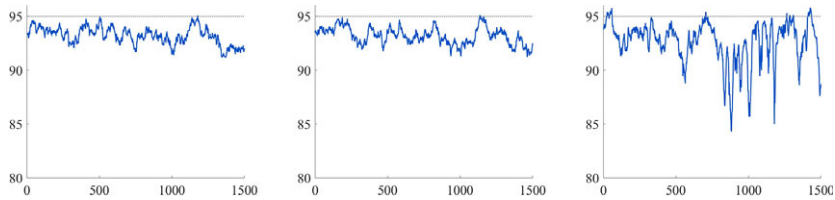
Type 2. Five hundred single observations are missing at randomly selected times.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press



(a) $\beta_{1,t}$ (b) $\beta_{2,t}$ (c) $\beta_{3,t}$

FIGURE 7. Robust 95% confidence bands for the time-varying parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ in Model 6.4: $n = 1,500$, bandwidth $H = n^{0.5}$. Single replication.



(a) $\beta_{1,t}$ (b) $\beta_{2,t}$ (c) $\beta_{3,t}$

FIGURE 8. Coverage rates (in %) of robust confidence intervals for the time-varying parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ in Model 6.4: $n = 1,500$, bandwidth $H = n^{0.5}$.

TABLE 3. Robust OLS estimation in Model 6.1 with block missing data (Type 1)

Parameters	Bias	RMSE	CP	CP_{st}	SD
β_1	−0.00818	0.04983	94.60	74.60	0.04915
β_2	0.00356	0.03875	94.00	67.90	0.03859
β_3	0.00246	0.03840	93.80	70.00	0.03832

Tables 3 and 4 report robust and standard estimation results for Model 6.1 with a fixed parameter. Table 3 shows that block missing data (Type 1) do not lead to noticeable changes in bias, RMSE and SD, and the coverage rate for robust confidence intervals remains around 95%. At the same time, the coverage rate CP_{st} of the standard confidence intervals is substantially distorted.

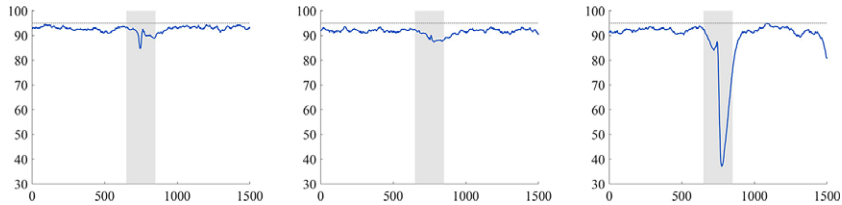
Table 4 shows that randomly missing data do not affect the coverage rate of robust confidence intervals which remains close to the nominal 95%, while the coverage rate of the standard confidence intervals drops to around 65%. This emphasizes the flexibility of the ROLS estimation of the fixed parameter in the presence of block or randomly missing data.

Figures 9–11 report the estimation results for Model 6.3 with the time-varying parameter β_t.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

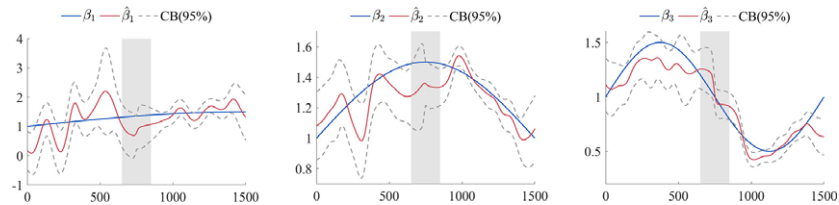
TABLE 4. Robust OLS estimation in Model 6.1 with randomly missing data (Type 2)

Parameters	Bias	RMSE	CP	CP_{st}	SD
β_1	−0.00567	0.05732	94.30	66.60	0.05704
β_2	0.00144	0.04251	95.20	63.50	0.04249
β_3	0.00289	0.04128	94.80	64.70	0.04118



(a) $\beta_{1,t}$ (b) $\beta_{2,t}$ (c) $\beta_{3,t}$

FIGURE 9. Coverage rates (in %) of robust confidence intervals for the time-varying parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ in Model 6.3 with block missing data (Type 1), $n = 1,500$, bandwidth $H = n^{0.5}$.



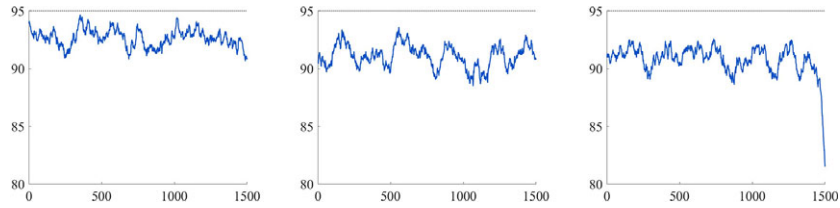
(a) $\beta_{1,t}$ (b) $\beta_{2,t}$ (c) $\beta_{3,t}$

FIGURE 10. Robust 95% confidence bands for the time-varying parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ in Model 6.3 with block missing data (Type 1), $n = 1,500$, bandwidth $H = n^{0.5}$. Single replication.

Figure 9 shows the coverage rates in time-varying robust estimation with block missing data (Type 1, shaded region) for $t = 1, \ldots, 1,500$. The coverage is close to the nominal 95%, with some distortion for parameters $\beta_{1,t}$ and $\beta_{2,t}$ and a larger distortion for $\beta_{3,t}$ within the shaded region. The distortion peaks at the center of the block, as expected. Although the width of the missing data block, 200, exceeds the bandwidth $H = n^{0.5} = 39$ used in estimating β_t, the coverage distortion seems to be offset by the smooth down-weighting of the data, and the performance of the robust time-varying OLS estimation exceeds expectations.

Figure 10 reports the path of the estimator $\widehat{\beta}_{kt}$ and the point-wise robust confidence intervals, for a single simulation. The robust confidence intervals become wider in the shaded region, which likely explains the satisfactory coverage performance during that period.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press



(a) $\beta_{1,t}$ (b) $\beta_{2,t}$ (c) $\beta_{3,t}$

FIGURE 11. Coverage rates (in %) of robust confidence intervals for the time-varying parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ in Model 6.3, 500 randomly missing data, $n = 1,500$, bandwidth $H = n^{0.5}$.

TABLE 5. Robust OLS estimation in the $AR(2)$ model (57)

Parameters	Bias	RMSE	CP	CP_{st}	SD
β_1	−0.00808	0.05250	94.9	92.3	0.05187
β_2	0.00104	0.04183	94.5	75.0	0.04182
β_3	0.00356	0.03091	94.8	88.8	0.03070

Figure 11 shows that randomly missing data (Type 2) do not distort the robust time-varying OLS estimation. For all three parameters and time periods t, the coverage rate is close to the nominal. Overall, the robust estimation of time-varying parameters does not appear to be affected by randomly missing data.

6.4. Estimation of a Stationary AR(p) Model

We assess the performance of the robust and standard procedures in the case of a stationary AR(2) model:

$$y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 y_{t-2} + \varepsilon_t, \quad \beta = (\beta_1, \beta_2, \beta_3)' = (0.5, 0.4, 0.3)', \tag{57}$$

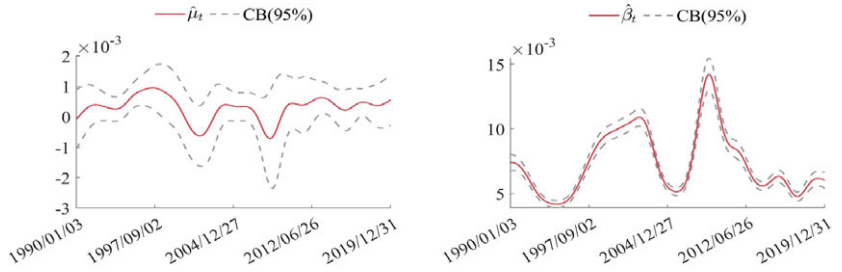
where $\varepsilon_t = e_t e_{t-1}$, $e_t \sim i.i.d.\mathcal{N}(0,1)$ is a stationary m.d. noise. The regressors $z_t = (z_{1,t}, z_{2,t}, z_{3,t})' = (1, y_{t-1}, y_{t-2})'$ include an intercept and the two past lags of y_t. By Theorem 5.1, the parameter β can be estimated by using the robust estimation method.

Table 5 shows that the coverage rate for the ROLS estimation is close to the nominal 95%, while the standard OLS estimation exhibits extensive coverage distortion for β_2 and β_3.

7. EMPIRICAL EXPERIMENT

In this section, we analyze the structure and dynamics of daily S&P 500 log returns, r_t, from 2 January 1990 to 31 December 2019 (sample size $n = 7,558$). We employ robust regression estimation to assess whether the returns r_t can be modeled using

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press



(a) Confidence bands for μ_t (b) Confidence bands for $\beta_{1t} = h_t E|\varepsilon_t|$

FIGURE 12. Robust 95% confidence bands for μ_t in model (58) and $\beta_{1t} = h_t E|\varepsilon_t|$ in model (59), $n = 7,558$, $H = n^{0.6}$.

a time-varying regression model of the form

$$r_t = \mu_t + u_t, \quad u_t = h_t \varepsilon_t, \tag{58}$$

where $\{\varepsilon_t\}$ is an i.i.d.$(0,1)$ noise, and the time-varying mean μ_t and scale factor h_t are independent of $\{\varepsilon_t\}$. Our objective is to estimate the time-varying mean μ_t, the scale factor h_t, and to test for the absence of autocorrelation in the centered absolute residuals $\widetilde{u}_t = h_t(|\varepsilon_t| - E|\varepsilon_t|)$, thereby assessing the fit of the model (58) to the data.

If returns r_t follow the model (58) with i.i.d. noise ε_t, then $\widetilde{u}_t$ are uncorrelated for $t \neq s$:

$$\operatorname{cov}(\widetilde{u}_t, \widetilde{u}_s) = E\big[h_t h_s \operatorname{cov}(|\varepsilon_t|, |\varepsilon_s|)\big] = 0.$$

Conversely, if the noise ε_t exhibits ARCH effects (stationary conditional heteroskedasticity), the sequence $|\varepsilon_t|$ becomes autocorrelated, and the null hypothesis of uncorrelated centered absolute residuals $\widetilde{u}_t$ would be rejected.

We estimate the time-varying mean μ_t using the time-varying OLS estimator with bandwidths $H = n^{0.4}, n^{0.5}, \ldots, n^{0.7}$. Figure 12(a) shows the estimated path of $\widehat{\mu}_t$ and the associated 95% confidence intervals for bandwidth $H = n^{0.6}$ indicating that μ_t is very likely to change over time.

Assumptions on model (58) imply that

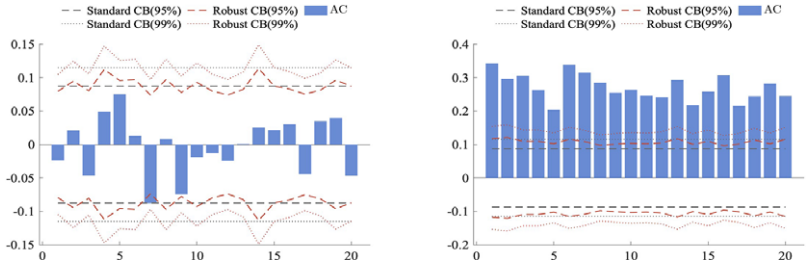
$$|u_t| = |r_t - \mu_t| = h_t|\varepsilon_t| = h_t E|\varepsilon_t| + h_t(|\varepsilon_t| - E|\varepsilon_t|).$$

Therefore, $|\widehat{u}_t| = |r_t - \widehat{\mu}_t| \sim h_t E|\varepsilon_t| + h_t(|\varepsilon_t| - E|\varepsilon_t|)$ and thus $y_t = |\widehat{u}_t|$ follows a time-varying regression model of the form

$$y_t = \beta_{1t} + \widetilde{u}_t, \quad \widetilde{u}_t = g_t \eta_t, \tag{59}$$

where $\beta_{1t} = h_t E|\varepsilon_t|$ represents a time-varying intercept, $g_t = h_t$ denotes the scale factor, and $\eta_t = |\varepsilon_t| - E|\varepsilon_t|$ is an i.i.d. noise. Hence, β_{1t} can be consistently estimated using the time-varying OLS estimator $\widehat{\beta}_{1t}$. Figure 12(b) displays the estimated path of $\widehat{\beta}_{1t}$ and the corresponding 95% confidence intervals for $\beta_{1t} =$

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press



(a) Testing at individual lag: $\widehat{\widehat{u}}_j$ (b) Testing at individual lag: $\widehat{u}_j^*$

FIGURE 13. Robust and standard tests for the absence of correlation in the subsample of residuals $\widehat{\widehat{u}}_j$, $\widehat{u}_j^*$, $j \in [500, 1000]$, $H = n^{0.6}$, significance level 5%.

$h_t E|\varepsilon_t|$ with bandwidth $H = n^{0.6}$, revealing pronounced time variation in the scale factor h_t.

Figure 13(a) reports test results for zero correlation at lags $k = 1, \ldots, 20$ in the residual sequence $\widehat{\widehat{u}}_t = y_t - \widehat{\beta}_{1t}$. We employ the standard and robust test procedures developed in Giraitis et al. (2024). Given that the sample size is large ($n = 7,558$) and β_{1t} is estimated non-parametrically with bandwidth $H = n^{0.6}$, we restrict the correlation analysis to the subsample $j \in [500, 1000]$. Both tests provide no evidence of significant correlation within this subsample, suggesting that the model (58) fits the returns r_t well during this time period.

The same is not likely to be true if $r_t^* = r_t - \widehat{\mu}_t$ follows a GARCH(1,1) process, as confirmed by the following experiment. We fit a GARCH(1,1) model to the demeaned returns $r_t^* = r_t - \widehat{\mu}_t$:

$$r_t^* = \sigma_t \varepsilon_t, \quad \sigma_t^2 = 1.563 \times 10^{-6} + 0.88913\sigma_{t-1}^2 + 0.096974 r_{t-1}^{*2}.$$

We generate a simulated GARCH(1,1) sample $r_{g1}^*, \ldots, r_{gn}^*$, apply the regression model (59) to the absolute values $y_t^* = |r_{gt}^*|$, and compute the residuals, $\widehat{u}_t^* = y_t^* - \widehat{\beta}_{1t}$. Figure 13(b) shows that both standard and robust tests detect significant correlation in the residuals $\widehat{u}_t^*$, confirming the presence of conditional heteroskedasticity in the simulated GARCH data.

8. CONCLUSION

The ROLS and time-varying OLS estimation and inference methods developed in this article offer considerable flexibility for modeling economic and financial data. They allow for general heterogeneity in regression components and for structural change in regression coefficients over time. Moreover, the generalization of the structure of regressors and error terms further expands the range of empirical settings to which the ROLS regression framework can be applied. In particular, the article develops asymptotic theory for general regression models with stochastic

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

regressors, possibly including a time-varying mean, and provides data-based robust standard errors that enable the construction of confidence intervals for the regression parameters. The Monte Carlo analysis demonstrates the strong performance of the robust estimation approach under complex settings, and confirms the consistency and asymptotic normality property of the proposed estimators.

COMPETING INTEREST STATEMENT

The authors declare that no competing interests exist.

FUNDING STATEMENT

The authors declare that no specific funding has been received for this article.

SUPPLEMENTARY MATERIAL

Proofs of all results and some additional simulations are provided in Giraitis, L., Kapetanios, G., Li, Y., & Ventouri, A. (2025). Supplement to "Unlocking the Regression Space," Econometric Theory Supplementary Material. To view, please visit: https://doi.org/10.1017/S0266466626100395.

REFERENCES

Angrist, J. D., & Pischke, J. S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2), 3–30.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66, 47–78.

Bardet, J.-M., & Wintenberger, O. (2009). Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *Annals of Statistics*, 37, 2730–2759.

Boldea, O., Cornea-Madeira, A., & Hall, A. (2019). Bootstrapping structural change tests. *Journal of Econometrics*, 213(2), 359–397.

Brockwell, P., & Davis, R. (1991). *Time series: Theory and methods*. Springer.

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society, Series B*, 37(2), 149–163.

Carriero, A., Clark, T. E., Marcellino, M., & Mertens, E. (2024). Addressing COVID-19 outliers in BVARs with stochastic volatility. *The Review of Economics and Statistics*, 106(5), 1403–1417.

Cattaneo, M., Jansson, M., & Newey, W. (2018). Inference in linear regression models with many covariates and heteroscedasticity. *Journal of the American Statistical Association*, 113(523), 1350–1361.

Chan, J. C. C., Koop, G., & Xuewen, Y. (2024). Large order-invariant Bayesian VARs with stochastic volatility. *Journal of Business and Economic Statistics*, 42(2), 825–837.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28, 591–605.

Chronopoulos, I., Giraitis, L., & Kapetanios, G. (2022). Choosing between persistent and stationary volatility. *Annals of Statistics*, 50(6), 3466–3483.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

Cogley, T., & Sargent, T. J. (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, 8(2), 262–302.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Annals of Statistics*, 25(1), 1–37.

Dahlhaus, R., & Giraitis, L. (1998). On the optimal segment length for parameter estimates for locally stationary time series. *Journal of Time Series Analysis*, 19(6), 629–655.

Dahlhaus, R., & Richter, S. (2023). Adaptation for nonparametric estimators of locally stationary processes. *Econometric Theory*, 39, 1123–1153.

Dahlhaus, R., Richter, S., & Wu, W. (2019). Towards a general theory for locally stationary processes. *Bernoulli*, 25(2), 1013–1044.

Dalla, V., Giraitis, L., & Robinson, P. M. (2020). Asymptotic theory for time series with changing mean and variance. *Journal of Econometrics*, 219, 281–313.

Dendramis, Y., Giraitis, L., & Kapetanios, G. (2021). Estimation of time-varying covariance matrices for large datasets. *Econometric Theory*, 37(6), 1100–1134.

Doukhan, P., & Wintenberger, O. (2008). Weakly dependent chains with infinite memory. *Stochastic Processes and their Applications*, 118(11), 1997–2013.

Eicker, F. (1963). Asymptotic normality and consistency of the least squares estimators for families of linear regressions. *Annals of Mathematical Statistics*, 34, 447–456.

Enders, C. K. (2022). *Applied missing data analysis*. Guilford Publications.

Eraker, B., Johannes, M., & Polson, N. (2003). The impact of jumps in volatility and returns. *Journal of Finance*, 58(3), 1269–1300.

Fan, J., & Zhang, W. (1999). Statistical estimation in varying coefficient models. *Annals of Statistics*, 27(5), 1491–1518.

Fu, Z., Hong, Y., Su, L., & Wang, X. (2023). Specification tests for time-varying coefficient models. *Journal of Econometrics*, 235, 720–744.

Georgiev, I., Harvey, D., Leybourn, S., & Robert Taylor, A. (2018). Testing for parameter instability in predictive regression models. *Journal of Econometrics*, 204, 101–118.

Giraitis, L., Kapetanios, G., & Yates, T. (2014). Inference on stochastic time-varying coefficient models. *Journal of Econometrics*, 179(1), 46–65.

Giraitis, L., Kapetanios, G., & Yates, T. (2018). Inference on multivariate heteroscedastic time varying random coefficient models. *Journal of Time Series Analysis*, 39(2), 129–149.

Giraitis, L., Li, Y., & Phillips, P. C. B. (2024). Robust inference on correlation under general heterogeneity. *Journal of Econometrics*, 240(1), 105691.

Giraitis, L., Taniguchi, M., & Taqqu, M. S. (2018). Estimation pitfalls when the noise is not iid. *Japanese Journal of Statistics and Data Science*, 1, 59–80.

Hall, A., Han, S., & Boldea, O. (2012). Inference regarding multiple structural changes in linear models with endogenous regressors. *Journal of Econometrics*, 270(2), 281–302.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57, 357–384.

Hansen, B. (2000). Testing for structural change in conditional models. *Journal of Econometrics*, 97(1), 93–115.

Hidalgo, J., Lee, Y., & Seo, M.-H. (2019). Robust inference for threshold regression models. *Journal of Econometrics*, 210, 291–309.

Hu, Z., Kasparis, I., & Wang, Q. (2025). Time-varying parameter regressions with stationary persistent data. *Econometric Theory*, 41(4), 817–843.

Jochmans, K. (2019). Heteroscedasticity-robust inference in linear regression models with many covariates. *Journal of the American Statistical Association*, 117(538), 887–896.

Karmakar, S., Richter, S., & Wu, W. B. (2022). Simultaneous inference for time-varying models. *Journal of Econometrics*, 227(2), 408–428.

Lai, T. L., & Wei, C. Z. (1982). Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *Annals of Statistics*, 10(1), 154–166.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press

Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review*, 73(1), 31–43.

Leamer, E. E. (2010). Tantalus on the road to Asymptopia. *Journal of Economic Perspectives*, 24(2), 31–46.

Li, D., Phillips, P. C. B., & Gao, J. (2020). Kernel-based inference in time-varying coefficient cointegrating regression. *Journal of Econometrics*, 215, 607–632.

Linton, O., & Xiao, Z. (2019). Efficient estimation of nonparametric regression in the presence of dynamic heteroskedasticity. *Journal of Econometrics*, 213, 608–631.

MacKinnon, J. G. (2012). Thirty years of heteroscedasticity-robust inference. In X. Chen, & N. R. Swanson (Eds.), *Recent advances and future directions in causality, prediction, and specification analysis* (pp. 437–461). Springer.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association*, 84(405), 223–230.

Ploberger, W., & Krämer, W. (1992). The CUSUM test with OLS residuals. *Econometrica*, 60, 271–285.

Priestley, M. B. (1965). Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society, Series B*, 27(2), 204–229.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, 72(3), 821–852.

Robinson, P. M. (1989). Nonparametric estimation of time-varying parameters. In P. Hackl (Ed.), *Economic structural change: Analysis and foresting* (pp. 253–264). Springer.

Robinson, P. M. (1991). Time-varying nonlinear regression. In P. Hackl, & A. H. Westland (Eds.), *Economic structural change: Analysis and forecasting* (pp. 179–190). Springer.

Sun, Y., Hong, Y., Lee, T.-H., Wang, S., & Zhang, X. (2021). Time-varying model averaging. *Journal of Econometrics*, 222, 974–992.

Terasvirta, T. (1998). Modelling economic relationships with smooth transition regressions. In A. Ullah, & D. E. A. Giles (Eds.), *Handbook of applied economic statistics* (pp. 507–552). Marcel Dekker.

Tong, H. (1990). *Non-linear time series: A dynamical system approach*. Oxford University Press.

Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Annals of Statistics*, 40(5), 2601–2633.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48, 817–838.

White, H. (2014). *Asymptotic theory for econometricians*. Academic Press.

Wu, W. (2005). M-estimation of linear models with dependent errors. *Annals of Statistics*, 25(2), 495–521.

Xu, K.-L., & Phillips, P. C. B. (2008). Adaptive estimation of autoregressive models with time-varying variances. *Journal of Econometrics*, 142, 265–280.

Zhang, T., & Wu, W. (2012). Inference of time-varying regression models. *Annals of Statistics*, 40(3), 1376–1402.

Zhang, T., & Wu, W. (2015). Time-varying nonlinear regression models: Nonparametric estimation and model selection. *Annals of Statistics*, 43(2), 741–768.

https://doi.org/10.1017/S0266466626100395 Published online by Cambridge University Press