mathematics

MDPI

Article

Identification and Estimation in Linear Models with Endogeneity Through Time-Varying Volatility

Shih-Tang Hwu

Department of Economics, California State Polytechnic University, 3801 W. Temple Ave., Pomona, CA 91768, USA; shwu@cpp.edu

Abstract: This paper proposes a novel control function approach to identify and estimate linear models with endogenous variables in the absence of valid instrumental variables. The identification strategy exploits time-varying volatility to address the multicollinearity problem that arises in conventional control function methods when instruments are weak. We establish the identification conditions and show that the proposed method is $\sqrt{T}$-consistent and asymptotically normal. We apply the proposed approach to estimate the elasticity of intertemporal substitution, a key parameter in macroeconomics. Using quarterly data on aggregate stock returns across eleven countries, we find that the data exhibit substantial time variation in volatility, supporting the identifying assumptions. The proposed method yields confidence intervals that are broadly consistent with the general findings in the literature and are substantially narrower than those obtained using weak-instrument robust methods.

Keywords: control function; nonparametric estimation; time-varying volatility; elasticity of intertemporal substitution

MSC: 91B84; 62G05; 62J05

check for updates

Academic Editor: Márcio Poletti Laurini

Received: 1 May 2025
Revised: 27 May 2025
Accepted: 31 May 2025
Published: 2 June 2025

Citation: Hwu, S.-T. Identification and Estimation in Linear Models with Endogeneity Through Time-Varying Volatility. *Mathematics* **2025**, *13*, 1849. https://doi.org/10.3390/math13111849

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In the linear regression framework, the instrumental variable (IV) is one of the most widely applied methods to tackle endogeneity in empirical research. A valid IV must satisfy three conditions: (1) it must be excluded from the structural equation; (2) it must be uncorrelated with the structural error term; (3) it must be strongly correlated with the endogenous regressors. If any of these conditions fail, the conventional IV estimator will produce biased and inconsistent results. In many empirical applications, it is often debatable whether the exclusion and the orthogonality conditions hold. Furthermore, even when these two conditions are satisfied, a vast literature has studied the consequences of weak instruments, as instruments are often only weakly correlated with the endogenous variables.

It is well known that in linear models, the IV method is equivalent to ordinary least squares (OLS) regression that includes the reduced form residuals as additional regressors to control for endogeneity, a method known as the control function (CF) approach (see, e.g., [1,2]). While the two approaches obtain numerically identical estimators under strong instruments, the CF method offers a distinct conceptual advantage by allowing researchers to model the source of endogeneity directly in the structural equation. This flexibility makes the CF approach particularly appealing in settings with conditional heteroskedasticity, where variation in the second moments can be exploited to aid identification.

However, when instruments are weak, the residuals from the reduced form equation become nearly collinear with the endogenous regressors, leading to identification failure

in the standard CF framework. In this paper, we propose a feasible CF approach that exploits time-varying volatility to address the identification issue in the absence of strong instruments. The idea is that heteroskedasticity generates variation in the scale of the reduced-form errors, which weakens their asymptotic collinearity with the endogenous variables. Earlier studies have applied similar logic in various scenarios. Ref. [3] formulates a model and shows that the parameters can be identified when there are multiple known regimes of heteroskedasticity under certain conditions. Ref. [4] explores the Markov-switching heteroskedascitiy to identify the correlation between trend inflation and inflation gap in an unobserved components model. In structural vector autoregressive framework, refs. [5,6] discuss how Markov-switching heteroskedasticity and stochastic volatility can aid identification.

The proposed method is developed under a general framework that allows for nonparametric time variation in volatility and does not require knowledge of the heteroskedasticity structure. We provide conditions under which the parameters of the structural model are identified and show that the proposed estimator is $\sqrt{T}$-consistent and asymptotically normal. Unlike existing methods that rely on known heteroskedastic regimes [3] or parametric specifications for the volatility process [4–6], our approach allows for a nonparametric volatility process subject only to mild smoothness conditions. We derive the asymptotic distribution of the estimator and discuss inference procedures that account for generated regressors and time-varying second moments. In contrast to the standard generalized method of moments (GMM), our framework allows for consistent estimation and testing under weak instruments. A test for endogeneity is also developed, which is applicable even without valid instruments.

We apply the proposed method to estimate the elasticity of intertemporal substitution (EIS), a central parameter in macroeconomics and asset pricing. The EIS governs the trade-off between current and future consumption and plays a critical role in determining asset prices. The empirical literature has found mixed results for the value of EIS. While [7] finds small values for the EIS, high values for EIS are required by models such as the long-run risk model of [8], which has been successful in explaining the equity premium and other asset pricing puzzles. As pointed out by [9,10], weak instruments may explain the divergence in the empirical literature. Motivated by the well-documented time-varying volatility in financial returns (e.g., [11]), we apply our proposed method to estimate the EIS using quarterly aggregate stock return data from eleven developed countries. We find that the confidence intervals obtained under our approach are much tighter than those derived from weak-instrument robust methods and broadly consistent with the literature.

This paper contributes to the literature on identification through heteroskedasticity by showing that smooth nonparametric volatility dynamics can serve as a source of identification when conventional instruments fail. It also contributes to the empirical macro-finance literature by providing new estimates of the EIS based on a control function approach that does not rely on instrument strength. The rest of the paper is organized as follows. Section 2 presents the model specification and the identification conditions. Section 3 discusses the estimation and inference under the proposed framework. Section 4 applies the method to the estimation of EIS. Section 5 concludes.

2. Model Specification and Identification

We consider the following structural model with endogenous regressors,

$$y_t = x_t'\beta_0 + u_t, \quad E(x_t u_t) \neq 0, \tag{1}$$

$$x_t = (I_p \otimes z_{t-1}')\delta_0 + v_t, \tag{2}$$

where y_t is the scalar dependent variable, observed over $t = 1, \ldots, T$. $x_t \in \mathbb{R}^p$ are the endogenous regressors, and $z_{t-1} \in \mathbb{R}^q$ are instruments observed at lag $t-1$. This specification allows for other strictly exogenous variables to be partialled out, following the approach in [12]. Our primary goal is to identify $\beta_0 \in \mathcal{B}$, with $\mathcal{B}$ compact.

We define the σ-field $\mathcal{F}_t$ generated by the history of observables $\{y_s, x_s, z_s\}_{s \le t}$, and introduce the following definitions:

$$\begin{aligned} \sigma_{ut}^2 &= Var(u_t|\mathcal{F}_{t-1}), \quad u_t^* = \frac{u_t}{\sigma_{ut}}, \\ \Omega_{vt} &= Var(v_t|\mathcal{F}_{t-1}), \quad v_t^* = \Omega_{vt}^{-\frac{1}{2}} v_t \end{aligned}$$

Assumption 1. *(Mixing and Moment Conditions):*

1. *The process u_t^* and v_{it}^* for $i = 1, \ldots, p$ are strong mixing (α-mixing) martingale difference sequences with respect to $\mathcal{F}_t$, with unit conditional variance.*
2. *There exist $\alpha_1 > 1$, $\alpha_{2i} > 1$, $M_1 > 0$ and $M_{2i} > 0$, such that $\sup_t E u_t^{*4\alpha_1} < M_1 < \infty$ and $\sup_t E v_{it}^{*4\alpha_{2i}} < M_{2i} < \infty$, for $i = 1, \ldots, p$.*

Assumption 1 assumes that both u_t^* and v_t^* have finite fourth moments. By this assumption and Lyapunov's inequality, $E\left(|u_t^*|^{4\xi}\right)$ and $E\left(|v_{it}^*|^{4\xi}\right)$ exist for all $\xi \le \min(\alpha_1, \alpha_{21}, \ldots, \alpha_{2p})$, as do all expectations involving up to any four combinations of u_t^* and v_{it}^*, for $i = 1, \ldots, p$. Note that by construction, $Var(u_t^*|\mathcal{F}_{t-1}) = Var(v_{it}^*|\mathcal{F}_{t-1}) = 1$.

To correct for endogeneity, we follow the control function approach. Conditional on $\mathcal{F}_{t-1}$, we project the structural error term u_t on the reduced-form error term,

$$\begin{aligned} u_t &= E(u_t|v_t) + w_t = v_t'[E(v_t v_t'|\mathcal{F}_{t-1})]^{-1} E(u_t v_t|\mathcal{F}_{t-1}) + w_t \\ &= v_t^{*\prime} E(u_t v_t|\mathcal{F}_{t-1}) + w_t \end{aligned} \tag{3}$$

Substituting (3) into (1), we obtain the following transformed model:

$$y_t = x_t'\beta_0 + v_t^{*\prime} E(u_t v_t|\mathcal{F}_{t-1}) + w_t \tag{4}$$

The transformed error term w_t plays a central role. The following lemma characterizes its orthogonality properties:

Lemma 1. *Under Assumption 1, we can show that*

$$\begin{aligned} &(1) \quad E(w_t|\mathcal{F}_{t-1}) = 0 \\ &(2) \quad E(x_t w_t|\mathcal{F}_{t-1}) = 0 \\ &(3) \quad E(v_t w_t|\mathcal{F}_{t-1}) = 0 \\ &(4) \quad E(v_t^* w_t|\mathcal{F}_{t-1}) = 0 \end{aligned}$$

The proof of Lemma 1 is in Appendix A. Lemma 1 shows that w_t is a martingale difference sequence and orthogonal to x_t, v_t, and v_t^*. Intuitively, w_t is orthogonal to v_t^* by the projection in (3). Since v_t^* is a transformation of v_t, and x_t is from the instruments and v_t, the orthogonality naturally extends to x_t and v_t. This orthogonality allows (4) to be used for consistent estimation, provided the model is identified.

However, when instruments are weak, the model in (4) is not always identifiable. We now formalize the identification strategy in this case.

Identification Strategy

When the instruments in (2) are weak, $\delta_0 \to 0$ as $T \to \infty$, and the regressors $x_t \to v_t$, causing the control function term v_t^* and x_t to become collinear. This undermines identification in (4). We formalize this situation with the following assumption:

Assumption 2.

1. *Orthogonality:* $\frac{1}{T}\sum_{t=2}^{T} z_{t-1}u_t \xrightarrow{p} E(z_{t-1}u_t) = \mathbf{0}_q$, $\frac{1}{T}\sum_{t=2}^{T} z_{t-1} \otimes v_t \xrightarrow{p} E(z_{t-1} \otimes v_t) = \mathbf{0}_{p\times q}$.
2. *Finite second moments:* $\sup_t E\left(z_t z_t'\right) < \infty$.
3. *Local-to-zero reduced-form coefficients:* $\delta_0 = \frac{C}{T^\nu}$, *where* $\nu \geq \frac{1}{2}$, $C \in \mathbb{R}^{pq}$.

Assumption 2(1) assumes that the instruments satisfy the conventional orthogonality condition. Assumption 2(2) states that the instruments have a finite second moment for all t. Both are standard in the literature. Assumption 2(3) captures a local-to-zero asymptotic framework, which is standard in the weak-instrument literature [13]. It indicates that the instruments are increasingly uninformative as the sample size grows. In practice, this means that the predictive power of z_{t-1} for x_t is low, and x_t approximates the reduced-form error v_t. Applied researchers can assess this by checking first-stage F-statistics, small eigenvalues, or instability in 2SLS estimates. Modeling this weakness explicitly allows us to show how time-varying volatility can restore identification.

Under Assumption 2, $x_t \approx v_t$. When volatility is constant over time, $v_t^* = \Omega_v^{-1/2} v_t$, both regressors in (4) become asymptotically collinear. In other words, the moment conditions derived from (4) collapse into a system of only p linearly independent equations for $2p$ parameters. Identification fails due to collinearity.

To overcome this, we allow volatility to vary smoothly with the following assumption:

Assumption 3.

1. $$\Omega_{vt} \equiv \Omega\left(\frac{t}{T}\right) = \begin{bmatrix} h_{11}\left(\frac{t}{T}\right) & h_{12}\left(\frac{t}{T}\right) & \dots & h_{1p}\left(\frac{t}{T}\right) \\ h_{21}\left(\frac{t}{T}\right) & h_{22}\left(\frac{t}{T}\right) & \dots & h_{2p}\left(\frac{t}{T}\right) \\ \vdots & \vdots & & \vdots \\ h_{p1}\left(\frac{t}{T}\right) & h_{p2}\left(\frac{t}{T}\right) & \dots & h_{pp}\left(\frac{t}{T}\right) \end{bmatrix},$$
   where $h_{ij}(\cdot) = h_{ji}(\cdot)$, $h_{ij}(\cdot)$ *are non-stochastic, measurable, uniformly bounded on the interval* $(0, 1]$, *and satisfy Lipschitz condition, for all* $i, j = 1, \dots, p$. $\inf_t h_{ii}\left(\frac{t}{T}\right) > 0$, *for all* $i = 1, \dots, p$.
2. $\Omega\left(\frac{t}{T}\right)$ *is not constant and nonsingular over* t.
3. $E(u_t v_{it}^* | \mathcal{F}_{t-1}) = \gamma_{i0}$ *where* γ_{i0} *are constant for all* t, $i = 1, \dots, p$.

When volatility varies smoothly over time, the standardization $v_t^* = \Omega_{vt}^{-1/2} v_t$ introduces time-varying heterogeneity that breaks the collinearity between x_t and v_t^*. This restores full rank in the moment conditions implied by (4), allowing point identification of (β_0, γ_0). The time-varying volatility structure assumed in Assumption 3 is consistent with a wide range of economic environments where conditional second moments evolve smoothly over time. In macro-finance, stochastic volatility is a well-documented feature of asset returns (e.g., [14,15]). For instance, persistent changes in macroeconomic uncertainty

or shifting investor sentiment can induce gradual fluctuations in return volatility. Our assumption accommodates such dynamics in a nonparametric form, allowing identification to be driven by economically plausible second-moment variation without requiring a specific structural model of volatility.

3. Estimation and Inference

Under Assumptions 1–3, the structural model in (4) takes the form

$$y_t = x_t'\beta_0 + v_t^{*\prime}\gamma_0 + w_t, \tag{5}$$

where w_t is a martingale difference sequence orthogonal to x_t and v_t^*, as established in Lemma 1.

Let $s_t = \left(x_t', v_t^{*\prime}\right)' \in \mathbb{R}^{2p}$, the population moment conditions are

$$E(s_t w_t) = 0$$

Because the number of moment conditions equals the number of parameters, the system is just-identified. The optimal GMM estimator reduces to OLS, and no weighting matrix is required. In this case, the moment conditions do not benefit from reweighting, and the weighting matrix is effectively the identity.

If the values of δ_0 and Ω_{vt} were known, the infeasible estimator could be written as

$$\begin{pmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{pmatrix} = (\sum_{t=1}^{T} s_t' s_t)^{-1} \sum_{t=1}^{T} s_t' y_t \tag{6}$$

Theorem 1. *Under Assumptions 1–3, and assuming δ_0 and Ω_{vt} are known, the estimator in (6) satisfies*

$$\sqrt{T}\begin{pmatrix} \widehat{\beta} - \beta_0 \\ \widehat{\gamma} - \gamma_0 \end{pmatrix} \xrightarrow{d} N(0, \Gamma_1^{-1}\Lambda_1\Gamma_1^{-1}),$$

where

$$\Gamma_1 = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E\left(s_t s_t'\right), \quad \Lambda_1 = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E\left(w_t^2 s_t' s_t\right)$$

The proof of Theorem 1 is given in Appendix A. Theorem 1 shows that the infeasible estimators are $\sqrt{T}$ consistent and asymptotically normally distributed when the true value of δ_0 and Ω_{vt} are known. It also demonstrates the problem of weak instruments in the absence of time-varying variance. In particular, when $\Omega_{vt} = \overline{\Omega}$, Γ_1 in Theorem 1 becomes

$$\Gamma_1 = \begin{pmatrix} \overline{\Omega} & \overline{\Omega}^{\frac{1}{2}} \\ \overline{\Omega}^{\frac{1}{2}} & 1 \end{pmatrix},$$

which is a singular matrix. Thus, the asymptotic variance–covariance matrix cannot be computed.

The estimators in Theorem 1 are infeasible in practice, as the true value of δ_0 and Ω_{vt} are unknown. We propose a feasible estimator in the following subsection.

3.1. Feasible Estimation Procedure

In practice, we construct a feasible estimator by estimating δ_0 and Ω_{vt}. Let $\widehat{v}_t = x_t - (I_p \otimes z_t)\widehat{\delta}$ be the first stage OLS residuals from estimating reduced-form Equation (2). We propose to use the generalized kernel proposed by [16] to estimate the time-varying volatility.

Let $b_{ij} \equiv b_{ij}(T) \to 0$ denote a bandwidth and let $K(\cdot)_{b_{ij}}$ denote a univariate generalized kernel function with the properties $K_{b_{ij}}(u,t) = 0$ if $u > t$ or $u < 1-t$; for all $t \in [0,1]$,

$$h^{-(j+1)} \int_{t-1}^{t} u^j K_{b_{ij}}(u,t)du = \begin{cases} 1, & \text{if } j = 0 \\ 0, & \text{if } 1 \le j \le r-1 \end{cases}$$

We call $K_{b_{ij}}(\cdot,\cdot)$ a univariate generalized kernel function of order r.

The following example is taken from [16]. Define

$$\mathcal{M}_{\theta,r}([a_1,a_2]) = \left\{ g \in Lip([a_1,a_2]), \int_{a_1}^{a_2} x^j g(x)dx = \begin{cases} 1, & \text{if } j = 0 \\ 0, & \text{if } 1 \le j \le r-1 \end{cases} \right\},$$

where $Lip([a_1,a_2])$ denotes the space of Lipschitz continuous functions on $[a_1,a_2]$. Define $K_+(\cdot,\cdot)$ and $K_-(\cdot,\cdot)$ as follows:

(i) The support of $K_+(x,q')$ is $[-1,q'] \times [0,1]$ and the support of $K_-(\cdot,\cdot)$ is $[-q',1] \times [0,1]$.

(ii) $K_+(\cdot,q') \in \mathcal{M}_{\theta,r}([-1,q'])$ and $K_-(\cdot,q') \in \mathcal{M}_{\theta,r}([-q',1])$. We note that $K_+(\cdot,1') = K_-(\cdot,1) = K(\cdot) \in \mathcal{M}_{\theta,r}([-1,1])$. Now let

$$K_h(u,t) = \begin{cases} K_+(u,1), & \text{if } h \le t \le 1-h \\ K_+(\frac{u}{h},\frac{t}{h}), & \text{if } 0 \le t \le h \\ K_-(\frac{u}{h},\frac{1-t}{h}), & \text{if } 1-h \le t \le 1 \end{cases}$$

We can show that $K_h(\cdot,\cdot)$ is a generalized kernel function of order r.

Assume the bandwidth parameter b_{ij} satisfies $O(T^{-\frac{1}{2}}) < b_{ij} < O(T^{-\frac{1}{5}}) \forall i,j = 1,2,\ldots,p$. The estimator $\widehat{v}_t^*$ can be expressed as

$$\widehat{v}_t^* = \widehat{\Omega}_{vt}^{-\frac{1}{2}} \widehat{v}_t \tag{7}$$

where

$$\widehat{\Omega}_{vt} = \begin{pmatrix} \widehat{h}_{11}\left(\frac{t}{T}\right) & \widehat{h}_{12}\left(\frac{t}{T}\right) & \cdots & \widehat{h}_{1p}\left(\frac{t}{T}\right) \\ \widehat{h}_{21}\left(\frac{t}{T}\right) & \widehat{h}_{22}\left(\frac{t}{T}\right) & \cdots & \widehat{h}_{2p}\left(\frac{t}{T}\right) \\ \vdots & \vdots & & \vdots \\ \widehat{h}_{p1}\left(\frac{t}{T}\right) & \widehat{h}_{p2}\left(\frac{t}{T}\right) & \cdots & \widehat{h}_{pp}\left(\frac{t}{T}\right) \end{pmatrix}$$

$$\widehat{h}_{ij}\left(\frac{t}{T}\right) = \sum_{\tau=1}^{T} w_{t\tau ij} \widehat{v}_{i\tau} \widehat{v}_{j\tau}, \quad w_{t\tau ij} = \left(\sum_{\tau=1}^{T} K_{b_{ij}}\left(\frac{t-\tau}{T}\right) \right)^{-1} K_{b_{ij}}\left(\frac{t-\tau}{T}\right)$$

Following [16,17], we can have the following convergence results under the above generalized kernel estimator,

$$\begin{aligned} &(1) \quad \widehat{v}_t - v_t = O_p\left(T^{-\frac{1}{2}}\right) \\ &(2) \quad \widehat{h}_{ij}(r) - h_{ij}^*(r) = O_p(b_{ij}^2) + O_p\left(T{b_{ij}}^{-\frac{1}{2}}\right), \text{ uniformly in } r \in [0,1] \\ &(3) \quad \widehat{v}_t^* - v_t^* = o_p\left(T^{-\frac{1}{4}}\right) \end{aligned} \tag{8}$$

A feasible estimator $(\widetilde{\beta}, \widetilde{\gamma})$ can be obtained by the following three-step approach.

Step 1: Estimate (1) by OLS, obtain a consistent estimator of v_t, $\widehat{v}_t$.

Step 2: Apply the generalized kernel estimator to estimate the time-varying variance–covariance matrix Ω_{vt}, and calculate $\widehat{v}_t^*$ as $\widehat{\Omega}_{vt}\widehat{v}_t$. To estimate the time-varying variance–

covariance matrix, we use a fourth-order Epanechnikov kernel for improved boundary performance and reduced bias, as recommended in [16]. The bandwidths $b_{ij}(T)$ were selected using leave-one-out cross-validation.

Step 3: Replace v_t^* in (5) with $\widehat{v}_t^*$ and estimate the new model by OLS.

Follow Lemma 5.1 in [18], the feasible approach can consistently estimate β and γ. To make correct inference on β, however, we need to account for the generated regressors problem due to the estimation of δ and Ω_{vt}. The following Corollary 1 provides the asymptotic distribution of the feasible estimators.

Corollary 1. *The feasible estimators* $(\widetilde{\beta}, \widetilde{\gamma})$ *have the following asymptotic distribution*

$$\sqrt{T}\begin{pmatrix}\widetilde{\beta}-\beta_0\\ \widetilde{\gamma}-\gamma_0\end{pmatrix} \xrightarrow{d} N(0, \Gamma_1^{-1}\Lambda_1^*\Gamma_1^{-1}),$$

where

$$\begin{aligned}
\Lambda_1^* &= \lim_{T\to\infty}\frac{1}{T}\sum_{t=2}^{T} E\left(w_t s_t' - G_\delta M^{-1} z_{t-1} v_t + \sum_{i=1}^{p}\sum_{i\le j}^{p} D_{i,j,t}\left[v_{it}v_{jt} - h_{ij}\left(\frac{t}{T}\right) - G_\delta M^{-1} z_{t-1} v_t\right]\right)\\
&\quad\times\left(w_t s_t' - G_\delta M^{-1} z_{t-1} v_t + \sum_{i=1}^{p}\sum_{i\le j}^{p} D_{i,j,t}\left[v_{it}v_{jt} - h_{ij}\left(\frac{t}{T}\right) - G_\delta M^{-1} z_{t-1} v_t\right]\right)'\\
G_\delta &= \lim_{T\to\infty}\frac{1}{T}\sum_{t=2}^{T} E\left(\frac{\partial(y_t - x_t\beta_0 - \gamma_0 v_t^*)s_t'}{\partial\delta}\right)\\
M &= \lim_{T\to\infty}\frac{1}{T}\sum_{t=2}^{T} E(z_{t-1}z_{t-1}')\\
D_{i,j,t} &= E\left(\frac{\partial(y_t - x_t\beta_0 - \gamma_0 v_t^*)s_t'}{\partial h_{ij}(\frac{t}{T})}\right)
\end{aligned}$$

Following [18,19], the exact form of Λ_1^* *includes adjustment terms involving the Jacobians of* s_t *with respect to* δ_0 *and* $h_{ij}(r)$, $i, j = 1, \ldots, p$.

3.2. Testing Endogeneity Without Instruments

Theorem 1 also provides a useful result for testing endogeneity. Given the specification in (5), $\gamma_i = \rho_i\sigma_u$; thus, when there exists no endogeneity, $\rho_i = \gamma_i = 0$. We can propose an extended version of the Hausman–Wu test [20,21] without valid instrument variables.

Corollary 2. *Under the same assumptions in Theorem 1, a test for endogeneity can be established as*

$$H_0 : \gamma_0 = 0,$$

where the test statistics $\sqrt{T}\widehat{\gamma}'(\Gamma_1^{-1}\Lambda_1\Gamma_1^{-1})_*^{-1}\widehat{\gamma} \overset{a}{\sim} \chi_p^2$, $(\Gamma_1^{-1}\Lambda_1\Gamma_1^{-1})_*$ *is the* $p \times p$ *bottom right block of the* $\Gamma_1^{-1}\Lambda_1\Gamma_1^{-1}$ *in Theorem 1.*

It is worth noting that the proposed test in Corollary 2 does not suffer from the generated regressors problem. This is because when $\rho = \gamma = 0$, (1) can be directly estimated by OLS and thus no extra uncertainty is introduced.

3.3. Inference and Robustness

The feasible estimator is $\sqrt{T}$-consistent and asymptotically normal. Since the model is just-identified, no GMM weighting matrix is required, and the estimator is semiparametrically efficient under the maintained assumptions. The use of generalized kernels ensures uniform convergence of the volatility estimates, even near boundaries. As discussed in [22–24], smooth time-varying volatility satisfies the regularity conditions necessary for

applying central limit theorems in heteroskedastic environments. This justifies the use of Wald-type inference.

Unlike conventional IV-GMM estimators, which are sensitive to weak instruments, our approach remains robust due to identification through time-varying volatility. The asymptotic distribution of our estimator is derived directly from moment conditions that remain informative even when the instruments are weak. As such, Wald-type inference remains valid under our framework, as long as the volatility process satisfies smoothness and regularity conditions. This distinction is critical in applications where traditional identification strategies are compromised.

4. Elasticity of Intertemporal Substitution: Consumption CAPM

The EIS is a central parameter in macroeconomics and asset pricing. It governs the intertemporal allocation of consumption and plays a crucial role in determining equilibrium asset returns. In recursive utility models, such as Epstein–Zin [25,26], the EIS appears as a structural determinant of households' willingness to substitute consumption over time in response to changes in the real return rate.

Empirically, the estimation of the EIS has produced mixed results. Using asset return data, studies such as [7] find small values of the EIS, while the long-run risk model in [8] requires a high EIS to match observed risk premia. This controversy has been partly attributed to the weakness of conventional instruments used to address endogeneity in return-predictability regressions (see [9,10]).

We apply the proposed control function approach to estimate the EIS using quarterly aggregate stock return data without valid instruments. This empirical setting is particularly well suited to the method for two reasons. First, the estimation of EIS involves a canonical predictive regression that is widely known to suffer from weak predictive power. Second, stock returns exhibit well-documented time-varying volatility, which is precisely the source of identification exploited by our framework. Thus, the EIS application offers both policy relevance and methodological transparency.

Following standard consumption-based asset pricing theory, we start from the log-linearized Euler equation implied by the Epstein–Zin utility specification:

$$E_t \Delta c_{t+1} = \tau_i + \psi E_t r_{i,t+1}, \tag{9}$$

where Δc_{t+1} is log consumption growth, $r_{i,t+1}$ is the log return on asset i, ψ is the EIS, and τ_i is a term that depends on the conditional variances and covariances of returns and consumption growth. We assume ψ to be constant over time to compare our results with [10]. Empirically, EIS could be time-varying; we leave this for future research. In empirical work, expectations are replaced by realized values, and we get the following regression:

$$\Delta c_{t+1} = \tau_i + \psi r_{i,t+1} + u_{i,t+1} \tag{10}$$

The variable $r_{i,t+1}$ represents the return on asset i. It may be correlated with the error term $u_{i,t+1}$, making it endogenous in the regression. This correlation can arise for several reasons. First, both consumption growth and asset returns are influenced by common shocks to economic fundamentals. Second, forward-looking behavior by investors can cause returns to respond to information that also affects consumption decisions. As a result, estimating (10) using ordinary least squares would lead to biased and inconsistent estimates of the EIS.

To address this issue, researchers typically use instrumental variable methods. These instruments are constructed from lagged economic variables that are assumed to be correlated with returns but uncorrelated with the error term. Standard choices include past

observations of interest rates, inflation, consumption growth, and the dividend–price ratio. However, as emphasized by [7,27], returns are difficult to predict in practice, and these instruments often have little explanatory power. This weak relationship between instruments and returns leads to unreliable estimates and very wide confidence intervals, which limit the usefulness of the results.

In line with the literature on weak instruments (e.g., [13]), we examined the first-stage relationship by estimating the reduced-form return prediction:

$$r_{i,t+1} = z_t'\delta + v_{i,t+1}, \tag{11}$$

where z_t is a set of instruments that consists of lagged values of nominal interest rates, inflation, consumption growth, and the log dividend–price ratio, each lagged twice to mitigate time aggregation bias, following [28]. The data are quarterly and span various periods from 1947 to 1999, depending on the country. Stock returns are constructed from major stock market indices for eleven developed countries, and log consumption growth is calculated from real per capita consumption expenditures (for a full description of the data set, see [7,29]). The third and fourth columns in Table 1 report corresponding F-statistics and partial R^2 values for each country. These statistics confirm the weakness of the instruments, with F-statistics well below conventional thresholds and low partial R^2 across the board. This supports the motivation for using a volatility-based identification strategy, as standard IV inference would likely suffer from low power and misleading coverage in such settings.

Table 1. Supporting evidence for the proposed identification strategy.

Country	Sample Period	F-Statistics	Partial R^2	$Corr(r_t, v_t^*)$
AUL	1970.3–1998.4	1.82	0.06	0.81
CAN	1970.3–1999.1	2.51	0.08	0.83
FR	1970.3–1998.3	3.1	0.10	0.81
GER	1979.1–1998.3	0.69	0.04	0.85
ITA	1971.4–1998.1	1.10	0.04	0.89
JAP	1970.3–1998.4	3.49	0.11	0.85
NTH	1977.3–1998.4	0.73	0.03	0.84
SWD	1970.3–1999.2	2.24	0.07	0.84
SWT	1976.2–1998.4	0.11	0.01	0.83
UK	1970.3–1999.1	2.62	0.09	0.80
USA	1947.3–1998.4	2.88	0.05	0.85

Note: This table reports the F-statistics and partial R^2 of the first-stage regressions and the correlations between r_t and v_t^*.

We estimate (10) using the proposed control function approach. The reduced-form residuals $\widehat{v}_{i,t+1}$ are obtained from (11), and standardized using an estimated conditional variance matrix $\widehat{\Omega}_{vt}$ constructed via the kernel estimator described in Section 3. The standardized residuals $\widehat{v}^*_{i,t+1} = \widehat{\Omega}_{vt}^{-1/2}\widehat{v}_t$ are then used as control functions in the second-stage regression,

$$\Delta c_{t+1} = \tau_i + \psi r_{i,t+1} + \gamma \widehat{v}^*_{i,t+1} + w_{i,t+1} \tag{12}$$

To estimate the time-varying volatility of returns, we apply the generalized kernel estimator described in Section 3.1. We use a fourth-order Epanechnikov kernel for improved boundary performance and reduced bias, as recommended in [16]. While the generalized kernel is theoretically more appealing, we found that the empirical results are robust to alternative kernel choices, including the quartic and Gaussian kernels. The bandwidths $b_{ij}(T)$ were selected using leave-one-out cross-validation using each country-specific sam-

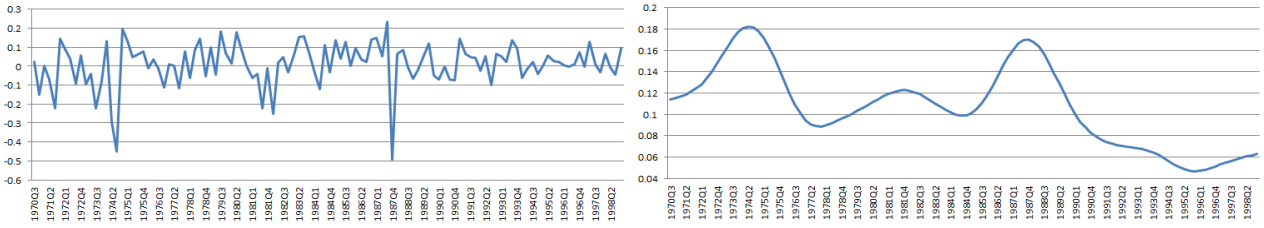
ple. The volatilities of returns exhibit significant time variation across countries, as shown in Figures 1 and 2. This provides empirical support for the identifying assumption of time-varying volatility. Furthermore, we calculate the correlation between between r_t and v_t^* to check whether the multicollinearity is avoided. The last column in Table 1 reports the corresponding correlations between r_t and v_t^* for each country. The correlations are all substantially below unity, consistent with the theoretical argument that time-varying volatility restores identification.

Table 2 reports confidence intervals for the EIS estimated using our control function method and compares them to intervals derived from three commonly used weak-instrument robust tests: the Anderson–Rubin [30], Kleibergen-LM [31], and Conditional Likelihood Ratio [32,33], as reported in [10]. In most cases, the confidence intervals obtained by three weak-instruments robust tests are extremely wide or unbounded, reflecting the limited power of conventional weak-IV methods. In contrast, our approach obtains narrower, informative intervals. The estimated EIS values are uniformly modest across countries, roughly in the -0.05 to 0.3 range, indicating a limited willingness of households to shift consumption over time. In economic terms, an EIS of this magnitude implies a strong preference for consumption smoothing. These findings are aligned with prior empirical findings that the EIS is typically low (e.g., [7,28]), standing in contrast to the much larger values (EIS well above 1) required by certain asset-pricing frameworks like long-run risk models [8]. Moreover, our approach delivers considerably tighter confidence intervals than standard weak-IV methods, reflecting a substantial gain in identification power. The improved precision suggests that by exploiting time-varying volatility, we can estimate the EIS with much greater confidence. These results demonstrate the empirical relevance of the theoretical framework developed in earlier sections.

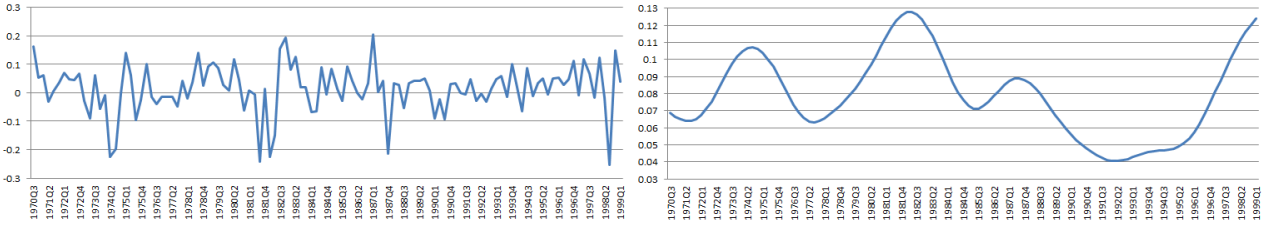
Table 2. Confidence interval based on the control function approach and the weak-instrument robust tests.

Country	Sample Period	CF	AR	LM	CLR
AUL	1970.3–1998.4	$[-0.11, 0.14]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
CAN	1970.3–1999.1	[0.03, 0.28]	[0.02, 4.03]	[0.05, 0.35]	[0.04, 0.41]
FR	1970.3–1998.3	$[-0.03, 0.15]$	$[-0.28, 0.20]$	$[-\infty, \infty]$	$[-0.16, 0.11]$
GER	1979.1–1998.3	$[-0.01, 0.19]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
ITA	1971.4–1998.1	$[-0.03, 0.17]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
JAP	1970.3–1998.4	$[-0.05, 0.19]$	$[-0.05, 0.32]$	$[-1.01, 0.20]$	$[-0.02, 0.21]$
NTH	1977.3–1998.4	[0.03, 0.19]	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
SWD	1970.3–1999.2	$[-0.01, 0.07]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
SWT	1976.2–1998.4	$[-0.04, 0.21]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
UK	1970.3–1999.1	[0.02, 0.29]	$[-0.51, -0.02]$	$[-\infty, \infty]$	$[-\infty, \infty]$
USA	1947.3–1998.4	[0.11, 0.27]	$[-0.21, -0.02]$	$[-\infty, \infty]$	$[-\infty, \infty]$

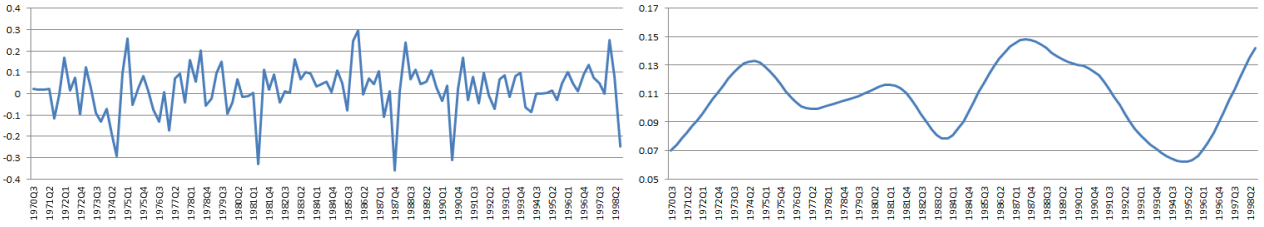
Note: This table reports the confidence interval of EIS estimated by the proposed approach and the weak-instrument robust tests. CF refers to the proposed estimator. AR refers to the Anderson–Rubin test. LM refers to the Kleibergen-LM test. CLR refers to the Conditional Likelihood Ratio test.



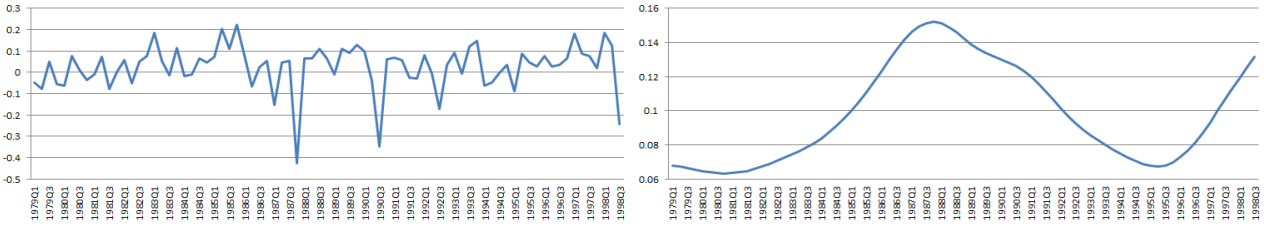
Australia: stock return (**left**) and estimated volatility (**right**)



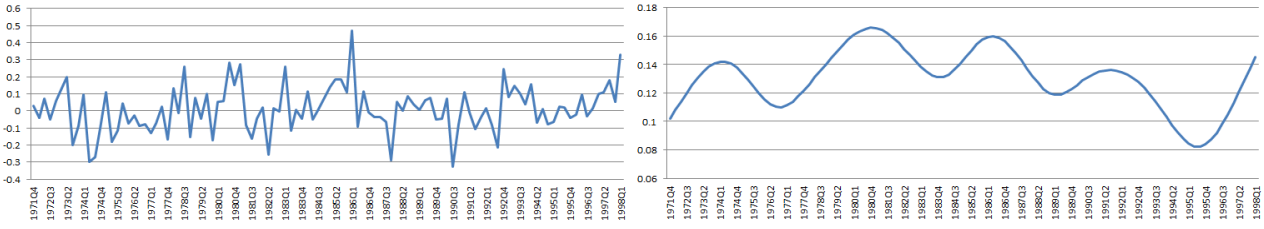
Canada: stock return (**left**) and estimated volatility (**right**)



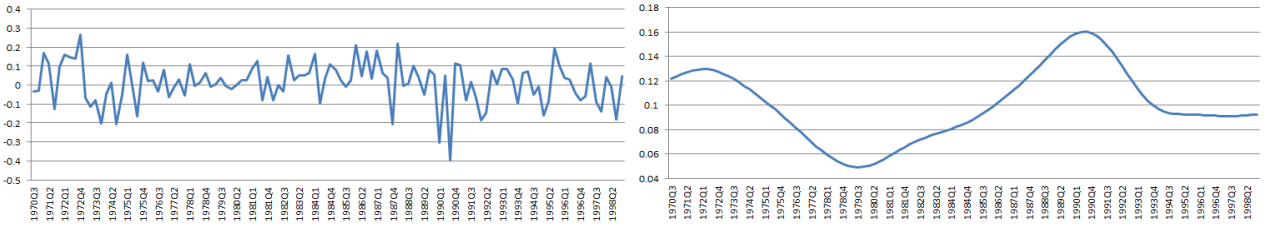
France: stock return (**left**) and estimated volatility (**right**)



Germany: stock return (**left**) and estimated volatility (**right**)

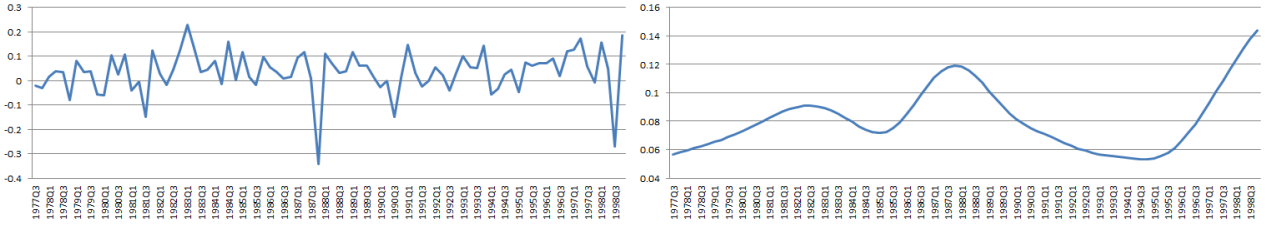


Italy: stock return (**left**) and estimated volatility (**right**)

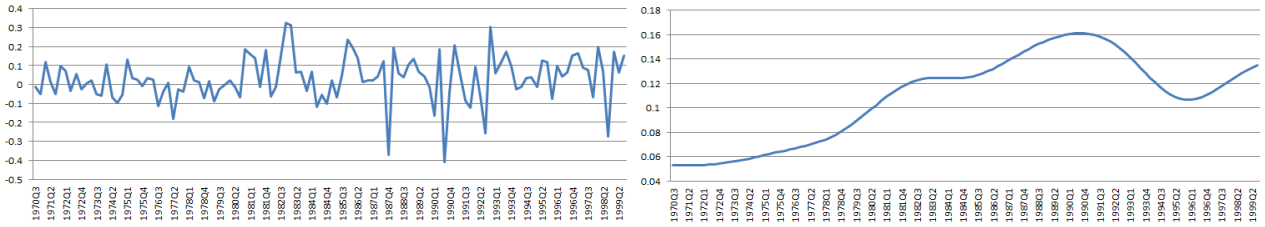


Japan: stock return (**left**) and estimated volatility (**right**)

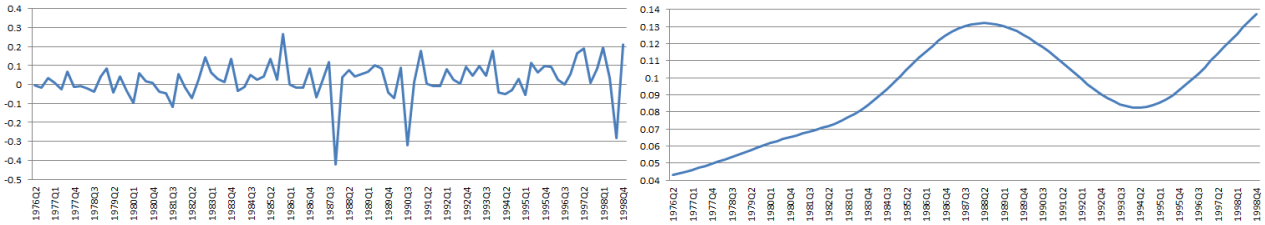
Figure 1. Aggregate stock return and estimated volatility.



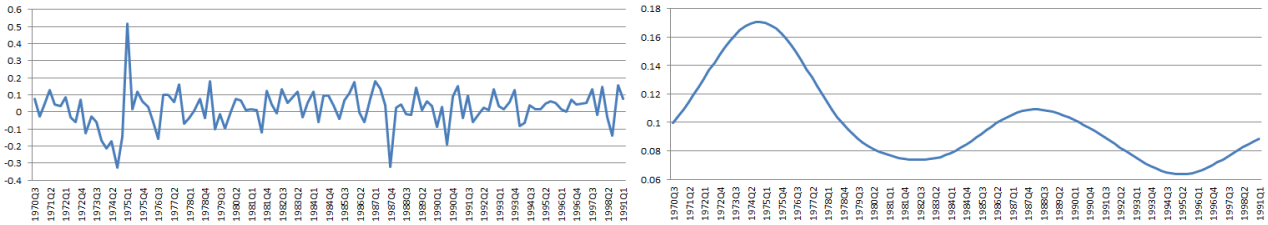
Netherlands: stock return (**left**) and estimated volatility (**right**)



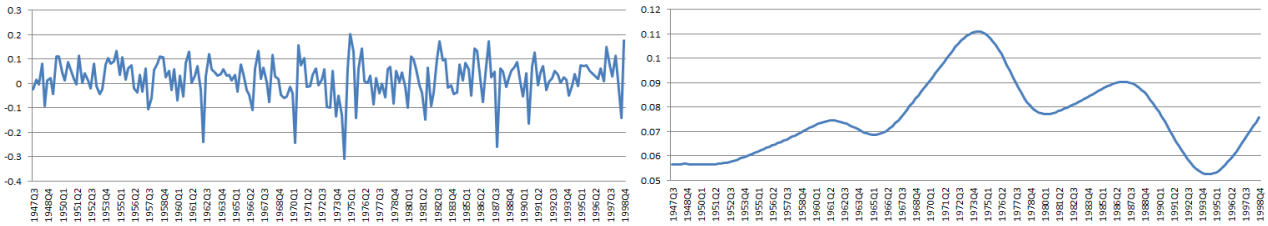
Sweden: stock return (**left**) and estimated volatility (**right**)



Switzerland: stock return (**left**) and estimated volatility (**right**)



United Kingdom: stock return (**left**) and estimated volatility (**right**)



United States: stock return (**left**) and estimated volatility (**right**)

Figure 2. Aggregate stock return and estimated volatility.

5. Concluding Remarks

This paper develops a feasible control function approach for identifying and estimating linear models with endogenous regressors without valid instruments. The method exploits time-varying volatility in the reduced-form errors as a source of identification, offering an alternative when conventional instrumental variable techniques are infeasible or unreliable.

The main contribution lies in demonstrating that conditional heteroskedasticity can restore identification even when valid instrumental variables are unavailable. By projecting the structural error term onto the reduced-form error terms and standardizing by the time-varying variance, the proposed estimator isolates exogenous variation that would

otherwise be obscured by collinearity. Unlike approaches that rely on strong instruments, known heteroskedastic regimes [3], or parametric volatility models [4–6], our framework allows for smooth, nonparametric variation in second moments and does not require external exclusion restrictions.

We formally establish that the estimator is $\sqrt{T}$-consistent and asymptotically normal. Importantly, we show that the estimator remains valid under weak identification and heteroskedasticity, provided the volatility process satisfies mild regularity conditions. We also derive a test for endogeneity that does not require valid instruments, extending the logic of the Hausman–Wu test to settings where standard IV methods fail.

In an empirical application, we estimate the EIS using quarterly stock return and consumption data from eleven developed countries. The data exhibit substantial time variation in return volatility, consistent with the identifying assumptions. The proposed method produces confidence intervals that are narrower and more stable than those derived from conventional weak-instrument robust procedures, while remaining consistent with existing findings in the literature.

The results demonstrate both the empirical relevance and practical applicability of using heteroskedasticity as a source of identification. The approach provides a flexible alternative to instrument-based methods in settings where external instruments are weak, contested, or unavailable. More broadly, this framework suggests that structured time variation in second moments can offer credible identification strategies in a range of economic applications.

Funding: This research received no external funding.

Data Availability Statement: Data are available in the public domain: https://drive.google.com/file/d/0BzR-ojpYuaFMaUJxd0RNSk84a00/view?resourcekey=0-XRMv1V1xJjM5JrOj-TB51Q (accessed on 20 May 2025).

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

Proof of Lemma 1. By construction, $w_t \equiv u_t - v_t^{*\prime}\gamma_t = u_t - v_t^{*\prime}\rho\sigma_{ut}$.

For (1), it follows immediately from Assumption 1.

For (2),

$$\begin{aligned} E(x_t w_t|\mathcal{F}_{t-1}) &= E\Big(((I_p \otimes z_{t-1}')\delta_0 + v_t)\Big(u_t - v_t^{*\prime}\rho\sigma_{ut}\Big)|\mathcal{F}_{t-1}\Big) \\ &= E\Big(v_t\Big(u_t - v_t^{*\prime}\rho\sigma_{ut}\Big)|\mathcal{F}_{t-1}\Big) \\ &= E(v_t u_t|\mathcal{F}_{t-1}) - E\Big(v_t v_t^{*\prime}|\mathcal{F}_{t-1}\Big)\rho\sigma_{ut} \\ &= \Big(\Omega_{vt}^{-\frac{1}{2}} - \Omega_{vt}^{-\frac{1}{2}} E\Big(v_t^* v_t^{*\prime}|\mathcal{F}_{t-1}\Big)\Big)\rho\sigma_{ut} \\ &= 0 \end{aligned}$$

where the second equation follows by Assumption 1, and the last follows $E\big(v_t^* v_t^{*\prime}|\mathcal{F}_{t-1}\big) = I_p$ by construction.

Proofs for (3) and (4) follow the same logic as (2) and thus are omitted. □

Proof of Theorem 1. The proof of Theorem 1 proceeds via a sequence of technical lemmas, presented below.

Lemma A1.

$$\frac{1}{T}\sum_{t=1}^{T} s_t' s_t \xrightarrow{p} \begin{pmatrix} \int \Omega & \int \Omega^{\frac{1}{2}} \\ \int \Omega^{\frac{1}{2}} & 1 \end{pmatrix}$$

Proof. To prove convergence of the sample second moment matrix, it suffices to establish the following:

(a) $\frac{1}{T}\sum_{t=1}^{T} x_t x_t' \xrightarrow{p} \int \Omega$.

(b) $\frac{1}{T}\sum_{t=1}^{T} x_t v_t^{*\prime} \xrightarrow{p} \int \Omega^{\frac{1}{2}}$.

For (a),

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} x_t x_t' = \frac{1}{T}\sum_{t=1}^{T}(I_p \otimes z_{t-1}')\delta\delta'(I_p \otimes z_{t-1}')' + \frac{1}{T}\sum_{t=1}^{T}(I_p \otimes z_{t-1}')\delta(\Omega^{\frac{1}{2}} v_t^*)' \\ + \frac{1}{T}\sum_{t=1}^{T}(\Omega^{\frac{1}{2}} v_t^*)\delta'(I_p \otimes z_{t-1}')' + \frac{1}{T}\sum_{t=1}^{T}(\Omega^{\frac{1}{2}} v_t^*)(\Omega^{\frac{1}{2}} v_t^*)'\end{aligned} \tag{A1}$$

For the first term, IVs (2) and (3) ensure the following convergence result,

$$\frac{1}{T}\sum_{t=1}^{T}(I_p \otimes z_{t-1}')\delta\delta'(I_p \otimes z_{t-1}')' = \frac{1}{T^2}\sum_{t=1}^{T}(I_p \otimes z_{t-1}')CC'(I_p \otimes z_{t-1}')' \xrightarrow{p} 0 \tag{A2}$$

For the second and third terms, we apply the orthogonality conditions implied by Assumption 2(1),

$$\frac{1}{T}\sum_{t=1}^{T}(I_p \otimes z_{t-1}')\delta(\Omega^{\frac{1}{2}} v_t^*)' \xrightarrow{p} 0 \tag{A3}$$

$$\frac{1}{T}\sum_{t=1}^{T}(\Omega^{\frac{1}{2}} v_t^*)\delta'(I_p \otimes z_{t-1}')' \xrightarrow{p} 0 \tag{A4}$$

For the fourth term, by the law of large numbers for martingale difference sequence and TV,

$$\frac{1}{T}\sum_{t=1}^{T}(\Omega^{\frac{1}{2}} v_t^*)(\Omega^{\frac{1}{2}} v_t^*)' \xrightarrow{p} \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E\left((\Omega^{\frac{1}{2}} v_t^*)(\Omega^{\frac{1}{2}} v_t^*)'\right) = \int \Omega$$

Thus, the desired result can be obtained by combining (A1) to (A4).

For (b),

$$\frac{1}{T}\sum_{t=1}^{T} x_t v_t^{*\prime} = \frac{1}{T}\sum_{t=1}^{T}(I_p \otimes z_{t-1}')v_t^{*\prime} + \frac{1}{T}\sum_{t=1}^{T}\Omega^{\frac{1}{2}} v_t^* v_t^{*\prime} \tag{A5}$$

The first term converges to 0 following the same argument in (a). For the second term,

$$\frac{1}{T}\sum_{t=1}^{T}\Omega^{\frac{1}{2}} v_t^* v_t^{*\prime} \xrightarrow{p} \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E\left(\Omega^{\frac{1}{2}} v_t^* v_t^{*\prime}\right) = \int \Omega^{\frac{1}{2}}$$

□

Lemma A2.

(a) $\frac{1}{T}\sum_{t=1}^{T}\left(v_t w_t^2 v_t' - E\left(v_t w_t^2 v_t'\right)\right) \xrightarrow{p} 0$

(b) $\frac{1}{T}\sum_{t=1}^{T}\left(v_t w_t^2 v_t^{*\prime} - E\left(v_t w_t^2 v_t^*\right)\right) \xrightarrow{p} 0$

(c) $\frac{1}{T}\sum_{t=1}^{T}\left(v_t^* w_t^2 v_t^{*\prime} - E\left(v_t^* w_t^2 v_t^{*\prime}\right)\right) \xrightarrow{p} 0$

Proof. For (a),

$$v_t w_t^2 v_t' = \begin{bmatrix} v_{1t}^2 w_t^2 & v_{1t}v_{2t}w_t^2 & \dots & v_{1t}v_{pt}w_t^2 \\ v_{2t}v_{1t}w_t^2 & v_{2t}^2 w_t^2 & \dots & v_{2t}v_{pt}w_t^2 \\ \vdots & \vdots & \dots & \vdots \\ v_{pt}v_{1t}w_t^2 & v_{pt}v_{2t}w_t^2 & \dots & v_{pt}^2 w_t^2 \end{bmatrix}$$

Since v_{it} and w_t are α-mixing by Theorem 14.1 in [34], thus $v_{it}v_{jt}$ and w_t^2 are also α-mixing and near-epoch dependent in L^2 norm on the α-mixing sequence $\{v_t^*, u_t^*\}$. So by Theorem 17.9 in [34], $v_{it}v_{jt}w_t^2$ is near-epoch dependent in L^1 norm on $\{v_t^*, u_t^*\}$. Let $\alpha = \min\{\alpha_1, \alpha_{21}, \dots, \alpha_{2p}\}$, it follows that

$$E\left[v_{it}v_{jt}w_t^2\right]^{\alpha} \leq E v_{it}^{4\alpha} E v_{jt}^{4\alpha} E w_t^{4\alpha} < \infty$$

by Assumption 1 and the Cauchy–Schwarz inequality. Thus, by the law of large numbers for L^1-mixingales [35], we have the desired results. Proofs for (b) and (c) follow the same logic and thus are omitted. □

Lemma A3.

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} s_t' w_t \xrightarrow{d} N(0, \Omega_2)$$

Proof. Following the Cramer–Wold theorem, it suffices to show that for any fixed vector $\lambda \in \mathbb{R}^{2p}$,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \lambda' s_t' w_t \xrightarrow{d} N(0, \lambda' \Omega_2 \lambda) \tag{A6}$$

for every fixed $2p \times 1$ vector $\lambda = (\lambda_1, \lambda_2)'$, where $\lambda_1 = (\lambda_{11}, \lambda_{12}, \dots, \lambda_{1p})$, $\lambda_2 = (\lambda_{21}, \lambda_{22}, \dots, \lambda_{2p})$.

First,

$$\begin{aligned} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \lambda' s_t' w_t &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\lambda_1' x_t w_t + \lambda_2' v_t^* w_t\right) \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\lambda_1' (I_p \otimes z_{t-1}') \delta w_t + \lambda_1' v_t w_t + \lambda_2' v_t^* w_t\right) \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\lambda_1' v_t w_t + \lambda_2' v_t^* w_t\right) + o_p(1) \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \lambda' s_t^{*\prime} w_t + o_p(1) \end{aligned}$$

where $s_t^* \equiv (v_t, v_t^*)$, and the third equation follows from Assumption 2(1) and 2(3).

Second, note that $\{\lambda' s_t^{*\prime} w_t, \mathcal{F}_t\}$ is a martingale difference sequence by Lemma 1. Then the following convergence results follow Lemma A2

$$\frac{1}{T} \sum_{t=1}^{T} \lambda' w_t^2 s_t^{*\prime} s_t^* \lambda \xrightarrow{p} \lambda' \Lambda \lambda$$

Moreover, we have

$$\begin{aligned} E(\lambda' s_t^* w_t)^{2\alpha} &= \|\lambda_1' v_t w_t + \lambda_2' v_t^* w_t\|_{2\alpha}^{2\alpha} \\ &\leq \left(\sum_{i=1}^{p} \|\lambda_{1i} v_{ti} w_t\|_{2\alpha} + \sum_{i=1}^{p} \|\lambda_{2i} v_{ti}^* w_t\|_{2\alpha}\right)^{2\alpha} \end{aligned} \tag{A7}$$

where the inequality follows from the Minkowski inequality. □

Each term in the bound (A7) is finite due to the assumed moment conditions and mixing properties,

$$\|\lambda_{1i} v_{ti} w_t\|_{2\alpha} = |\lambda_{1i}| (E(v_{it} w_t)^{2\alpha})^{\frac{1}{2\alpha}} \leq |\lambda_{1i}| (E v_{it}^{4\alpha} E w_t^{4\alpha})^{\frac{1}{4\alpha}} < \infty$$

$$\|\lambda_{2i} v_{ti}^* w_t\|_{2\alpha} = |\lambda_{2i}| (E(v_{it}^* w_t)^{2\alpha})^{\frac{1}{2\alpha}} \leq |\lambda_{2i}| (E v_{it}^{*\,4\alpha} E w_t^{4\alpha})^{\frac{1}{4\alpha}} < \infty$$

where the inequality follows from the Cauchy–Schwarz inequality.

So, we can conclude that (A7) is finite, by the central limit theorem for the martingale differences, (A6) holds, and then we have the desired results.

The proof of Theorem 1 follows Lemma A1–A3. □

References

1. Galvao, A.F.; Montes-Rojas, G. On the Equivalence of Instrumental Variables Estimators for Linear Models. *Econ. Lett.* **2015**, *134*, 13–15. [CrossRef]
2. Kim, C.-J.S. Dealing with Endogeneity in Regression Models with Dynamic Coefficients. *Found. Trends Econ.* **2008**, *3*, 165–266. [CrossRef]
3. Rigobon, R. Identification Through Heteroskedasticity. *Rev. Econ. Stat.* **2003**, *85*, 777–792. [CrossRef]
4. Hwu, S.-T.; Kim, C.-J. Estimating Trend Inflation Based on Unobserved Components Models: Is It Correlated with the Inflation Gap? *J. Money Credit Bank.* **2019**, *51*, 2305–2319. [CrossRef]
5. Bertsche, D.; Braun, R. Identification of Structural Vector Autoregressions by Stochastic Volatility. *J. Bus. Econ. Stat.* **2022**, *40*, 328–341. [CrossRef]
6. Hertwartz, H.; Lütkepohl, H. Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks. *J. Econ.* **2014**, *183*, 104–116.
7. Campbell, J.Y. Consumption-based asset pricing. *Handb. Econ. Financ.* **2003**, *1*, 803–887. [CrossRef]
8. Bansal, R.; Yaron, A. Risks for the long run: A potential resolution of asset pricing puzzles. *J Financ.* **2004**, *59*, 1481–1509. [CrossRef]
9. Neely, C.J.; Roy, A.; Whiteman, C.H. Risk aversion versus intertemporal substitution: A case study of identification failure in the intertemporal consumption capital asset pricing model. *J. Bus. Econ. Stat.* **2001**, *19*, 395–403. [CrossRef]
10. Yogo, M. Estimating the elasticity of intertemporal substitution when instruments are weak. *Rev. Econ. Stat.* **2004**, *86*, 797–810. [CrossRef]
11. Schwert, G.W. Why Does Stock Market Volatility Change Over Time? *J. Financ.* **1989**, *44*, 1115–1153. [CrossRef]
12. Hahn, J.; Hausman, J.A. A New Specification Test for the Validity of Instrumental Variables. *Econometrica* **2002**, *70*, 163–189. [CrossRef]
13. Staiger, D.; Stock, J.H. Instrumental variables regression with weak instruments. *Econometrica* **1997**, *65*, 557–586. [CrossRef]
14. Hull, J.; White, A. The Pricing of Options on Assets with Stochastic Volatilities. *J. Financ.* **1987**, *42*, 281–300. [CrossRef]
15. Harvey, A.C.; Ruiz, E.; Shephard, N. Multivariate Stochastic Variance Models. *Rev. Econ. Stud.* **1994**, *61*, 247–264. [CrossRef]
16. Müller, H.-G. Smooth optimum kernel estimators near endpoints. *Biometrika* **1991**, *78*, 521–530. [CrossRef]
17. Feng, Y. Simultaneously Modeling Conditional Heteroskedasticity and Scale Change. *Econ. Theory* **2004**, *20*, 563–596. [CrossRef]
18. Newey, W.K. The Asymptotic Variance of Semiparametric Estimators. *Econometrica* **1994**, *62*, 1349–1382. [CrossRef]
19. Pagan, A. Econometric issues in the analysis of regressions with generated regressors. *Int. Econ. Rev.* **1984**, *65*, 557–586. [CrossRef]
20. Hausman, J. Specification tests in econometrics. *Econometrica* **1978**, *46*, 1251–1271. [CrossRef]
21. Wu, D. Alternative tests of independence between stochastic regressors and disturbances. *Econometrica* **1973**, *41*, 733–750. [CrossRef]
22. Cavaliere, G. Unit root tests under time-varying variance shifts. *Econ. Rev.* **2004**, *23*, 259–292. [CrossRef]
23. Phillips, P.C.B.; Xu, K.-L. Inference in Autoregression under Heteroskedasticity. *J. Time Ser. Anal.* **2006**, *27*, 289–308. [CrossRef]
24. Xu, K.-L.; Phillips, P.C.B. Adaptive Estimation of Autoregressive Models with Time-Varying Variances. *J. Econ.* **2008**, *142*, 265–280. [CrossRef]
25. Epstein, L.G.; Zin, S.E. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* **1989**, *57*, 937–969. [CrossRef]
26. Epstein, L.G.; Zin, S.E. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An Empirical Analysis. *J. Political Econ.* **1991**, *99*, 263–286. [CrossRef]
27. Cochrane, J.H. The Dog That Did Not Bark: A Defense of Return Predictability. *Rev. Financ. Stud.* **2008**, *21*, 1533–1575. [CrossRef]
28. Hall, R.E. Intertemporal Substitution in Consumption. *J. Political Econ.* **1991**, *96*, 339–357. [CrossRef]

29. Campbell, J. Data Appendix for 'Asset Prices, Consumption, and the Business Cycle'. *Unpublished Manuscript*. 1999. Available online: https://campbell.scholars.harvard.edu//publications/data-appendix-asset-prices-consumption-and-business-cycle (accessed on 25 May 2025).
30. Anderson, T.W.; Rubin, H. Estimation of the parameters of a single equation in a complete system of stochastic equations. *Ann. Math. Stat.* **1949**, *20*, 46–63. [CrossRef]
31. Kleibergen, F. Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica* **2002**, *70*, 1781–1803. [CrossRef]
32. Moreira, M.J. A conditional Likelihood ratio test for structure models. *Econometrica* **2002**, *71*, 1027–1048. [CrossRef]
33. Moreira, M.J. Test with Correct size when instruments can be arbitrarily weak. *J Econ.* **2007**, *152*, 131–140. [CrossRef]
34. Davidson, J. *Stochastic Limit Theory*; Oxford University Press: Oxford, UK, 1994.
35. Andrew, D.W.K. Law of Large Numbers for Dependent Non-Identical Distributed Random Variables. *Econ. Theory* **1988**, *4*, 458–467. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.