INTERNATIONAL MONETARY FUND

Robust Inference Via Heteroskedasticity in Linear Models

Prepared by Omer Faruk Akbal and Max-Sebastian Dovi

WP/26/100

2026
MAY

WORKING PAPER

 WP/26/100

IMF Working Paper
Research Department

Robust Inference Via Heteroskedasticity in Linear Models
Prepared by Omer Faruk Akbal and Max-Sebastian Dovi

Authorized for distribution by Petia Topalova
May 2026

ABSTRACT: We study inference via heteroskedasticity in linear models commonly used for macroeconomic policy analysis, where covariate endogeneity must often be addressed with limited time and data. Our framework nests standard heteroskedasticity-based approaches, allows for new non-nested restrictions, and does not require ex-ante regime labelling. We propose an easily implementable weak-identification-robust test and derive sufficient conditions for its validity. Simulation results show good size and power properties in a wide range of settings. Empirical applications to the fuel-price passthrough in Sierra Leone, the effect of remittances on consumption in the Philippines, and exchange-rate passthroughs in many countries illustrate the versatility and scalability of our approach.

RECOMMENDED CITATION: Akbal, O. and M. Dovi. 2026. *Robust Inference Via Heteroskedasticity in Linear Models*. IMF Working Paper WP/26/100. Washington, DC: International Monetary Fund.

JEL Classification Numbers:	C26, C12, C22, E31, F31, F24 .
Keywords:	Heteroskedasticity; weak identification; Anderson-Rubin; two-step inference
Authors' email addresses:	farukakbal@gmail.com mdovi@imf.org

WORKING PAPERS

Robust Inference Via Heteroskedasticity in Linear Models[1]

Prepared by Omer Faruk Akbal[2] and Max-Sebastian Dovi[3]

[1] This work is supported by the Macroeconomic Research in Low-Income Countries program of the U.K.'s Foreign, Commonwealth and Development Office (FCDO) and the Macroeconomic Research on Climate Change and Emerging Risks in Asia program of the Ministry of Economy and Finance of the Government of Korea. The views expressed in this paper are those of the authors and should not be attributed to the IMF, its Executive Board, or IMF management. We thank seminar participants at the 2025 'Sail the Macro' conference, as well as seminar participants at the International Monetary Fund, in the Research Department (RES) and the Asia Pacific Department (APD), for their comments.

[2] International Monetary Fund, farukakbal@gmail.com.

[3] International Monetary Fund, mdovi@imf.org.

1 Introduction

Macroeconomic policy work requires timely assessment of causal relationships among many variables. This understanding is important for advising policymakers on how to influence outcomes. However, the many different policies that need to be evaluated coupled with the fast pace of current events often limit the time available for such assessments.

The linear model is the workhorse in applied macroeconometric policy research, but key explanatory variables are often endogenous. Linear models are popular in policy work because they can be interpreted and communicated easily (for recent examples see, e.g., ECB [2025], IMF [2025]). However, simultaneity, mismeasurement or omitted variables often cause explanatory variables to correlate with the model's error term, making OLS inconsistent.

Heteroskedasticity-based methods offer a promising way to address endogeneity in policy settings, but weak-identification robust techniques should be used. These methods achieve identification without requiring external data, which is attractive in policy settings for two reasons. First, time, data, and resources are often limited. Second, not requiring external data for identification implies that the same specification can be used for different units (e.g., countries), which improves comparability. At the same time, robustness to weak identification is important due to usual pre-testing concerns, compounded by the fact that diagnostic tests are not routinely reported in policy work. Moreover, in multi-country settings, even strong ex ante identification arguments for a single country may not hold uniformly across all countries, so weak-identification-robust methods provide a safeguard that avoids both pre-testing and multiple-testing distortions. [1]

We provide a framework for inference via heteroskedasticity in the standard linear model that nests several existing setups and provides new and non-nested restric-

[1] We refer to Lewis [2022] for a discussion of weak identification arising from higher-order moments of U.S. monetary policy shocks.

tions.[2] We show that our setup, adapted from Lewbel's (2012), includes Rigobon's (2003) and a version of Lewis's (2021) as a special case. Importantly, and similarly to Lewis [2021], our approach does not require researchers to manually label regimes ex-ante, which may be impractical in many policy settings, although these can be exploited if available. Furthermore, we show that it generates two identifying restrictions not foreseen by either. First, in Rigobon's (2003) setup, we show that imposing homoskedasticity on one of the errors or suitably restricting simultaneity makes it possible to weaken the assumption of uncorrelated error terms. Second, we show that our setting addresses mismeasurement-induced endogeneity.

We propose an easily implementable weak identification robust test and show by simulation that it performs well in existing setups, as well as in the new ones we consider. We achieve robustness to weak identification by exploiting Anderson and Rubin [1949] arguments and provide precise conditions sufficient for our test's asymptotic validity. Deriving the relevant test statistic and establishing its limiting distribution involves two technical complications. First, the variance of the test statistic has to be adjusted to account for the presence of generated regressors. Second, existing tools often used in the i.i.d. two-step estimation literature (e.g., Newey and McFadden [1994]) need to be adapted to the case of non-independently and non-identically distributed data. Simulations show that our proposed test has good size and power properties in Rigobon's (2003) setup, a setup similar to Lewis's (2021), Rigobon's (2003) setup with correlated errors, as well as a setup featuring mismeasured regressors.

We consider three macroeconometric applications that exploit different identifying restrictions covered by our approach.[3] First, we consider the passthrough of fuel

[2]As implied by Footnote 5 in Section 3 below, our discussion of identifying restrictions also applies to SVARs. However, since our proposed inference procedure is for coefficients in a linear model–and not objects that are typically of inferential interest in Structural Vector Autoregressions (SVAR) contexts, such as impulse response functions–we do not provide a detailed survey of the vast literature on identifying SVARs via higher moments. We refer to Lewis [2024] for a recent comprehensive review.

[3]While our focus is on macroeconometrics in policy settings, our approach can also find applicability in academic ones. For instance, limited-information inference on key linear structural

prices to headline inflation in Sierra Leone in Rigobon's (2003) framework where the timing of fuel price formula adjustments are used to restrict coefficients, and the error terms are potentially correlated. Second, we assess the effect of remittances–which are often mismeasured–on consumption growth in the Philippines by exploiting correlation between host countries' economic conditions and remittance volatility. Third, we consider the exchange-rate passthrough to inflation in a large panel of countries, illustrating the scalability of our method when manual labelling is infeasible.

The rest of this paper is organised as follows. Section 2 presents the general model we consider. Section 3 discusses how our setup relates to existing ones. Section 4 presents our proposed test and sufficient conditions for its asymptotic validity. Section 5 provides simulation evidence. Section 6 presents the empirical results. Section 7 concludes. All the proofs can be found in the appendix.

Some notation and useful definitions. We let $0_{m\times n}$ denote the $m \times n$ matrix of zeros. For any $a \in \mathbb{R}$, $\lfloor a \rfloor$ is the largest $b \in \mathbb{N}$ such that $b \leq a$. For any sequence of random vectors $\{U_t\}_{t=1}^T$, we let $\bar{\mathbb{E}}[U_t] \equiv \mathbb{E}[\frac{1}{T}\sum_{t=1}^T U_t]$. C denotes universal positive constants that can differ across instances (even within the same display). $||\cdot||_2$ denotes the Euclidean norm, and $||\cdot||_\infty$ denotes the maximum norm. For a sequence of random vectors $\{U_t\}$, let $\mathcal{B}_U(n, n+m)$ denote the Borel σ-field generated by $\{U_t, t = n, \ldots, n+m\}$, and denote the strong mixing coefficient by $\alpha_{mixing}(b) = \sup_a \sup_{G\in\mathcal{B}_U(-\infty,a), H\in\mathcal{B}_U(a+b,\infty)} |\mathbb{P}(G \cap H) - \mathbb{P}(G)\mathbb{P}(H)|$. We follow the convention of calling the sequence of random vectors $\{U_t\}$ as 'strong mixing of size $-d$' if $\alpha_{mixing}(b) = O(b^{-d-\omega})$ for some $\omega > 0$.

macroeconomic relations has become a popular approach to assess macroeconomic models (see, e.g., Galì and Gertler [1999], Kleibergen and Mavroeidis [2009], Mavroeidis [2005, 2010], Mirza and Storjohann [2014], Mavroeidis et al. [2014], Ascari et al. [2021]).

2 Model

We consider the linear model

$$y_t = \beta' Y_t + \delta' X_t + \varepsilon_t, \tag{2.1}$$

for $t = 1, \ldots, T$, where y_t is the scalar outcome variable, Y_t is a $g \times 1$ vector of endogenous variables, X_t is a $p \times 1$ vector of exogenous variables, β and δ are coefficient vectors, and ε_t is an error term such that $\bar{\mathbb{E}}[X_t \varepsilon_t] = 0_{p\times 1}$. The goal is to conduct inference on β, by inverting the test

$$H_0 : \beta = \beta_0 \text{ vs. } H_1 : \beta \neq \beta_0, \tag{2.2}$$

for a pre-determined β_0. The $(1-\alpha)$ confidence set can be constructed by collecting the values of β_0 for which the null hypothesis in (2.2) is not rejected at the $\alpha \in (0,1)$ level of significance.

We introduce the auxiliary projection

$$Y_t = \xi' Q_t + v_t, \tag{2.3}$$

where Q_t is a $k_Q \times 1$ vector that includes the elements of X_t as well as potentially other variables (but is also allowed to only include the elements of X_t), ξ is a coefficient matrix with ξ_j denoting column j of ξ, and v_t is an error term with $v_{t,j}$ denoting its j^{th} component, such that $\bar{\mathbb{E}}[Q_t v_{t,j}] = 0_{k_Q \times 1}$ for $j = 1, \ldots, g$ (by construction). For notational convenience, we let $\zeta = [\xi_1', \ldots, \xi_g']'$.

Our validity condition is taken from Lewbel [2012]. Letting Z_t be a $k_Z \times 1$ vector that could (but does not have to) include components of (or be equal to) the vector Q_t or X_t and $\bar{\mu} = \bar{\mathbb{E}}[Z_t]$, we impose the 'IV validity condition' that

$$\bar{\mathbb{E}}[(Z_t - \bar{\mu}) v_{t,j} \varepsilon_t] = 0_{k_Z \times 1}, \tag{2.4}$$

for all $j = 1, \ldots, g$.

A straightforward generalisation (see Appendix) of the case of a single endogenous regressor considered in Lewbel [2012] implies that β is identified if in addition to (2.4), the following 'IV strength condition' is satisfied:

$$\bar{\mathbb{E}}[v_t v_t' \otimes (Z_t - \bar{\mu})] \quad \text{has full rank.} \tag{2.5}$$

We use the term 'strength' to distinguish this rank/relevance requirement—ensuring the Lewbel-type instruments are informative about the endogenous regressors—from the validity condition (2.4), which is an orthogonality/exogeneity restriction. We state (2.5) to clarify when point identification would obtain, but throughout the paper we do not impose identification assumptions.

3 Discussion of identifying restrictions

In this section, we discuss how our setup and the restrictions we impose compare to other approaches in the literature.

In Section 3.1, we show that the setting and restrictions imposed in standard regime-based identification schemes such as in Rigobon [2003] and Lewis [2022] imply ours for a particular choice of Z_t and Q_t. We also show that a different choice of Z_t and Q_t allows us to relax some of the restrictions in these setups at the expense of strengthening others. Finally, we argue that our approach is advantageous because it is immediately generalisable to multiple regimes.

In Section 3.2 we consider two policy-relevant macroeconometric settings that satisfy our restrictions without requiring pre-labelled regimes. First, we consider a version of the model in Rigobon [2003] where the variances of the error terms depend on other variables. This can be seen as a 'continuous regimes' generalisation, which is relevant in policy settings where time and resource constraints may make manually identifying regimes impractical. We show that this approach is related but not equivalent to the one in Prono [2014]. Second, we show that our setting can address endogeneity caused by mismeasured variables, which is an issue often encountered in macroeconometric analyses of EMDEs.

In some of the examples in this section, we set Q_t and Z_t equal to functions of ε_t. While this section deals with identification and not estimation, we nevertheless note that this is feasible in our approach, since ε_t is identified under the null hypothesis in (2.2). The simulations in Section 5 confirm that setting Q_t and Z_t equal to functions of the error term under the null hypothesis yields non-trivial power. Furthermore, for simplicity–and because some other methods in the literature do not explicitly foresee multiple endogenous variables–we consider the case of a single endogenous variable. We note that in this case, the strength condition in (2.5) reduces to $\bar{\mathbb{E}}[v_t^2(Z_t - \bar{\mu})] \neq 0_{k_Z \times 1}$.

3.1 Regime-based identification via heteroskedasticity

To compare our approach to regime-based approaches, we consider the basic Rigobon [2003] setting (cf. his Proposition 1)

$$
\begin{aligned}
y_t &= \beta Y_t + \varepsilon_t \\
Y_t &= \vartheta\, y_t + \tilde{v}_t \\
S_t &\in \{0, 1\},
\end{aligned}
\tag{3.6}
$$

where β and ϑ are scalar coefficients with $\beta\vartheta \neq 1$,[4] ε_t is a mean-zero error term, $\tilde{v}_t$ is a mean-zero error term, S_t is an indicator variable describing two regimes, and y_t and Y_t are assumed to be mean-zero random variables. This setup is also considered in Lewis [2022],[5] whose focus is SVAR identification and inference. The error terms

[4]Rigobon [2003] does not explicitly state this last requirement, but implicitly imposes it in the display immediately following his Equation (2). Lewis [2022] explicitly imposes this requirement by assuming that H in his notation is invertible.

[5]The equivalence between the model in (3.6) and Equation (1) in Lewis [2022] can be seen by setting η_{1t}, η_{2t}, H_{12}, H_{21}, ε_{1t}, and ε_{2t} in the notation of Lewis [2022] equal to y_t, Y_t, β, ϑ, ε_t, and $\tilde{v}_t$ in (3.6), respectively.

are assumed to satisfy, for all $t = 1, \ldots, T$,

$$\mathbb{E}[\varepsilon_t | S_t = 0] = \mathbb{E}[\varepsilon_t | S_t = 1] = \mathbb{E}[\tilde{v}_t | S_t = 0] = \mathbb{E}[\tilde{v}_t | S_t = 1] = 0. \tag{3.7}$$

and for some finite non-random $c_{t,0}$ and $c_{t,1}$, for all $t = 1, \ldots, T$,

$$\mathbb{E}[\varepsilon_t \tilde{v}_t | S_t = 0] = c_{t,0} \quad \mathbb{E}[\varepsilon_t \tilde{v}_t | S_t = 1] = c_{t,1}, \tag{3.8}$$

In the basic model considered in Rigobon [2003, Proposition 1] and Lewis [2022, Assumption 1] $c_{t,0} = c_{t,1} = 0$ for all $t = 1, \ldots, T$.

The variances of the error terms are given, for all $t = 1, \ldots, T$, by

$$\mathbb{E}[\varepsilon_t^2 | S_t = 0] = \sigma^2_{\varepsilon,0}, \quad \mathbb{E}[\varepsilon_t^2 | S_t = 1] = \sigma^2_{\varepsilon,1}, \quad \mathbb{E}[\tilde{v}_t^2 | S_t = 0] = \sigma^2_{\tilde{v},0}, \quad \mathbb{E}[\tilde{v}_t^2 | S_t = 1] = \sigma^2_{\tilde{v},1},$$

with $\mathbb{P}(S_t = 0) = \rho_t$ and $\mathbb{P}(S_t = 1) = 1 - \rho_t$ for $0 < \rho_t < 1$.

Under this setup, the parameters β and ϑ are identified if

$$\sigma^2_{\varepsilon,0} \sigma^2_{\tilde{v},1} \neq \sigma^2_{\varepsilon,1} \sigma^2_{\tilde{v},0}. \tag{3.9}$$

We consider two ways to map the regime-based setup into ours. In the first, we impose $c_{t,0} = c_{t,1} = 0$ and set Q_t (in our notation) equal to ε_t and Z_t (in our notation) equal to S_t. We show that in this setup, our validity and strength conditions are satisfied if those in Rigobon [2003] and Lewis [2022] are. Furthermore, we also argue that our approach is advantageous because it allows for a straightforward extension to multiple regimes. In the second, we set both Q_t and Z_t (in our notation) equal to S_t. We show that this setup makes it possible to relax the assumption of uncorrelated disturbances ε_t and $\tilde{v}_t$ at the expense of requiring that ε_t^2 be uncorrelated with $Z_t - \bar{\mu}$ or imposing that $\vartheta = 0$.

3.1.1 Setting $Q_t = \varepsilon_t$, $Z_t = S_t$ and imposing $c_{t,0} = c_{t,1} = 0$

Solving for Y_t in (3.6), and equating it to our projection in (2.3) with $Q_t = \varepsilon_t$ yields

$$Y_t = \frac{\vartheta}{1-\beta\vartheta}\varepsilon_t + \frac{1}{1-\beta\vartheta}\tilde{v}_t = \xi\varepsilon_t + v_t.$$

Since $\mathbb{E}[\varepsilon_t \tilde{v}_t] = 0$ for all $t = 1, \ldots, T$,

$$\xi := \frac{\bar{\mathbb{E}}[Y_t\varepsilon_t]}{\bar{\mathbb{E}}[\varepsilon_t^2]} = \frac{\frac{\vartheta}{1-\beta\vartheta}\bar{\mathbb{E}}[\varepsilon_t^2] + \frac{1}{1-\beta\vartheta}\bar{\mathbb{E}}[\tilde{v}_t\varepsilon_t]}{\bar{\mathbb{E}}[\varepsilon_t^2]} = \frac{\vartheta}{1-\beta\vartheta},$$

so that

$$v_t = \frac{1}{1-\beta\vartheta}\tilde{v}_t.$$

Substituting $Z_t = S_t$ and v_t into our validity condition (2.4) yields

$$\bar{\mathbb{E}}[(Z_t - \bar{\mu})v_t\varepsilon_t] = \frac{1}{1-\beta\vartheta}\bar{\mathbb{E}}[(S_t - \bar{\mu})\tilde{v}_t\varepsilon_t] = 0, \tag{3.10}$$

where the second equality follows from (3.8). Substituting $Z_t = S_t$ and v_t into our strength condition (2.5) yields for $\bar{\rho} = \frac{1}{T}\sum_{t=1}^{T}\rho_t$,

$$\begin{aligned}
\bar{\mathbb{E}}[(Z_t - \bar{\mu})v_t^2] &= \frac{1}{(1-\beta\vartheta)^2}\bar{\mathbb{E}}[(S_t - \bar{\mu})\tilde{v}_t^2] \\
&= \frac{1}{(1-\beta\vartheta)^2}\bar{\mathbb{E}}[\mathbb{E}[(S_t - \bar{\mu})\tilde{v}_t^2|S_t]] \\
&= \frac{1}{(1-\beta\vartheta)^2}\frac{1}{T}\sum_{t=1}^{T}\left(-\rho_t(1-\bar{\rho})\mathbb{E}\left[\tilde{v}_t^2\middle| S_t = 0\right] + (1-\rho_t)\bar{\rho}\mathbb{E}\left[\tilde{v}_t^2\middle| S_t = 1\right]\right) \\
&= \frac{\bar{\rho}(1-\bar{\rho})}{(1-\beta\vartheta)^2}\left(\sigma_{\tilde{v},1}^2 - \sigma_{\tilde{v},0}^2\right),
\end{aligned} \tag{3.11}$$

where the third inequality follows since $\bar{\mu} = 1 - \bar{\rho}$. Thus,

$$\bar{\mathbb{E}}[(Z_t - \bar{\mu})v_t^2] \neq 0 \iff \sigma_{\tilde{v},0}^2 \neq \sigma_{\tilde{v},1}^2, \tag{3.12}$$

which is equivalent to the relevant restriction in (3.9).[6]

Therefore, (3.10) and (3.12) imply that setting $Q_t = \varepsilon_t$ and $Z_t = S_t$ in the Rigobon [2003] (or Lewis [2022]) setting satisfies our validity and strength conditions. Furthermore, our setup allows for a straightforward extension to the case of multiple regimes. For instance, the case of three regimes $S_t \in \{0, 1, 2\}$ can be accommodated by setting $Z_t = e_{2,S_t}$ for $S_t \neq 0$, where $e_{d,j}$ denotes the $d \times 1$ vector whose element j is equal to unity, and all other elements equal to zero.

3.1.2 Setting $Q_t = Z_t = S_t$ with $c_{t,0} = c_{t,1} = c_t \neq 0$

We now show how our validity and strength conditions can be satisfied even when $c_{t,0} = c_{t,1} = c_t := \mathbb{E}[\varepsilon_t \tilde{v}_t]$, where $\mathbb{E}[\varepsilon_t \tilde{v}_t]$ is non-random, potentially non-zero, and can vary with t. This corresponds to the case where the error terms are allowed to comove, but this comovement is not systemtically related to the regime S_t.

Solving for Y_t in (3.6), and equating it to our projection in (2.3) with $Q_t = Z_t = S_t$ yields

$$Y_t = \frac{\vartheta}{1-\beta\vartheta}\varepsilon_t + \frac{1}{1-\beta\vartheta}\tilde{v}_t = \xi S_t + v_t.$$

By (3.7), for all $t = 1, \ldots, T$, $\mathbb{E}[S_t \varepsilon_t] = \mathbb{E}[S_t \tilde{v}_t] = 0$. Thus,

$$\xi := (\bar{\mathbb{E}}[S_t^2])^{-1}\bar{\mathbb{E}}[Y_t S_t] = \frac{1}{1-\bar{\rho}}\left(\frac{\vartheta}{1-\beta\vartheta}\bar{\mathbb{E}}[\varepsilon_t S_t] + \frac{1}{1-\beta\vartheta}\bar{\mathbb{E}}[\tilde{v}_t S_t]\right) = 0,$$

so that

$$v_t = \frac{\vartheta}{1-\beta\vartheta}\varepsilon_t + \frac{1}{1-\beta\vartheta}\tilde{v}_t.$$

[6] We would have to strengthen this condition to the one in (3.9) if we also wanted to identify ϑ (separately from β).

Substituting $Z_t = S_t$ and v_t into our validity condition (2.4) yields

$$\begin{aligned}\bar{\mathbb{E}}[(Z_t - \bar{\mu})v_t\varepsilon_t] &= \frac{\vartheta}{1-\beta\vartheta}\bar{\mathbb{E}}[(S_t - \bar{\mu})\varepsilon_t^2] + \frac{1}{1-\beta\vartheta}\bar{\mathbb{E}}[(S_t - \bar{\mu})\tilde{v}_t\varepsilon_t] \\ &= \frac{\vartheta}{1-\beta\vartheta}\bar{\mathbb{E}}[(S_t - \bar{\mu})\varepsilon_t^2] + \frac{1}{1-\beta\vartheta}\bar{\mathbb{E}}[(S_t - \bar{\mu})\mathbb{E}[\varepsilon_t\tilde{v}_t|S_t]] \\ &= \frac{\vartheta}{1-\beta\vartheta}\bar{\mathbb{E}}[(S_t - \bar{\mu})\varepsilon_t^2].\end{aligned} \tag{3.13}$$

Substituting v_t into our strength condition (2.5) yields

$$\begin{aligned}\bar{\mathbb{E}}[(Z_t - \bar{\mu})v_t^2] &= \frac{\vartheta^2}{(1-\beta\vartheta)^2}\bar{\mathbb{E}}[(S_t - \bar{\mu})\varepsilon_t^2] + \frac{1}{(1-\beta\vartheta)^2}\bar{\mathbb{E}}[(S_t - \bar{\mu})\tilde{v}_t^2] \\ &\quad + \frac{2\vartheta}{(1-\beta\vartheta)^2}\bar{\mathbb{E}}[(S_t - \bar{\mu})\varepsilon_t\tilde{v}_t] \\ &= \frac{\vartheta^2}{(1-\beta\vartheta)^2}\bar{\mathbb{E}}[(S_t - \bar{\mu})\varepsilon_t^2] + \frac{1}{(1-\beta\vartheta)^2}\bar{\mathbb{E}}[(S_t - \bar{\mu})\tilde{v}_t^2] \\ &= \frac{\bar{\rho}(1-\bar{\rho})}{(1-\beta\vartheta)^2}\left((\vartheta^2\sigma_{\varepsilon,1}^2 + \sigma_{\tilde{v},1}^2) - (\vartheta^2\sigma_{\varepsilon,0}^2 + \sigma_{\tilde{v},0}^2)\right),\end{aligned}$$

where the second equality follows from (3.13) and $\bar{\mu} = 1 - \bar{\rho}$, and the third from (3.11) combined with analogous derivations for $\bar{\mathbb{E}}[(S_t - \bar{\mu})\varepsilon_t^2]$. Thus,

$$\bar{\mathbb{E}}[(Z_t - \bar{\mu})v_t^2] \neq 0 \iff \vartheta^2\sigma_{\varepsilon,0}^2 + \sigma_{\tilde{v},0}^2 \neq \vartheta^2\sigma_{\varepsilon,1}^2 + \sigma_{\tilde{v},1}^2. \tag{3.14}$$

Our validity condition in (3.13) and strength condition in (3.14) can be jointly satisfied in several ways. First, if $\vartheta = 0$ – precluding contemporaneous feedback from y_t to Y_t – our validity condition is satisfied even if errors comove, and our strength condition is satisfied if $\sigma_{\tilde{v},1}^2 \neq \sigma_{\tilde{v},2}^2$. Second, if the variance of ε_t^2 does not change across regimes, our validity condition is satisfied even if errors comove, and our strength condition requires only $\sigma_{\tilde{v},1}^2 \neq \sigma_{\tilde{v},2}^2$. Thus, satisfying our validity and strength conditions with $Q_t = Z_t = S_t$ leads to restrictions that are not nested by those in the standard regime-based approach.

Having discussed general identification conditions when $c_{t,0} = c_{t,1} = c_t$, we provide a plausible example where these hold. Suppose ε_t and $\tilde{v}_t$ can be decomposed into common and idiosyncratic terms,

$$\varepsilon_t = \varepsilon_t^{id} + a_t u_t, \quad \tilde{v}_t = \tilde{v}_t^{id} + b_t u_t, \tag{3.15}$$

where for all $t = 1, \ldots, T$, $\mathbb{E}[\varepsilon_t^{id}|S_t] = \mathbb{E}[\tilde{v}_t^{id}|S_t] = \mathbb{E}[u_t|S_t] = \mathbb{E}[\varepsilon_t^{id}\tilde{v}_t^{id}|S_t] = 0$, and a_t, b_t are non-random coefficients. $c_{t,0} = c_{t,1} = c_t$ is satisfied if for all $t = 1, \ldots, T$, $\mathbb{E}[u_t^2|S_t] = \mathbb{E}[u_t^2]$ since then

$$\mathbb{E}[\varepsilon_t \tilde{v}_t|S_t] = \mathbb{E}[a_t b_t u_t^2|S_t] = a_t b_t \mathbb{E}[u_t^2] =: c_t.$$

The two validity conditions discussed in this section – $\vartheta = 0$ or $\sigma^2_{\varepsilon,0} = \sigma^2_{\varepsilon,1}$ – remain sufficient in this setup, and the strength condition in both these cases reduces to $\mathbb{E}[(\tilde{v}_t^{id})^2|S_t = 0] \neq \mathbb{E}[(\tilde{v}_t^{id})^2|S_t = 1]$. Intuitively, the identifying assumption is that the regimes shift the variance of idiosyncratic shocks but not that of common shocks. In Section 6.1, we provide an example based on the fuel-price passthrough in Sierra Leone.

3.2 Non-regime-based identification via heteroskedasticity

We consider two settings in which our identifying assumptions hold without pre-labelled regimes.[7,8] First, we consider a version of the Rigobon [2003] model where the variance of the error terms depends on other variables. Second, we show that our setup can address endogeneity caused by mismeasured variables.

For the first non-regime-based setting, consider the same structural model as in (3.6), but, rather than assuming exogenous and pre-labelled regimes, we assume that a mean-zero variable Z_t exists such that

$$\bar{\mathbb{E}}[Z_t \tilde{v}_t \varepsilon_t] = 0.$$

Setting $Q_t = \varepsilon_t$, analogous derivations leading to (3.10) imply that our validity condition in (2.4) is satisfied, i.e., $\bar{\mathbb{E}}[Z_t v_t \varepsilon_t] = 0$. Similarly, analogous derivations leading to (3.11) imply that our strength condition is satisfied if $\bar{\mathbb{E}}[Z_t \tilde{v}_t^2] \neq 0$. Thus, rather than requiring the variances to differ only across exogenously labelled regimes,

[7]We note that these settings are analogous to those considered in Lewbel [2012]. The main contributions of our paper relative to the approach in Lewbel [2012] lie in robustifying it to weak identification and explicitly allowing for heteroskedastic dependent data in deriving limiting results. We discuss these contributions after the presentation of our test in Section 4.

[8]Prono [2014] considers the same linear model that we do. In our notation, the restrictions are

$$\begin{aligned} \mathbb{E}\left[\begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix} \middle| \mathcal{F}_{t-1} \right] &= 0 \\ \mathbb{E}\left[\begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix}' \middle| \mathcal{F}_{t-1} \right] &= \begin{bmatrix} h_{1,1,t} & h_{1,2,t} \\ h_{2,1,t} & h_{2,2,t} \end{bmatrix} \\ h_{1,1,t} &= \omega_{1,1} + a_{1,1}\varepsilon_{t-1}^2 + b_{1,1}h_{1,1,t-1} \\ h_{1,2,t} &= \omega_{1,2} + a_{1,2}\varepsilon_{t-1}v_{t-1} + b_{1,2}h_{1,2,t-1} \\ h_{2,2,t} &= \omega_{2,2} + a_{2,2}v_{t-1}^2 + b_{2,2}h_{2,2,t-1}, \end{aligned} \tag{3.16}$$

where $\omega_{i,j}$, $a_{i,j} > 0$, and $b_{i,j} \geq 0$ are coefficients with $i, j = 1, 2$, and $\mathcal{F}_{t-1} := \mathcal{B}_{\{[y,Y,X,Z]'\}}(-\infty, t-1)$ is the Borel σ-field generated by $\{[y_s, Y_s, X_s, Z_s]', s \leq t-1\}$. Our and Prono's identification restrictions are not nested. On one hand, our approach is more general because it does not impose a GARCH assumptions on the error terms. On the other hand, there are DGPs and choices of Z_t that satisfy the GARCH restrictions but violate our moment condition in (2.4). Indeed, suppose that $h_{1,2,t}$ in (3.16) is correctly specified, and that Z_t depends on $\varepsilon_{t-1}v_{t-1}$. Then our identifying restriction (2.4) will generally not hold since $\mathbb{E}[v_t\varepsilon_t|Z_t] \neq 0$.

the variances of $\tilde{v}_t$ are allowed to depend in time-varying ways on Z_t.

The most immediate way in which the above framework can be operationalised is by setting Z_t equal to an observed variable that is uninformative of the co-movement of the error terms but informative of their time-varying variance. This can be interpreted as identifying a variable that tracks regimes in a 'continuous' way. Alternatively, if the additional restriction is imposed that $\mathbb{E}[\tilde{v}_t|\varepsilon_t] = 0$ for all $t = 1, \ldots, T$, it is possible to set $Z_t = \varepsilon_t^2$: the validity condition is satisfied since $\bar{\mathbb{E}}[(\varepsilon_t^2 - \bar{\mathbb{E}}[\varepsilon_t^2])\tilde{v}_t] = \bar{\mathbb{E}}[(\varepsilon_t^2 - \bar{\mathbb{E}}[\varepsilon_t^2])\mathbb{E}[\tilde{v}_t|\varepsilon_t]] = 0$ and the strength condition is satisfied if $\bar{\mathbb{E}}[\varepsilon_t^2\tilde{v}_t^2] \neq \bar{\mathbb{E}}[\varepsilon_t^2]\bar{\mathbb{E}}[\tilde{v}_t^2]$. We note that this approach recovers identification restrictions similar to those of Lewis [2021].

For the second non-regime-based setting, we consider the model

$$\begin{aligned} y_t &= Y_t\beta + X_t\delta + \tilde{\varepsilon}_t \\ Y_t^* &= Y_t + \tilde{v}_t, \end{aligned} \tag{3.17}$$

where Y_t^* is the observed value of the mismeasured mean-zero variable Y_t, X_t is a scalar mean-zero exogenous covariate, $\tilde{\varepsilon}_t$ is a mean-zero error term independent of Y_t and X_t, and $\tilde{v}_t$ is assumed to be a classical mean-zero measurement error independent of X_t, Y_t, and $\tilde{\varepsilon}_t$. Consider the projection

$$Y_t = \breve{\xi}X_t + \breve{v}_t,$$

where $\breve{v}_t$ is a mean-zero projection error with $\bar{\mathbb{E}}[X_t\breve{v}_t] = 0$ by construction, and $\breve{\xi}$ is a scalar coefficient. Substituting the above display into the expression for Y_t^* in (3.17) yields

$$Y_t^* = \breve{\xi}X_t + \breve{v}_t + \tilde{v}_t.$$

Substituting out the unobserved Y_t leaves the endogenous system

$$\begin{aligned} y_t &= Y_t^*\beta + X_t\delta + \varepsilon_t \\ Y_t^* &= \breve{\xi}X_t + v_t, \end{aligned}$$

where $\varepsilon_t = \tilde{\varepsilon}_t - \beta\tilde{v}_t$ and $v_t = \breve{v}_t + \tilde{v}_t$. Setting $Q_t = Z_t = X_t$ satisfies our validity condition in (2.4) since for all $t = 1, \ldots, T$, $\mathbb{E}[X_t\varepsilon_t v_t] = \mathbb{E}[X_t(\tilde{\varepsilon}_t - \beta\tilde{v}_t)(\breve{v}_t + \tilde{v}_t)] = \mathbb{E}[X_t\tilde{\varepsilon}_t\breve{v}_t] + \mathbb{E}[X_t\tilde{\varepsilon}_t\tilde{v}_t] - \beta\mathbb{E}[X_t\tilde{v}_t\breve{v}_t] - \beta\mathbb{E}[X_t\tilde{v}_t^2] = 0$.[9] Our strength condition is satisfied if $\bar{\mathbb{E}}[X_t v_t^2] = \bar{\mathbb{E}}[X_t\breve{v}_t^2] \neq 0$,[10] which can be interpreted as a conditional-heteroskedasticity requirement: if the conditional second moment $\bar{\mathbb{E}}[\tilde{v}_t^2|X_t]$ varies with X_t in a way that is not orthogonal to X_t, then $\bar{\mathbb{E}}[X_t\tilde{v}_t^2] = \bar{\mathbb{E}}[X_t\mathbb{E}[\tilde{v}_t^2|X_t]] \neq 0$.

4 The HetAR test

4.1 Definition of the HetAR test

We follow the standard AR arguments presented in Chernozhukov and Hansen [2008, pp. 69]. Let $y_{t,0} = y_t - \beta_0' Y_t$. Under the null $H_0 : \beta = \beta_0$, the structural equation (2.1) implies $y_{t,0} = \delta' X_t + \varepsilon_t$, and the IV validity condition yields $\bar{\mathbb{E}}[(Z_t - \bar{\mu})v_{t,j}\varepsilon_t] = 0$ for all $j = 1, \ldots, g$. Equivalently, in the population linear projection of $y_{t,0}$ onto $[X_t', (Z_t - \bar{\mu})' v_{t,1}, \ldots, (Z_t - \bar{\mu})' v_{t,g}]'$, the coefficients on $(Z_t - \bar{\mu})v_{t,j}$ are zero. Hence, writing the corresponding auxiliary regression as

$$y_{t,0} = \delta' X_t + \sum_{j=1}^{g} \alpha^{(j)\prime}[(Z_t - \bar{\mu})v_{t,j}] + \varepsilon_t,$$

we have $\alpha^{(j)} = 0$ for all $j = 1, \ldots, g$ under the null $H_0 : \beta = \beta_0$. Therefore, a joint test of $\alpha^{(j)} = 0$ for all $j = 1, \ldots, g$ provides a test of $H_0 : \beta = \beta_0$. We note that the strength condition in (2.5) is not imposed, although $\alpha = [\alpha^{(1)\prime}, \ldots, \alpha^{(g)\prime}]'$ will only be different from zero under alternatives if (2.5) holds, implying that our test will control size regardless of identification strength but will only be powerful when (2.5) holds. We illustrate these properties in the simulations section.

[9] This follows because $\mathbb{E}[X_t\tilde{\varepsilon}_t\breve{v}_t] = 0$ since $\tilde{\varepsilon}_t$ is independent of X_t and Y_t, $\mathbb{E}[X_t\tilde{\varepsilon}_t\tilde{v}_t] = 0$ since $\tilde{v}_t$ is independent of X_t and $\tilde{\varepsilon}_t$, $\mathbb{E}[X_t\tilde{v}_t\breve{v}_t] = 0$ since $\tilde{v}_t$ is independent of X_t and Y_t, and $\mathbb{E}[X_t\tilde{v}_t^2] = 0$ since $\tilde{v}_t$ is independent of X_t and X_t has mean zero.

[10] This follows because $\bar{\mathbb{E}}[X_t v_t^2] = \bar{\mathbb{E}}[X_t(\breve{v}_t + \tilde{v}_t)^2] = \bar{\mathbb{E}}[X_t\breve{v}_t^2] + \bar{\mathbb{E}}[X_t\tilde{v}_t^2] + 2\bar{\mathbb{E}}[X_t\breve{v}_t\tilde{v}_t]$, $\bar{\mathbb{E}}[X_t\tilde{v}_t^2] = 0$ since $\tilde{v}_t$ is independent of X_t and X_t is mean zero, and $\bar{\mathbb{E}}[X_t\breve{v}_t\tilde{v}_t] = 0$ since $\tilde{v}_t$ is independent of X_t and Y_t.

Substituting the sample analogues, and letting $W_t([\hat{\mu}', \hat{\zeta}']') = [X_t', (Z_t - \hat{\mu})' v_{t,1}(\hat{\xi}), \ldots, (Z_t - \hat{\mu})' v_{t,g}(\hat{\xi})]'$ and $\gamma = [\delta', \alpha']'$ yields

$$y_{t,0} = \gamma' W_t([\hat{\mu}', \hat{\zeta}']') + \epsilon_t, \tag{4.18}$$

where $\epsilon_t = \gamma'(W_t - W_t([\hat{\mu}', \hat{\zeta}']')) + \varepsilon_t$, and $W_t = [X_t', (Z_t - \bar{\mu})' v_{t,1}, \ldots, (Z_t - \bar{\mu})' v_{t,g}]'$. Let $\hat{\gamma}$ be the OLS estimator

$$\hat{\gamma} = \left(\frac{1}{T} \sum_{t=1}^{T} W_t([\hat{\mu}', \hat{\zeta}']') W_t'([\hat{\mu}', \hat{\zeta}']') \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} W_t([\hat{\mu}', \hat{\zeta}']') y_{t,0} = [\hat{\delta}', \hat{\alpha}']'.$$

We note that $\hat{\gamma}$ depends on the null hypothesis through $y_{t,0}$, but suppress this dependence in our notation for convenience.

Then, setting $R = [0_{k_Z g \times p} \quad I_{k_Z g}]$, the proposed test rejects H_0 at the q significance level whenever the following statistic exceeds the $1 - q$ quantile of a χ^2 distribution with $k_Z g$ degrees of freedom:

$$HetAR = T \hat{\gamma}' R' (R \hat{\Psi} R')^{-1} R \hat{\gamma},$$

where $\hat{\Psi}$ is a suitable estimator of the covariance matrix of $\sqrt{T}(\hat{\gamma} - \gamma)$, say Ψ. We note that $\hat{\Psi}$ has to be recalculated for each null hypothesis being tested since it depends on $y_{t,0}$.

To account for potential generated-regressor issues,[11] we propose to use the modified

[11] A sufficient condition under which generated regressors do not need to be accounted for in variance estimation is that $\bar{\mathbb{E}}[\nabla_\zeta (Z_t - \bar{\mu}) v_{t,j}(\zeta) \varepsilon_t] = 0_{k_Z \times k_Q}$, where the dependence of $v_{t,j}$ on ζ is made explicit (see, e.g., the discussion surrounding Equation 6.8 in Wooldridge [2010]). However, this condition does not plausibly hold for several examples in Section 3. In the example considered in Section 3.1.1, $\bar{\mathbb{E}}[\nabla_\zeta (Z_t - \bar{\mu}) v_{t,j}(\zeta) \varepsilon_t] = \bar{\mathbb{E}}[(S_t - \bar{\rho}) \varepsilon_t^2] = \bar{\rho}(1 - \bar{\rho})(\sigma_{\varepsilon,0}^2 - \sigma_{\varepsilon,1}^2)$. This will only be zero when $\sigma_{\varepsilon,0}^2 = \sigma_{\varepsilon,1}^2$, which would preclude heteroskedasticity in both error terms across the two regimes. Whether this condition holds for the examples considered in Section 3.2 depends on the choice of Q_t and Z_t. If $Q_t = \varepsilon_t$ and Z_t is a mean-zero variable (the first example), then the condition reduces to $\bar{\mathbb{E}}[\varepsilon_t^2 Z_t] = 0$. While possible, it is not in the spirit of the 'continuous regime' interpretation, since this condition, together with the identifying condition, imposes the restriction that Z_t is informative of $\tilde{v}_t^2$ but not ε_t^2. If $Z_t = \varepsilon_t^2$ and $Q_t = \varepsilon_t$ (second example), this condition

Newey and West [1987] estimator

$$\hat{\Psi} = \left(\frac{1}{T}\sum_{t=1}^{T} W_t([\hat{\mu}', \hat{\zeta}']')W_t'([\hat{\mu}', \hat{\zeta}']')\right)^{-1} \hat{V} \left(\frac{1}{T}\sum_{t=1}^{T} W_t([\hat{\mu}', \hat{\zeta}']')W_t'([\hat{\mu}', \hat{\zeta}']')\right)^{-1},$$

where $\hat{V}$ is the term accounting for the generated regressors given in (A.1) in the Appendix.

4.2 Asymptotic properties of the HetAR test

We make the following assumptions to derive the limiting distribution of $HetAR$ under the null hypothesis in (2.2).

Assumption 4.1. *Let* $W_t = [X_t', (Z_t-\bar{\mu})'v_{t,1}, \ldots, (Z_t-\bar{\mu})'v_{t,g}]'$ *and* $\mathcal{Z}_t = [X_t'\varepsilon_t, (\varepsilon_t v_t \otimes (Z_t - \bar{\mu}) + \left(\bar{\mathbb{E}}[\varepsilon_t v_t] \otimes I_{k_Z}\right)(Z_t - \bar{\mu}) - \left(I_g \otimes \bar{\mathbb{E}}[\varepsilon_t (Z_t - \bar{\mu})Q_t']\right)(v_t \otimes M_{Q,T}^{-1} Q_t))']'$.

(a) $\{[X_t', \varepsilon_t, Q_t', Z_t', vec(v_t)']'\}$ *is a strong-mixing sequence of size* $-2r/(r-2)$, $r > 2$.

(b) $\bar{\mathbb{E}}[Q_t v_t'] = 0_{k_Q \times g}$, $\bar{\mathbb{E}}[(Z_t - \bar{\mu})v_{t,j}\varepsilon_t] = 0_{k_Z \times 1}$ *for all* $j = 1, \ldots, g$, $\bar{\mathbb{E}}[X_t'\varepsilon_t] = 0_{p\times 1}$.

(c) $\max_{1\le t\le T}\max_{1\le j\le k_Z} \mathbb{E}[|Z_{t,j}^8|^r] < \infty$, $\max_{1\le t\le T}\max_{1\le j\le k_Q} \mathbb{E}[|Q_{t,j}^8|^r] < \infty$ $\max_{1\le t\le T}\max_{1\le l\le g} \mathbb{E}[|v_{t,l}^8|^r] < \infty$, $\max_{1\le t\le T}\max_{1\le j\le p} \mathbb{E}[|X_{t,j}^4|^r] < \infty$, *and* $\max_{1\le t\le T} \mathbb{E}[|\varepsilon_t^4|^r] < \infty$.

(d) The following matrices are uniformly positive definite: $M_{Q,T} \equiv \bar{\mathbb{E}}[Q_t Q_t']$, $M_{W,T} \equiv \bar{\mathbb{E}}[W_t W_t']$, $V_{\mathcal{Z},Z} \equiv var(\frac{1}{\sqrt{T}}\sum_{t=1}^T \mathcal{Z}_t)$.

(e) The following exist: $\lim_{T\to\infty} M_{Q,T}$, $\lim_{T\to\infty} V_{\mathcal{Z},T}$, $\lim_{T\to\infty}\bar{\mu}$, $\lim_{T\to\infty}\bar{\mathbb{E}}[\varepsilon_t v_t]$, $\lim_{T\to\infty}\bar{\mathbb{E}}[\varepsilon_t (Z_t - \bar{\mu})Z_t]$.

(f) $m \to \infty$ *as* $T \to \infty$ *and* $m = o(T^{1/4})$.

reduces to $\bar{\mathbb{E}}[(\varepsilon_t^2 - \mathbb{E}[\varepsilon_t^2])\varepsilon_t^2] = 0$, which generally does not hold. We note that this condition holds in the setup in Section 3.1.2 and the mismeasured regressor setup in Section 3.2. In the former, the condition reduces to $\bar{\mathbb{E}}[(S_t - \bar{\rho})S_t\varepsilon_t] = 0$, which holds under the setup. In the latter, X_t^2 is independent of $\tilde{\varepsilon}_t$ and $\tilde{v}_t$, so that the condition $\bar{\mathbb{E}}[X_t^2\varepsilon_t] = 0$ holds. To provide a test that can be used across all examples discussed in Section 3, we use the adjusted variance estimator throughout.

Assumption 4.1. (a) is a weak dependence condition that allows for different types of time-series data and does not require stationarity. This covers ARMA processes under standard regularity conditions (see Mokkadem [1988] for precise conditions), as well as linear and non-linear GARCH models (see Carrasco and Chen [2002] for precise conditions). We refer to Chen et al. [2016, Section 2.1] for additional examples and references. Assumption 4.1. (b) contains the validity condition. Assumption 4.1. (c) contains sufficient moment conditions, and we note that weaker ones may be available at the expense of less transparent conditions on the dependency structure of the data. Assumption 4.1. (d) contains standard non-degeneracy conditions. Assumption 4.1. (f) is a standard assumption in HAC variance estimation. In practice, we set $m = \lfloor T^{1/5} \rfloor$. Assumption 4.1. (e) is imposed for simplicity and can be replaced with more general conditions at the expense of additional notation.

We note that Assumption 4.1 presents one convenient set of assumptions sufficient for size control under time series and heteroskedasticity. We note that alternative assumptions – including trading off the strength of our dependence and alternative NW bandwidths such as those discussed in Lazarus et al. [2018] – are possible.

We are now in a position to state the asymptotic distribution of the $HetAR$ test under the null hypothesis.

Theorem 4.1. *Suppose Assumption 4.1 holds. Then, under the null hypothesis in* (2.2)*,* $HetAR \xrightarrow{d} \chi^2_{k_{Zg}}$.

We note that our approach provides a two-fold extension of the generated-regressor-corrected TSLS estimator of Lewbel [2012]. First, we do not impose any identification assumption. Second, whereas the generated-regressor-corrected TSLS result in Lewbel [2012] relies on (unconditionally) i.i.d. or strictly stationary data,[12] our result remains valid under heteroskedasticity of arbitrary form and weak mixing conditions. We achieve this by suitably adapting the results on two-step estimation in Newey and McFadden [1994], and note that these results may be of independent interest.

[12] See the discussion immediately following Lemma 2.4 in Newey and McFadden [1994].

5 Simulations

To illustrate our approach in the regime-based case, we consider the following DGP for $\ell = 0, 1$:

$$
\begin{aligned}
y_t &= Y_t\beta + \varepsilon_t \\
Y_t &= \vartheta y_t + \tilde{v}_t \\
\varepsilon_t &= \varepsilon_t^{id} + a_t u_t \\
\tilde{v}_t &= \tilde{v}_t^{id} + b_t u_t \\
\begin{bmatrix} \varepsilon_t^{id} \\ \tilde{v}_t^{id} \end{bmatrix} \Bigg| S_t = \ell &\overset{i.i.d.}{\sim} \mathcal{N}\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\varepsilon^{id},\ell} & 0 \\ 0 & \sigma^2_{\tilde{v}^{id},\ell} \end{bmatrix}\right] \\
u_t &\overset{i.i.d.}{\sim} \mathcal{N}[0, \sigma_u^2] \\
\mathbb{P}(S_t = 0) &= \rho,
\end{aligned}
$$

where $\beta = 0.5$, $\sigma_u^2 = 0.01$, and $\rho = 0.5$. To illustrate the standard regime-based approach discussed in Section 3.1.1, we set $\sigma^2_{\varepsilon,0} = 0.05$, $\sigma^2_{\varepsilon,1} = 0.04$, $\sigma^2_{\tilde{v}^{id},0} = 0.05$, $\sigma^2_{\tilde{v}^{id},1} = 0.01$, $a_t = b_t = 0$, $\vartheta = 0.5$, $Q_t = \varepsilon_t$ and $Z_t = S_t$ (DGP-1). To illustrate the robustness of our test to weak identification, we set $\sigma^2_{\varepsilon,0} = \sigma^2_{\varepsilon,1} = 0.05$, $\sigma^2_{\tilde{v}^{id},0} = \sigma^2_{\tilde{v}^{id},1} = 0.05$, $a_t = b_t = 0$, $\vartheta = 0.5$, $Q_t = \varepsilon_t$ and $Z_t = S_t$ (DGP-2). To illustrate the first of the two new regime-based restrictions discussed in Section 3.1.2, we set $\sigma^2_{\varepsilon,0} = 0.05$, $\sigma^2_{\varepsilon,1} = 0.04$, $\sigma^2_{\tilde{v}^{id},0} = 0.05$, $\sigma^2_{\tilde{v}^{id},1} = 0.01$, $a_t = b_t = 1$, $\vartheta = 0$, and $Q_t = Z_t = S_t$ (DGP-3). To illustrate the second, we set $\sigma^2_{\varepsilon,0} = \sigma^2_{\varepsilon,1} = 0.05$, $\sigma^2_{\tilde{v}^{id},0} = 0.05$, $\sigma^2_{\tilde{v}^{id},1} = 0.01$, $a_t = b_t = 1$, $\vartheta = 0.5$, and $Q_t = Z_t = S_t$ (DGP-4).

To illustrate our approach in the first non-regime case (DGP-5), we consider the following DGP:

$$
\begin{aligned}
y_t &= Y_t\beta + \varepsilon_t \\
Y_t &= \vartheta y_t + \tilde{v}_t \\
\varepsilon_t &\sim \mathcal{N}[0, \sigma_\varepsilon^2] \\
\tilde{v}_t &= \varrho_{\tilde{v}} \log(\varepsilon_t^2)\eta_t, \quad \eta_t \overset{i.i.d.}{\sim} \mathcal{N}[0, 1],
\end{aligned}
$$

where $\beta = \vartheta = 0.5$, $\sigma_\varepsilon^2 = 0.05$, $\varrho_{\tilde{v}} = 0.01$ and η_t is independent of ε_t. In applying our test, we set $Q_t = y_t - Y_t\beta_0$, and $Z_t = (y_t - Y_t\beta_0)^2$.

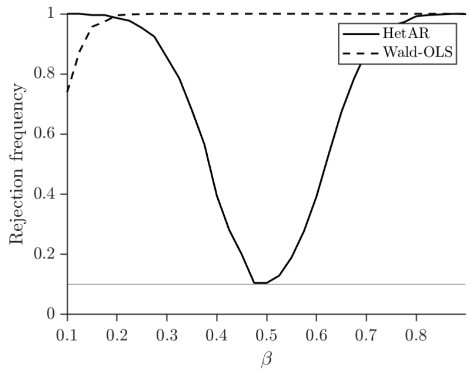
To illustrate our approach in the 'mismeasured regressors' case (DGP-6), we consider the following DGP:

$$\begin{aligned}
y_t &= Y_t^*\beta + X_t\delta + \varepsilon_t \\
Y_t^* &= \breve{\xi} X_t + v_t \\
\varepsilon_t &= \tilde{\varepsilon}_t - \beta\tilde{v}_t \\
v_t &= \tilde{v}_t + \breve{v}_t \\
[X_t, \tilde{\varepsilon}_t, \tilde{v}_t]' &\overset{i.i.d.}{\sim} \mathcal{N}[0_{3\times 1}, \text{diag}[\sigma_X^2, \sigma_{\tilde{\varepsilon}}^2, \sigma_{\tilde{v}}^2]] \\
\breve{v}_t &= \exp\left(\varrho_{X,\breve{v}} X_t\right)\eta_t, \quad \eta_t \overset{i.i.d.}{\sim} \mathcal{N}[0,1],
\end{aligned}$$

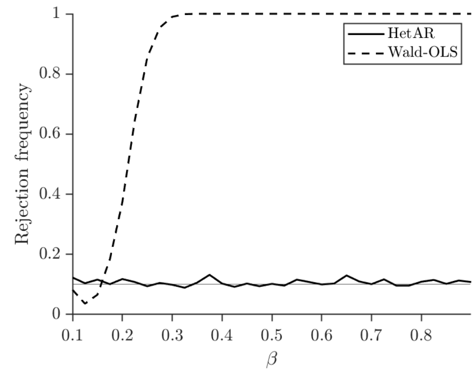
where $\beta = 0.5$, $\delta = \breve{\xi} = 1$, $\sigma_X^2 = 0.1$, $\sigma_{\tilde{\varepsilon}}^2 = 0.05$, $\sigma_{\tilde{v}}^2 = 0.4$, $\varrho_{X,\breve{v}} = 1$, and η_t is independent of $[X_t, \tilde{\varepsilon}_t, \tilde{v}_t]'$. In applying our test, we set $Q_t = Z_t = X_t$.

We only consider power curves, noting that the size properties are reported on the point on the curve where $\beta = \beta_0$. To give a sense of the degree of endogeneity in the DGPs, we also report power curves from a Wald test of the null hypothesis from an OLS regression of y_t on Y_t (and X_t in DGP-6) calculated using a Newey and West [1987] variance estimator with the same number of lags as for the HetAR test, which we label Wald-OLS. Throughout, we consider 1,000 Monte Carlo replications and $T \in \{200, 500\}$.

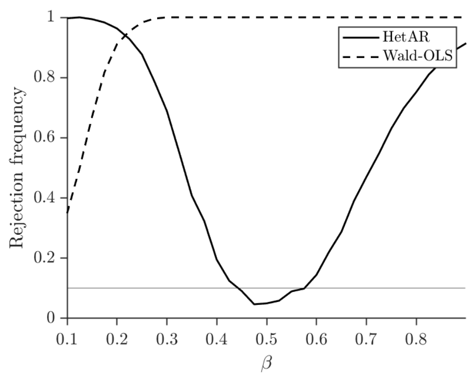
The results in Figure 1 show that our proposed test has good size properties in all DGPs, including when identification fails in DGP-2. It also has good power properties in all DGPs where β is identified, although we note some excessive conservativeness in DGP-4. This confirms the discussion of how our restrictions relate to others in the literature and the versatility of our proposed framework. The Online Appendix contains the results for $T = 200$.



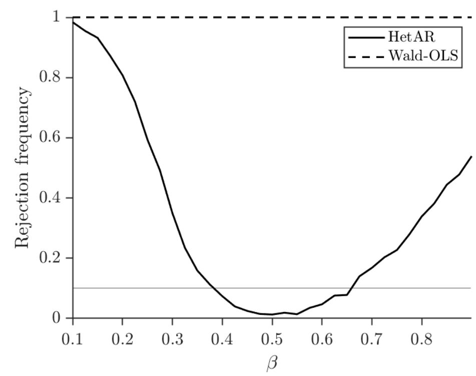
(a) DGP-1



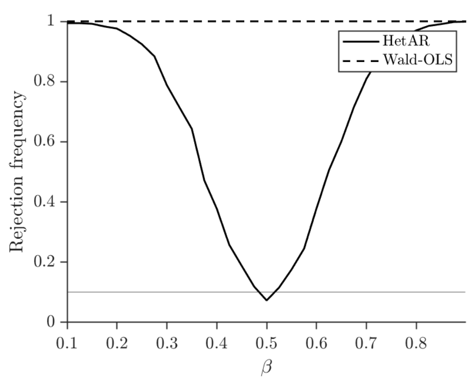
(b) DGP-2



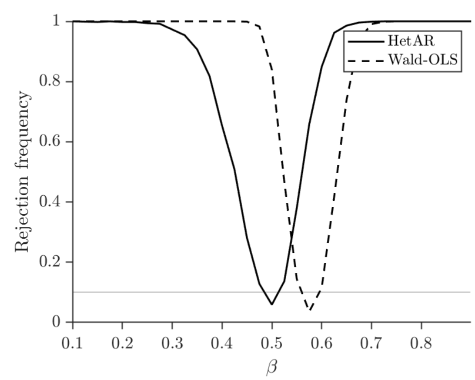
(c) DGP-3



(d) DGP-4



(e) DGP-5



(f) DGP-6

Figure 1: Power curves for nominal test of size 10%, $H_0 : \beta = 0.5$, $T = 500$.

6 Empirical applications

We now turn to three policy-relevant applications. First, we estimate the passthrough of fuel prices to headline inflation in Sierra Leone as an example of regime-based identification. Second, we consider the effect of remittances on consumption growth in the Philippines, using our method to address the likely mismeasurement of remittance data. Lastly, we illustrate the scalability of our approach by estimating the exchange-rate passthrough to inflation in a multi-country setup.

For the first two applications, we show the full p-value profile obtained by inverting the test of the null hypothesis. This lets the reader recover potentially asymmetric $(1-q)$ confidence sets $\{\beta_0 : p\text{-value}(\beta_0) \geq q\}$ for any value of $q \in (0,1)$. Because Figures 2 and 3 plot $1 - p$-value on the vertical axis, the $(1-q)$ confidence set corresponds to the values of β_0 for which the curve lies below $1-q$. The horizontal line at 0.9 therefore marks the 90% confidence set.

6.1 Passthrough of fuel prices to headline inflation in Sierra Leone

Quantifying the contemporaneous passthrough of administered fuel–price adjustments to inflation is policy–relevant for calibrating buffer mechanisms and for program design that seeks to protect the budget while containing inflation risks.

We consider the setup in (3.6), and exploit the new identifying restrictions of our method and institutional features of fuel-price setting in Sierra Leone to achieve identification while allowing the error terms in (3.6) to be correlated. y_t is Sierra Leone's monthly headline CPI, Y_t its monthly domestic fuel–price index and $Z_t = Q_t = S_t \in \{0,1\}$ indexes 'low' (administered fuel prices half a standard deviation below the mean) versus 'high' (administered fuel prices above the 'low' threshold) global oil–price volatility regimes[13]. Under the legal and regulatory framework ad-

[13]CPI data are publicly available at the IMF IFS database (International Monetary Fund, 2025). The fuel price series is confidential and cannot be redistributed.

ministered by the Petroleum Regulatory Agency (PRA), monthly pump prices are set via a published formula that maps landed costs, taxes/levies, distribution margins, and the exchange rate into retail caps [International Monetary Fund, 2024]. Because S_t primarily captures landed-cost volatility – an external process to which Sierra Leone is a price taker – and because retail pump prices are set mechanically via the PRA pricing formula rather than discretionarily in response to contemporaneous domestic conditions, the regime label is plausibly predetermined. These institutional features imply that Y_t is predetermined by the pricing rule and there is no same–period feedback from y_t to Y_t, i.e., $\vartheta = 0$. At the same time, macroeconomic shocks—such as exchange-rate and broad inflationary pressures—can contemporaneously affect both CPI and inputs to the fuel-price formula, implying $\mathbb{E}[\varepsilon_t \tilde{v}_t \mid S_t]$ may be nonzero. We interpret this comovement as arising from a common domestic component that is not systematically related to the regime label S_t, which captures variance shifts in fuel-price innovations through global landed-cost volatility, so that $\mathbb{E}[\varepsilon_t \tilde{v}_t \mid S_t] = \mathbb{E}[\varepsilon_t \tilde{v}_t] = c_t$. With $\vartheta = 0$, our validity condition in (3.13) continues to hold even when $c_t \neq 0$. Moreover, large landed–cost swings in 'high' windows imply a larger variance in $S_t = 1$, so our strength requirement in (3.14), $\sigma^2_{\tilde{v},1} \neq \sigma^2_{\tilde{v},2}$, is plausibly satisfied.

The confidence sets implied by our method are similar to the ones implied by OLS (Figure 2), although we find some evidence that OLS may underestimate the passthrough from fuel prices to inflation, potentially reflecting endogeneity generated by policy tightening following shocks in global fuel prices.

6.2 Remittance and consumption growth in the Philippines

Remittances are an important source of income in the Philippines, standing at roughly nine percent of GDP. Thus, estimating their effect is important to assess household welfare and gauge implications for fiscal revenue. However, due to the many formal and informal remittance channels available, remittances are often mismeasured, making OLS inconsistent.

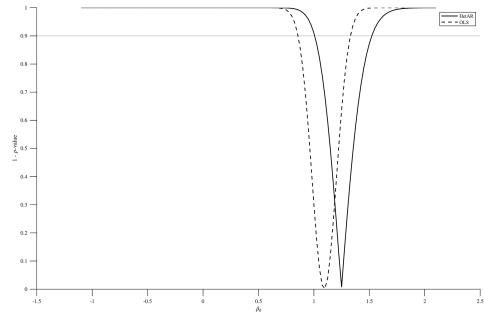


Figure 2: Passthrough of fuel-price to CPI inflation in Sierra Leone.

We address the endogeneity caused by potentially mismeasured remittance data using the setup in (3.17). y_t is quarterly consumption growth, Y_t true quarterly remittance growth, Y_t^* personal quarterly remittance growth as taken from the Bangko Sentral ng Pilipinas, and $Y_t^* = Y_t + \tilde{v}_t$ with $\tilde{v}_t$ independent of Y_t and $\tilde{\varepsilon}_t$. Given that the largest share of remittances to the Philippines comes from the US, we set X_t equal to the quarterly US unemployment rate of foreign-born workers[14]. Thus, our validity condition is satisfied under the plausible assumption that the unemployment rate of foreign-born workers in the US is uncorrelated with the measurement error of remittances to the Philippines. Our strength condition is satisfied if unusually strong or weak unemployment among foreign-born workers in the US generate wider dispersion in Filipino migrants' earnings and remitting behaviour (e.g., some remit windfalls immediately while others smooth).

While the OLS estimator fails to find a statistically meaningful impact of remittance on consumption growth, our approach suggests a significantly positive effect of remittance on consumption growth (Figure 3). This is consistent with mismeasurement-induced endogeneity, which causes OLS to suffer from attenuation bias.

[14]Philippines consumption growth, overseas remittances, and United States foreign born employment data can be downloaded from International Monetary Fund [2025], Bangko Sentral ng Pilipinas [2025], and Federal Reserve Bank of St. Louis [2025].

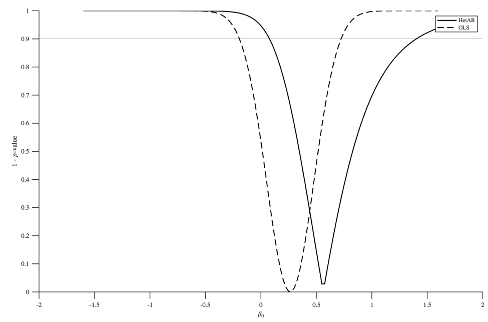


Figure 3: Effect of remittance on consumption growth in the Philippines.

6.3 Exchange rate passthrough in a large number of countries

Studying exchange rate passthrough is crucial for understanding how external shocks feed into domestic prices, and for informing monetary and exchange rate policy. Furthermore, for timely cross-country surveillance, it is advantageous to use a common approach that does not require manual regime labelling. The former is important to ensure that cross-country results are comparable and not an artefact of differing identification strategies or specifications. The latter is important because timely analysis may not afford the time needed for manual regime labelling.

We address the likely endogeneity of inflation and exchange rate growth using the setup in (3.6) as operationalised for the 'continuous regime' case in Section 3.2. y_t is the quarterly year-on-year change in CPI inflation and Y_t is the year-on-year percentage depreciation of the US$/local currency exchange rate[15]. Following López-Villavicencio and Mignon [2017], Jašová et al. [2019], Cheikh and Zaied [2020], we add lagged GDP growth[16] and lagged inflation change as controls up to their sec-

[15]GDP, CPI inflation, and exchange rates data are publicly available at the IMF IFS database (International Monetary Fund, 2025).

[16]Some studies use an output gap measure instead of GDP growth and an import price index instead of the lagged exchange rate as control variables. We prefer GDP growth and the nominal USD exchange rate to avoid cross-country differences in output gap estimation methods, and additional data-availability constraints.

ond lags, X_t. We set Q_t and Z_t equal to ε_t, and conduct inference by projecting out X_t. Our validity condition requires $\bar{\mathbb{E}}[Z_t \tilde{v}_t \varepsilon_t \mid Z_t] = 0$, which imposes that, conditional on lagged GDP growth and inflation, the residual inflation ε_t is not systemically co-moving with the idiosyncratic component of exchange-rate shocks. This exclusion restriction is plausible because lagged and current inflation, exchange-rate movements, and lagged GDP growth jointly summarize information about the business-cycle state and the monetary policy response, so the remaining error term is not expected to be systematically related to the idiosyncratic exchange-rate shock. Our strength condition requires $\bar{\mathbb{E}}[Z_t \tilde{v}_t^2] \neq 0$, which is satisfied if lagged GDP growth is informative about the volatility of exchange-rate shocks. This assumption is also reasonable, as lagged GDP growth is informative about the state of the business cycle and, hence, about the intensity of exchange-rate shocks, making it a relevant predictor of their conditional volatility.

Figure 4 shows the estimated contemporaneous response of inflation to a one percentage point depreciation in the local currency, where we report the CUEs as the point estimates. Countries shown in light gray do not publish quarterly GDP and are excluded from the analysis. For countries shown in dark grey, no null hypothesis can be rejected at the 68% level, which we interpret as the effect not being significantly different from zero. This includes countries that have historically pegged their exchange rate to the US$ (or have used the US$ as legal tender) as well as most advanced economies. The highest passthroughs are found in EMDEs. These results are consistent with the conventional view that passthroughs are higher in economies with weaker nominal anchors, higher and more volatile inflation, and lower in advanced economies and economies with pegged exchange rates.

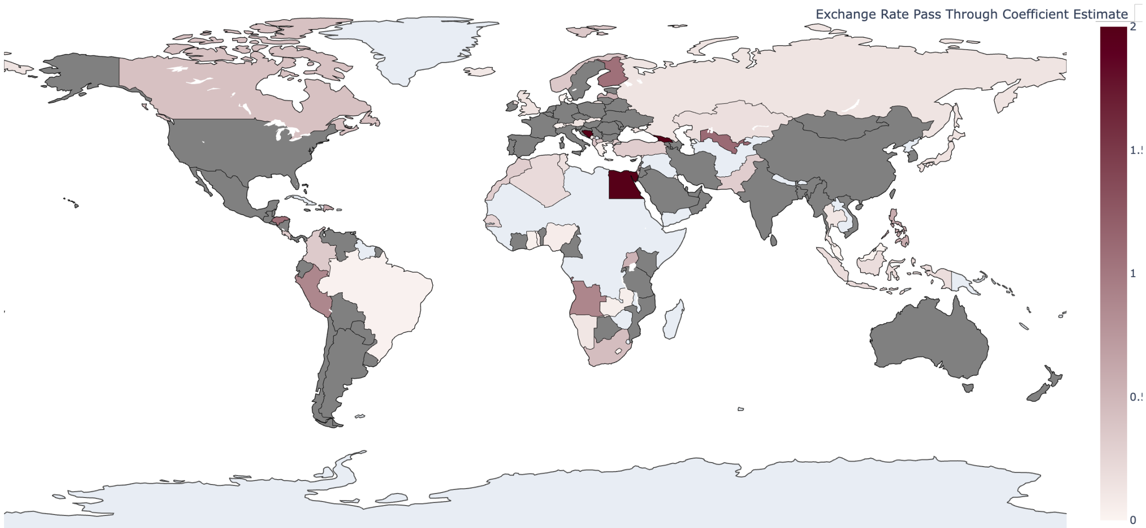


Figure 4: Exchange rate passthrough coefficients by countries.

7 Conclusion

We developed a practical framework for inference via heteroskedasticity in linear models. By nesting and extending existing approaches, we showed how identification can be achieved without external instruments, without manual regime labelling, and in the presence of mismeasured regressors. We proposed an easily implementable Anderson–Rubin type test, derived conditions for its asymptotic validity in non-i.i.d. settings, and documented good size and power in simulations. Three applications—fuel price passthrough to inflation in Sierra Leone, the effect of remittances on consumption growth in the Philippines, and exchange rate passthrough in a large cross-country panel—illustrated the versatility and scalability of our approach.

References

Theodore Anderson and Herman Rubin. Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *The Annals of Mathematical Statistics*, 20(1):46–63, 1949.

Guido Ascari, Leandro Magnusson, and Sophocles Mavroeidis. Empirical Evidence on the Euler Equation for Consumption in the US. *Journal of Monetary Economics*, 117:129–152, 2021.

Bangko Sentral ng Pilipinas. Overseas Filipinos' Cash Remittances. `https://www.bsp.gov.ph/SitePages/Statistics/Statistics.aspx`, 2025.

Marine Carrasco and Xiaohong Chen. Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models. *Econometric Theory*, 18(1):17–39, 2002.

Nidhaleddine Ben Cheikh and Younes Ben Zaied. Revisiting the pass-through of exchange rate in the transition economies: New evidence from new eu member states. *Journal of International Money and Finance*, 100:102093, 2020.

Xiaohong Chen, Qiman Shao, Weibiao Wu, and Lihu Xu. Self-normalized Cramér-type Moderate Deviations Under Dependence. *The Annals of Statistics*, 44(4): 1593–1617, 2016.

Victor Chernozhukov and Christian Hansen. The reduced form: A simple approach to inference with weak instruments. *Economics Letters*, 100(1):68–71, 2008.

ECB. Consumer confidence and household consumption decisions. *ECB Economic Bulletin, Issue 5/2025*, 2025.

Federal Reserve Bank of St. Louis. Federal Reserve Economic Data (FRED): Unemployment Rate, Foreign Born. `https://fred.stlouisfed.org/`, 2025.

Jordi Galì and Mark Gertler. Inflation Dynamics: A Structural Econometric Analysis. *Journal of Monetary Economics*, 44(2):195–222, 1999.

IMF. The rise of the silver economy: Global implications of population aging. *World Economic Outlook*, 2025.

International Monetary Fund. Sierra leone: 2024 article iv consultation and requests for ecf reviews. Technical report, IMF, 2024. URL `https://www.elibrary.imf.org/view/journals/002/2024/321/002.2024.issue-321-en.pdf`.

International Monetary Fund. International Financial Statistics. `https://data.imf.org/`, 2025.

Martina Jašová, Richhild Moessner, and Elod Takats. Exchange rate pass-through: What has changed since the crisis? *International Journal of Central Banking*, 15 (3):27–58, 2019.

Frank Kleibergen and Sophocles Mavroeidis. Weak Instrument Robust Tests in GMM and the New Keynesian Phillips Curve. *Journal of Business & Economic Statistics*, 27(3):293–311, 2009.

Eben Lazarus, Daniel J Lewis, James H Stock, and Mark W Watson. Har inference: Recommendations for practice. *Journal of Business & Economic Statistics*, 36(4): 541–559, 2018.

Arthur Lewbel. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic statistics*, 30(1): 67–80, 2012.

Daniel J Lewis. Identifying shocks via time-varying volatility. *The Review of Economic Studies*, 88(6):3086–3124, 2021.

Daniel J Lewis. Robust inference in models identified via heteroskedasticity. *Review of Economics and Statistics*, 104(3):510–524, 2022.

Daniel J Lewis. Identification based on higher moments. 2024.

Antonia López-Villavicencio and Valerie Mignon. Exchange rate pass-through in emerging countries: Do the inflation environment, monetary policy regime and central bank behavior matter? *Journal of International Money and Finance*, 79: 20–38, 2017.

Sophocles Mavroeidis. Identification Issues in Forward-Looking Models Estimated by GMM, with an Application to the Phillips Curve. *Journal of Money, Credit and Banking*, 37(3):421–448, 2005.

Sophocles Mavroeidis. Monetary Policy Rules and Macroeconomic Stability: Some New Evidence. *American Economic Review*, 100(1):491–503, 2010.

Sophocles Mavroeidis, Mikkel Plagborg-Møller, and James Stock. Empirical Evidence on Inflation Expectations in the New Keynesian Phillips Curve. *Journal of Economic Literature*, 52(1):124–188, 2014.

Harun Mirza and Lidia Storjohann. Making Weak Instrument Sets Stronger: Factor-Based Estimation of Inflation Dynamics and a Monetary Policy Rule. *Journal of Money, Credit and Banking*, 46(4):643–664, 2014.

Abdelkader Mokkadem. Mixing Properties of ARMA Processes. *Stochastic Processes and Their Applications*, 29(2):309–315, 1988.

Whitney Newey and Kenneth West. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3): 703–708, 1987. doi: 10.2307/1913610.

Whitney K Newey and Daniel McFadden. Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4:2111–2245, 1994.

Todd Prono. The role of conditional heteroskedasticity in identifying and estimating linear triangular systems, with applications to asset pricing models that include a mismeasured factor. *Journal of Applied Econometrics*, 29(5):800–824, 2014.

Roberto Rigobon. Identification through heteroskedasticity. *Review of Economics and Statistics*, 85(4):777–792, 2003.

Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data.* MIT press, 2010.

Appendix A: Proofs of Results

This appendix contains the proof of Theorem 4.1 and the identification result for the case of multiple endogenous variables. The online appendix contains the lemmas and their proofs. We note that some of the lemmas assume different mixing rates, but we note that the mixing rate in Assumption 4.1. (a) implies the ones considered in this appendix since for $r > 2$, $\frac{2r}{r-2} \geq \frac{r}{r-2} \geq \frac{r}{r-1}$.

Identification

Re-write the structural equation and the reduced-form projection as

$$\begin{aligned} y_t &= \beta'(\zeta' Q_t) + \delta' X_t + W_{1,t} \\ Y_t &= \zeta' Q_t + W_{2,t}. \end{aligned}$$

where $W_{1,t} = \varepsilon_t + \beta' v_t$ and $W_{2,t} = v_t$. By the instrument validity condition (2.4),

$$\begin{aligned} \bar{\mathbb{E}}[v_t \otimes (Z_t - \bar{\mu})\varepsilon_t] = 0 &\iff \bar{\mathbb{E}}[W_{2,t} \otimes (Z_t - \bar{\mu})(W_{1,t} - \beta' W_{2,t})] = 0 \\ &\iff \bar{\mathbb{E}}[W_{2,t} \otimes (Z_t - \bar{\mu})W_{1,t}] - \bar{\mathbb{E}}[W_{2,t}W_{2,t}' \otimes (Z_t - \bar{\mu})]\beta = 0, \end{aligned}$$

so that β is identified whenever $\bar{\mathbb{E}}[W_{2,t}W_{2,t}' \otimes (Z_t - \bar{\mu})]$ has full rank.

Expression for $\hat{V}$

$$\hat{V} = \frac{1}{T}\sum_{t=1}^{T} \hat{\mathcal{Z}}_t \hat{\mathcal{Z}}_t' + \frac{1}{T}\sum_{\tau=1}^{m} w_\tau \sum_{t=\tau+1}^{T} (\hat{\mathcal{Z}}_t \hat{\mathcal{Z}}_{t-\tau}' + \hat{\mathcal{Z}}_{t-\tau}\hat{\mathcal{Z}}_t'), \quad w_\tau = 1 - \frac{\tau}{m+1}, \tag{A.1}$$

where

$$\hat{\mathcal{Z}}_t = \begin{bmatrix} X_t \hat{\epsilon}_t([\hat{\mu}', \hat{\zeta}']') \\ \hat{\epsilon}_t([\hat{\mu}', \hat{\zeta}']')\hat{v}_t(\hat{\zeta}) \otimes (Z_t - \hat{\mu}) + \left(\hat{\varrho}([\hat{\mu}', \hat{\zeta}']', \hat{\zeta}) \otimes I_{k_Z}\right)(Z_t - \hat{\mu}) - \left(I_g \otimes \hat{\theta}([\hat{\mu}', \hat{\zeta}']', \hat{\mu})\right)(\hat{v}_t(\hat{\zeta}) \otimes \hat{M}_Q^{-1} Q_t) \end{bmatrix},$$

$\hat{\epsilon}_t([\hat{\mu}', \hat{\zeta}'] = y_{t,0} - \hat{\gamma}' W_t([\hat{\mu}', \hat{\zeta}']')$, $\hat{\varrho}([\hat{\mu}', \hat{\zeta}']', \hat{\zeta}) = \frac{1}{T}\sum_{t=1}^{T} v_t(\hat{\zeta})\hat{\epsilon}_t([\hat{\mu}', \hat{\zeta}']')$, $\hat{\theta}([\hat{\mu}', \hat{\zeta}']', \hat{\mu}) = \frac{1}{T}\sum_{t=1}^{T} \hat{\epsilon}_t([\hat{\mu}', \hat{\zeta}']')(Z_t - \hat{\mu})Q_t'$, and $\hat{M}_Q = \frac{1}{T}\sum_{t=1}^{T} Q_t Q_t'$.

Proof of Theorem 4.1

The re-scaled and centred OLS estimator can be expressed as

$$\sqrt{T}(\hat{\gamma} - \gamma) = \underbrace{\left(\frac{1}{T}\sum_{t=1}^{T} W_t([\hat{\mu}', \hat{\zeta}']') W_t'([\hat{\mu}', \hat{\zeta}']')\right)^{-1}}_{I} \underbrace{\frac{1}{\sqrt{T}}\sum_{t=1}^{T} W_t([\hat{\mu}', \hat{\zeta}']')\left[(W_t - W_t([\hat{\mu}', \hat{\zeta}']'))'\gamma + \varepsilon_t\right]}_{II}.$$

By Lemma 1.1, $I - M_{W,T}^{-1} = o_p(1)$, and by Lemma 1.2, $[\hat{\mu}', \hat{\zeta}']' - [\bar{\mu}', \zeta'] = o_p(1)$. Thus, letting $\tilde{\lambda} = [\tilde{\mu}', \tilde{\xi}_1', \ldots, \tilde{\xi}_g']'$ and $\lambda = [\bar{\mu}', \xi_1', \ldots, \xi_g']'$, a mean-value expansion gives

$$II = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} W_t \varepsilon_t + \frac{1}{T}\left[\sum_{t=1}^{T} \underbrace{\varepsilon_t \nabla_{\tilde{\lambda}} \,|_{\tilde{\lambda}=\lambda} W_t(\tilde{\lambda}) - W_t \gamma' \nabla_{\tilde{\lambda}} \,|_{\tilde{\lambda}=\lambda} W_t(\tilde{\lambda})}_{II_{a,t}}\right] \underbrace{\sqrt{T}([\hat{\mu}', \hat{\zeta}']' - \lambda)}_{II_b} + o_p(1).$$

Letting $\mathfrak{z}_{t,j} = (Z_{t,j} - \tilde{\mu}_j)$, $\nabla_{\tilde{\lambda}} W_t(\tilde{\lambda})$ is given by

$$\nabla_{\tilde{\lambda}} W_t(\tilde{\lambda}) = \begin{bmatrix}
\frac{\partial X_{t,1}}{\partial \tilde{\mu}_1} & \ldots & \frac{\partial X_{t,1}}{\partial \tilde{\mu}_{k_Z}} & \frac{\partial X_{t,1}}{\partial \tilde{\zeta}_{1,1}} & \ldots & \frac{\partial X_{t,1}}{\partial \tilde{\zeta}_{1,k_Q}} & \ldots & \frac{\partial X_{t,1}}{\partial \tilde{\zeta}_{g,1}} & \ldots & \frac{\partial X_{t,1}}{\partial \tilde{\zeta}_{g,k_Q}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial X_{t,p}}{\partial \tilde{\mu}_1} & \ldots & \frac{\partial X_{t,p}}{\partial \tilde{\mu}_{k_Z}} & \frac{\partial X_{t,p}}{\partial \tilde{\zeta}_{1,1}} & \ldots & \frac{\partial X_{t,p}}{\partial \tilde{\zeta}_{1,k_Q}} & \ldots & \frac{\partial X_{t,p}}{\partial \tilde{\zeta}_{g,1}} & \ldots & \frac{\partial X_{t,p}}{\partial \tilde{\zeta}_{g,k_Q}} \\
\frac{\partial \mathfrak{z}_{t,1} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\mu}_1} & \ldots & \frac{\partial \mathfrak{z}_{t,1} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\mu}_{k_Z}} & \frac{\partial \mathfrak{z}_{t,1} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\zeta}_{1,1}} & \ldots & \frac{\partial \mathfrak{z}_{t,1} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\zeta}_{1,k_Q}} & \ldots & \frac{\partial \mathfrak{z}_{t,1} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\zeta}_{g,1}} & \ldots & \frac{\partial \mathfrak{z}_{t,1} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\zeta}_{g,k_Q}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial \mathfrak{z}_{t,k_Z} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\mu}_1} & \ldots & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\mu}_{k_Z}} & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\zeta}_{1,1}} & \ldots & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\zeta}_{1,k_Q}} & \ldots & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\zeta}_{g,1}} & \ldots & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,1}(\tilde{\zeta})}{\partial \tilde{\zeta}_{g,k_Q}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial \mathfrak{z}_{t,1} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\mu}_1} & \ldots & \frac{\partial \mathfrak{z}_{t,1} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\mu}_{k_Z}} & \frac{\partial \mathfrak{z}_{t,1} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\zeta}_{1,1}} & \ldots & \frac{\partial \mathfrak{z}_{t,1} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\zeta}_{1,k_Q}} & \ldots & \frac{\partial \mathfrak{z}_{t,1} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\zeta}_{g,1}} & \ldots & \frac{\partial \mathfrak{z}_{t,1} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\zeta}_{g,k_Q}} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial \mathfrak{z}_{t,k_Z} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\mu}_1} & \ldots & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\mu}_{k_Z}} & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\zeta}_{1,1}} & \ldots & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\zeta}_{1,k_Q}} & \ldots & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\zeta}_{g,1}} & \ldots & \frac{\partial \mathfrak{z}_{t,k_Z} v_{t,g}(\tilde{\zeta})}{\partial \tilde{\zeta}_{g,k_Q}}
\end{bmatrix}.$$

Notice that for all $j = 1, \ldots, k_Z$, $l = 1, \ldots, p$, $m = 1, \ldots, g$, and $s = 1, \ldots, k_Q$,

$$\frac{\partial X_{t,l}}{\partial \tilde{\mu}_j} = \frac{\partial X_{t,l}}{\partial \tilde{\xi}_{m,s}} = 0.$$

For all $j, l = 1, \ldots, k_Z$ and $m = 1, \ldots, g$

$$\frac{\partial \mathfrak{z}_{t,j} v_{t,m}(\tilde{\zeta})}{\partial \tilde{\mu}_l} = \begin{cases} -v_{t,m}(\tilde{\zeta}) & \text{if } j = l \\ 0 & \text{otherwise,} \end{cases}$$

and for $j = 1, \ldots, k_Z$, $l = 1, \ldots, k_Q$, $m = 1, \ldots, g$, and $q = 1, \ldots, g$

$$\frac{\partial \mathfrak{z}_{t,j} v_{t,m}(\tilde{\zeta})}{\partial \tilde{\xi}_{q,l}} = \begin{cases} -Q_{t,l}(Z_{t,j} - \tilde{\mu}_j) & \text{if } m = q \\ 0 & \text{otherwise.} \end{cases}$$

Thus,

$$\begin{aligned} \nabla_{\tilde{\lambda}} \left.\right|_{\tilde{\lambda}=\lambda} W_t(\tilde{\lambda}) &= \begin{bmatrix} 0_{p \times k_Z} & 0_{p \times 1} & 0_{p \times 1} & \cdots & 0_{p \times 1} \\ v_{t,1} I_{k_Z} & -(Z_t - \bar{\mu})Q_t' & 0_{k_Z \times 1} & \ddots & 0_{k_Z \times 1} \\ v_{t,2} I_{k_Z} & 0_{k_Z \times 1} & -(Z_t - \bar{\mu})Q_t' & \ddots & 0_{k_Z \times 1} \\ \vdots & \vdots & \ddots & \ddots & 0_{k_Z \times 1} \\ v_{t,g} I_{k_Z} & 0_{k_Z \times 1} & \cdots & 0_{k_Z \times 1} & -(Z_t - \bar{\mu})Q_t' \end{bmatrix} \\ &= \begin{bmatrix} 0_{p \times k_Z} & 0_{p \times g k_Q} \\ v_t \otimes I_{k_Z} & -I_g \otimes (Z_t - \bar{\mu})Q_t' \end{bmatrix}. \end{aligned}$$

Recall that under the null, $\gamma = [\delta', \alpha']' = [\delta', 0'_{gk_Z \times 1}]'$. Thus,

$$\gamma' \nabla_{\tilde{\lambda}} |_{\tilde{\lambda}=\lambda} W_t(\tilde{\lambda}) = [\delta', 0'_{gk_Z\times 1}] \begin{bmatrix} 0_{p\times k_Z} & 0_{p\times k_Q} & 0_{p\times k_Q} & \dots & 0_{p\times k_Q} \\ v_{t,1} I_{k_Z} & -(Z_t - \bar{\mu})Q'_t & 0_{k_Z\times k_Q} & \ddots & 0_{k_Z\times k_Q} \\ v_{t,2} I_{k_Z} & 0_{k_Z\times k_Q} & -(Z_t - \bar{\mu})Q'_t & \ddots & 0_{k_Z\times k_Q} \\ \vdots & \vdots & \ddots & \ddots & 0_{k_Z\times k_Q} \\ v_{t,g} I_{k_Z} & 0_{k_Z\times k_Q} & \dots & 0_{k_Z\times k_Q} & -(Z_t - \bar{\mu})Q'_t \end{bmatrix}$$

$$= 0_{1\times(k_Z+gk_Q)},$$

implying $W_t \gamma' \nabla_{\tilde{\lambda}} |_{\tilde{\lambda}=\lambda} W_t(\tilde{\lambda}) = 0_{(p+gk_Z)\times(k_Z+gk_Q)}$. Hence,

$$II_{a,t} = \varepsilon_t \nabla_{\tilde{\lambda}} |_{\tilde{\lambda}=\lambda} W_t(\tilde{\lambda}) = \begin{bmatrix} 0_{p\times k_Z} & 0_{p\times gk_Q} \\ \varepsilon_t v_t \otimes I_{k_Z} & -\varepsilon_t I_g \otimes (Z_t - \bar{\mu})Q'_t \end{bmatrix},$$

and by Lemma 1.4 and Lemma 1.5, $\frac{1}{T}\sum_{t=1}^T II_{a,t} - \mathbb{E}[\frac{1}{T}\sum_{t=1}^T II_{a,t}] = o_p(1)$ (where the lemmas are applied elementwise).

By Lemma 1.7, $\frac{1}{\sqrt{T}}\sum_{t=1}^T Q_t v_{t,j} = O_p(1)$ for all $j = 1, \dots, g$, implying together with Lemma 1.2 that

$$II_b = \frac{1}{\sqrt{T}} \sum_{t=1}^T \begin{bmatrix} Z_t - \bar{\mu} \\ M_{Q,T}^{-1} Q_t v_{t,1} \\ \vdots \\ M_{Q,T}^{-1} Q_t v_{t,g} \end{bmatrix} + o_p(1), \tag{A.2}$$

where, as in the statement of Lemma 1.2, $M_{Q,T} \equiv \mathbb{E}[\frac{1}{T}\sum_{t=1}^T Q_t Q'_t]$.

Letting $G_T \equiv \bar{\mathbb{E}}[\frac{1}{T}\sum_{t=1}^{T} II_{a,t}]$, we hence have

$$II = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} W_t\varepsilon_t + \frac{1}{\sqrt{T}}\sum_{t=1}^{T} G_T \begin{bmatrix} Z_t - \bar{\mu} \\ M_{Q,T}^{-1}Q_t v_{t,1} \\ \vdots \\ M_{Q,T}^{-1}Q_t v_{t,g} \end{bmatrix} + o_p(1)$$

$$= \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \underbrace{\begin{bmatrix} X_t\varepsilon_t \\ \varepsilon_t v_t \otimes (Z_t - \bar{\mu}) + \left(\bar{\mathbb{E}}[\varepsilon_t v_t] \otimes I_{k_Z}\right)(Z_t - \bar{\mu}) - \left(I_g \otimes \bar{\mathbb{E}}[\varepsilon_t (Z_t - \bar{\mu})Q_t']\right)(v_t \otimes M_{Q,T}^{-1}Q_t) \end{bmatrix}}_{\mathcal{Z}_t}$$

$$+ o_p(1),$$

so that

$$\sqrt{T}(\hat{\gamma} - \gamma) = M_{W,T}^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_t + o_p(1).$$

By Lemma 1.6, for $V_{\mathcal{Z},T} \equiv \text{var}(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_t)$, $V_{\mathcal{Z},T}^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_t \xrightarrow{d} \mathcal{N}[0, I_{(p+k_Z g)\times(p+k_Z g)}]$, so that

$$\sqrt{T}(\hat{\gamma} - \gamma) = M_{W,T}^{-1}V_{\mathcal{Z},T}^{1/2}V_{\mathcal{Z},T}^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{Z}_t + o_p(1),$$

and hence

$$M_{W,T}V_{\mathcal{Z},T}^{-1/2}\sqrt{T}(\hat{\gamma} - \gamma) \xrightarrow{d} \mathcal{N}[0, I_{(p+k_Z g)\times(p+k_Z g)}].$$

By Lemma 1.8, $V_{\mathcal{Z},T}$ can be consistently estimated by $\hat{V}$, so that the required result follows.

PUBLICATIONS

Robust Inference Via Heteroskedasticity in Linear Models

Working Paper No. WP/2026/100