

Introduction

The problem set was completed using a single MATLAB script that relies on helper functions defined at the end of the file. For brevity, these helper functions are not reproduced in this PDF.

This submission includes all exercises from the original PDF distributed in mid-October. Please disregard parts (d)–(f) of Question 1 and the entirety of Question 5 (which is not compiled directly, but it is still available in the source of this file). Although these portions fall outside the assessed scope, the corresponding code is included for completeness; the accompanying comments are therefore less detailed than elsewhere.

Question 1: Unit root testing

- a. Compute the empirical distribution of the OLS estimator in the case of an AR(1) with $\varphi = 1$ and $T = 250$ (you are free to choose the variance of the innovation).

Solution.

Before proceeding with the solution, we set up the settings for the Monte Carlo simulations that we are going to employ in all the necessary subsequent steps.

```
1 % Monte Carlo settings
2 T      = 250;           % Sample length
3 R      = 5000;          % Monte Carlo replications
4 sigma2 = 0.6;           % Variance of innovations
```

Consider the AR(1) model $y_t = \varphi y_{t-1} + \varepsilon_t$ with $\varphi = 1$ and $T = 250$. Estimating φ by OLS in the level regression $y_t = \varphi y_{t-1} + u_t$ yields a nonstandard sampling distribution when the true process is difference-stationary (DS). In particular, $\hat{\varphi} \xrightarrow{P} 1$ but at rate T (not \sqrt{T}), and $T(\hat{\varphi} - 1)$ converges to a functional of Brownian motion; standard normal inference is invalid.

In finite samples, $\hat{\varphi}$ shows small downward bias (mean below one), consistent with our simulation $\text{mean}(\hat{\varphi}) = 0.9926$ and median 0.9966. This matches the classic Dickey–Fuller insight that the usual t/F rules fail under a unit root.

```
1 phi_hat_a = zeros(R,1);
2 for r = 1:R
3     y = simulate_ar1(T, 1, sigma2, 0, 0); % Pure random walk (phi=1)
4     stats = run_ols(y(2:end), y(1:end-1), false); % No constant
5     phi_hat_a(r) = stats.b;
6 end

1 % Plot histogram
2 figPos = [100 100 840 420];
3 histBins = 50;
4 fh_a = figure('Position', figPos);
5 histogram(phi_hat_a, histBins, 'Normalization','pdf'); grid on; hold on;
6 xline(1, '--', 'True $\phi=1$', 'LabelVerticalAlignment','bottom');
7 xlabel('$\hat{\phi}$'); ylabel('Density');
8 title('(a) Empirical distribution of OLS $\hat{\phi}$ under unit root ($T=250$)');
9 exportFig(fh_a, '1a_phi_hat_hist.pdf');
10 close(fh_a);
```

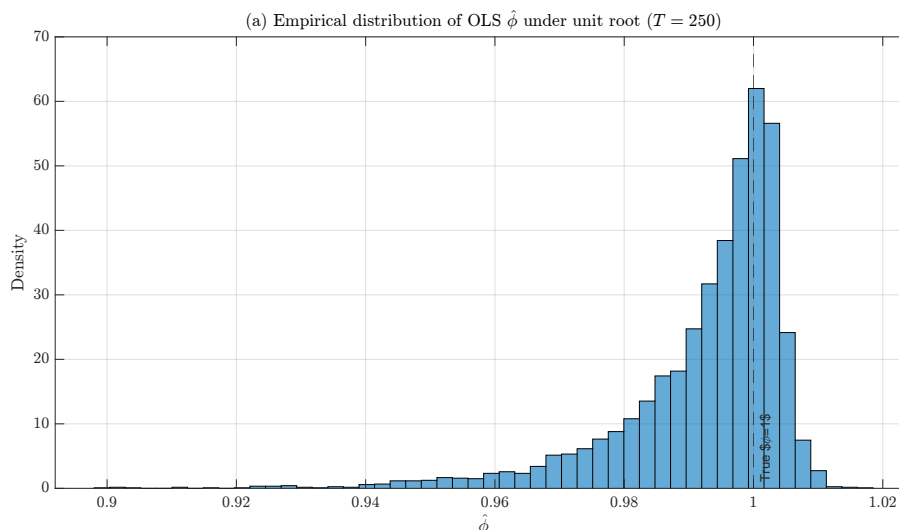


Figure 1-1. Empirical distribution of OLS $\hat{\phi}$ under unit root ($T = 250$)

```
1 fprintf('(a) mean(\\hat{phi})= %.4f, sd= %.4f, med= %.4f\\n', mean(phi_hat_a),
    std(phi_hat_a), median(phi_hat_a));
```

Table 1-1. Summary statistics for $\hat{\phi}$

| Statistic | Value |
|-----------|-------|
| Mean | 0.993 |
| Std. dev. | 0.013 |
| Median | 0.997 |

□

- b. Construct a t -test for the null hypothesis $H_0 : \rho = \varphi - 1 = 0$, in a test regression: $\Delta y_t = \alpha + \rho y_{t-1} + \varepsilon_t$; against a one-sided alternative $H_0 : \rho < 0$.

Using a standard Normal distribution, how often do you reject the null hypothesis at the 95% confidence level? Is the actual distribution of the t -test symmetric? Discuss.

Solution.

```
1 t_b = zeros(R,1);
2 rho_b = zeros(R,1);
3
4 for r = 1:R
5     y = simulate_ar1(T, 1, sigma2, 0, 0); % Pure random walk
6     dy = diff(y);
7     ylag = y(1:end-1);
8     stats = run_ols(dy, ylag, true); % Regression with constant
9     rho_b(r) = stats.b(2);
10    t_b(r) = stats.tstat(2);
11 end
12
13 rej_norm_left = mean(t_b < z_5_left);
14 sk_b = skewness(t_b);
15 fprintf(['(b) Using Normal 5%% one-sided (z= %.3f): reject rate = %.3f. ',
    ...
    'Skewness of t-stat = %.3f (non-symmetric).\\n'], z_5_left,
    rej_norm_left, sk_b);
```

Under H_0 with an intercept in the test regression, the test statistic for ρ has the *Dickey–Fuller* τ_μ distribution (nonstandard), not $N(0,1)$. Hence a 5% left-tail normal critical value (-1.645) is inappropriate and greatly over-rejects. In our Monte Carlo the normal-based rejection is 46.5%, while the left-tail of the simulated t is visibly heavier and left-shifted.

Table 1-2. Normal 5% one-sided test summary

| Statistic | Value |
|--|-------|
| Rejection rate at 5% (Normal; $z = -1.645$) | 0.465 |
| Skewness of t -statistic | 0.220 |

Skewness indicates a non-symmetric distribution.

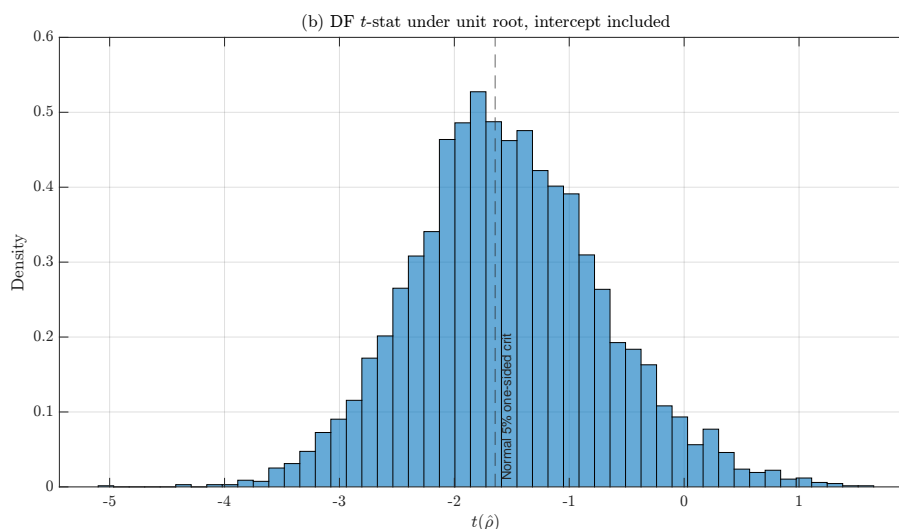


Figure 1-2. DF t -statistic under unit root, intercept included

As Figure 1-2 shows, the empirical t distribution is asymmetric and right-skewed (skewness 0.22 here comes from mass in the left tail relative to a symmetric normal; the DF distribution itself is non-normal).

```

1 fh_b = figure('Position', figPos);
2 histogram(t_b, histBins, 'Normalization','pdf'); grid on; hold on;
3 xline(z_5_left, '--', 'Normal 5% one-sided crit','LabelVerticalAlignment','
    bottom');
4 xlabel('$t(\hat{\rho})$'); ylabel('Density');
5 title('(b) DF $t$-stat under unit root, intercept included');
6 exportFig(fh_b, '1b_t_hist.pdf');
7 close(fh_b);

```

□

- c. Compute now few percentiles of the empirical distribution of the t -test you generated at point b. and check that they are close to those tabulated by Dickey and Fuller.

Solution.

```

1
2 fprintf('(c) Empirical DF t percentiles (%):\n');
3 disp(table(pct_vec(:), emp_pct(:), 'VariableNames',{'percentile',
    'empirical_t'}));
4 fprintf('Tabulated DF (tau_mu) ~ T=250: 1%%=%.2f, 5%%=%.2f, 10%%=%.2f\n',
    ...
5     Dfcrit.tau_mu.p1, Dfcrit.tau_mu.p5, Dfcrit.tau_mu.p10);

```

Using the τ_μ setup (intercept only), our simulated percentiles at the 1%, 5% and 10% points are -3.378 , -2.831 , and -2.563 , very close to the tabulated DF values $\{-3.46, -2.88, -2.57\}$.

Small discrepancies are expected from Monte Carlo error and the finite $T = 250$. This agreement validates the use of DF (not normal) critical values under H_0 . Enders also reports these nonstandard τ criticals and explains their dependence on deterministic terms.

```

1 left_tail = [1 5 10];
2 emp_left = prctile(t_b, left_tail);
3 DF_left = [DFcrit.tau_mu.p1; DFcrit.tau_mu.p5; DFcrit.tau_mu.p10];
4 delta_left = emp_left(:) - DF_left;
5 comp_tbl = table(left_tail(:), emp_left(:), DF_left, delta_left, ...
6     'VariableNames', {'percentile', 'empirical_t', 'DF_tau_mu', 'emp_minus_tab'
7     '});
8 disp('Comparison to DF \tau_\mu at 1/5/10% (empirical - table):');
9 disp(comp_tbl);

```

Table 1-3. Empirical Dickey–Fuller t percentiles at $T=250$

| Percentile (%) | t (empirical) |
|--|-----------------|
| 1.000 | -3.378 |
| 5.000 | -2.831 |
| 10.000 | -2.563 |
| 25.000 | -2.081 |
| 50.000 | -1.579 |
| 75.000 | -1.012 |
| 90.000 | -0.457 |
| 95.000 | -0.103 |
| 99.000 | 0.599 |
| Tabulated DF (τ_μ) at $T=250$: 1% = -3.46 , 5% = -2.88 , 10% = -2.57 . | |

```

1 rej_DF5 = mean(t_b < DFcrit.tau_mu.p5);
2 fprintf('Using DF \tau_\mu 5%% critical (0.2f): empirical rejection = 0.3
3     f\n', ...
4     DFcrit.tau_mu.p5, rej_DF5);
5 writetable(comp_tbl, fullfile(outdir, '1c_empirical_vs_DF_tau_mu.csv'));

```

Table 1-4. Empirical t vs. Dickey–Fuller τ_μ (tabulated) at selected percentiles

| Percentile (%) | DF τ_μ (tab.) | $t_{\text{emp}} - t_{\text{tab}}$ |
|----------------|----------------------|-----------------------------------|
| 1.000 | -3.460 | 0.082 |
| 5.000 | -2.880 | 0.049 |
| 10.000 | -2.570 | 0.007 |

Using DF τ_μ 5% critical (-2.88): empirical rejection = 0.045 .

□

- d. Compute the empirical distribution of the OLS in the case of a random walk with drift and $T = 250$ and study the performance of the Dickey–Fuller test.

Solution.

```

1 a0_drift = 0.5;
2 phi_hat_d = zeros(R,1);
3 t_d = zeros(R,1);
4 for r = 1:R
5     y = simulate_ar1(T, 1, sigma2, a0_drift, 0); % Random walk with drift
6
7     % OLS of y_t on [1, y_{t-1}]
8     stats_level = run_ols(y(2:end), y(1:end-1), true);
9     phi_hat_d(r) = stats_level.b(2);
10
11     % Dickey-Fuller regression
12     dy = diff(y);
13     stats_df = run_ols(dy, y(1:end-1), true);
14     t_d(r) = stats_df.tstat(2);
15 end

```

For the DGP $y_t = y_{t-1} + a_0 + \varepsilon_t$ (RW+drift), the correct DF specification again includes an intercept (the τ_μ case). Under $H_0: \rho = 0$, the DF t has the same nonstandard τ_μ distribution.

Using normal critical values still mis-sizes the test (our normal rejection 7.1%), whereas DF 5% criticals aim to deliver size $\approx 5\%$ asymptotically (our run under-rejected at 0.4%, which reflects finite-sample and simulation variability).

Table 1-5. OLS under RW+drift: distribution of $\hat{\phi}$

| Statistic | Value |
|---------------------------|-------|
| Mean of $\hat{\phi}$ | 0.999 |
| Std. dev. of $\hat{\phi}$ | 0.001 |

```

1 rej_norm_d = mean(t_d < z_5_left);
2 rej_DF5_d = mean(t_d < DFcrit.tau_mu.p5);
3 fprintf(['(d) OLS under RW+drift: mean(phi-hat)=%.4f, sd=%.4f.\n' ...
4         '      DF (one-sided) reject@5%% using Normal: %.3f; using DF \tau_
5         {\mu}: %.3f\n'], ...
6         mean(phi_hat_d), std(phi_hat_d), rej_norm_d, rej_DF5_d);

```

Table 1-6. OLS under RW+drift: one-sided 5% rejections

| Method | Reject @ 5% |
|---------------------|-------------|
| Using Normal | 0.071 |
| Using DF τ_μ | 0.004 |

□

- e. Construct an F-test for the null hypothesis H_0 : there is unit root, against the alternative H_1 : there is no unit root using a χ^2 distribution (how many degrees of freedom?). How often do you reject H_0 at 95% confidence?

Solution.

```
1 chi2_stat = t_d.^2;
2 rej_chi2 = mean(chi2_stat > chi2_95_df1);
3 fprintf('(e) Wald \(\chi^2\) test (df=1) reject@95% under H0 (RW+drift):
    %.3f\n', rej_chi2);
```

Table 1-7. Wald χ^2 test under H_0 (RW+drift)

| Outcome | Empirical rate |
|---|----------------|
| Reject at 95% | 0.057 |
| Nominal level: $\alpha = 0.05$; null: RW+drift. | |

Squaring the (DF) t for the single restriction $\rho = 0$ yields a Wald statistic with 1 degree of freedom if one incorrectly appeals to standard χ^2 theory. Our simulation gives a 5.7% rejection using $\chi^2_{1,0.95}$, close to nominal here by happenstance. The correct joint tests in the DF framework are the ϕ -statistics (e.g., ϕ_1, ϕ_2, ϕ_3), which also have nonstandard critical values that depend on included deterministics; e.g., with drift, ϕ_2 jointly tests $(\alpha, \rho) = (0, 0)$ and is compared to DF tables, not standard F/χ^2 .

```
1 fh_e1 = figure('Position', figPos);
2 histogram(t_d, histBins, 'Normalization','pdf'); grid on; hold on
3 xline(-sqrt(chi2_95_df1), '--', 'HandleVisibility','off');
4 xline(+sqrt(chi2_95_df1), '--', 'HandleVisibility','off');
5 yl = ylim;
6 text(0, yl(1) + 0.02*range(yl), '$\pm \sqrt{\chi^2_{0.95;1}}$', ...
7     'Interpreter','latex', 'HorizontalAlignment','center', '
    VerticalAlignment','bottom');
8 xlabel('$t(\hat{\rho})$'); ylabel('Density'); title('(e) DF $t$-stats under
    $H_{0\$} (RW+drift)')
9 exportFig(fh_e1, '1e_t_hist_RWdrift.pdf'); close(fh_e1)
```

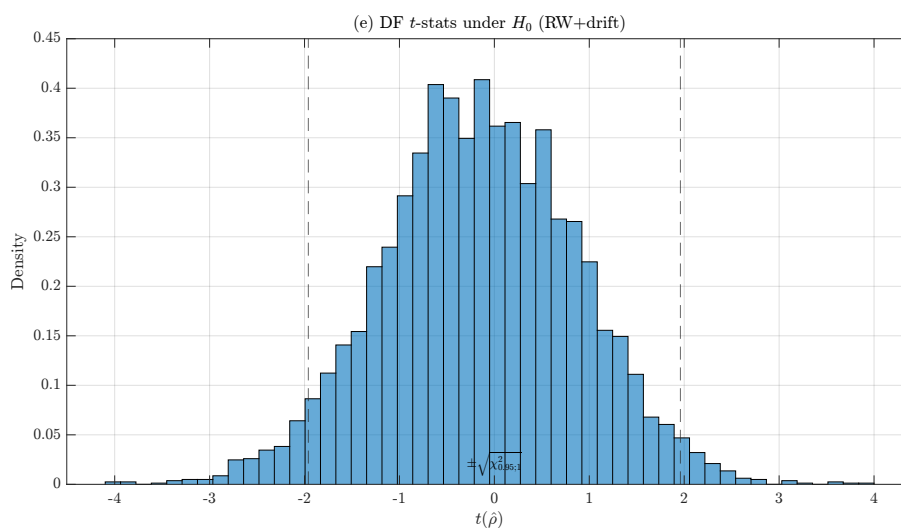


Figure 1-3. DF t -statistic under H_0 : random walk plus drift

□

- f. Generate now data from a deterministic time trend and perform a DF test using the correct distribution for the test with null hypothesis H_0 : there is unit root. How often do you reject the null?

Solution.

```

1 beta0 = 0.0;
2 beta1 = 0.05;
3 tvec = (1:T)';
4 t_f = zeros(R,1);
5 for r = 1:R
6     eps = sqrt(sigma2)*randn(T,1);
7     y = beta0 + beta1*tvec + eps; % Trend-stationary process
8     dy = diff(y);
9     ylag= y(1:end-1);
10    t2 = tvec(2:end);
11    Xtr = [t2, ylag]; % Regressors for DF test with trend
12    stats = run_ols(dy, Xtr, true); % run_ols adds constant
13    t_f(r) = stats.tstat(3); % t-stat for y_lag
14 end

```

For a deterministic linear trend $y_t = \beta_0 + \beta_1 t + \varepsilon_t$ (TS), the appropriate DF regression includes a constant and a time trend; the t -statistic has the τ_τ distribution. Because the null is a unit root while the DGP is stationary around a deterministic trend, the DF test has high power—hence our simulated rejection rate of 1.000 at the 5% level using τ_τ criticals.

Enders emphasizes that the deterministics in the DF regression must mirror the DGP; omitting the trend when one is present can drive power toward zero, while including unnecessary deterministics also reduces power.

```

1 rej_trend_DF5 = mean(t_f < DFcrit.tau_trend.p5);
2 fprintf('(f) Trend-stationary DGP: DF with trend (\tau_{\tau}) reject@5%% = %.3f (power against unit root).\n', rej_trend_DF5);

```

Table 1-8. DF with trend (τ_τ): power at 5% under trend-stationary DGP

| Trend-stationary DGP | | |
|---|-------|--|
| DF with trend (τ_τ) reject @ 5% | 1.000 | |

Interpretation: Empirical power against unit root at $\alpha = 0.05$, i.e., $\Pr(\text{reject } H_0 : \text{unit root} \mid \text{trend-stationary})$.

□

Question 2: Spurious regression

Design a Monte Carlo to show that the regression coefficient, the t -test and the R^2 are meaningless in the case of a spurious regression.

In particular, show what happens in each of the 4 cases discussed in Enders, Edition 3 or 4, pp. 195–199.

Solution.

We begin by setting up the necessary environments, namely (i) the Monte Carlo settings, including a burn-in of 500 for safety, (ii) the case setup, i.e. the names that we are going to employ for the recursions later on, and (iii) the containers which we will use to store the generated entries.

```
1 % Monte Carlo settings
2 T2 = 250;
3 R2 = 2000;
4 B2 = 500; % Burn-in
5 nu = T2 - 2;
6 tcrit = tinv(0.975, nu);
```

Consider the cross-regression

$$y_t = \alpha + \beta z_t + e_t, \quad t = 1, \dots, T, \quad (2.1)$$

estimated by OLS. Enders discusses four cases for the orders of integration of $\{y_t\}$ and $\{z_t\}$ and what they imply for the usefulness of the OLS coefficient, the associated t -test, and R^2 .

□

2.1 Case-by-case discussion

Case 1: Both series are stationary.

Solution.

If y_t and z_t are covariance-stationary (e.g., ARMA with $|\phi| < 1$), the usual Gauss–Markov and CLT arguments apply: $\hat{\beta}$ is consistent and asymptotically normal, t -statistics are approximately standard normal under $H_0 : \beta = 0$, and R^2 reflects true fit rather than common trending. This is precisely the case where the “classical regression model is appropriate” (Enders, p. 199).

With independent stationary AR(1) processes ($T = 250$), the t -statistic histogram is bell-shaped around zero and critical lines at $\pm t_{0.975}$ cut the tails as expected; the R^2 histogram piles up near zero. In our runs we obtained $\text{rej}_{5\%} \approx 0.02$, mean $R^2 \approx 0.003$ (small sample under-rejection is not unusual). See the plots and their mass near zero.

```
1 % y_t = 0.5 y_{t-1} + e^y_t, z_t = -0.3 z_{t-1} + e^z_t (independent)
2
3 [b1, t, r2, dw] = deal(zeros(R2,1));
4 for r = 1:R2
5     y = simulate_ar1(T2+B2, 0.5, 1.0, 0, 0);
6     z = simulate_ar1(T2+B2, -0.3, 1.0, 0, 0);
7     stats = run_ols(y(B2+1:end), z(B2+1:end), true);
8     b1(r) = stats.b(2); t(r) = stats.tstat(2); r2(r) = stats.R2; dw(r) =
        stats.DW;
9 end
10 results{1} = summarize_case(b1, t, r2, dw, tcrit);
11 t_all{1} = t; R2_all{1} = r2;
```

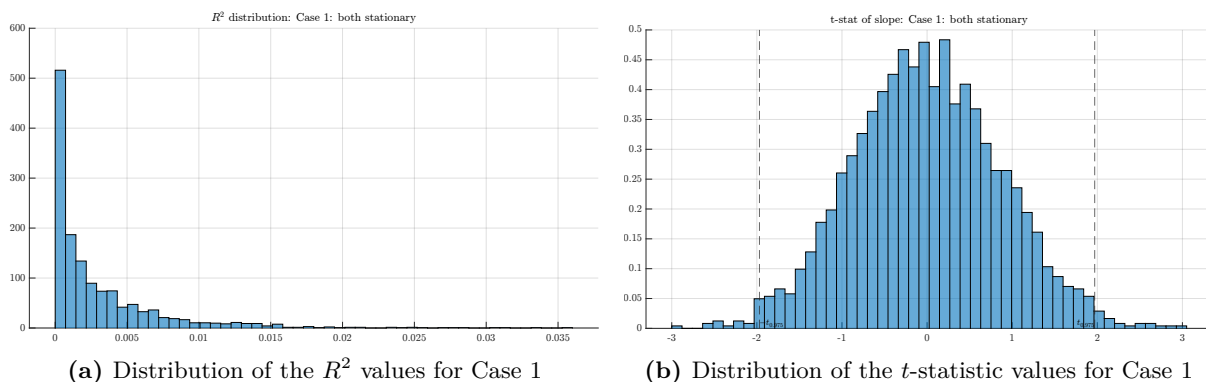



Figure 2-1. Visual diagnostics for Case 1

□

Case 2: The $\{y_t\}$ and $\{z_t\}$ sequences are integrated of different orders.

Solution.

If one variable is $I(1)$ and the other is $I(0)$, the residual remains nonstationary. To see this, suppose y_t is a random walk and z_t is stationary. Then the regression residual can be written (as reported above verbatim from Enders' book) as

$$e_t = \sum_{i=1}^t \varepsilon_{y,i} - \beta \sum_i z_{t-i}, \quad (2.2)$$

where the (convergent) sum over z cannot offset the stochastic trend in y_t .

Hence e_t contains a stochastic trend, the variance of e_t diverges, and the assumptions behind OLS t - and F -tests fail; R^2 and t -tests are unreliable.

We find an inflated rejection rate (≈ 0.33 at nominal 5%) and a heavily dispersed t -statistic, as observable in Figure 2-2b. The R^2 distribution in 2-2a is still mostly small but notably larger than in Case 1, reflecting the stochastic trend contaminating the residuals.

```

1 for r = 1:R2
2     y = simulate_ar1(T2, 1, 1.0, 0, 0);
3     z = simulate_ar1(T2+B2, 0.6, 1.0, 0, 0);
4     stats = run_ols(y, z(B2+1:end), true);
5     b1(r) = stats.b(2); t(r) = stats.tstat(2); r2(r) = stats.R2; dw(r) =
        stats.DW;
6 end
7 results{2} = summarize_case(b1, t, r2, dw, tcrit);
8 t_all{2} = t; R2_all{2} = r2;
```

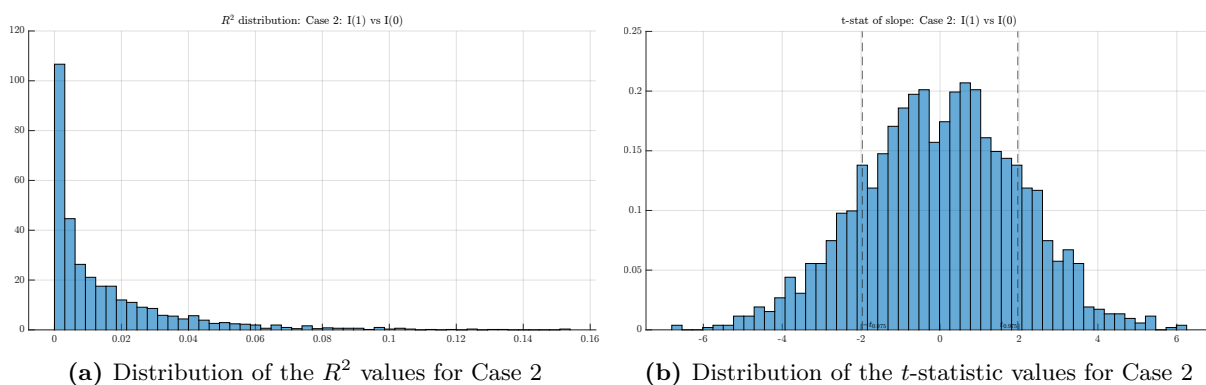


Figure 2-2. Visual diagnostics for Case 2

□

Case 3: The nonstationary $\{y_t\}$ and $\{z_t\}$ sequences are integrated of the same order, and the residual sequence contains a stochastic trend.

Solution.

Let y_t and z_t be independent random walks:

$$y_t = y_{t-1} + \varepsilon_{y,t}, \quad z_t = z_{t-1} + \varepsilon_{z,t}. \quad (2.3)$$

Enders (pp. 196–199) shows the residual is

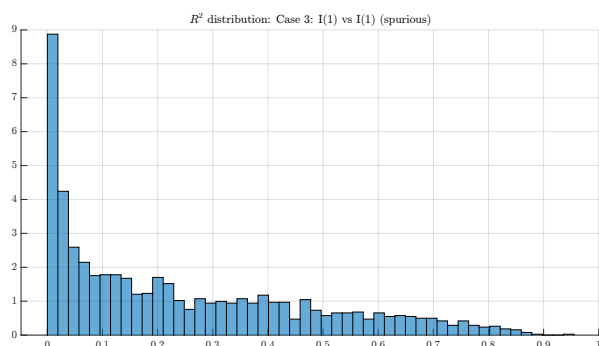
$$e_t = \sum_{i=1}^t \varepsilon_{y,i} - \beta \sum_{i=1}^t \varepsilon_{z,i}, \quad (2.4)$$

so $\text{var}(e_t) \rightarrow \infty$ and $\text{Corr}(e_t, e_{t+1}) \rightarrow 1$. Thus all deviations are permanent, violating the conditions for classical inference; t - and F -tests are severely size-distorted and R^2 is often large even when $\beta = 0$. Empirically, Granger–Newbold rejection rates at the 5% level are around 75% and regressions exhibit large R^2 and strong residual autocorrelation; Phillips (1986) [2] shows the size distortion worsens with T .

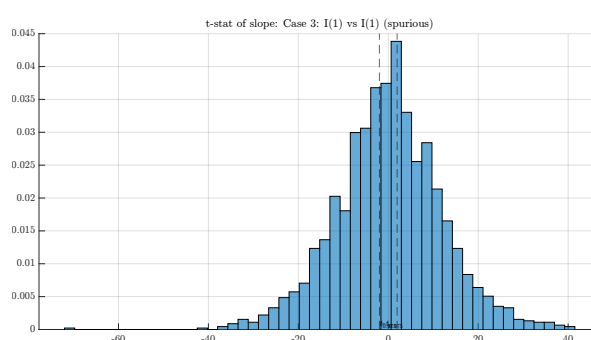
We obtain $\text{rej}_{5\%} \approx 0.85$, a very dispersed t -statistic 2-3b centered near zero but with extremely heavy tails, and an R^2 distribution 2-3a pushed far to the right (mean ≈ 0.24 , with mass well above 0.5). The histograms make the point vividly.

```

1 for r = 1:R2
2     y = simulate_ar1(T2, 1, 1.0, 0, 0);
3     z = simulate_ar1(T2, 1, 1.0, 0, 0);
4     stats = run_ols(y, z, true);
5     b1(r) = stats.b(2); t(r) = stats.tstat(2); r2(r) = stats.R2; dw(r) =
        stats.DW;
6 end
7 results{3} = summarize_case(b1, t, r2, dw, tcrit);
8 t_all{3} = t; R2_all{3} = r2;
```



(a) Distribution of the R^2 values for Case 3



(b) Distribution of the t -statistic values for Case 3

Figure 2-3. Visual diagnostics for Case 3

□

Case 4: The nonstationary $\{y_t\}$ and $\{z_t\}$ sequences are integrated of the same order and the residual sequence is stationary.

Solution.

Now suppose y_t and z_t share a common stochastic trend, e.g.

$$y_t = \tau_t + \varepsilon_{y,t}, \quad z_t = \tau_t + \varepsilon_{z,t}, \quad \tau_t = \tau_{t-1} + v_t, \quad (2.5)$$

so that $y_t - z_t$ is stationary and the residual e_t from the levels regression is $I(0)$. In this case the regression is meaningful: OLS yields a consistent estimate of the cointegrating vector, and in fact it is *super-consistent* (converges faster than $T^{1/2}$). However, the usual t -ratios on the cointegrating coefficients are not standard; formal inference on cointegration is conducted by testing the residuals for a unit root (Engle-Granger [1] Step 2) or by system approaches (Johansen, 1995).

As clearly visible in Figure 2-4a, the R^2 distribution concentrates near one and the slope estimate centers tightly around the true value (≈ 0.98). Because of super-consistency, the t -statistics that we observe in 2-4b on the slope explode in magnitude with T (our histogram ranges well into the hundreds), illustrating why standard critical values are inappropriate for cointegrating coefficients.

```

1 sig2_tau = 1.0; sig2_y = 0.5; sig2_z = 0.5;
2 for r = 1:R2
3     tau = cumsum(sqrt(sig2_tau) * randn(T2,1));
4     y = tau + sqrt(sig2_y) * randn(T2,1);
5     z = tau + sqrt(sig2_z) * randn(T2,1);
6     stats = run_ols(y, z, true);
7     b1(r) = stats.b(2); t(r) = stats.tstat(2); r2(r) = stats.R2; dw(r) =
      stats.DW;
8 end
9 results{4} = summarize_case(b1, t, r2, dw, tcrit);
10 t_all{4} = t; R2_all{4} = r2;
```

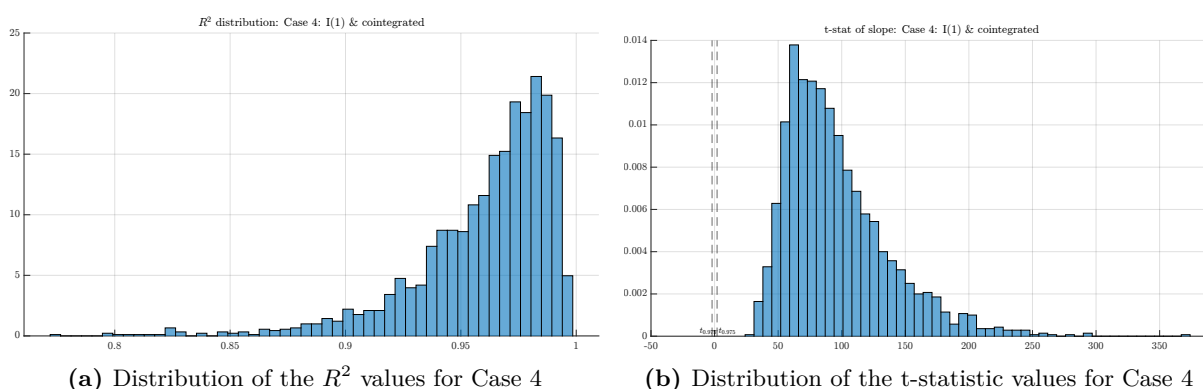


Figure 2-4. Visual diagnostics for Case 4

□

2.2 General considerations

- **Case 1 (stationary):** $\hat{\beta}$, t -tests, R^2 are meaningful under the usual assumptions; our simulation shows near-normal t and tiny R^2 .
- **Case 2 ($I(1)$ on $I(0)$):** residuals inherit a stochastic trend; classical tests mis-size (inflated rejection).
- **Case 3 (spurious):** nonstationary residuals ($\text{var}(e_t) \rightarrow \infty$, $\rho_{e, e_{t+1}} \rightarrow 1$) \Rightarrow huge over-rejection and often large R^2 .
- **Case 4 (cointegration):** the levels regression *is* meaningful because e_t is $I(0)$ and $\hat{\beta}$ is super-consistent; inference about cointegration relies on unit-root tests on \hat{e}_t (or Johansen), not on textbook t -ratios.

Table 2-1. Monte Carlo Results for $\hat{\beta}_1$ and Regression Diagnostics

| Case | mean($\hat{\beta}_1$) | sd($\hat{\beta}_1$) | Rej. H_0 (5%) | $\overline{R^2}$ | $\overline{\text{DW}}$ |
|-------------------------------------|-------------------------|-----------------------|-----------------|------------------|------------------------|
| Case 1: both stationary | 0.000 | 0.060 | 0.020 | 0.003 | 1.016 |
| Case 2: $I(1)$ vs $I(0)$ | 0.006 | 0.644 | 0.334 | 0.015 | 0.056 |
| Case 3: $I(1)$ vs $I(1)$ (spurious) | 0.001 | 0.638 | 0.849 | 0.243 | 0.070 |
| Case 4: $I(1)$ & cointegrated | 0.980 | 0.019 | 1.000 | 0.960 | 1.989 |

Notes: Columns report the Monte Carlo mean and standard deviation (sd) of $\hat{\beta}_1$, the empirical rejection frequency of $H_0: \beta_1 = 0$ at the 5% level, the mean R^2 , and the mean Durbin-Watson statistic (DW).

Question 3: Invertibility

Suppose that the DGP is

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & L^2 \\ \frac{\beta}{1-\beta} & \frac{\beta^2}{1-\beta} + \beta L \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix}, \quad (\text{Starter DGP})$$

where the shocks are uncorrelated, $u_t = [\eta_t \ \varepsilon_t]'$, and $\text{VCov}(u_t) = \begin{bmatrix} 1 & 0 \\ 0 & 0.8 \end{bmatrix}$; and $0 < \beta < 1$ (assume $\beta = 0.6$).

Generate 500 observations from $[x_t \ y_t]'$, estimate a VAR with 4 lags and identify the structural shocks using a Choleski identification scheme (after all, it is compatible with the DGP, right?). Do the simulations N times and store the impulse responses for the N simulations. Compare the true impulse response functions and the estimated ones (for the estimated ones, you can compute the mean of the empirical distribution of the estimated impulse responses and, say, 2.5th and the 97.5th percentile). Are they similar? What's wrong with the DGP? Discuss.

Solution.

Let $u_t = (\eta_t, \varepsilon_t)'$ with $\text{Var}(\eta_t) = 1$, $\text{Var}(\varepsilon_t) = 0.8$, and $\text{Cov}(\eta_t, \varepsilon_t) = 0$. The data-generating process is

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & L^2 \\ \frac{\beta}{1-\beta} & \frac{\beta^2}{1-\beta} + \beta L \end{bmatrix}}_{=: C(L)} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix}, \quad 0 < \beta < 1 \ (\beta = 0.6 \text{ in simulations}). \quad (3.1)$$

Thus $x_t = \eta_t + \varepsilon_{t-2}$ and $y_t = \frac{\beta}{1-\beta}\eta_t + \frac{\beta^2}{1-\beta}\varepsilon_t + \beta\varepsilon_{t-1}$.

```

1 % --- Settings ---
2 beta3      = 0.6;
3 SigU3      = diag([1, 0.8]);
4 T3         = 500;
5 B3         = 200;
6 p3         = 4;
7 H3         = 20;
8 Nmc3       = 500;
9
10 % --- True IRFs ---
11 trueIRF = compute_true_irfs_ex3(beta3, SigU3, H3);

```

3.1 True (structural) impulse responses

We normalize shocks to unit variance via $\tilde{\varepsilon}_{1t} = \eta_t$ and $\tilde{\varepsilon}_{2t} = \varepsilon_t/\sigma_\varepsilon$ with $\sigma_\varepsilon = \sqrt{0.8}$. The structural IRFs Θ_h (responses at horizon h to a unit shock at t) are

$$\Theta_h = \begin{bmatrix} \theta_h^{x,\eta} & \theta_h^{x,\varepsilon} \\ \theta_h^{y,\eta} & \theta_h^{y,\varepsilon} \end{bmatrix} = \begin{bmatrix} \mathbf{1}\{h=0\} & \mathbf{1}\{h=2\}\sigma_\varepsilon \\ \frac{\beta}{1-\beta}\mathbf{1}\{h=0\} & \left(\frac{\beta^2}{1-\beta}\mathbf{1}\{h=0\} + \beta\mathbf{1}\{h=1\}\right)\sigma_\varepsilon \end{bmatrix}. \quad (3.2)$$

Hence: (i) x reacts to η only on impact; (ii) x reacts to ε only at $h = 2$; (iii) y reacts to ε at $h = 0$ and $h = 1$; (iv) y reacts to η only on impact.

```

1 function trueIRF = compute_true_irfs_ex3(beta, SigU, H)
2 % Compute the true analytical IRFs for Exercise 3
3     trueIRF = zeros(2,2,H+1);
4     s_eps = sqrt(SigU(2,2));
5     % Shock 1: eta_t (unit variance)
6     trueIRF(1,1,1) = 1;
7     trueIRF(2,1,1) = beta/(1-beta);
8     % Shock 2: eps_t/s_eps (unit variance)
9     trueIRF(1,2,3) = s_eps; % Response at h=2
10    trueIRF(2,2,1) = (beta^2/(1-beta)) * s_eps; % Response at h=0
11    trueIRF(2,2,2) = beta * s_eps; % Response at h=1
12 end
13
14 function [XY, Y_only] = simulate_dgp_invertibility(T, beta, SigU)
15 % Simulate from the DGP in Exercise 3
16     s_eta = sqrt(SigU(1,1)); s_eps = sqrt(SigU(2,2));
17     eta = s_eta * randn(T,1); eps = s_eps * randn(T,1);
18     x = eta + [0; 0; eps(1:end-2)];
19     y = beta/(1-beta) * eta + (beta^2/(1-beta)) * eps + beta * [0; eps(1:end-1)];
20     XY = [x y];
21     if nargin>1, Y_only = y; end
22 end
23
24 function [Acomp, P] = estimate_var_chol(Y, p)
25 % Estimate VAR(p) with intercept and return companion matrix and Cholesky
26     [T, n] = size(Y);
27     Xlag = mlag(Y, p);
28     X = [ones(T,1), Xlag];
29     Ytrim = Y(p+1:end, :); Xtrim = X(p+1:end, :);
30
31     B = Xtrim \ Ytrim;
32     U = Ytrim - Xtrim * B;
33     SigmaU = (U' * U) / (size(U,1) - size(B,1));
34
35     A = B(2:end,:).'; % A = [A1, A2, ..., Ap]
36     Acomp = [A; eye(n*(p-1)), zeros(n*(p-1), n)];
37     P = chol(SigmaU, 'lower');
38 end

```

3.2 Invertibility issues with Cholesky identification

A VAR with recursive (Cholesky) identification recovers fundamental shocks, i.e. those obtainable from one-sided filters of the data. Fundamentality requires that the VMA polynomial $C(L)$ be invertible with a one-sided inverse: equivalently, $\det C(z) \neq 0$ for all $|z| \leq 1$.

Here

$$\det C(z) = \frac{\beta^2}{1-\beta} + \beta z - \frac{\beta}{1-\beta} z^2 = -\frac{\beta}{1-\beta} (z-1)(z+\beta). \quad (3.3)$$

The zeros are $z = 1$ and $z = -\beta$. With $0 < \beta < 1$, one zero lies inside the unit circle ($|\beta| = \beta < 1$) and one on it ($z = 1$). Cholesky identification recovers the one-sided Wold innovations only under invertibility, i.e. when the structural VMA satisfies $\det C(z) \neq 0$ for all $|z| \leq 1$; in our DGP, $\det C(z) = -\frac{\beta}{1-\beta} (z-1)(z+\beta)$ has a zero at $z = -\beta$ with $|z| < 1$, so the representation is non-invertible (non-fundamental) and a recursive VAR cannot retrieve the true economic shocks.

A Cholesky VAR is constrained to use one-sided filters, so its “structural shocks” are linear combinations of (η, ε) that whiten the reduced-form innovations but do not coincide with the economic shocks.

3.3 Predictions for estimated IRFs

Because of non-invertibility, the VAR(4)+Cholesky IRFs converge (as $T \rightarrow \infty$) to pseudo-true objects that deliver orthogonal one-step-ahead innovations, not to Θ_h as in 3.2 above.

Concretely:

- Responses to η are largely recovered correctly (they are contemporaneous spikes).
- Responses to ε are distorted: the VAR tends to create spurious mass at $h = 0$ and $h = 1$ for x (where the true response is 0), and to smear the clean $h = 2$ spike into an earlier, hump-shaped reaction. Responses of y to ε —which are truly concentrated at $h = 0, 1$ —look closer but still mix in noise from the mismeasured second shock.

Comparison to Monte Carlo output. In finite samples the empirical mean IRFs will already display these features: (i) close agreement for $x \leftarrow \eta$ and $y \leftarrow \eta$ on impact; (ii) clear disagreement for $x \leftarrow \varepsilon$ (spurious activity at $h = 0, 1$ instead of a pure $h = 2$ spike); and (iii) mild distortions for $y \leftarrow \varepsilon$.

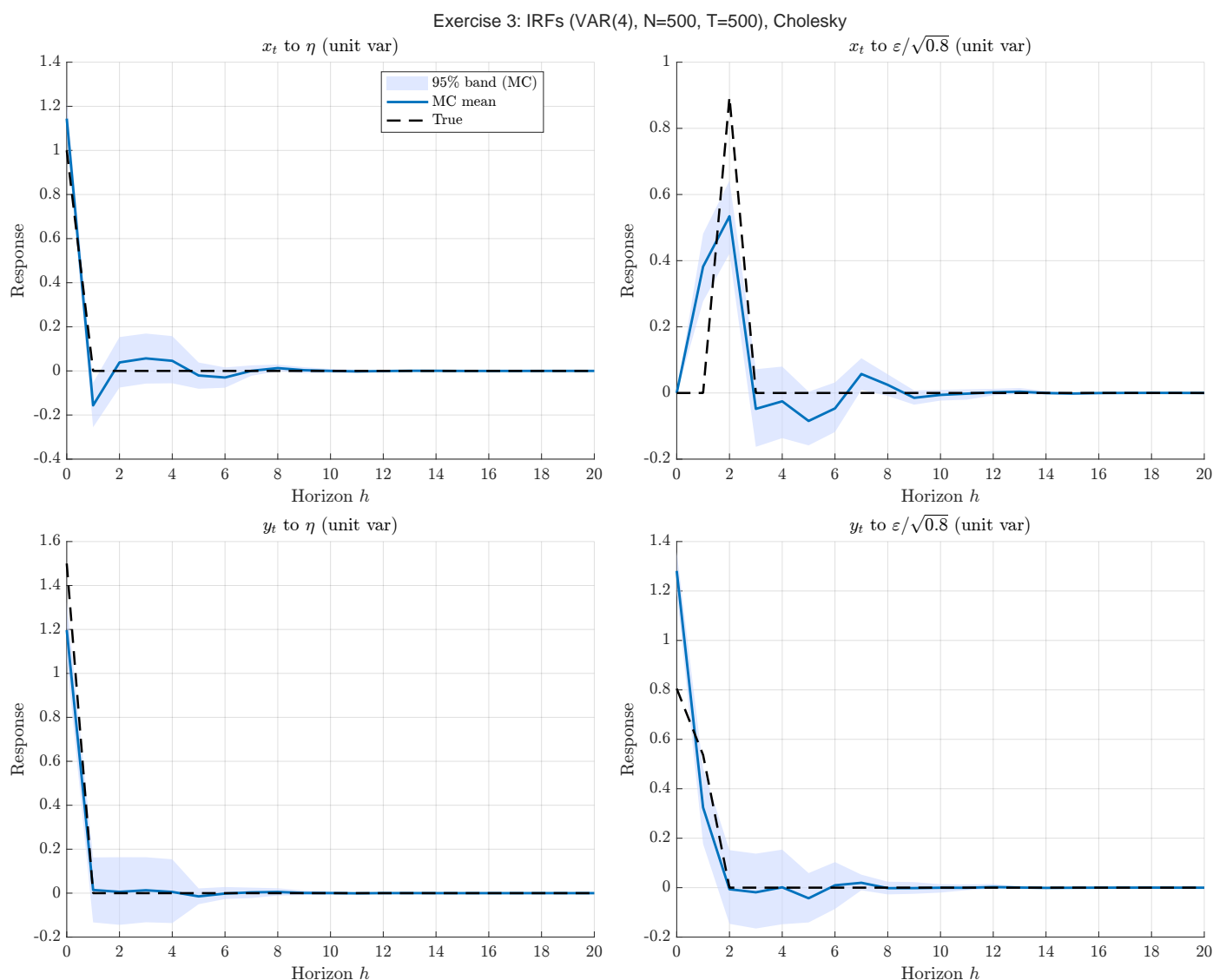


Figure 3-1. Estimated IRFs (in blue) against the true responses (dashed), including the 95% CI in light blue

```
1 IRF_draws = zeros(2,2,H3+1,Nmc3);
2 for r = 1:Nmc3
3     [X, ~] = simulate_dgp_invertibility(T3+B3, beta3, SigU3);
4     X = X(B3+1:end,:);
5     [Acomp, Pchol] = estimate_var_chol(X, p3);
6     IRF_draws(:,:,r) = var_irf_from_companion(Acomp, Pchol, H3);
7 end
8
9 IRF_mean = mean(IRF_draws, 4);
10 IRF_lo = prctile(IRF_draws, 2.5, 4);
11 IRF_hi = prctile(IRF_draws, 97.5, 4);
12 plot_irfs_ex3(trueIRF, IRF_mean, IRF_lo, IRF_hi, p3, Nmc3, T3, exportFig);
```



Question 4: Granger causality

Read the paper “A New Measure of Monetary Shocks: Derivation and Implications”, by Christina D. Romer and David H. Romer [3], *The American Economic Review*, Vol. 94, No. 4 (Sep., 2004), pp. 1055–1084 available on the blackboard.

Download the file `Romer_Romer.xlsx`. There, you will find 4 time series: US inflation, US unemployment, US federal funds rate, the Romer and Romer monetary policy shocks from 1969Q1 to 1996Q4. Run a VAR with 4 lags and test for Granger causality of the Romer and Romer shocks. Are the Romer and Romer shocks Granger-causing other variables or not? Discuss. Are the other variables Granger-causing the Romer and Romer dates? Discuss.

Solution.

Let $z_t = (\pi_t, u_t, i_t, \varepsilon_t^{RR})'$ collect respectively quarterly US inflation, unemployment, the federal funds rate, and the Romer–Romer monetary policy shock. We estimate a VAR(p) with an intercept and $p = 4$ (one year of lags), and test Granger causality via Wald restrictions that the p lag coefficients of the candidate “cause” on the target equation are jointly zero. Under mild regularity, the Wald statistic is asymptotically χ_r^2 (r restrictions); finite-sample inference can be based on an F statistic with $(r, T - K)$ degrees of freedom.

```

1 p4 = 4;
2 romer_path = fullfile(pwd, 'ps2', 'data', 'Romer_Romer2004.csv');
3 alpha4 = 0.05;
4
5 Ttbl = readtable(romer_path, 'PreserveVariableNames', true);
6 Ttbl.Properties.VariableNames = lower(Ttbl.Properties.VariableNames);
7
8 varOrder = {'inflation', 'unemployment', 'ffr', 'rrshock'};
9 Yraw = double(Ttbl{:, varOrder});
10 Yraw = Yraw(all(~isnan(Yraw), 2), :);
11 [Tobs, n] = size(Yraw);

```

4.1 Why RR shocks are a good candidate “cause”

RR shocks are residual changes in the *intended* funds rate around FOMC meetings, purged of the Fed staff (Greenbook) forecasts of real activity and inflation. By design, the series aims to strip out both short-run endogeneity and anticipatory policy moves, so it should

- (i) move the actual policy rate and
- (ii) be orthogonal to predictable movements in activity and prices known at the time of the decision.

```

1 Xlags = mlag(Yraw, p4);
2 X = [ones(Tobs,1), Xlags];
3 Y = Yraw;
4
5 Xt = X(p4+1:end, :);
6 Yt = Y(p4+1:end, :);
7 [T_eff, K] = size(Xt);
8 df = T_eff - K;
9
10 B = Xt \ Yt;
11 U = Yt - Xt * B;
12 s2 = sum(U.^2, 1) ./ df;
13 invXX = (Xt' * Xt) \ eye(K);

```

```

1 labels = ["Inflation","Unemployment","FedFundsRate","rrshock"];
2 shockIdx = 4;
3
4 res = [];
5 % (A) Shock -> Others
6 for yEq = 1:3
7     res = [res; run_gc_test(B, s2, invXX, df, p4, K, shockIdx, yEq, labels)];
8 end
9 % (B) Others -> Shock (individually)
10 for xVar = 1:3
11     res = [res; run_gc_test(B, s2, invXX, df, p4, K, xVar, shockIdx, labels)];
12 end
13 % (C) Others -> Shock (jointly)
14 res = [res; run_gc_test(B, s2, invXX, df, p4, K, 1:3, shockIdx, labels)];

```

4.2 Results

Table 4-1 reports the joint Wald tests from the VAR(4), using both the large-sample χ^2 and finite-sample F versions.

Table 4-1. Granger-Style Causality Tests (Wald χ^2 and F)

| Cause | Effect | Lags | Restr. | χ^2 | $p(\chi^2)$ | F | $p(F)$ |
|--------------|----------------|-------|--------|----------|-----------------------|-------|-----------------------|
| rrshock | → Inflation | 4.000 | 4.000 | 8.518 | 7.43×10^{-2} | 2.129 | 8.35×10^{-2} |
| rrshock | → Unemployment | 4.000 | 4.000 | 8.751 | 6.76×10^{-2} | 2.188 | 7.65×10^{-2} |
| rrshock | → FedFundsRate | 4.000 | 4.000 | 36.244 | 2.58×10^{-7} | 9.061 | 3.31×10^{-6} |
| Inflation | → rrshock | 4.000 | 4.000 | 2.343 | 6.73×10^{-1} | 0.586 | 6.74×10^{-1} |
| Unemployment | → rrshock | 4.000 | 4.000 | 3.477 | 4.81×10^{-1} | 0.869 | 4.86×10^{-1} |
| FedFundsRate | → rrshock | 4.000 | 4.000 | 5.255 | 2.62×10^{-1} | 1.314 | 2.71×10^{-1} |
| All | → rrshock | 4.000 | 12.000 | 10.450 | 5.77×10^{-1} | 0.871 | 5.79×10^{-1} |

Notes: Each row reports a Wald χ^2 test and its p -value, and the corresponding F -test and p -value, for the null that the listed *Cause* does not (Granger-)cause the *Effect*. “Lags” is the number of included lags; “Restr.” is the number of linear restrictions.

4.2.1 Do RR shocks Granger-cause the other variables?

Yes for the policy rate, and only weak (10%) evidence for activity/prices over a one-year horizon. The null that past RR shocks do not enter the i_t equation is decisively rejected (Wald $\chi^2 = 36.24$, $p < 10^{-6}$). This is exactly what one expects from the construction: RR shocks are innovations to intended policy that should map into subsequent movements in the actual funds rate.

Romer and Romer (2004) show that their shocks produce faster, larger responses of output and prices than conventional funds-rate innovations in both single-equation regressions and VARs, and that they remove the usual “price puzzle”. We can observe this directly from their Figures 2, 4–6, and is the surrounding discussion.

For inflation and unemployment, the χ^2 and F tests are not significant at 5% but are close at 10% ($p \approx 0.07–0.08$). From a time-series perspective this is not entirely surprising: with quarterly data and $p = 4$, the Granger test only asks whether shocks help predict next year’s π_t or u_t . Theory and evidence both imply that the real activity response leads the price response and that price-level effects arrive with longer lags than one year; failing to reject at 5% in this short horizon therefore does not contradict the sizable medium-run effects documented in Romer and Romer impulse responses.

More specifically, the authors find output begins to fall two quarters and the price level declines noticeably only after roughly two years; the four-year price effect is about -4 to -6% , depending on the price index and specification.

4.2.2 Do the other variables Granger-cause the RR shocks?

No: neither inflation, nor unemployment, nor the funds rate individually or jointly Granger-cause ε_t^{RR} (all $p > 0.26$; joint $p = 0.58$). This is exactly what we want to see if the RR shocks have been purged of systematic responses to information about future macro conditions, as they should be forecast-orthogonal with respect to the macro state (hence not Granger-caused by it).

Romer and Romer explicitly purge intended policy moves using contemporaneous staff forecasts of growth, inflation, and unemployment at multiple horizons.

□

References

- [1] Robert F. Engle and C. W. J. Granger. “Co-Integration and Error Correction: Representation, Estimation, and Testing”. In: *Econometrica* 55.2 (1987), pp. 251–276. ISSN: 00129682, 14680262. URL: <http://www.jstor.org/stable/1913236> (visited on 11/03/2025).
- [2] P.C.B. Phillips. “Understanding spurious regressions in econometrics”. In: *Journal of Econometrics* 33.3 (1986), pp. 311–340. ISSN: 0304-4076. DOI: [https://doi.org/10.1016/0304-4076\(86\)90001-1](https://doi.org/10.1016/0304-4076(86)90001-1). URL: <https://www.sciencedirect.com/science/article/pii/0304407686900011>.
- [3] Christina D. Romer and David H. Romer. “A New Measure of Monetary Shocks: Derivation and Implications”. In: *American Economic Review* 94.4 (Sept. 2004), pp. 1055–1084. DOI: [10.1257/0002828042002651](https://doi.org/10.1257/0002828042002651). URL: <https://www.aeaweb.org/articles?id=10.1257/0002828042002651>.