You need to submit your code, data and output for all exercises, and you need to write your results as if it is a paper, that is, STATA or MATLAB outputs are not sufficient: you need to tabulate and report your results formally. The code should be documented and should run so that it is exactly repeatable. Interpret your results in every case. You also need to define any extra notation you introduce. **Results with no discussion nor interpretation and any notation you introduce that is not defined will be penalized.**

This question is based on three recent papers: Inoue, Rossi and Wang (2022), Andrews, Stock and Sun (2019), and Lee, McCrary, Moreira and Porter (2022), all posted on Brightspace.

A. (30%) Consider the IV regression $t = 1, .., T$

$$y_t = \beta Y_t + \gamma X_t + u_t \quad (1)$$
$$Y_t = \delta_X X_t + \delta_Z Z_t + v_t \quad (2)$$

where Y_t refers to the single right-hand-side endogenous regressor and Z_t refers to the single instrument. The model is thus for simplicity exactly identified. X_t consists of

included exogenous variables. Assume that u_t and v_t are contemporaneously correlated but for simplicity, the data is otherwise *i.i.d.* normal, that is $(u_t, v_t) \overset{iid}{\sim} N(0, \Sigma)$ where Σ is 2×2. To answer part A, it may be very helpful to write the above system in matrix form. It is not compulsory, but you need to provide explicit formulas, so the matrix notation may help. Do not forget that you need to define any extra notation you introduce; there will be penalties for any notation that is not defined. For example, you cannot take it for granted that Y will refer to the vector of Y_ts. If this is what you want as a notation, you need clear definitions.

1. Write the reduced form of this model, which will consist in a system of two equations describing the regression of all endogenous variables on all exogenous variables in the system. Do not forget to provide the distribution of the errors of this reduced form. Next verify the following claim on page 8 of the paper by Andrews, Stock and Sun (2019): "the IV model implies that ... the IV coefficient is simply the constant of proportionality between the reduced-form coefficient and the fist-stage parameter". Recall of course that this paper and the model you are given do not share the name notation (but the parallel must be clear to you).

2. Provide the formulas for the OLS estimation of the reduced form parameters, and their variance covariance matrix. To do this, you can refer to the lecture notes on seemingly unrelated regression posted on brightspace.

3. Provide the formulas to derive a confidence set for β in (1) using the 2SLS method.

4. Using the formulas derived in question 2 and your lecture notes, show (provide formulas with derivations) how you can apply the Fieller method to derive a confidence set for β in (1).

5. Using your lecture notes and Dufour and Jasiak (2001), derive the formulas to obtain the Anderson-Rubin confidence sets for β in (1).

6. Do you expect the methods above to give the same results? Discuss. Which of the above methods are robust to identification; explain what this means.

7. **BONUS (up to 5%).** Lee, McCrary, Moreira and Porter (2022) suggest that a modified t-test can be applied for inference on β in (1). Explain what this would involve in your case, and how you can derive a confidence interval for β using this t-statistic.

B. (25%) Using the New Keynesian Phillips curve equations as in Inoue, Rossi and Wang (2022)

$$\pi_t = \lambda s_t + \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1} + u_t \tag{3}$$

apply the methods you developed in part A to answer their research question: has λ decreased towards zero over time? To do this, derive all the above confidence sets for λ over 3 sub-samples: before 1984, 1984-2007, 2008 till the longest data span you can find. A data set is provided for you [from the paper Antoine, Khalaf, Kichian and Lin (2022)], but it is recommended that you try to expand it. Furthermore, for this question, you need to use for the survey-based expected inflation variable "the three-quarter-ahead forecast of mean PGDP inflation from the Survey of Professional Forecasters" as in Inoue, Rossi and Wang (2022). This is because you are asked (in this sub-question) to make the following simplifying assumption: (a) the only endogenous variable in (3) is s_t, and (b) u_t is *i.i.d.** *So to be clear, in this sub-question, use an observed expectation series and assume i.i.d. errors.* With the confidence sets for λ that you obtain over the sub-samples, what do you learn about its evolution using the different methods? You can write your own code or use the Stata "weakiv" code [Finlay and Magnusson (2009)]. Discuss your results, and provide a summative perspective. Be very clear about your choice of instruments. No restrictions are proposed so chose whatever you want as long as you document and justify your choice.

C. (25%) Now relax the simplifying assumptions as follows: (a) allow inflation expectations as well as s_t to be endogenous, (b) relax the *i.i.d.* error assumption. Obtain the 2SLS and Anderson-Rubin confidence sets again, with HAC standard errors. For the Anderson-Rubin case, you will obtain a joint region for λ and γ_f . You can code your own grid search and report graphically, or use the Stata "weakiv" code [Finlay and Magnusson (2009)]. Graphical response regions are a good reporting option. With these confidence sets for λ over the sub-samples, what do you learn about its evolution? Are results very different from sub-question B? Discuss. Be very clear about your choice of instruments.

References

[1] Andrews I, Stock J. H. and L. Sun (2019). *Weak Instruments in IV Regression :*
Theory and Practice. Annual Review of Economics 11, 727-753.

[2] Antoine B., Khalaf L., Kichian M. and Z. Lin (2022). *Simulation-based matching*

inference with applications to DSGE models. The Journal of Business and Economic Statistics. https://doi.org/10.1080/07350015.2021.2019046

[3] Doko Tchatoka, F. and Dufour, J.-M. (2014). Identification-robust inference for endogeneity parameters in linear structural models. The Econometrics Journal, 17(1):165–187.

[4] Dufour J.-M., Flachaire E. and L. Khalaf (2019). *Reliable inference for inequality measures with heavy-tailed distribution.* The Journal of Business and Economic Statistics 37, 457-470.

[5] Dufour J.-M. and J. Jasiak (2001). *Finite Sample Limited Information Inference Methods for Structural Equations and Models with Generated Regressors.* International Economic Review 42, 815-843.

[6] Finlay K. and L. M. Magnusson (2009). *Implementing weak-instrument robust tests for a general class of instrumental-variables models.* The Stata Journal 9, Number 3, pp. 398–421. With update in The Stata Journal (2011), 11, p. 325.

[7] Inoue A., Rossi B. and Y. Wang (2022). Has the Phillips Curve Flattened and Why? Working paper.

[8] Lee D. S., McCrary J., Moreira M. J. and J. R. Porter (2022). *Valid t-ratio for inference for IV.* NBER Working Paper 29124.

[9] Mikusheva, A., and B. P. Poi. (2006). *Tests and confidence sets with correct size when instruments are potentially weak*. Stata Journal 6: 335–347.