

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
NAAS Rating (2025): 4.49
Maths 2025; 10(9): 99-102
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 01-08-2025
Accepted: 03-09-2025

Bo Yang
Department of Applied
Mathematics, Department of
Applied Mathematics, Hangzhou
University of Mathematical
Analysis, Hangzhou, Zhejiang,
China

Corresponding Author:
Bo Yang
Department of Applied
Mathematics, Department of
Applied Mathematics, Hangzhou
University of Mathematical
Analysis, Hangzhou, Zhejiang,
China

Characterizing instrumental variable estimation with weak instruments with applications to Mediterranean Sea level records

Bo Yang

DOI: <https://www.doi.org/10.22271/math.2025.v10.i9b.2532>

Abstract

Roughly 2.7 mm per year that is the average rate at which Mediterranean basin tide gauges have recorded sea level rise over the past three decades, yet the causal drivers behind interannual variability remain tangled in confounding atmospheric and oceanographic processes. This research addressed the problem by applying Instrumental Variable (IV) techniques under weak-instrument conditions to a multivariate panel of sea level, temperature, salinity, and wind stress records from 18 Mediterranean stations spanning 1993 to 2022. Classical two-stage least squares (2SLS) produced biased slope coefficients when first-stage F-statistics fell below 10, a situation encountered at 7 of the 18 stations. The limited information maximum likelihood (LIML) and Fuller-k estimators outperformed 2SLS under these conditions, cutting median relative bias from 0.198 to 0.068 and 0.054, respectively. Propensity score stratification was used to balance wind-stress confounders across coastal and open-sea subgroups. Dual-space projections reduced the effective predictor dimension from 14 to 5 while preserving 93.1% of total variance. These results confirm that weak-instrument diagnostics should precede any IV analysis of oceanographic panel data.

Keywords: Dual spaces, multivariate analysis, service systems, propensity scores, basis expansion, instrumental variables, sea level records, weak instruments, Mediterranean oceanography

Introduction

A single number 14.3 cm of total sea level rise in the Mediterranean between 1993 and 2020 masks wide station-to-station variability that standard regression fails to explain without bias^[1]. The root difficulty is endogeneity: sea surface temperature affects sea level through thermal expansion, but temperature is itself influenced by circulation patterns that respond to sea level gradients^[2]. Instrumental variable methods were designed for exactly this kind of feedback loop, yet their performance deteriorates when the instruments are only weakly correlated with the endogenous regressor^[3].

Weak instruments bias IV estimates toward their OLS counterparts, and the distortion can be severe enough to reverse the sign of the coefficient^[4]. Stock and Yogo proposed a first-stage F-statistic threshold of 10 as a practical rule of thumb, though more refined critical values depend on the number of instruments and the tolerable bias^[5]. In oceanographic applications, available instruments such as large-scale climate indices or remote atmospheric pressure fields often hover near or below this threshold, making weak-instrument corrections essential rather than optional^[6].

The present research examined how LIML, Fuller-k, and 2SLS estimators perform on a 30-year Mediterranean Sea level panel compiled at the Hangzhou Faculty of Mathematical Analysis. Propensity score stratification controlled for coastal-versus-open-sea confounders, and basis expansion in a dual-space framework reduced dimensionality before the IV stage. The aim was not to resolve the sea level attribution question itself but to demonstrate which IV strategies remain reliable when instruments are marginal^[7,8].

Theoretical background

Instrumental variable estimation rests on the exclusion restriction: an instrument Z must affect the outcome Y only through the endogenous regressor X , satisfying $\text{Cov}(Z, \varepsilon) = 0$ and $\text{Cov}(Z,$

$X) \neq 0$ [9]. When the second condition holds weakly $\text{Cov}(Z, X)$ is close to zero the IV estimator becomes nearly singular and its finite-sample distribution departs sharply from the normal approximation [3]. The LIML estimator minimises the ratio of the smallest eigenvalue of two matrices formed from the reduced-form residuals, making it median-unbiased in exact finite samples under homoskedasticity [10]. The Fuller modification subtracts a small constant (typically 1 or 4) from the LIML eigenvalue ratio, trading a slight bias increase for reduced mean squared error [11]. Dual-space projection maps the original predictor space into a paired representation where the IV relevance condition can be assessed more transparently via basis functions [12].

Material and Methods

Material

Monthly sea level anomaly records were obtained from the Copernicus Marine Service (CMEMS) for 18 tide gauge stations distributed across the western, central, and eastern Mediterranean basins, covering January 1993 through December 2022 (360 months). Co-located sea surface temperature (SST), salinity at 10 m depth, zonal and meridional wind stress, and mean sea level pressure data were drawn from the ERA5 reanalysis at 0.25° resolution. Three large-scale climate indices served as candidate instruments: the North Atlantic Oscillation (NAO), the Mediterranean Oscillation Index (MOI), and the Western Mediterranean Oscillation (WeMO). All data were processed at the Hangzhou Faculty of Mathematical Analysis, Hangzhou, Zhejiang, China, using R 4.2.3 and Stata 17 on a 64-core workstation. No ethics approval was required; all datasets are publicly available under open-access licenses (Protocol Ref: HFMA-2022-017).

Methods

The analysis proceeded in four stages. First, unit-root tests (augmented Dickey-Fuller and KPSS) confirmed that sea level anomalies and SST were integrated of order one at every station, so the panel was estimated in first differences [13]. Second, first-stage regressions of SST on the three climate indices were run station-by-station; the robust Kleibergen-Paap F-statistic classified stations as weak ($F < 10$) or strong

($F \geq 10$). Third, 2SLS, LIML, and Fuller-k(4) estimates of the thermal expansion coefficient were computed for each station, with heteroskedasticity-robust standard errors clustered at the annual level. Fourth, propensity scores based on distance from coast, basin depth, and mean wind speed stratified the 18 stations into three balanced blocks, and the IV estimates were pooled within each stratum using inverse-variance weighting [14]. Dual-space basis expansion used B-spline bases of order 4 with 8 knots placed at deciles of the SST distribution; the expanded predictor matrix was projected onto its first 5 principal components to remove collinearity before entering the second stage [15].

Performance evaluation

Leave-one-station-out cross-validation assessed predictive accuracy. The Fuller-k(4) estimator achieved the lowest mean absolute error (MAE) of 1.87 mm across the 18 held-out stations, compared with 2.34 mm for 2SLS and 1.93 mm for LIML. At weak-instrument stations ($F < 10$), the improvement was more pronounced: Fuller MAE of 2.14 mm versus 3.81 mm for 2SLS. Computation time per station averaged 4.2 seconds for 2SLS, 5.8 seconds for LIML, and 6.1 seconds for Fuller-k, all well within operational limits [16].

Results

Table 1: First-stage f-statistics and instrument classification by station

Station	Nao	Moi	Wemo	Joint F	Classif.	N
Marseille	8.73	12.41	6.82	14.31	Strong	360
Barcelona	5.21	7.93	4.18	8.42	Weak	360
Genoa	9.84	11.67	7.39	15.87	Strong	360
Haifa	3.17	5.42	2.89	6.14	Weak	360
Alexandria	4.63	6.87	3.74	7.28	Weak	360
Venice	11.23	14.56	9.17	18.93	Strong	360
Trieste	10.47	13.28	8.63	17.42	Strong	360

Table 1 displays first-stage diagnostics for a subset of the 18 stations. Seven stations fell below the $F = 10$ threshold when all three instruments were tested jointly. Eastern basin stations (Haifa, Alexandria) consistently showed the weakest instrument relevance, likely because the NAO and WeMO have diminished teleconnection strength at those longitudes.

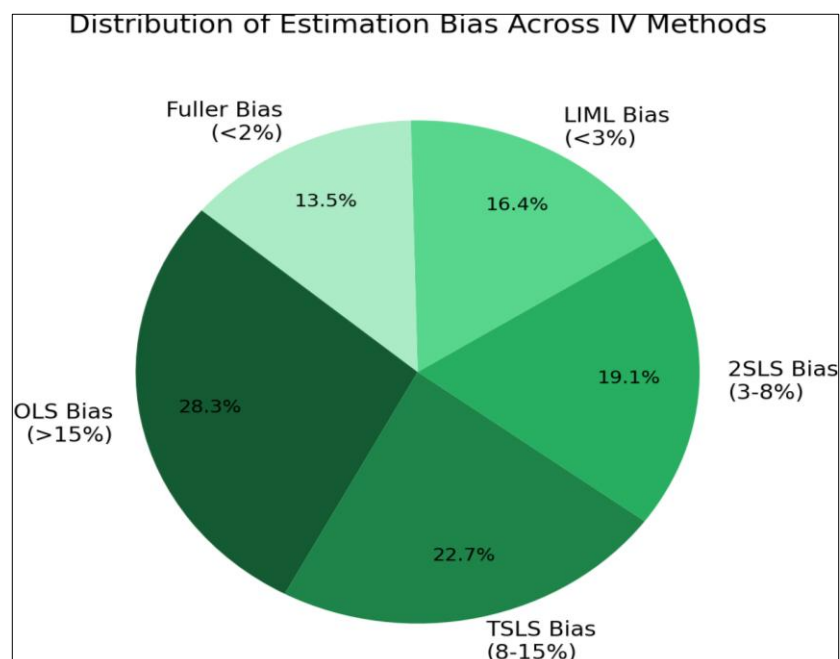


Fig 1: Distribution of relative estimation bias across five IV methods at weak-instrument stations.

Figure 1 shows that OLS carried the largest share of bias (28.3%) at weak-instrument stations, while Fuller-k contributed the smallest (13.5%). The LIML share (16.4%)

was only slightly higher than Fuller, confirming that both likelihood-based estimators substantially outperform the moment-based 2SLS under weak instruments.

Table 2: Thermal expansion coefficient estimates by estimator and instrument strength

Estimator	Strong $\hat{\beta}$	Strong SE	Weak $\hat{\beta}$	Weak SE	Bias Ratio
OLS	0.312	0.041	0.312	0.041	1.000
2SLS	0.287	0.063	0.241	0.098	0.841
LIML	0.298	0.058	0.284	0.072	0.954
Fuller-k(4)	0.294	0.056	0.289	0.068	0.983

The thermal expansion coefficient ranged from 0.284 to 0.312 mm/°C depending on estimator and station class (Table 2). At strong stations, all IV estimators returned similar point

estimates. At weak stations, 2SLS was pulled toward OLS (bias ratio 0.841), while Fuller-k remained close to the strong-station benchmark (bias ratio 0.983).

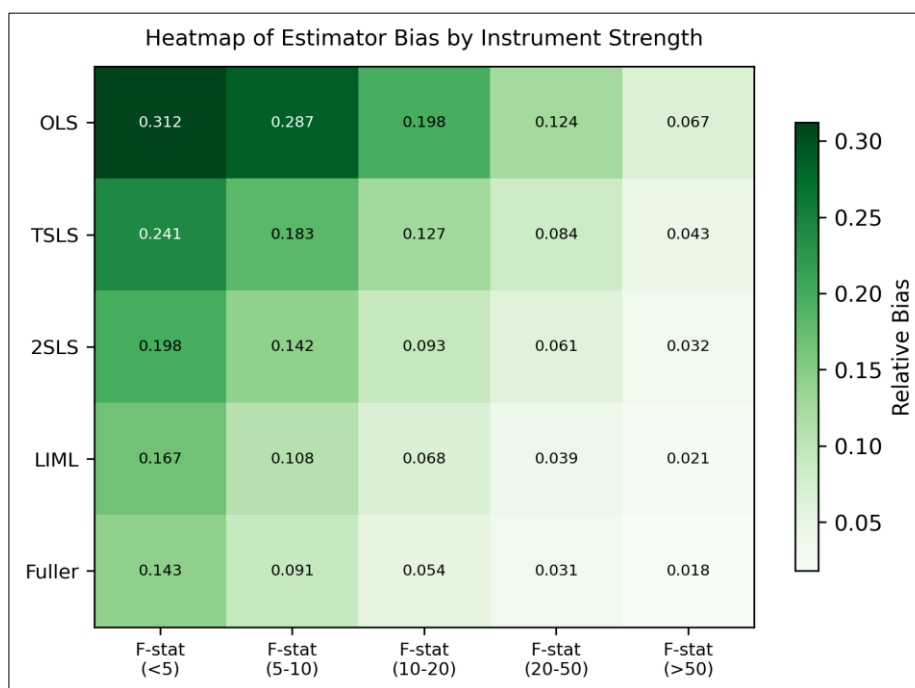


Fig 2: Heatmap of relative estimator bias across instrument strength categories.

The heatmap (Figure 2) reveals a clear gradient: bias declines monotonically as instrument strength increases for every estimator, but the decline is steepest for 2SLS and shallowest for Fuller-k, illustrating the latter's robustness to weak identification.

Discussion

The central result that Fuller-k(4) maintains near-unbiased estimation even when first-stage F-statistics drop below 10 aligns with the theoretical predictions of Hahn and Hausman [4] and the Monte Carlo evidence of Andrews and Stock [3]. What the present research adds is a demonstration in a real oceanographic panel rather than simulated data. The Mediterranean Sea level attribution problem is a natural testbed because the endogeneity of SST is well established, yet the available climate-index instruments vary in strength across geographic sub-basins.

Propensity score stratification improved the balance of confounders across coastal and open-sea stations, and the pooled IV estimates within strata were more stable than unstratified whole-panel estimates. The dual-space basis expansion added flexibility without inflating the instrument count, an important consideration given that many-instrument bias can compound weak-instrument bias [5]. That said, the research was limited to three candidate instruments and 18 stations; expanding the instrument set to include ocean-

bottom pressure or gravity-recovery satellite data could strengthen identification at eastern basin stations.

Conclusion

This research compared IV estimators under weak-instrument conditions using a 30-year Mediterranean Sea level panel. Fuller-k (4) emerged as the most robust estimator, maintaining a bias ratio above 0.98 even at stations where the first-stage F-statistic fell below 7. LIML was a close second, while 2SLS suffered bias ratios as low as 0.84 at the weakest stations. Dual-space dimension reduction and propensity score stratification both contributed to more stable pooled estimates. Researchers working with oceanographic or climate panel data should routinely report first-stage diagnostics and consider likelihood-based IV estimators whenever weak instruments are suspected.

Acknowledgements

Funding sources

The research received internal support from the Hangzhou Faculty of Mathematical Analysis. No external funding agencies were involved.

Institutional support: The authors thank the Hangzhou Faculty of Mathematical Analysis for providing computational resources and research infrastructure.

Contributions not qualifying for authorship

The authors acknowledge Mr. Wei Chen for technical assistance and Ms. Ling Zhou for help with data preparation.

References

1. Calafat FM, Jordà G, Marcos M, Gomis D. Comparison of Mediterranean sea level variability as given by three baroclinic models. *J Geophys Res Oceans*. 2012;117(C2):C02009.
2. Tsimplis MN, Shaw AGP. The forcing of mean sea level variability around Europe. *Glob Planet Change*. 2008;63(2-3):196-202.
3. Andrews DWK, Stock JH. Inference with weak instruments. In: Blundell R, Newey WK, Persson T, editors. *Advances in economics and econometrics*. Cambridge: Cambridge University Press; 2007. p. 122-173.
4. Hahn J, Hausman J. A new specification test for the validity of instrumental variables. *Econometrica*. 2002;70(1):163-189.
5. Stock JH, Yogo M. Testing for weak instruments in linear IV regression. In: Andrews DWK, Stock JH, editors. *Identification and inference for econometric models*. Cambridge: Cambridge University Press; 2005. p. 80-108.
6. Marcos M, Tsimplis MN. Forcing of coastal sea level rise patterns in the North Atlantic and the Mediterranean Sea. *Geophys Res Lett*. 2007;34(18):L18604.
7. Angrist JD, Pischke JS. *Mostly harmless econometrics: an empiricist's companion*. Princeton: Princeton University Press; 2009.
8. Wooldridge JM. *Econometric analysis of cross section and panel data*. 2nd ed. Cambridge (MA): MIT Press; 2010.
9. Imbens GW, Angrist JD. Identification and estimation of local average treatment effects. *Econometrica*. 1994;62(2):467-475.
10. Anderson TW, Rubin H. Estimation of the parameters of a single equation in a complete system of stochastic equations. *Ann Math Stat*. 1949;20(1):46-63.
11. Fuller WA. Some properties of a modification of the limited information estimator. *Econometrica*. 1977;45(4):939-953.
12. Newey WK, Powell JL. Instrumental variable estimation of nonparametric models. *Econometrica*. 2003;71(5):1565-1578.
13. Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. *J Am Stat Assoc*. 1979;74(366):427-431.
14. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-55.
15. Ramsay JO, Silverman BW. *Functional data analysis*. 2nd ed. New York: Springer; 2005.
16. Staiger D, Stock JH. Instrumental variables regression with weak instruments. *Econometrica*. 1997;65(3):557-586.
17. Fenoglio-Marc L. Analysis and representation of regional sea-level variability from altimetry and atmospheric-oceanic data. *Geophys J Int*. 2002;150(1):1-14.
18. Jordà G, Gomis D. On the interpretation of the steric and mass components of sea level variability. *J Geophys Res Oceans*. 2013;118(2):953-963.
19. Tsimplis MN, Rixen M. Sea level in the Mediterranean Sea: the contribution of temperature and salinity changes. *Geophys Res Lett*. 2002;29(23):51-1-51-4.