MR5005304 62G10 62G09 62G20 62J05 62R10

Dorn, Manuela (D-BAYR-NDM); **Birke, Melanie** (D-BAYR-NDM);
Jentsch, Carsten (D-DTU-NDM)

Testing exogeneity in the functional linear regression model. (English. English summary)

Statist. Sinica **36** (2026), *no.* 1, 1–30. *ISBN*

Consider a functional linear regression model

$$Y=\int_{[0,1]}X(t)\beta(t)dt+U=\langle\beta,X\rangle+U,$$

where Y is a real-valued random variable, X is a functional regressor, β is an unknown slope function, and U is a real-valued error term. It is often of interest and importance to test whether X is exogenous in this model. In this article, a test for the null $H_0:\mathbb{E}\{X(t)U\}=0$ for all $t\in[0,1]$ (i.e., X is *exogenous*) against the alternative $H_1:\mathbb{E}\{X(t)U\}\neq 0$ for at least one $t\in[0,1]$ (i.e., X is *endogenous*) is proposed.

The proposed test statistic can be constructed by two estimators for β. The estimator under exogeneity, denoted by $\widehat{\beta}$, is consistent only if X is exogenous and inconsistent otherwise. On the other hand, the estimator under endogeneity, denoted by $\widehat{\beta}_{IV}$, is built on a functional instrumental variable $W(t)$ satisfying $\mathbb{E}\{W(t)U\}=0$ for all $t\in[0,1]$. It is consistent under both exogeneity and endogeneity but usually less efficient. The setup so far is reminiscent of the Hausman test, which is popularly applied as a test for exogeneity in linear regression models. It turns out, however, that the Hausman test cannot be directly transferred to the functional linear regression model because the L_2-distance between $\widehat{\beta}$ and $\widehat{\beta}_{IV}$ has no proper limiting distribution.

The proposed test statistic is grounded on the sum of squared differences of projections of $\widehat{\beta}$ and $\widehat{\beta}_{IV}$. It is demonstrated that the test statistic is asymptotically normal under the null and consistent under the alternative. However, the asymptotic test is not recommended; to do so, many nuisance parameters must be estimated and substituting their consistent estimates into the test statistic is likely to deteriorate its power properties. Instead, a residual-based bootstrap inference procedure is investigated, and the first-order validity of the bootstrap-based test is established. The bootstrap procedure can avoid estimating nuisance parameters, and it is indicated in simulations that the bootstrap test is robust to the choice of the user-specific regularization parameter.

Masayuki Hirukawa

[References]

1. Alj, A., Azrak, R. and Mélard, G. (2014). On conditions in central limit theorems for martingale difference arrays. *Economics Letters* **123**, 305–307. MR3202259
2. Braakmann, N. and Wildman, J. (2016). Reconsidering the effect of family size on labour supply: The twin problems of the twin birth instrument. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* **179**, 1093–1115. MR3555011
3. Cardot, H., Mas, A. and Sarda, P. (2007). CLT in functional linear regression models. *Probability Theory and Related Fields* **138**, 325–361. MR2299711
4. Chiou, J.-M., Yang, Y.-F. and Chen, Y.-T. (2016). Multivariate functional linear regression and prediction. *Journal of Multivariate Analysis* **146**, 301–312. Special Issue on Statistical Models and Methods for High or Infinite Dimensional Spaces. MR3477667
5. Cuesta-Albertos, J. A., García-Portugués, E., Febrero-Bande, M. and González-Manteiga, W. (2019). Goodness-of-fit tests for the functional linear model based on randomly projected empirical processes. *The Annals of Statistics* **47**, 439–467.

MR3909938
6. DasGupta, A. (2008). *Asymptotic Theory of Statistics and Probability.* Springer, New York. MR2664452
7. Dorn, M. (2021). Tests auf Exogenität im Funktionalen Linearen Regressionsmodell Unter Schwacher Stationarität. Ph.D. Thesis. Universität Bayreuth, Bayreuth. DOI: `10.15495/EPub UBT 00005597.`
8. Florens, J.-P. and Van Bellegem, S. (2015). Instrumental variable estimation in functional linear models. *Journal of Econometrics* **186**, 465–476. MR3343797
9. Gaenssler, P., Strobel, J. and Stute, W. (1978). On central limit theorems for martingale triangular arrays. *Acta Mathematica Academiae Scientiarum Hungaricae* **31**, 205–216. MR0471030
10. García-Portugués, E., Álvarez-Liébana, J., Álvarez-Pérez, G. and González-Manteiga, W. (2020). Goodness-of-fit tests for functional linear models based on integrated projections. In *Functional and High-Dimensional Statistics and Related Fields. Selected Papers Presented at the 5th International Workshop on Functional and Operatorial Statistics, IWFOS 2021, Brno, Czech Republic, June 23–25, 2021*, 107–114. Springer, Cham. MR4365786
11. García-Portugués, E., González-Manteiga, W. and Febrero-Bande, M. (2014). A goodness-of-fit test for the functional linear model with scalar response. *Journal of Computational and Graphical Statistics* **23**, 761–778. MR3224655
12. Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Application.* Academic Press, New York. MR0624435
13. Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica* **46**, 1251–1271. MR0513692
14. Hörmann, S. and Jammoul, F. (2023). Prediction in functional regression with discretely observed and noisy covariates. *Computational Statistics & Data Analysis* **178**, 107600. MR4482911
15. Johannes, J. (2008). Nonparametric estimation in functional linear model. In *Functional and Operatorial Statistics*, 215–221. Physica-Verlag HD, Heidelberg. MR2490352
16. Johannes, J. (2009). Nonparametric estimation in functional linear models with second order stationary regressors. *arXiv:0901.4266.* MR2490352
17. Johannes, J. (2016). Functional linear instrumental regression under second order stationarity. *arXiv:1603.01649.*
18. Mair, B. A. and Ruymgaart, F. (1996). Statistical inverse estimation in Hilbert scales. *SIAM Journal on Applied Mathematics* **56**, 1424–1444. MR1409127
19. Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *The Annals of Statistics* **21**, 255–285. MR1212176
20. Müller, H.-G. and Stadtmüller, U. (2005). Generalized functional linear models. *The Annals of Statistics* **33**, 774–805. MR2163159
21. Neubauer, A. (1988). An a posteriori parameter choice for Tikhonov regularization in the presence of modeling error. *Applied Numerical Mathematics* **4**, 507–519. MR0964300
22. Reiss, P. C. (2016). Just how sensitive are instrumental variable estimates? *Foundations and Trends® in Accounting* **10**, 204–237.
23. Ruymgaart, F., Wang, J., Wei, S.-H. and Yu, L. (2011). Some asymptotic theory for functional regression and classification. *Advances in Decision Sciences* **2011**, article ID 485974. DOI: `10.1155/2011/485974`. MR2874610
24. Seong, D. and Seo, W.-K. (2021). Functional instrumental variable regression with an application to estimating the impact of immigration on native wages. *arXiv:2110.12722.* MR5021265

25. Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, Hoboken. MR0595165
26. Tsybakov, A. B. (2004). *Introduction à l'estimation Non-paramétrique*. Springer, Paris. MR2013911
27. Wu, D.-M. (1974). Alternative tests of independence between stochastic regressors and disturbances: Finite sample results. *Econometrica* **42**, 529–546. MR0443237
28. Zhou, H., Yao, F. and Zhang, H. (2022). Functional linear regression for discretely observed data: From ideal to reality. *Biometrika* **110**, 381–393. MR4588359

Note: This list reflects references listed in the original paper as accurately as possible with no attempt to correct errors.