DOI: 10.1111/rssb.12523

Chaohua Dong, Jiti Gao and Oliver Linton's contribution to the Discussion of 'Assumption-lean inference for generalised linear model parameters' by Vansteelandt and Dukes

Chaohua Dong[1] | **Jiti Gao**[2] | **Oliver Linton**[3]

[1]Zhongnan University of Economics and Law, Wuhan, China

[2]Monash University, Melbourne, Australia

[3]University of Cambridge, Cambridge, UK

Correspondence
Oliver Linton, University of Cambridge, Cambridge, UK.
Email: obl20@outlook.com

The title of this paper is ironically self-fulfilling, since there are almost no meaningful assumptions made throughout! The starting point is that we have some plausible semi-parametric model, which is a special case of a more general non-parametric model, but we wish to allow for misspecification and in particular define an estimand that is meaningful in the non-parametric model and that specialises in the plausible model to a slope coefficient. However, since the estimator that is proposed is not the semi-parametric efficient estimator of that slope coefficient under the semi-parametric model, we wonder what is the role of the model at all? The theory side of it seems to assume in Theorem 2, for example that $E(Y|A, L)$ is consistently estimated in L_2 under the full unrestricted -parametric setting. But if that is possible, then why bother with the model? The authors talk casually about machine learning methods being used to estimate $E(Y|A, L)$, but if that is a silver bullet, then who needs the model? The model embodies some structure around A, L but the discussion is focussed away from the dimensionality of L, which is a big reason why one might want a structured model such as additivity (Linton & Nielsen, 1995) or the partial linear model (Robinson, 1988). Perhaps it would help if a full model was written down for the effect of high-dimensional L. Perhaps the point is that the parameter of interest is only defined in terms of low-dimensional conditional expectations, but this does not appear to be the case in the sense that high-dimensional smoothing is employed in (a) of p.14, which is then projected down by conditional expectation onto L, but if A is binary, then this has not reduced dimensionality at all, the dimensionality issue sits in L and what structure is assumed about its effect on Y.

The authors say that the usual choice of smoothing parameter is not tuned to the estimand. This point has been made in Linton (1995) who derived an optimal bandwidth for estimation of slope parameters and Wald statistics in the partially linear model based on local polynomial estimators; the optimal rates are indeed different from those that minimise the mean squared error of the non-parametric regression involved.

An alternative approach is to use sieve methods throughout and penalisation. Of course this questions whether it is necessary to pay too much attention to the approximating model. For example, suppose that we let X_i be the $(2dK + 1) \times 1$ vector containing A_i and basis functions $\psi_k(L_{ji})$ and $A_i\psi_k(L_{ji})$ (if interactions between L_j and A are important) for $k = 1, \dots, K$ and $j = 1, \dots, d$, and $i = 1, \dots, n$, and let $\theta \in \mathbb{R}^{2dK+1}$ minimise

$$\left\| \frac{1}{n} \sum_{i=1}^{n} m(Y_i, \theta^\intercal X_i) \right\| + \text{pen}_\lambda(\theta), \tag{1}$$

where m is a large vector of (possibly non-linear) moment condition, while pen_λ is a penalty function such as SCAD or LASSO. Dong et al. (2018) establish, as a special case, the consistency and asymptotic normality of the estimators in (1) and provide consistent inference methods when the dimensionality of X_i is diverging and a smooth penalty like SCAD is used.

REFERENCES

Dong, C., Gao, J. & Linton, O. (2018) High dimensional semiparametric moment restriction models. To Appear in Journal of Econometrics. Available from: https://core.ac.uk/reader/186326795

Linton, O. (1995) Second order approximation in the partially linear regression model. *Econometrica*, 63, 1079–1112.

Linton, O. & Nielsen, J. (1995) A kernel method of estimating nonparametric structured regression based on marginal integration. *Biometrika*, 82, 93–100.

Robinson, P.M. (1988) Root-N-consistent semiparametric regression. *Econometrica*, 56, 931–954.

How to cite this article: Dong, C., Gao, J. & Linton, O. (2022) Chaohua Dong, Jiti Gao and Oliver Linton's contribution to the Discussion of 'Assumption-lean inference for generalised linear model parameters' by Vansteelandt and Dukes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 84, 707–708. Available from: https://doi.org/10.1111/rssb.12523