

# Copula-based regression models with responses missing at random: A unified approach

Shigeyuki Hamori   Kaiji Motegi   Zheng Zhang  
Kobe University   Kobe University   Renmin University of China

## 1 Motivation

The existing literature of copula-based regression models typically focuses on either conditional mean regression or conditional quantile regression, and assumes complete data. This paper unifies the conditional mean regression of Noh, El Ghouch, and Bouezmarni (2013), the conditional quantile regression of Noh, El Ghouch, and Van Keilegom (2015), and other interesting regressions such as asymmetric least squares by formulating a general loss function. Furthermore, we relax the rather unrealistic assumption of complete data by allowing the response variable to be missing at random (MAR).

## 2 Main results

A semiparametric estimator of the regression curve is proposed. When the response variable is MAR, a proper weight should be assigned to each individual in order to achieve consistent inference. The proper weights are characterized by the inverse propensity score function, and we replace them with calibration weights. The calibration approach is originally proposed by Chan, Yam, and Zhang (2016) and recently applied to the estimation of copula models with data missing at random by Hamori, Motegi, and Zhang (2019).

The consistency and asymptotic normality of the estimated regression curve are proved. We show via Monte Carlo simulations that the proposed approach performs well, and it dominates the benchmark equal-weight approach. When the response variable is MAR, the proposed estimator is consistent while the equal-weight estimator has large bias.

## References

- CHAN, K. C. G., S. C. P. YAM, AND Z. ZHANG (2016): “Globally Efficient Non-Parametric Inference of Average Treatment Effects by Empirical Balancing Calibration Weighting,” *Journal of the Royal Statistical Society, Series B*, 78, 673–700.
- HAMORI, S., K. MOTEGI, AND Z. ZHANG (2019): “Calibration estimation of semiparametric copula models with data missing at random,” *Journal of Multivariate Analysis*, 173, 85–109.
- NOH, H., A. EL GHOUGH, AND T. BOUEZMARNI (2013): “Copula-based regression estimation and inference,” *Journal of the American Statistical Association*, 108(502), 676–688.
- NOH, H., A. EL GHOUGH, AND I. VAN KEILEGOM (2015): “Semiparametric conditional quantile estimation through copula-based multivariate models,” *Journal of Business & Economic Statistics*, 33(2), 167–178.