Statistica Sinica **36** (2026), 1-23
doi:https://doi.org/10.5705/ss.202023.0271

A QUASI SYNTHETIC CONTROL METHOD FOR NONLINEAR MODELS WITH HIGH-DIMENSIONAL COVARIATES

Zongwu Cai[1], Ying Fang[2], Ming Lin[*2] and Zixuan Wu[2]

[1] *University of Kansas and* [2] *Xiamen University*

Abstract: To make the conventional synthetic control method more flexible to estimate the average treatment effect (ATE), this article proposes a quasi synthetic control method for nonlinear models under the index model framework with possible high-dimensional covariates, together with a suggestion of using the minimum average variance estimation (MAVE) method to estimate parameters and the LASSO-type procedure to choose high-dimensional covariates. We derive the asymptotic distribution of the proposed ATE estimators for both finite and diverging dimensions of covariates. A properly designed Bootstrap method is proposed to obtain confidence intervals and its theoretical justification is provided. When the dimension of covariates is greater than the sample size, we suggest using the robust version of sure independence screening procedure based on the distance correlation to first reduce the dimensionality and then apply the MAVE approach to estimate parameters. Finally, Monte Carlo simulation studies are conducted to examine the finite sample performance of our proposed estimators and Bootstrap procedure. In addition, an empirical application to reanalyzing data from the National Supported Work Demonstration demonstrates the practical usefulness of our proposed method.

Key words and phrases: Average treatment effect, bootstrap inference; index model, semiparametric estimation, synthetic control method.

1. Introduction

When evaluating the impact of policy interventions, one of the main challenges lies in estimating unknown counterfactual outcomes. With observable covariates, a natural idea is to construct an outcome regression model. In practice, the classic linear regression model is usually inadequate or even incorrect. To fully capture the relationship between the covariates and the outcomes, researchers suggest using the nonparametric model which can avoid the risk of model misspecification. However, the nonparametric model is challenged by the so-called *curse of dimensionality*. Therefore, as a combination of the parametric and nonparametric models, the semiparametric model has been conceived to overcome the aforementioned limitations.

*Corresponding author. E-mail: linming50@xmu.edu.cn