!

DISCUSSION PAPERS IN ECONOMICS

W Pa N . 24-02

"""""""""0&1(23"4)5"
6 7!9()18%5"&:"; &<&)=>&"?&@<>()"

"
"
"
"

#$%&'()"*+,"-.-/"

"

"
"
"
"
"

Department of Economics

"

University of Colorado Boulder
Boulder, Colorado 80309

! "#$%&'()"*+,"-.-/,"0&1(23"4)5"

Robust Inference when Nuisance Parameters may be Partially Identified with Applications to Synthetic Controls

Joseph Fry[y]

[y]University of Colorado at Boulder, 256 UCB, 80309 Boulder, Colorado, United States

E-mail address: joseph.fry@colorado.edu

October 13, 2024

Abstract

When conducting inference for the average treatment effect on the treated with a Synthetic Control Estimator, the vector of control weights is a nuisance parameter which is often

1 Introduction

In the context of method of moments estimation, if the moment conditions jointly identify a subvector of the parameters, then a standard estimation method such as General Method of Moments (GMM) is generally a consistent estimator for that subvector. However, if the remaining subvector is unidenti ed, then this remaining parameter cannot be consistently estimated and as a result the estimates for the identi ed subvector are not asymptotically normal (see, for example, Andrews and Cheng (2012)), which signi cantly complicates inference. Additionally, if a nuisance parameter is at the boundary of the parameter space or close to the boundary relative to the sample size, then this can also result in our estimates for the parameter of interest not being asymptotically normal (see, for example, Andrews (1999) and Geyer (1994)). Lastly, if the full vector is high dimensional, this can complicate standard asymptotic normality results even when the subvector we would like to perform inference on is low dimensional. I propose an estimation method that aims to simultaneously overcome these complications to obtain an estimate for an identi ed parameter of interest that is asymptotically normal even when a nuisance parameter is partially identi ed, on or near the boundary of the parameter space, and, in some cases, high dimensional.

The procedure I propose can be decomposed into three steps. In the rst step, regularized estimates of all parameters are found by minimizing a penalty function subject to the constraint that the sample moment conditions are close to zero. The primary purpose of this penalty function is to make the estimated nuisance parameter converge to a unique element of the identi ed set. We must therefore choose which element of the identi ed set we would like our estimate to converge to. Since this choice a ects the asymptotic variance of our subsequent estimate of the parameter of interest, I base the penalty function on an estimate of the asymptotic variance as a function of the nuisance parameter, provided that this asymptotic variance function can be estimated su ciently accurately. In cases where the asymptotic variance cannot be accurately estimated, such as when using time-series or panel data and the degree of temporal dependence is high relative to the number of time periods,

I discuss alternative ways of choosing the penalty function in Section 3. The second step in my procedure is to use Neyman orthogonalization to construct a set of moment conditions that are orthogonal with respect to the nuisance parameter. This involves introducing a new parameter that is chosen to make the derivative of the moment conditions with respect to the original nuisance parameter equal to zero. After the estimated nuisance parameters have been plugged into the orthogonal moment conditions, the third step is to use these moment conditions to re-estimate the parameter of interest. In Section 2, I give two suggestions of how this can be done and provide asymptotic normality results for both methods.

My primary application of this approach is as a Synthetic Control Estimator (SCE) which gives an asymptotically normal estimate of the average treatment e ect on the treated (ATT). Several other inference methods have been proposed for SCEs. The placebo method of Abadie et al. (2010) is commonly used in practice and, as previously noted (see Abadie et al. (2010) and Abadie et al. (2015)), corresponds to a traditional Fisher Randomization Test when treatment is randomly assigned. While this would mean that this test have exact size from a design-based perspective, this condition is unreali01(elaceb)-27(o)-298(metho)-27(d ectivased persp70(a)-471(Ss.Abadie)-361(et)-360(al.)]TJ 0 g 0 G [-361(()]TJ 0.36472 0.410(Ch)-315n

where the parameter of interest is identified, we can focus on conducting the estimation in a way that makes inference simpler for that parameter (by making its estimated values asymptotically normal) and possibly in a way that makes the estimates of that parameter more precise (by minimizing its asymptotic variance). On the other hand, methods that conduct inference on the whole identified set have the advantage that they can be used when both the parameter of interest and the nuisance parameter are partially identified.

This work also extends the literature on the Neyman Orthogonalized Score. While the technique dates back to Neyman (1959), several more recent papers have used it as a way to achieve asymptotic normality after obtaining an estimate of a high-dimensional nuisance parameter using a regularization penalty or machine learning technique (e.g., Belloni et al. (2018), Ning and Liu (2017), Chernozhukov et al. (2015), Belloni et al. (2014), and Chernozhukov et al. (2018)). Many of these methods estimate the nuisance parameter with LASSO. While this can be a powerful technique when the nuisance parameter is high dimensional but has a sparse, point-identified value, if the nuisance parameter is partially identified, the LASSO penalty may often be insufficient for the estimates to converge to a specific vector. I help extend this literature by showing how the Neyman Orthogonalized Score can be applied in cases where the nuisance parameter is partially identified. The literature on optimal instruments is also related, and in particular, Singh et al. (2020) take a similar approach as I do here, where they choose the function of the instruments that minimizes the estimated asymptotic variance for the parameter of interest. In section 5, I give an example of how my method can be applied to the optimal choice of instruments.

In Section 2, I discuss how a set of moment conditions can be combined with initial regularized estimates of the parameters to create the orthogonal moment conditions and then how these orthogonal moment conditions can be used to estimate the parameter of interest. I discuss this first to show what properties we want the regularized estimates of the parameters to satisfy, and how the limiting values of the estimated nuisance parameters influence the asymptotic variance for the estimated parameter of interest. In Section 3,

I then show when the regularized estimates satisfy these conditions and discuss how to adjust the procedure to handle cases where the asymptotic variance cannot be consistently estimated. In Section 4, I show when the high level conditions in the previous sections are satis ed for a SCE under a linear factor model structure. I apply this SCE by replicating the work of Andersson (2019) estimating the e ect of Sweden's carbon pricing policies on CO2 emissions. My method nds that their results are statistically signi cant at commonly