

# The Wild Bootstrap with a “Small” Number of “Large” Clusters\*

Ivan A. Canay	Andres Santos
Department of Economics	Department of Economics
Northwestern University	U.C.L.A.
iacanay@northwestern.edu	andres@econ.ucla.edu

Azeem M. Shaikh  
Department of Economics  
University of Chicago  
amshaikh@uchicago.edu

May 4, 2018

## Abstract

This paper studies the properties of the wild bootstrap-based test proposed in [Cameron et al. \(2008\)](#) in settings with clustered data. [Cameron et al. \(2008\)](#) provide simulations that suggest this test works well even in settings with as few as five clusters, but existing theoretical analyses of its properties all rely on an asymptotic framework in which the number of clusters is “large.” In contrast to these analyses, we employ an asymptotic framework in which the number of clusters is “small,” but the number of observations per cluster is “large.” In this framework, we provide conditions under which the limiting rejection probability of an un-Studentized version of the test does not exceed the nominal level. Importantly, these conditions require, among other things, certain homogeneity restrictions on the distribution of covariates. We further establish that the limiting rejection probability of a Studentized version of the test does not exceed the nominal level by more than an amount that decreases exponentially with the number of clusters. We study the relevance of our theoretical results for finite samples via a simulation study.

KEYWORDS: Wild bootstrap, Clustered Data, Randomization Tests.

---

\*We thank Colin Cameron, Patrick Kline, Simon Lee, Magne Mogstad and Ulrich Mueller for helpful comments. The research of the first author was supported by National Science Foundation Grant SES-1530534. The research of the third author was supported by National Science Foundation Grants DMS-1308260, SES-1227091, and SES-1530661. We thank Max Tabord-Meehan for excellent research assistance.