SUKJIN HAN

Home Address:
100 York St., Apt. 10R
New Haven, CT 06511

Office Address:
Department of Economics
Yale University
28 Hillhouse Avenue
PO Box 208268
New Haven, CT 06520-8268
Fax: (203) 432-2128

Telephone: (203) 623-6736

E-mail: sukjin.han@yale.edu

Webpage: https://sites.google.com/site/sukjinhanwebpage

Citizenship: Korea, Republic of (F-1 Visa)

Gender: Male

Fields of Concentration:

Econometric Theory
Applied Econometrics

Desired Teaching:

Econometrics and Applied Econometrics
Probability and Statistics
Intermediate Microeconomics
Labor Economics

Comprehensive Examinations Completed:

May 2008, (Oral) Econometrics, Labor Economics
May 2007, (Written) Microeconomics, Macroeconomics

Dissertation Title:

Weak Instruments in Nonparametric and Nonlinear Models with Endogeneity

Committee:

Professor Donald Andrews (co-chair)
Professor Edward Vytlacil (co-chair)
Professor Xiaohong Chen
Professor Yuichi Kitamura

Expected Completion Date: May 2012

Degrees:

Ph.D., Economics, Yale University (2012, expected)
M.Phil., Economics, Yale University (2009)
M.A., Economics, Yale University (2009)
B.A. (*with highest honors*), Economics, Seoul National University (2005)

Fellowships, Honors and Awards:

Carl Arvid Anderson Prize Fellowship, Cowles Foundation (2011-2012)
Dissertation Fellowship, Yale University (2012)
National Science Foundation Graduate Research Fellowship (2008)
Cowles Foundation Fellowship, Yale University (2006-2010)
University Fellowship, Yale University (2006-2010)
University Summer Fellowship, Yale University (2007, 2008)
Research Fellowship, Department of Economics, Seoul National University (2005-2006)
Teaching Fellowship, Department of Economics, Seoul National University (2005)
Development Fund Full Scholarship, Seoul National University (2002-2004)

Teaching Experience:

(Graduate) Econometrics I, Fall 2010
(Undergraduate) Introductory Microeconomics, Spring 2010 — Evaluations
(Undergraduate) Econometrics and Data Analysis I, Fall 2009 — Evaluations
(Graduate) Econometrics II, Spring 2009

Research Experience:

Summer Internship, Office of Economic Analysis, Securities and Exchange Commission (2007)
Research Assistantship, Korea Fixed Income Research Institute (2005-2006)

Publications:

"Invalidity of the Bootstrap and the *m* out of *n* Bootstrap for Confidence Interval Endpoints Defined by Moment Inequalities," with Donald Andrews, *Econometrics Journal* (2009), Volume 12, pp. S172–S199

Working Papers:

"Nonparametric Triangular Simultaneous Equations Models with Weak Instruments," *Job Market Paper*
"Identification in Bivariate Probit Models with Endogeneity" with Edward Vytlacil
"Inference in Bivariate Probit Models with Weak Instruments"
"The Secular Trend in the Standard of Living During Industrialization in Britain: Comments on Komlos (1993)"

Stata Command:

"CQIV" with Victor Chernozhukov, Ivan Fernandez-Val, and Amanda Kowalski

Conference Presentation:

CIREQ Conference on Inference with Incomplete Models, Montréal, October 2008

Referee Service:

Journal of Applied Econometrics

Language: English (fluent), Korean (native), French (intermediate)

References:

Professor Donald W. K. Andrews
Cowles Foundation
Yale University
P.O. Box 208281
New Haven, CT 06520-8281
Phone: (203) 432-3698
Fax: (203) 432-6167
E-mail: donald.andrews@yale.edu

Professor Edward Vytlacil
Cowles Foundation
Yale University
P.O. Box 208281
New Haven, CT 06520-8281
Phone: (203) 436-3994
Fax: (203) 432-6167
E-mail: edward.vytlacil@yale.edu

Professor Xiaohong Chen
Cowles Foundation
Yale University
P.O. Box 208281
New Haven, CT 06520-8281
Phone: (203) 432-5852
Fax: (203) 432-6167
E-mail: xiaohong.chen@yale.edu

Professor Yuichi Kitamura
Cowles Foundation
Yale University
P.O. Box 208281
New Haven, CT 06520-8281
Phone: (203) 432-3699
Fax: (203) 432-6167
E-mail: yuichi.kitamura@yale.edu

Dissertation Abstract

Instrumental variables (IVs) are widely used to identify and estimate causal relationships in models with endogenous explanatory variables. Having relevant instruments is crucial for accurate inference. An extensive literature has studied the consequences of weak instruments in linear simultaneous equations models, with weak instruments defined as instruments that are only weakly correlated with the endogenous explanatory variables. One might conjecture that the problem of weak instruments becomes even more important in nonparametric or limited dependent variables models with endogeneity, as more flexible or complex models generally require stronger identification power, and hence plausibly stronger instruments. Despite the problem's importance and the growing popularity of such models, there has been no study focused on weak instruments in nonparametric settings and only few in nonlinear settings.

This dissertation conducts identification analyses and studies weak instruments in several nonparametric and nonlinear simultaneous equations models. It proposes methods for estimation and, in some cases, inference when instruments are possibly weak. In the first chapter, I consider a nonparametric triangular model and define nonparametric weak instruments by deriving a suitable identification condition and "localizing it." I propose a penalized series estimator, where I use penalization to alleviate the effect of the weak instruments, and derive the weak instrument asymptotics of the estimator.

The second and third chapters consider bivariate probit models with endogenous binary regressors. I conduct identification analyses, define weak instruments, and derive weak instrument asymptotics of the maximum likelihood estimator. I establish the asymptotic size of tests and confidence sets, and I introduce an inferential method that is robust to weak instruments.

I. Identification and Estimation in Nonparametric Triangular Models with Weak Instruments

In this chapter, I generalize the concept of weak instruments and analyze the effect of weak instruments on identification, estimation and inference in a nonparametric triangular model. The motivation for the model and problem can be illustrated with the example of Angrist and Lavy (1999), who study the effect of class size on students' test scores. Class size is endogenous and they use a rule on maximum class size as an instrument. While Angrist and Lavy impose a linear model, a more flexible functional form is appealing to allow the marginal effect of class size on test scores to be different across class-size levels. One may worry, however, that the instrument might be weak when the effect is estimated nonparametrically. With Angrist and Lavy's sample, the first stage F-statistic is 191.66, which indicates that the instrument is very strong based on conventional criteria for linear models. In simulations, however, I show that when one estimates a nonparametric structural function, even instruments that appear to be strong in a conventional sense may be considered weak, in that the IV estimator does poorly in terms of mean squared error compared to the least squares estimator that ignores endogeneity. This result raises the need for greater attention to weak instruments in nonparametric settings.

This chapter proceeds with a triangular model fully nonparametric in both the outcome and first-stage equations. I follow the control function approach for the identification and estimation of the structural function. I first obtain a novel identification result based on which I establish "weak identification." With a mild support condition, I show that a particular rank condition is necessary and sufficient for identification, so that a "slight violation" of it can result in weak identification. This rank condition for identification is substantially weaker than the sufficient condition previously established in the literature.

I then define nonparametric weak instruments, which generalizes the existing definition of weak instruments of a linear first stage. In particular, I consider a sequence of first-stage regression functions that converges at a certain rate to a function with no identification power, namely, a constant function. The rate effectively measures the strength of the instruments and later appears in my local asymptotic results.

Interestingly, the weak instrument problem can be seen as a multicollinearity problem and is related to the ill-posed inverse problem in the literature. Given this insight, I introduce a penalization scheme in estimation as a regularization method for weak instruments, and derive the asymptotic properties of the resulting penalized series estimator. My results on the rate of convergence of the estimator imply that weak instruments affect bias and variance "symmetrically," unlike the common wisdom that collinearity results in imprecise estimates but does not introduce bias. Consistency and asymptotic normality are achieved provided that the instrument is only "mildly weak" and that the penalty shrinks fast enough.

This chapter provides useful implications for applied researchers who consider various nonparametric models: IV estimates can be misleading even when instruments are conventionally strong; the effects of weak instruments cannot be alleviated by the choice of the order of series; simulation results suggest that penalization can reduce variance significantly and, in some cases, bias as well; and lastly, the strength of instruments can be improved by considering a nonparametric first stage, so that the nonlinear relationship between the endogenous variable and instruments is fully exploited.

As an empirical application, I consider the class size example mentioned above. I calculate penalized series estimates for the class-size effect with different choices of tuning parameters. I also contrast estimates based on linear versus nonparametric first stage and provide evidence that the instrument is actually weak in a nonparametric sense. I compare my "control function estimates" with the "nonparametric IV estimates" of Horowitz (2011) which is subject to the ill-posed inverse problem.

II. Identification in Bivariate Probit Models with Endogeneity (Joint with Edward Vytlacil) and Inference with Weak Instruments

These two chapters of my dissertation study the identification problem in bivariate probit (BVP) models with endogenous binary regressors, and develop inferential methods that are robust to weak instruments. The outcome equation is a probit model that includes a binary endogenous regressor, which is explained by another probit equation in the first stage. For example, this model is used in the schooling literature to determine the effect of attending catholic high schools on college attendance; attendance in catholic schools is endogenous and so instruments are introduced. Although BVP models have been used extensively in applied research, relatively little research has been done on identification analysis and weak instruments for these models.

The first part starts with deriving identification conditions in a simple BVP model. Identification is achieved without the strong support conditions for exogenous variables like those of Freedman and Sekhon (2008). We find that nonzero coefficients of excluded instruments are sufficient for local identification, and a compact parameter space guarantees global identification. We then extend the model in several interesting directions, one of which is to a model with common exogenous regressors in both equations but *without* excluded instruments in the first stage. We find that identification is still achieved here, due to the strong distributional assumption of a BVP model.

With the same class of models, the second part is to introduce weak instruments as local to zero coefficients of the excluded instruments or of the common exogenous regressors, and to find weak instrument asymptotics of maximum likelihood estimators. In doing so, I develop a reparametrization scheme that allows me to adapt the framework of Andrews and Cheng (2011). I verify their assumptions for asymptotic theory and establish consistency, lack-of-consistency and asymptotic distributions of the estimators under a full range of strength of instruments. Based on these results, I establish the asymptotic size of tests and confidence sets in a uniform sense. I also introduce an inferential method that is robust to weak instruments.