行政院國家科學委員會專題研究計畫 成果報告

條件動差限制下的弱認定問題及其一致性估計方法
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 97-2410-H-004-179-
執 行 期 間 ：97 年 11 月 01 日至 98 年 07 月 31 日
執 行 單 位 ：國立政治大學經濟學系

計畫主持人：徐士勛

計畫參與人員：博士班研究生-兼任助理人員：謝子雄
博士班研究生-兼任助理人員：徐兆璿
博士班研究生-兼任助理人員：曾憲政

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 98 年 10 月 29 日

行政院國家科學委員會補助專題研究計畫 ■ 成 果 報 告 □期中進度報告

條件動差限制下的弱認定問題及其一致性估計方法

計畫類別：■ 個別型計畫 □ 整合型計畫
計畫編號：NSC 97-2410-H-004 -179 -
執行期間： 97 年 11 月 1 日至 98 年 7 月 31 日

計畫主持人：徐士勛
共同主持人：
計畫參與人員：謝子雄，徐兆璿，曾憲政

成果報告類型(依經費核定清單規定繳交)：■精簡報告 □完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：

中 華 民 國 98 年 10 月 29 日

摘要

近年來, 模型中待估參數可能產生的「弱認定」問題逐漸受到學術界的關注。一般而言, 推斷這個問題的存在與否都是依據「非條件動差限制式」而來。在此研究計畫中, 我首先舉例說明當理論所推衍出的模型爲「條件動差限制式」時, 文獻上依據所選擇有限個數的「非條件動差限制式」來推論參數弱認定問題的方式是不適當的。這是因爲如果我們要表徵「條件動差限制式」中所有關於待估參數的訊息, 我們必須能同時考慮無窮多條其所衍生出的「非條件動差限制式」才行。針對此問題, 我首先直接考慮在「條件動差限制式」下參數的「弱認定」問題, 並會對相關的研究文獻作一完整探討。在此研究計畫的第二部份中, 我將會利用「條件動差限制式」所推衍出無窮多條「非條件動差限制式」建立一個可行的一般化的動差估計式。此估計式被證明具有一致性且其相關極限理論性質也都建立。在計畫的最後一部份, 我藉由蒙地卡羅模擬來檢驗估計式的小樣本性質。在「弱認定」的分析架構下, 此計畫所提出的方法和所建立的極限性質在文獻上具有一定的貢獻且應會有相當的參考價值。

關鍵詞:弱認定, 條件動差限制式, 一般化的動差估計式

Abstract

In recent years, the “weak identification” problem of the parameters of interest in models has drawn much attention. Typically, this problem is identified by some “unconditional” moment restrictions. In the first part of this project, I show that, given the conditional moment restrictions, to identify weak identification only based on some selected finite unconditional moments may be misleading. This is because all of the information in the conditional moment restrictions is accounted for *only when* we consider infinitely many implied unconditional moments. In this project, I identify the weak identification problem directly based on the conditional moments, and I review the weak identification problems in detail in the literature. In the second part of this project, I construct a GMM type estimator based on a continuum of unconditional moments that reveals all the information of the parameters of interest contained in conditional moment restrictions. The proposed estimator is shown to be consistent and the corresponding theoretical properties are also established. The last part of this project examines finite sample performance of the proposed estimators, via extensive Monte Carlo simulations. When identifying the weak identification problem by conditional moment restrictions, the proposed estimation approach and the asymptotics in this project are new in the literature.

Keywords: weak identification, conditional moment restrictions, generalized method of moment

1 Introduction

Moment restrictions usually arise in economic or financial models. In particular, conditional moment restriction plays a major role since we usually build models conditioned on some information sets in an uncertain environment. Therefore, how to "accurately" estimate the parameters of interest in the models via those moment restrictions is an important task. In recent years, many researchers are attracted to the "weak identification" problem in generalized method of moment (GMM) framework based on some "unconditional" moment restrictions induced from the conditional ones. Typically, the weak identification means that the selected (unconditional) moment restrictions do not provide enough information about the parameters of interest. In linear instrumental variables regression models, this problem corresponds to the well-known "weak instruments" when the instruments are weakly correlated with the endogenous variables. In the presence of weak identification (or weak instruments), the conventional first-order asymptotic analysis is misleading, and the resulting inference becomes unstable. Given a set of unconditional moment restrictions, there have been many researches focus on the issues related to weak identification, Staiger and Stock (1997), Stock and Wright (2000), Hahn and Hausman (2003), and Stock, Wright and Yogo (2002) are only a few examples.

Given the conditional moment restrictions, however, it is not a good way to consider the identification problem of the parameters only based on a finite set of selected unconditional moment restrictions. This is because all of the information in the conditional moment restrictions is accounted for *only when* we consider infinitely many implied unconditional moments. Without any prior knowledge, the finite set of selected unconditional moments may contain less information about the parameters and then we may conclude that the parameters are weak identified by applying some well-established approaches in the literature, even the parameters actually are *well identified* by the conditional moment restrictions. From this viewpoint, the weak identification problem may be overstated if we infer based on some finitely implied unconditional moment restrictions. To my best knowledge, no work on the weak identification problem emphasizes this point in the literature.

When identifying the problem of weak identification based on the conditional moment restrictions directly, it follows that any induced unconditional moment restriction is "weak" to identify the parameters of interest. How to construct a consistent estimator is thus the main part of this project. In the framework of linear instrumental variables regression, Chao and Swanson (2005) show that as the number of weak instruments increases to the infinity at some suitable rate, there may be a consistent estimator. It indicates that we may still have enough information for identifying the parameters and thus the consistent estimation is achievable even when instruments are all "weak". Their work motivates us to consider an infinite number of induced unconditional moment restrictions in the weak identification problem.

On the other hand, when the parameters of interest are well identified by the conditional moment restrictions, many consistent (and efficient) estimation methods have been proposed in the literature. In particular, trying to take all induced unconditional moment restrictions into account in estimation is popular. For example, Bierens(1990) consider a set of unconditional moment restrictions generating by the exponentials of the conditioning variables. Donald et. al. (2003) form the unconditional moments by using the power series and splines. Hsu and Kuan (2008) consider a continuum set of unconditional moments based on generically comprehensively revealing functions of Stinchcombe and White (1998) and further construct a class of consistent estimators by introducing Fourier analysis. It should be noted that, however, the asymptotic properties of these approaches are valid only when the parameters of interest are well identified by the conditional moment restrictions.

2 Weak Identification

Let us consider a conditional moment restriction as

$$\mathbb{E}[\boldsymbol{h}(Y_t, \theta_o)|Z_t] = 0, \qquad \text{with probability one (w.p.1)}, \tag{1}$$

where $\boldsymbol{h}(\cdot)$ is an $s \times 1$ vector of known functions, Y_t is data variable, Z_t is conditioning variable with dimension one, and $\theta_o \in \Theta$ is a $q \times 1$ vector of unknown parameters of interest.

Given this conditional moment restriction (1), the weak identification problem is typically addressed as follows. Selecting some finite-dimensional vector of measurable functions of Z_t, $\boldsymbol{f}(Z_t)$ say, the conditional moment restriction (1) implies a set of unconditional moment restrictions

$$\mathbb{E}[\boldsymbol{h}(Y_t, \theta_o) \otimes \boldsymbol{f}(Z_t)] = 0, \tag{2}$$

by the law of iterated expectations. If $\mathbb{E}[\boldsymbol{h}(Y_t, \theta) \otimes \boldsymbol{f}(Z_t)]$ is nearly zero for $\theta \neq \theta_o$, then θ_o is thought as being weak identified; see e.g., Stock and Wright (2000), Stock, Wright and Yogo (2002) and Wright (2003).

As stated above, to address the weak identification problem based on the unconditional moment restriction (2) is "risky" since this set of unconditional moments does not contain all of the information about θ_o. We may draw the wrong inference from this set of unconditional moments. The following example points out this possibility.

Example: Assume the random variable Y satisfies the simple nonlinear model $\mathbb{E}[Y|X] = \exp(X\theta_o)$, where θ_o is an unknown parameter of interest and X follows a normal distribution with a zero mean and an nonzero variance. The corresponding conditional moment restriction is then given by

$\mathbb{E}[Y - \exp(X\theta_o)|X] = 0$. Based on this conditional restriction, we consider three valid, implied unconditional moment restrictions to illustrate the identification problem, they are, respectively, (i) $\mathbb{E}[Y - \exp(X\theta)] = 0$; (ii) $\mathbb{E}[(Y - \exp(X\theta))X] = 0$; and (iii) $\mathbb{E}[(Y - \exp(X\theta))X^2] = 0$. Some calculations show that restrictions (i) and (iii) can not help to identify θ_o since $-\theta_o$ also satisfies these two restrictions. Instead, moment restriction (ii) gives the unique solution θ_o and thus is helpful to identify θ_o. It means that if only the unconditional moment restrictions (i) and (iii) are considered, we may conclude that θ_o cant not be identified (or θ_o is weak identified) by that conditional moment restriction.

As a result, in this project, I say the parameters are weak identified if it satisfies the following definition:

Definition (Weak Identification)
θ_o is weak identified in (1) if $\mathbb{E}[\boldsymbol{h}(Y_t, \theta)|Z_t]$ is nearly zero for $\theta \neq \theta_o, \theta \in \Theta$.

3 The Proposed Approaches

Given that θ_o is weak identified in the conditional moment restriction (1), I will consider a class of sets of induced unconditional moment restrictions in this proposed estimation approach. Note that not matter what θ_o is well or weak identified in conditional moment restriction (1), it is necessary to take all induced unconditional moments, which are infinitely many, into account in order to contain the equivalent information about θ_o in the original conditional moments. This point of view has been emphasized by Donald et al. (2003) and Hsu and Kuan (2008.)

In this project, I consider a continuum of unconditional moments based on the generically comprehensive revealing functions of Z_t. Any real analytic function but not a polynomial is generically comprehensive revealing, so this class of functions is not particular. Let $\boldsymbol{\tau} := [\tau_0, \tau_1]' \in \mathbb{R}^2$ and $\mathcal{A}$ denote the affine transformation of Z_t such that $\mathcal{A}(Z_t, \boldsymbol{\tau}) = \tau_0 + \tau_1 Z_t$, and function G is generically comprehensive revealing, we have

$$\mathbb{E}\left[\boldsymbol{h}(Y_t, \theta_o) \otimes G\big(\mathcal{A}(Z_t, \boldsymbol{\tau})\big)\right] = 0, \text{ for almost all } \boldsymbol{\tau} \in \mathcal{T} \subset \mathbb{R}^2, \tag{3}$$

where $\mathcal{T}$ has nonempty interior. This continuum of unconditional moment restrictions reveals all the information of θ_o contained in the conditional one. The details can be found in Stinchcombe and White (1998) and Hsu and Kuan (2008). For any z in the complex vector space, $\bar{z}$ denotes the complex conjugate of z, and the magnitude $|z|^2 = z'\bar{z}$. Following the approach in Hsu and Kuan (2008), I jointly consider all moment restrictions indexed by $\boldsymbol{\tau}$ in (3) via computing L_2-norm by the integration and invoking Paserval's Theorem. Apart from a scaling factor, we obtain

$$\int_{\mathcal{T}} \left| \mathbb{E}\left[\boldsymbol{h}(Y_t, \theta_o) \otimes G\big(\mathcal{A}(Z_t, \boldsymbol{\tau})\big)\right]\right|^2 d\boldsymbol{\tau} = \sum_{k_0, k_1 = -\infty}^{\infty} \left|\Psi_{t,G,\boldsymbol{k}}(\theta_o)\right|^2 = 0,$$

where

$$\begin{aligned}
\Psi_{t,G,\boldsymbol{k}}(\theta) &= \int_{\mathcal{T}} \mathbb{E}\left[\boldsymbol{h}(Y_t, \theta) \otimes G(\mathcal{A}(Z_t, \boldsymbol{\tau}))\right] \exp\left(-i\boldsymbol{k}'\boldsymbol{\tau}\right) d\boldsymbol{\tau} \\
&= \mathbb{E}\left[\boldsymbol{h}(Y_t, \theta) \otimes \int_{\mathcal{T}} G(\mathcal{A}(Z_t, \boldsymbol{\tau})) \exp\left(-i\boldsymbol{k}'\boldsymbol{\tau}\right) d\boldsymbol{\tau}\right] \\
&:= \mathbb{E}\left[\boldsymbol{h}(Y_t, \theta) \otimes \gamma_{t,G,\boldsymbol{k}}(\theta)\right]
\end{aligned}$$

is a Fourier coefficient induced by projecting the function $\mathbb{E}\left[\boldsymbol{h}(Y_t, \theta) \otimes G(\mathcal{A}(Z_t, \boldsymbol{\tau}))\right]$ along the exponential Fourier series, and $\gamma_{t,G,\boldsymbol{k}}$ can be viewed as the function which incorporates the full continuum of the original unconditional moment restrictions indexed by $\boldsymbol{\tau}$. By imposing the sample counterpart, we obtain the estimator

$$\hat{\theta}(K_T) = \underset{\theta \in \Theta}{\operatorname{argmin}} \sum_{k_0, k_1 = -K_T}^{K_T} \left| \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{h}(y_t, \theta) \otimes \gamma_{t,G,\boldsymbol{k}}(\theta) \right|^2, \tag{4}$$

where K_T is some positive integer and needs to grow with the sample size T to ensure proper asymptotic properties. This estimator (4) is first proposed by Hsu and Kuan (2008). However, as stated before, Hsu and Kuan (2008) deal with the conditional moments that θ_o is well identified. In subsequent section, the corresponding asymptotics for the proposed estimator will be established when θ_o is weakly identified. It is an extension of the work in Hsu and Kuan (2008).

4 The Asymptotic Properties

For $\boldsymbol{k} = (k_0, k_1)$, $k_0, k_1 = -K_T, -K_T + 1, \cdots, 0, 1, 2, \cdots, K_T - 1, K_T$, let

$$g_{t\boldsymbol{k}}(\theta) = \boldsymbol{h}(y_t, \theta) \otimes \gamma_{t,G,\boldsymbol{k}}(\theta), \quad \bar{g}_{T\boldsymbol{k}}(\theta) = \frac{1}{T} \sum_{t=1}^{T} g_{t\boldsymbol{k}}(\theta),$$

and $\bar{g}_T(\theta)$ be the corresponding column $s \times (2K_T + 1)^2$-vector of $\bar{g}_{T\boldsymbol{k}}(\theta)$. Then we represent the objective function, $Q_T(\theta)$ denoted, in (4) as

$$
\begin{aligned}
Q_T(\theta) &= \sum_{k_0,k_1=-K_T}^{K_T} \left| \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{h}(y_t, \theta) \otimes \gamma_{t,G,\boldsymbol{k}}(\theta) \right|^2 \\
&= \sum_{k_0,k_1=-K_T}^{K_T} \left| \frac{1}{T} \sum_{t=1}^{T} g_{t\boldsymbol{k}}(\theta) \right|^2 = \sum_{k_0,k_1=-K_T}^{K_T} \left[\frac{1}{T} \sum_{t=1}^{T} g_{t\boldsymbol{k}}(\theta) \right]' \overline{\left[\frac{1}{T} \sum_{t=1}^{T} g_{t\boldsymbol{k}}(\theta) \right]} \\
&= \sum_{k_0,k_1=-K_T}^{K_T} \frac{1}{T^2} \left[\sum_{t \neq \tau} g_{t\boldsymbol{k}}(\theta)' \overline{g_{\tau\boldsymbol{k}}(\theta)} + \sum_{t=1}^{T} g_{t\boldsymbol{k}}(\theta)' \overline{g_{t\boldsymbol{k}}(\theta)} \right] \\
&= \sum_{k_0,k_1=-K_T}^{K_T} |\bar{g}_{T\boldsymbol{k}}|^2 = |\bar{g}_T(\theta)|^2 = \bar{g}_T(\theta)' \overline{\bar{g}_T(\theta)}.
\end{aligned}
$$

We define the population versions of $\bar{g}_{T\boldsymbol{k}}(\theta)$ and $\bar{g}_T(\theta)$, respectively, as $\bar{m}_{T\boldsymbol{k}}(\theta) = \mathbb{E}[\bar{g}_{T\boldsymbol{k}}(\theta)]$ and $\bar{m}_T = \mathbb{E}[\bar{g}_T(\theta)]$. For any $\boldsymbol{k}$, the difference between $\bar{g}_{T\boldsymbol{k}}$ and $\bar{m}_{T\boldsymbol{k}}$ multiplied by $\sqrt{T}$ is defined as

$$
\zeta_{T\boldsymbol{k}}(\theta) = \sqrt{T}[\bar{g}_{T\boldsymbol{k}}(\theta) - \bar{m}_{T\boldsymbol{k}}(\theta)],
$$

and the corresponding column $s \times (2K_T + 1)^2$-vector of $\zeta_{T\boldsymbol{k}}(\theta)$ is defined as $\zeta_T(\theta)$. Besides, for each $\boldsymbol{k}$ and t, we define $\xi_{T\boldsymbol{k}}(\theta, t)$ be the difference between $\boldsymbol{h}(y_t, \theta) \otimes \gamma_{t,G,\boldsymbol{k}}(\theta)$ and its expectation $\mathbb{E}[\boldsymbol{h}(y_t, \theta) \otimes \gamma_{t,G,\boldsymbol{k}}(\theta)]$, that is

$$
\xi_{T\boldsymbol{k}}(\theta, t) = \boldsymbol{h}(y_t, \theta) \otimes \gamma_{t,G,\boldsymbol{k}}(\theta) - \mathbb{E}[\boldsymbol{h}(y_t, \theta) \otimes \gamma_{t,G,\boldsymbol{k}}(\theta)],
$$

and the corresponding column $s \times (2K_T + 1)^2$-vector of $\xi_{T\boldsymbol{k}}(\theta)$ is defined as $\xi_T(\theta, t)$.

The following conditions are imposed as Hsu and Kuan (2008) and Han and Phillips (2006):

(A1) The observed data $(y_t, Z_t)'$, $t = 1, 2, \ldots, T$ are independent realization of $(Y, Z)'$.

(A2) For each $\theta \in \Theta$, $\boldsymbol{h}(., \theta)$ is measurable, and for each $y \in \mathbb{R}^r$, $\boldsymbol{h}(y, \cdot)$ is continuous on Θ. Also, θ_o in Θ is the unique solution to $\mathbb{E}[\boldsymbol{h}(Y_t, \theta)|Z_t] = 0$.

(A3) $\mathbb{E}[\boldsymbol{h}(Y_t, \theta)|Z_t]$ is nearly zero for $\theta \neq \theta_o$, $\theta \in \Theta$.

(A4) The eigenvalues of $\mathbb{E}[\xi_T(\theta, t)\xi_T(\theta, t)]'$ are bounded from above for all $\theta \in \Theta$, for all t, and for all K_T and T.

(A5) $\delta_T(\theta) := K_T^{-2}\mathbb{E}[|\zeta_T(\theta)|^2] \to \delta(\theta)$ uniformly in $\theta \in \Theta$.

(A6) There is a sequence of positive numbers c_T such that
$\gamma_T(\theta) := c_T^{-1}|\bar{m}_T(\theta)|^2 \to \gamma(\theta)$ uniformly in $\theta \in \Theta$.

(A7) $\alpha_T := K_T^2/(Tc_T) \to \alpha \in [0, \infty)$ and $Tc_T \to \infty$ as $T \to \infty$.

(A1) and (A2) are standard conditions when θ_o is well identified by the conditional moment restriction; (A3) assumes the situation that the conditional moment restriction is weak; (A4)–(A7), which are also introduced in Han and Phillips (2006), regulate the related sequences while deriving the asymptotics of the proposed estimator. Given (A1), (A5), (A6) and (A7), we have the following decomposition as in Han and Phillips (2006):

$$\begin{aligned}
\mathbb{E}[T^{-1}Q_T(\theta)] &= \mathbb{E}[|\sqrt{T}\bar{g}_T(\theta)|^2] \\
&= \mathbb{E}\left[\sum_{k_0,k_1=-K_T}^{K_T} \frac{1}{T}\left(\sum_{t\neq\tau} g_{t\boldsymbol{k}}(\theta)'\overline{g_{\tau\boldsymbol{k}}(\theta)} + \sum_{t=1}^{T} g_{t\boldsymbol{k}}(\theta)'\overline{g_{t\boldsymbol{k}}(\theta)}\right)\right] \\
&= \mathbb{E}\left[\sum_{k_0,k_1=-K_T}^{K_T} \frac{1}{T}\left(\sum_{t\neq\tau} g_{t\boldsymbol{k}}(\theta)'\overline{g_{\tau\boldsymbol{k}}(\theta)}\right) + \sum_{k_0,k_1=-K_T}^{K_T} \frac{1}{T}\left(\sum_{t=1}^{T} g_{t\boldsymbol{k}}(\theta)'\overline{g_{t\boldsymbol{k}}(\theta)}\right)\right] \\
&= |\sqrt{T}\bar{m}_T(\theta)|^2 + \mathbb{E}[|\zeta_T(\theta)|^2] \\
&= Tc_T\gamma_T(\theta) + K_T^2\delta_T(\theta) \\
&= Tc_T\left[\gamma_T(\theta) + \alpha_T\delta_T(\theta)\right].
\end{aligned}$$

Given this decomposition and the conditions (A5), (A6) and (A7), we have the following property of the proposed estimator.

Theorem 4.1 *(Consistency)*
Given (A1)-(A7), $\hat{\theta}(K_T) \xrightarrow{P} \theta_o$, if either (i) both $\gamma(\theta)$ and $\delta(\theta)$ are minimized at θ_o; or (ii) $\gamma(\theta)$ is minimized uniquely at θ_o and $K_T = o(\sqrt{Tc_T})$ as $T \to \infty$.

Denote $f_T(\theta) = c_T^{-1}\bar{g}_T(\theta)'\bar{g}(\theta)$, $\bar{f}_T(\theta) = \gamma_T(\theta) + \alpha_T\delta_T(\theta)$, and $W_T(\cdot) = f_T(\cdot) - \bar{f}_T\cdot$, then we have a local linear approximation of $W_T(\cdot)$:

$$\begin{aligned}
(Tc_T)^{1/2}\nabla w_T(\theta) =& 2T^{-1/2}\sum_{t=1}^{T} c_T^{-1/2}\nabla\{\bar{m}_T(\theta)'\xi_T(\theta, t)\} \\
&+ \alpha_T^{1/2}K_T^{-1/4}\sum_{\boldsymbol{k}}\nabla\{\zeta_{T\boldsymbol{k}}^2(\theta) - \mathbb{E}[\zeta_{t\boldsymbol{k}}^2(\theta)]\}.
\end{aligned}$$

Then the following two conditions is needed to establish the asymptotic distributions of the proposed estimator.

(A8) $\nabla \bar{f}_T(\cdot)$ permits a linear approximation

$$\nabla \bar{f}_T(\theta) = \nabla \bar{f}_T(\theta_o) + V_T(\bar{\theta}) + (\theta - \theta_o)$$

with $V_T(\theta_T) \to V(\theta_o)$ if $\theta_T \to \theta_o$, where $V(\theta_o)$ is nonsingular.

(A9) At $\theta_o \in \Theta$,

$$2T^{-1/2} \sum_{t=1}^{T} c_T^{-1/2} \nabla \{\bar{m}_T(\theta)' \xi_T(\theta, t)\} \xrightarrow{d} N(0, A).$$

Theorem 4.2 *(Asymptotic Normality)*
Given (A1)–(A9), $K_T = o(\{Tc_T\}^{1/4})$ as $T \to \infty$, then

$$(TC_T)^{1/2}(\hat{\theta}(K_T) - \theta_o) \xrightarrow{d} N(V^{-1}AV^{-1}).$$

Note that the growth rates of K_T in these two Theorems heavily depends on c_T, which measures the strength of signal. These results are different from the well identified case in Hsu and Kuan (2008).

5 Simulations

In this section, I focus on the finite-sample performance of the proposed estimator when the parameter of interest is weak identified. Two simulations are constructed, one is linear model and the other is nonlinear. Our comparison is based on the bias, standard error (SE), and mean squared error (MSE) of these estimators, and percentage changes when $\mathcal{K}_\mathrm{T}$ increases.

5.1 Linear model

The model specification is:

$$Y = \theta_o Z + \epsilon,$$

and $Z = pX + \nu$, with

$$\begin{bmatrix} \epsilon \\ \nu \end{bmatrix} \sim N\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right),$$

where $\theta_o = 1$, $\rho = 0.5$, and $X \sim N(0, 1)$ is independent of ϵ and ν. Sample size $T = 200$. Given this specification, $\mathbb{E}(\epsilon|X) = 0$. In the setting, p is determined by the equality $p^2 = R^2(1 - R^2)^{-1}$, where the theoretical R^2 is equal to 0.0001 and 0.3. When $R^2 = 0.001$, this is actually the weak identified case, and $R^2 = 0.3$ represents the well identified case. In this experiment, we compare the proposed estimator with 2-Step-least-square (2SLS) estimator. The simulation results are collected in Table 1.

In Table 1, we observe that the bias and SE of 2SLS estimator are very large when θ_o is weakly identified in the case with $R^2 = 0.0001$, and the proposed estimator $\hat{\theta}(K_T)$ outperforms the 2SLS estimator very much. When θ_o is well identified ($R^2 = 0.3$) the performance of 2SLS estimator is better, but the difference between these two estimators is not quite large.

5.2 Nonlinear model

The model specification is:

$$Y = \theta_o^2 Z + \theta_o Z^2 + \epsilon,$$

and $Z = pX + \nu$, with

$$\begin{bmatrix} \epsilon \\ \nu \end{bmatrix} \sim N\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right),$$

where $\theta_o = 1.25$, $\rho = 0.9$, and $X \sim N(0, 1)$ is independent of ϵ and ν. Sample size $T = 500$. Given this specification, $\mathbb{E}(\epsilon|X) = 0$. In this experiment, we compare the proposed estimator with nonlinear-least-square (NLS) estimator. The simulation results are collected in Table 2.

Based on this result, we observe that the proposed estimator outperforms the NLS estimator when $R^2 = 0.0001$. Even in the case that $R^2 = 0.3$, the proposed estimator still has a smaller bias than the NLS estimator.

6 Concluding Remarks

This project is concerned with consistent estimation of conditional moment restrictions when the parameters of interests are weakly identified. There are some contributions in this project. First, I show that weak identification may be overstated in conditional moment restrictions by an example. Second, I clearly address the problem of weak identification based on the conditional moment restrictions, and construct a consistent estimation approach. To this end, I follow the approach proposed in Hsu and

Kuan (2008). I systematically generate the sets of infinitely many weak unconditional moment restrictions, by using a set of generically comprehensively revealing functions of conditioning variables. Based on these sets of unconditional moment restrictions, I construct a GMM type estimator by using Fourier analysis, cf. Hsu and Kuan (2008). Besides, I establish the corresponding theoretical properties based on the framework in Han and Phillips (2006). I also examine finite sample performance of the proposed estimators, via extensive Monte Carlo simulations. In sum, this project reviews the weak identification problems in the literature and theoretically extend Hsu and Kuan's (2008) work to the case that the parameters are weak identified in conditional moment restrictions. Besides, this work also shows how to find a large number of weak unconditional moments and guides us to construct a consistent estimator in practice.

Moreover, based ont the framework of this project, we may construct jackknife type estimator as

$$\hat{\theta}_J(K_T) = \operatorname*{argmin}_{\theta\in\Theta} \sum_{k_0,k_1=-K_T}^{K_T} \frac{1}{T^2}\left[\sum_{t\neq\tau} g_{t\boldsymbol{k}}(\theta)'\overline{g_{\tau\boldsymbol{k}}(\theta)}\right]. \tag{5}$$

This jackknife estimator may have better finite sample properties since it removes the second order bias, cf. Chao and Swanson (2005) and Newey and Windmeijer(2008). In this framework, the jackknife estimator (5) is new in the literature. All properties of this jackknife estimator are left to future research.

References

Angrist, J. D. and Kruger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, **106**, 974–1014.

Bierens, H. J. (1990). A consistent conditional moment test of functional form, *Econometrica*, **58**, 1443–1458.

Chao, J. C., and N. R. Swanson (2005). Consistent estimation with large number of weak instruments, *Econometrica*, **73**, 1673–1692.

Donald, S. G., G. W. Imbens, and W. K. Newey (2003). Empirical likelihood estimation and consistent tests with conditional moment restrictions, *Journal of Econometrics*, **117**, 55-93.

Hahn, J. and J. Hausman (2003). Weak instruments: diagnosis and cures in empirical econometrics, *American Economic Review*, **93**, 118–125.

Han, C. and P. C. B. Phillips (2006). GMM with many moment conditions, *Econometrica*, **74**, 147–192.

Hsu, S-H. and C-M. Kuan (2008). Estimation of conditional moment restrictions without assuming parameter identifiability in the implied unconditional moments, working paper.

Newey, W. K. and F. Windemeijer (2008). GMM with many weak moment conditions, working paper.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments, *Econometrica* **65**, 557–586.

Stinchcombe, M. B. and H. White (1998). Consistent specification testing with nuisance parameters present only under the alternative, *Econometric Theory*, **14**, 295–325.

Stock, J. H. and J. H. Wright (2000). GMM with weak identification, *Econometrica*, **68**, 1055–1096.

Stock, J. H., Wright, J. H. and M. Yogo (2002). A survey of weak instruments and weak identification in generalized method of moments, *Journal of Business and Economic Statistics*, **20**, 518–529.

Wright, J. H. (2003). Detecting lack of identification in GMM, *Econometric Theory*, **19**, 322-330.

Table 1: The performance of $\hat{\theta}(K_T)$ with various K_T.

	$R^2 = 0.0001$					
K_T	Bias	Bias(+%)	SE	SE(+%)	MSE	MSE(+%)
1	0.51557	0.00000	2.47658	0.00000	6.39801	0.00000
2	0.51348	-0.40552	2.42898	-1.92188	6.16242	-3.68229
3	0.51278	-0.13611	2.41875	-0.42132	6.11210	-0.81653
4	0.51243	-0.06786	2.41534	-0.14059	6.09531	-0.27473
5	0.51222	-0.04077	2.41395	-0.05775	6.08836	-0.11400
6	0.51208	-0.02726	2.41331	-0.02656	6.08512	-0.05318
7	0.51198	-0.01954	2.41300	-0.01294	6.08351	-0.02644
8	0.51190	-0.01472	2.41284	-0.00634	6.08269	-0.01341
9	0.51185	-0.01148	2.41277	-0.00291	6.08230	-0.00656
10	0.51180	-0.00922	2.41275	-0.00106	6.08212	-0.00281
15	0.51166	-0.02783	2.41284	0.00365	6.08240	0.00458
20	0.51158	-0.01404	2.41297	0.00565	6.08299	0.00960
2SLS	1.83313		77.23494		5967.40257	
	$R^2 = 0.3$					
K_T	Bias	Bias(+%)	SE	SE(+%)	MSE	MSE(+%)
1	-0.02930	0.00000	0.28524	0.00000	0.08220	0.00000
2	-0.02763	-5.71355	0.27111	-4.95383	0.07425	-9.67729
3	-0.02711	-1.87063	0.26622	-1.80301	0.07159	-3.57488
4	-0.02686	-0.90181	0.26378	-0.91626	0.07029	-1.82383
5	-0.02672	-0.52515	0.26233	-0.55077	0.06952	-1.09798
6	-0.02663	-0.34205	0.26137	-0.36646	0.06901	-0.73108
7	-0.02657	-0.23991	0.26069	-0.26099	0.06865	-0.52087
8	-0.02652	-0.17733	0.26018	-0.19515	0.06838	-0.38956
9	-0.02648	-0.13629	0.25978	-0.15135	0.06818	-0.30217
10	-0.02646	-0.10797	0.25947	-0.12077	0.06801	-0.24113
15	-0.02637	-0.31750	0.25854	-0.35981	0.06752	-0.71747
20	-0.02633	-0.15530	0.25807	-0.17873	0.06728	-0.35666
2SLS	-0.00453		0.11004		0.01213	

Table 2: The performance of $\hat{\theta}(K_T)$ with various K_T.

	$R^2 = 0.0001$					
K_T	Bias	Bias(+%)	SE	SE(+%)	MSE	MSE(+%)
1	-0.00381	0.00000	0.04919	0.00000	0.00243	0.00000
2	-0.00389	1.97901	0.04959	0.81099	0.00247	1.64270
3	-0.00392	0.92784	0.04977	0.36435	0.00249	0.73695
4	-0.00394	0.52738	0.04987	0.20356	0.00250	0.41155
5	-0.00396	0.33774	0.04993	0.12916	0.00251	0.26108
6	-0.00397	0.23407	0.04998	0.08900	0.00251	0.17989
7	-0.00397	0.17151	0.05001	0.06495	0.00252	0.13128
8	-0.00398	0.13096	0.05004	0.04945	0.00252	0.09996
9	-0.00398	0.10322	0.05006	0.03889	0.00252	0.07861
10	-0.00399	0.08342	0.05007	0.03138	0.00252	0.06342
15	-0.00400	0.25490	0.05012	0.09553	0.00253	0.19317
20	-0.00400	0.12978	0.05014	0.04849	0.00253	0.09803
NLS	0.23839		0.01130		0.05696	
	$R^2 = 0.3$					
K_T	Bias	Bias(+%)	SE	SE(+%)	MSE	MSE(+%)
1	-0.00007	0.00000	0.02296	0.00000	0.00053	0.00000
2	-0.00007	-6.98614	0.02270	-1.13857	0.00052	-2.26428
3	-0.00006	-3.41847	0.02260	-0.46138	0.00051	-0.92068
4	-0.00006	-1.96558	0.02254	-0.24547	0.00051	-0.49037
5	-0.00006	-1.26232	0.02251	-0.15136	0.00051	-0.30251
6	-0.00006	-0.87480	0.02249	-0.10235	0.00051	-0.20461
7	-0.00006	-0.64030	0.02247	-0.07371	0.00050	-0.14738
8	-0.00006	-0.48823	0.02246	-0.05557	0.00050	-0.11111
9	-0.00006	-0.38426	0.02245	-0.04337	0.00050	-0.08672
10	-0.00006	-0.31011	0.02244	-0.03477	0.00050	-0.06954
15	-0.00006	-0.94020	0.02242	-0.10462	0.00050	-0.20915
20	-0.00006	-0.47776	0.02240	-0.05246	0.00050	-0.10490
NLS	0.01696		0.01168		0.00042	

成果自評

In this project, I obtain the following results.

1. Demonstrating that the weak identification problem may be overstated in conditional moment restriction models by some examples.
2. Constructing an GMM type estimator.
3. Establishing the corresponding asymptotics under Han and Phillips's(2006) framework.
4. Simulation results of the proposed estimators.
5. Providing a possible extension based on the approach used in this project.

In this project, some contributions to this topic are as follows.

- In this project, we clearly learn the problem of weak identification based on the conditional moment restrictions.
- This project extends the work of Hsu and Kuan (2008) to the case that parameters of interest is weakly identified. And the properties of the proposed estimator are also uncovered.
- When addressing the problem of weak identification based on the conditional moment restrictions directly, this project shows how to find a large number of weak unconditional moments and guides us to construct a consistent estimator in practice.