LOCAL IDENTIFICATION OF NONPARAMETRIC AND SEMIPARAMETRIC MODELS

by

Xiahong Chen, Victor Chernozhukov, Sokbae Lee and Whitney Newey

COWLES FOUNDATION PAPER NO. 1415

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

2014

http://cowles.econ.yale.edu/

ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 82, No. 2 (March, 2014), 785–809

LOCAL IDENTIFICATION OF NONPARAMETRIC AND SEMIPARAMETRIC MODELS

XIAOHONG CHEN
Yale University, New Haven, CT 06520, U.S.A.

VICTOR CHERNOZHUKOV
MIT, Cambridge, MA 02142, U.S.A.

SOKBAE LEE
Seoul National University, Seoul, 151-742, Respublic of Korea

WHITNEY K. NEWEY
MIT, Cambridge, MA 02142, U.S.A.

Econometrica, Vol. 82, No. 2 (March, 2014), 785–809

LOCAL IDENTIFICATION OF NONPARAMETRIC AND SEMIPARAMETRIC MODELS

BY XIAOHONG CHEN, VICTOR CHERNOZHUKOV, SOKBAE LEE, AND WHITNEY K. NEWEY[1]

In parametric, nonlinear structural models, a classical sufficient condition for local identification, like Fisher (1966) and Rothenberg (1971), is that the vector of moment conditions is differentiable at the true parameter with full rank derivative matrix. We derive an analogous result for the *nonparametric*, nonlinear structural models, establishing conditions under which an infinite dimensional analog of the full rank condition is sufficient for local identification. Importantly, we show that additional conditions are often needed in nonlinear, nonparametric models to avoid nonlinearities overwhelming linear effects. We give restrictions on a neighborhood of the true value that are sufficient for local identification. We apply these results to obtain new, primitive identification conditions in several important models, including nonseparable quantile instrumental variable (IV) models and semiparametric consumption-based asset pricing models.

KEYWORDS: Identification, local identification, nonparametric models, asset pricing.

1. INTRODUCTION

THERE ARE MANY IMPORTANT MODELS IN ECONOMETRICS that give rise to conditional moment restrictions. These restrictions often take the form

$$\mathrm{E}\big[\rho(Y, X, \alpha_0)|W\big] = 0,$$

where $\rho(Y, X, \alpha)$ has a known functional form but α_0 is unknown. Parametric models (i.e., models when α_0 is finite dimensional) of this form are well known from the work of Hansen (1982), Chamberlain (1987), and others. Nonparametric versions (i.e., models when α_0 is infinite dimensional) are motivated by the desire to relax functional form restrictions. Identification and estimation of linear nonparametric conditional moment models have been studied by Newey and Powell (2003), Hall and Horowitz (2005), Blundell, Chen, and Kristensen (2007), Darolles, Fan, Florens, and Renault (2011), and others.

The purpose of this paper is to derive identification conditions for α_0 when ρ may be nonlinear in α and for other nonlinear nonparametric models. Nonlinear models are important. They include models with conditional quantile restrictions, as discussed in Chernozhukov and Hansen (2005) and

[1]The National Science Foundation (Grants SES-0838161 and SES-1132399), European Research Council (ERC-2009-StG-240910-ROMETA), and the National Research Foundation of Korea (NRF-2011-327-B00073) provided financial support for this paper. Helpful comments were provided by a co-editor, three anonymous referees, D. Andrews, D. Chetverikov, K. Evdokimov, J. P. Florens, L. Hansen, J. Heckman, S. Hoderlein, H. Ichimura, T. Komarova, O. Linton, A. Santos, and participants in seminars at June 2010 Cowles Conference, June 2010 Beijing Conference, June 2010 LSE Conference, and December 2010 EC2 conference.

 DOI: 10.3982/ECTA9988

Chernozhukov, Imbens, and Newey (2007), and various economic structural and semiparametric models, as further discussed below. In this paper, we focus on conditions for local identification of these models. It may be possible to extend these results to provide global identification conditions.

In parametric models, there are easily interpretable rank conditions for local identification, as shown in Fisher (1966) and Rothenberg (1971). We give a pair of conditions that are sufficient for parametric local identification from solving a set of equations. They are (a) pointwise differentiability at the true value, and (b) the rank of the derivative matrix is equal to the dimension of the parameter α_0. We find that the nonparametric case is different. Differentiability and the nonparametric version of the rank condition may not be sufficient for local identification. We suggest a restriction on the neighborhood that does give local identification, via a link between curvature and an identification set. We also give more primitive conditions for Hilbert spaces, that include interesting econometric examples. In addition, we consider semiparametric models, providing conditions for identification of a finite dimensional Euclidean parameter. These conditions are based on "partialling out" the nonparametric part and allow for identification of the parametric part even when the nonparametric part is not identified.

The usefulness of these conditions is illustrated by three examples. One example gives primitive conditions for local identification of the nonparametric endogenous quantile models, where primitive identification conditions had only been given previously for discrete regressors. Another example gives sufficient conditions for local identification of a semiparametric consumption capital asset pricing model. The third example, given in the Supplemental Material (Chen, Chernozhukov, Lee, and Newey (2014)), gives conditions for local identification of a semiparametric index model with endogeneity, including conditions for identification of parametric components when nonparametric components need not be identified.

In relation to previous literature, in some cases the nonparametric rank condition is a local version of identification conditions for linear conditional moment restriction models that were considered in Newey and Powell (2003). Chernozhukov, Imbens, and Newey (2007) also suggested differentiability and a rank condition for local identification but did not recognize the need for additional restrictions on the neighborhood. Florens and Sbai (2010) gave local identification conditions for games, but their conditions do not apply to the kind of conditional moment restrictions that arise in instrumental variable settings and are a primary subject of this paper.

Section 2 presents general nonparametric local identification results and relates them to sufficient conditions for identification in parametric models. Section 3 gives more primitive conditions for Hilbert spaces and applies them to the nonparametric endogenous quantile model. Section 4 provides conditions for identification in semiparametric models, and Section 5 discusses the semiparametric asset pricing example. Section 6 briefly concludes. Appendices A and B give proofs for Sections 2 and 3. The proofs for Sections 4 and 5 and some additional identification results are given in the Supplemental Material (Chen et al. (2014)).

2. NONPARAMETRIC MODELS

2.1. *The Setting and Definition of Local Identification*

To help explain the nonparametric results and give them context, we give a brief description of sufficient conditions for local identification in parametric models. Let α be a $p \times 1$ vector of parameters and $m(\alpha)$ a $J \times 1$ vector of functions with $m(\alpha_0) = 0$ for the true value α_0. Also let $|\cdot|$ denote the Euclidean norm in either $\mathbb{R}^p$ or $\mathbb{R}^J$ depending on the context. We say that α_0 is locally identified if there is a neighborhood of α_0 such that $m(\alpha) \neq 0$ for all $\alpha \neq \alpha_0$ in the neighborhood. Let m' denote the derivative of $m(\alpha)$ at α_0 when it exists. Sufficient conditions for local identification can be stated as follows:

If $m(\alpha)$ is differentiable at α_0 and $\operatorname{rank}(m') = p$, *then α_0 is locally identified.*

This statement is proved in Appendix A. Here, the sufficient conditions for parametric local identification are pointwise differentiability at the true value α_0 and the rank of the derivative equal to the number of parameters.

In order to extend these conditions to the nonparametric case, we need to modify the notation and introduce structure for infinite dimensional spaces. Let α denote a function with true value α_0 and $m(\alpha)$ a function of α with $m(\alpha_0) = 0$. Conditional moment restrictions are an important example where $\rho(Y, X, \alpha)$ is a finite dimensional residual vector depending on an unknown function α and $m(\alpha) = \mathrm{E}[\rho(Y, X, \alpha)|W]$. We impose some mathematical structure by assuming that $\alpha \in \mathcal{A}$, a Banach space with norm $\|\cdot\|_{\mathcal{A}}$, and $m(\alpha) \in \mathcal{B}$, a Banach space with a norm $\|\cdot\|_{\mathcal{B}}$, that is, $m : \mathcal{A} \mapsto \mathcal{B}$. The restriction of the model is that $\|m(\alpha_0)\|_{\mathcal{B}} = 0$. The notion of local identification we consider is the following:

DEFINITION: α_0 is locally identified on $\mathcal{N} \subseteq \mathcal{A}$ if $\|m(\alpha)\|_{\mathcal{B}} > 0$ for all $\alpha \in \mathcal{N}$, with $\alpha \neq \alpha_0$.

This local identification concept is more general than the one introduced by Chernozhukov, Imbens, and Newey (2007). Note that local identification is defined on a set $\mathcal{N}$ in $\mathcal{A}$. Often there exists an $\varepsilon > 0$ such that $\mathcal{N}$ is a subset of an open ball

$$\mathcal{N}_\varepsilon \equiv \{\alpha \in \mathcal{A} : \|\alpha - \alpha_0\|_{\mathcal{A}} < \varepsilon\}.$$

It turns out that it may be necessary for $\mathcal{N}$ to be strictly smaller than an open ball $\mathcal{N}_\varepsilon$ in $\mathcal{A}$, as discussed below.

2.2. *Local Identification via Full Rank Conditions*

The nonparametric version of the derivative will be a bounded (i.e., continuous) linear map $m' : \mathcal{A} \mapsto \mathcal{B}$. Under the conditions we give, m' will be a Gâteaux derivative at α_0, that can be calculated as

$$m'h = \frac{\partial}{\partial t} m(\alpha_0 + th)\bigg|_{t=0} \tag{2.1}$$

for $h \in \mathcal{A}$ and t a scalar. Sometimes we also require that, for any $\delta > 0$, there is $\varepsilon > 0$ with

$$\frac{\|m(\alpha) - m(\alpha_0) - m'(\alpha - \alpha_0)\|_{\mathcal{B}}}{\|\alpha - \alpha_0\|_{\mathcal{A}}} < \delta$$

for all $\alpha \in \mathcal{N}_\varepsilon$. This is Fréchet differentiability of $m(\alpha)$ at α_0 (which implies that the linear map $m' : \mathcal{A} \mapsto \mathcal{B}$ is continuous). Fréchet differentiability of estimators that are functionals of the empirical distribution is known to be too strong, but is typically satisfied in local identification analysis, as shown by our examples.

In parametric models, the rank condition is equivalent to the null space of the derivative matrix being zero. The analogous nonparametric condition is that the null space of the linear map m' is zero, as follows:

ASSUMPTION 1—Rank Condition: *There is a set $\mathcal{N}'$ such that $\|m'(\alpha - \alpha_0)\|_{\mathcal{B}} > 0$ for all $\alpha \in \mathcal{N}'$ with $\alpha \neq \alpha_0$.*

This condition is familiar from identification of a linear conditional moment model where $Y = \alpha_0(X) + U$ and $\mathrm{E}[U|W] = 0$. Here $\rho(Y, X, \alpha) = Y - \alpha(X)$, so that $m(\alpha) = \mathrm{E}[Y - \alpha(X)|W]$ and $m'h = -\mathrm{E}[h(X)|W]$. In this case, Assumption 1 requires that $\mathrm{E}[\alpha(X) - \alpha_0(X)|W] \neq 0$ for any $\alpha \in \mathcal{N}'$ with $\alpha - \alpha_0 \neq 0$. For $\mathcal{N}' = \mathcal{A}$, this is the completeness condition discussed in Newey and Powell (2003). Andrews (2011) has recently shown that if X and W are continuously distributed, there are at least as many instruments in W as regressors in X, and the conditional distribution of X given W is unrestricted (except for a mild regularity condition), then the completeness condition holds generically, in a sense defined in that paper. In Section 3, we also give a genericity result for a different range of models. For this reason, we think of Assumption 1 with $\mathcal{N}' = \mathcal{A}$ as a weak condition when there are as many continuous instruments W as the endogenous regressors X, just as it is in a parametric linear instrumental variables model with unrestricted reduced form. It is also an even weaker condition if some conditions are imposed on the deviations, so in the statement of Assumption 1 we allow it to hold only on $\mathcal{N}' \subset \mathcal{A}$. For example, if we restrict $\alpha - \alpha_0$ to be a bounded function of X, then, in linear conditional moment restriction models, Assumption 1 only requires that the conditional distribution of X given W be bounded complete, which is known to hold for even more distributions than does completeness. This makes Assumption 1 even more plausible in models where α_0 is restricted to be bounded, such as in Blundell, Chen, and Kristensen (2007). See, for example, Mattner (1993), Chernozhukov and Hansen (2005), D'Haultfoeuille (2011), and Andrews (2011) for discussions of completeness and bounded completeness.

Fréchet differentiability and the rank condition are not sufficient for local identification in an open ball $\mathcal{N}_\varepsilon$ around α_0, as we further explain below. One condition that can be added to obtain local identification is that $m' : \mathcal{A} \mapsto \mathcal{B}$ is onto.

THEOREM 1: *If $m(\alpha)$ is Fréchet differentiable at α_0, the rank condition is satisfied on $\mathcal{N}' = \mathcal{N}_\varepsilon$ for some $\varepsilon > 0$, and $m' : \mathcal{A} \mapsto \mathcal{B}$ is onto, then α_0 is locally identified on $\mathcal{N}_{\tilde{\varepsilon}}$ for some $\tilde{\varepsilon}$ with $0 < \tilde{\varepsilon} \leq \varepsilon$.*

This result extends previous nonparametric local identification results by only requiring pointwise Fréchet differentiability at α_0, rather than continuous Fréchet differentiability in a neighborhood of α_0. This extension may be helpful for showing local identification in nonparametric models, because conditions for pointwise Fréchet differentiability are simpler than for continuous differentiability in nonparametric models.

Unfortunately, the assumption that m' is onto is too strong for many econometric models, including many nonparametric conditional moment restrictions. An onto m' implies that m' has a continuous inverse, by the Banach Inverse Theorem (Luenberger (1969, p. 149)). The inverse of m' may not be continuous for nonparametric conditional moment restrictions, as discussed in Newey and Powell (2003). Indeed, the discontinuity of the inverse of m' is a now well known ill-posed inverse problem that has received much attention in the econometrics literature; for example, see the survey of Carrasco, Florens, and Renault (2007). Thus, in many important econometric models, Theorem 1 cannot be applied to obtain local identification.

It turns out that α_0 may not be locally identified on any open ball in ill-posed inverse problems, as we show in an example below. The problem is that, for infinite dimensional spaces, $m'(\alpha - \alpha_0)$ may be small when $\alpha - \alpha_0$ is large. Consequently, the effect of nonlinearity, which is related to the size of $\alpha - \alpha_0$, may overwhelm the identifying effect of nonzero $m'(\alpha - \alpha_0)$, resulting in $m(\alpha)$ being zero for α close to α_0.

We approach this problem by restricting the deviations $\alpha - \alpha_0$ to be small when $m'(\alpha - \alpha_0)$ is small. The restrictions on the deviations will be related to the nonlinearity of $m(\alpha)$ via the following condition:

ASSUMPTION 2: *There are $L \geq 0$, $r \geq 1$, and a set $\mathcal{N}''$ such that, for all $\alpha \in \mathcal{N}''$,*

$$\left\| m(\alpha) - m(\alpha_0) - m'(\alpha - \alpha_0) \right\|_{\mathcal{B}} \leq L \|\alpha - \alpha_0\|_{\mathcal{A}}^r.$$

This condition is general. It includes the linear case where $\mathcal{N}'' = \mathcal{A}$ and $L = 0$. It also includes Fréchet differentiability, where $r = 1$, L is any positive number, and $\mathcal{N}'' = \mathcal{N}_\varepsilon$ for any sufficiently small $\varepsilon > 0$. Cases with $r > 1$ are analogous to Hölder continuity of the derivative in finite dimensional spaces, with $r = 2$ corresponding to twice continuous Fréchet differentiability. We would only have $r > 2$ when the second derivative is zero. This condition applies to many interesting examples, as we will show in the rest of the paper. The term $L\|\alpha - \alpha_0\|_{\mathcal{A}}^r$ represents a magnitude of nonlinearity that is allowed for $\alpha \in \mathcal{N}''$. The following result uses Assumption 2 to specify restrictions on α that are sufficient for local identification.

THEOREM 2: *If Assumption 2 is satisfied, then α_0 is locally identified on $\mathcal{N} = \mathcal{N}'' \cap \mathcal{N}'''$ with $\mathcal{N}''' = \{\alpha : \|m'(\alpha - \alpha_0)\|_{\mathcal{B}} > L\|\alpha - \alpha_0\|_{\mathcal{A}}^r\}$.*

The strict inequality in $\mathcal{N}'''$ is important for the result. It does exclude α_0 from $\mathcal{N}$, but that works because local identification specifies what happens when $\alpha \neq \alpha_0$. This result includes the linear case, where $L = 0$, $\mathcal{N}'' = \mathcal{A}$, and $\mathcal{N} = \mathcal{N}''' = \mathcal{N}'$. It also includes nonlinear cases where only Fréchet differentiability is imposed, with $r = 1$ and L equal to any positive constant. In that case, $\mathcal{N}'' = \mathcal{N}_\varepsilon$ for some ε small enough and $\alpha \in \mathcal{N}'''$ restricts $\alpha - \alpha_0$ to a set where the inverse of m' is continuous by requiring that $\|m'(\alpha - \alpha_0)\|_{\mathcal{B}} > L\|\alpha - \alpha_0\|_{\mathcal{A}}$. In general, by $L\|\alpha - \alpha_0\|_{\mathcal{A}}^r \geq 0$, we have $\mathcal{N}''' \subseteq \mathcal{N}'$ for $\mathcal{N}'$ from Assumption 1, so the rank condition is imposed by restricting attention to the $\mathcal{N}$ of Theorem 2. Here, the rank condition is still important, since if it is not satisfied on some interesting set $\mathcal{N}'$, Theorem 2 cannot give local identification on an interesting set $\mathcal{N}$.

Theorem 2 forges a link between the curvature of $m(\alpha)$ as in Assumption 2 and the identification set $\mathcal{N}$. An example is a scalar α and twice continuously differentiable $m(\alpha)$ with bounded second derivative. Here, Assumption 2 will be satisfied with $r = 2$, $L = \sup_\alpha |d^2 m(\alpha)/d\alpha^2|/2$, and $\mathcal{N}''$ equal to the real line, where $|\cdot|$ denotes the absolute value. Assumption 1 will be satisfied with $\mathcal{N}'$ equal to the real line as long as $m' = dm(\alpha_0)/d\alpha$ is nonzero. Then $\mathcal{N}''' = \{\alpha : |\alpha - \alpha_0| < L^{-1}|m'|\}$. Here, $L^{-1}|m'|$ is the minimum distance α must go from α_0 before $m(\alpha)$ can "bend back" to zero. In nonparametric models, $\mathcal{N}'''$ will be an analogous set.

When $r = 1$, the set $\mathcal{N}'''$ will be a linear cone with vertex at α_0, which means that if $\alpha \in \mathcal{N}'''$, then so is $\lambda\alpha + (1 - \lambda)\alpha_0$ for $\lambda > 0$. In general, $\mathcal{N}'''$ is not convex, so it is not a convex cone. For $r > 1$, the set $\mathcal{N}'''$ is not a cone, although it is star shaped around α_0, meaning that for any $\alpha \in \mathcal{N}'''$, we have $\lambda\alpha + (1 - \lambda)\alpha_0 \in \mathcal{N}'''$ for $0 < \lambda \leq 1$.

Also, if $r > 1$, then for any $L > 0$ and $1 \leq r' < r$, there is $\delta > 0$ such that

$$\mathcal{N}_\delta \cap \left\{\alpha : \left\|m'(\alpha - \alpha_0)\right\|_{\mathcal{B}} > L\|\alpha - \alpha_0\|_{\mathcal{A}}^{r'}\right\} \subseteq \mathcal{N}_\delta \cap \mathcal{N}'''.$$

In this sense, $\alpha \in \mathcal{N}'''$ as assumed in Theorem 2 is less restrictive the larger is r, that is, the local identification neighborhoods of Theorem 2 are "richer" the larger is r.

2.3. *Discussion of Assumptions 1 and 2*

Restricting the set of α to be smaller than an open ball can be necessary for local identification in nonparametric models, as we now show in an example. Suppose $\alpha = (\alpha_1, \alpha_2, \dots)$ is a sequence of real numbers. Let $(p_1, p_2, \dots)$ be probabilities, $p_j > 0$, $\sum_{j=1}^{\infty} p_j = 1$. Let $f(x)$ be a twice continuously differentiable function of a scalar x that is bounded with bounded second derivative. Suppose $f(x) = 0$ if and only if $x \in \{0, 1\}$ and $df(0)/dx = 1$. Let $m(\alpha) = (f(\alpha_1), f(\alpha_2), \dots)$ also be a sequence with $\|m(\alpha)\|_{\mathcal{B}} = (\sum_{j=1}^{\infty} p_j f(\alpha_j)^2)^{1/2}$. Then for $\|\alpha\|_{\mathcal{A}} = (\sum_{j=1}^{\infty} p_j \alpha_j^4)^{1/4}$, the function $m(\alpha)$ will be Fréchet differentiable at $\alpha_0 = 0$, with $m'h = h$. A fourth moment norm for α, rather than a second moment norm, is needed to make $m(\alpha)$ Fréchet differentiable under the second moment norm for $m(\alpha)$. Here, the map m' is not onto, even though it is the identity, because the norm on $\mathcal{A}$ is stronger than the norm on $\mathcal{B}$.

In this example, the value $\alpha_0 = 0$ is not locally identified by the equation $m(\alpha) = 0$ on any open ball in the norm $\|\alpha - \alpha_0\|_{\mathcal{A}}$. To show this result, consider α^k which has zeros in the first k positions and a 1 everywhere else, that is, $\alpha^k = (0, \dots, 0, 1, 1, \dots)$. Then $m(\alpha^k) = 0$, and for $\Delta^k = \sum_{j=k+1}^{\infty} p_j \longrightarrow 0$, we have $\|\alpha^k - \alpha_0\|_{\mathcal{A}} = (\sum_{j=1}^{\infty} p_j [\alpha_j^k]^4)^{1/4} = (\Delta^k)^{1/4} \longrightarrow 0$. Thus, we have constructed a sequence of α^k not equal to α_0 such that $m(\alpha^k) = 0$ and $\|\alpha^k - \alpha_0\|_{\mathcal{A}} \longrightarrow 0$.

We can easily describe the set $\mathcal{N}$ of Theorem 2 in this example, on which $\alpha_0 = 0$ will be locally identified. By the second derivative of f being bounded, Assumption 2 is satisfied with $\mathcal{N}'' = \mathcal{A}$, $r = 2$, and $L = \sup_a |\partial^2 f(a)/\partial a^2|/2$, where $L \geq 1$ by the fact that $f'(0) = 1$ and $f(0) = f(1) = 0$ (an expansion gives $0 = f(1) = 1 + 2^{-1} \partial^2 f(\bar{a})/\partial a^2$ for $0 \leq \bar{a} \leq 1$). Then,

$$\mathcal{N} = \left\{ \alpha = (\alpha_1, \alpha_2, \dots) : \left(\sum_{j=1}^{\infty} p_j \alpha_j^2 \right)^{1/2} > L \left(\sum_{j=1}^{\infty} p_j \alpha_j^4 \right)^{1/2} \right\}.$$

The sequence $(\alpha^k)_{k=1}^{\infty}$ given above will not be included in this set because $L \geq 1$. A simple subset of $\mathcal{N}$ (on which α_0 is locally identified) is $\{\alpha = (\alpha_1, \alpha_2, \dots) : |\alpha_j| < L^{-1} (j = 1, 2, \dots)\}$.

It is important to note that Theorems 1 and 2 provide sufficient, and not necessary, conditions for local identification. In fact, the conditions of Theorems 1 and 2 are sufficient for

$$\|m(\alpha) - m'(\alpha - \alpha_0)\|_{\mathcal{B}} \neq \|m'(\alpha - \alpha_0)\|_{\mathcal{B}}, \tag{2.2}$$

which implies $m(\alpha) \neq 0$, to hold on $\mathcal{N}$. The set where equation (2.2) holds may be larger than the set $\mathcal{N}$ of Theorems 1 or 2. We have focused on the set $\mathcal{N}$

of Theorems 1 or 2 because those conditions and the associated locally identified set $\mathcal{N}$ are relatively easy to interpret. Additional identification results related to equation (2.2) are discussed in the Supplemental Material (Chen et al. (2014)), under the heading of tangential cone conditions.

Assumption 1 may not be needed for identification in nonlinear models, although local identification is complicated in the absence of Assumption 1. Conditions may involve nonzero higher order derivatives. Such results for parametric models were discussed by, for example, Sargan (1983). Here, we focus on models where Assumption 1 is satisfied.

3. LOCAL IDENTIFICATION IN HILBERT SPACES

3.1. *Full Rank Condition in Hilbert Spaces*

The restrictions imposed on α in Theorem 2 are not very transparent. In Hilbert spaces, it is possible to give more interpretable conditions based on a lower bound for $\|m'(\alpha-\alpha_0)\|_{\mathcal{B}}^2$. Let $\langle\cdot,\cdot\rangle$ denote the inner product for a Hilbert space.

ASSUMPTION 3: $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ *and* $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ *are separable Hilbert spaces and either* (a) *there is a set* $\mathcal{N}'$, *an orthonormal basis* $\{\phi_1, \phi_2, \ldots\} \subseteq \mathcal{A}$, *and a bounded, positive sequence* $(\mu_1, \mu_2, \ldots)$ *such that, for all* $\alpha \in \mathcal{N}'$,

$$\left\|m'(\alpha-\alpha_0)\right\|_{\mathcal{B}}^2 \geq \sum_{j=1}^{\infty} \mu_j^2 \langle \alpha-\alpha_0, \phi_j\rangle^2;$$

or (b) m' *is a compact linear operator with positive singular values* $(\mu_1, \mu_2, \ldots)$.

The hypothesis in (b) that m' is a compact operator is a mild one when m' is a conditional expectation. Recall that an operator $m: \mathcal{A} \mapsto \mathcal{B}$ is compact if and only if it is continuous and maps bounded sets in $\mathcal{A}$ into relatively compact sets in $\mathcal{B}$. Under very mild conditions, $m(\alpha) = \mathrm{E}[\alpha(X)|W]$ is compact; see Zimmer (1990, Chapter 3), Kress (1999, Section 2.4), and Carrasco, Florens, and Renault (2007) for a variety of sufficient conditions. When m' in (b) is compact, there is an orthonormal basis $\{\phi_j : j = 1, \ldots\}$ for $\mathcal{A}$ with

$$\left\|m'(\alpha-\alpha_0)\right\|_{\mathcal{B}}^2 = \sum_{j=1}^{\infty} \mu_j^2 \langle \alpha-\alpha_0, \phi_j\rangle^2,$$

where μ_j^2 are the eigenvalues and ϕ_j the eigenfunctions of the operator m'^*m', so that condition (a) is satisfied, where m'^* denotes the adjoint of m'. The assumption that the singular values are all positive implies the rank condition holds for $\mathcal{N}' = \mathcal{A}$. Part (a) differs from part (b) by imposing a lower bound on $\|m'(\alpha-\alpha_0)\|_{\mathcal{B}}^2$ only over a subset $\mathcal{N}'$ of $\mathcal{A}$ and by allowing the basis $\{\phi_j\}$ to be different from the eigenfunction basis of the operator $m'^{*}m'$. In principle, this allows us to impose restrictions on $\alpha - \alpha_0$, like boundedness and smoothness, which could help Assumption 3(a) to hold. For similar assumptions in estimation context, see, for example, Chen and Reiß (2011) and Chen and Pouzo (2012).

It turns out that there is a precise sense in which the rank condition is satisfied for most data generating processes, if it is satisfied for one, in the Hilbert space environment here. The following result is related to but different than Andrews (2011). In this sense, the rank condition turns out to be generic. Let $\mathcal{A}$ and $\mathcal{B}$ be separable Hilbert spaces, and $\mathcal{N}' \subseteq \mathcal{A}$. Suppose that there exists at least one compact linear operator: $\mathcal{K} : \mathcal{A} \mapsto \mathcal{B}$ which is injective, that is, $\mathcal{K}\delta = 0$ for $\delta \in \mathcal{A}$ if and only if $\delta = 0$. This is an infinite dimensional analog of the order condition, that, for example, rules out $\mathcal{B}$ having smaller finite dimension than $\mathcal{A}$ (e.g., having fewer instruments than right-hand side endogenous variables in a linear regression model). The operator $m' : \mathcal{N}' \mapsto \mathcal{B}$ is generated by the nature as follows:

1. The nature selects a countable orthonormal basis $\{\phi_j\}$ of cardinality $N \leq \infty$ in $\mathcal{A}$ and an orthonormal set $\{\varphi_j\}$ of equal cardinality in $\mathcal{B}$.
2. The nature samples a bounded sequence of real numbers $\{\lambda_j\}$ according to a probability measure η whose each marginal is dominated by the Lebesgue measure on $\mathbb{R}$, namely, $\text{Leb}(A) = 0$ implies $\eta(\{\lambda_j \in A\}) = 0$ for any measurable $A \subset \mathbb{R}$ for each j.

Then the nature sets, for some scalar number $\kappa > 0$, and every $\delta \in \mathcal{N}'$,

$$m'\delta = \kappa\left(\sum_{j=0}^{N} \lambda_j \langle \phi_j, \delta \rangle \varphi_j\right). \tag{3.1}$$

This operator is properly defined on $\mathcal{N}' := \{\delta \in \mathcal{A} : m'\delta \in \mathcal{B}\}$.

LEMMA 3: (1) *In the absence of further restrictions on m', the algorithms obeying conditions* 1 *and* 2 *exist*. (2) *If m' is generated by any algorithm that obeys conditions* 1 *and* 2, *then the probability that m' is not injective over $\mathcal{N}'$ is zero*, *namely*, $\Pr_\eta\{\exists \delta \in \mathcal{N}' : \delta \neq 0$ *and* $m'\delta = 0\} = 0$. *Moreover*, *Assumption* 3 *holds with $\mu_j = |\kappa\lambda_j|$ with probability* 1 *under* η.

Genericity of the rank condition is further discussed in the Supplemental Material (Chen et al. (2014)).

3.2. *Local Identification in Hilbert Spaces*

In what follows, let $b_j = \langle \alpha - \alpha_0, \phi_j \rangle$, $j = 1, 2, \dots$ denote the Fourier coefficients for $\alpha - \alpha_0$, so that $\alpha = \alpha_0 + \sum_{j=1}^{\infty} b_j \phi_j$. Under Assumptions 2 and 3, we can characterize an identified set in terms of the Fourier coefficients.

THEOREM 4: *If Assumptions 2 and 3 are satisfied, then α_0 is locally identified on $\mathcal{N} = \mathcal{N}'' \cap \mathcal{N}'''$, where*

$$\mathcal{N}''' = \left\{\alpha = \alpha_0 + \sum_{j=1}^{\infty} b_j \phi_j : \sum_{j=1}^{\infty} \mu_j^2 b_j^2 > L^2 \left(\sum_{j=1}^{\infty} b_j^2\right)^r\right\}.$$

When $r = 1$, it is necessary for $\alpha \in \mathcal{N}'''$ that the Fourier coefficients b_j where μ_j^2 is small not be too large relative to the Fourier coefficients where μ_j^2 is large. In particular, when $r = 1$, any $\alpha \neq \alpha_0$ with $b_j = 0$ for all j with $\mu_j > L$ will not be an element of $\mathcal{N}'''$. When $r > 1$, we can use the Hölder inequality to obtain a sufficient condition for $\alpha \in \mathcal{N}'''$ that is easier to interpret.

COROLLARY 5: *If Assumptions 2 and 3 are satisfied, with $L > 0, r > 1$, then α_0 is locally identified on $\mathcal{N} = \mathcal{N}'' \cap \mathcal{N}'''$ where $\mathcal{N}''' = \{\alpha = \alpha_0 + \sum_{j=1}^{\infty} b_j \phi_j : \sum_{j=1}^{\infty} \mu_j^{-2/(r-1)} b_j^2 < L^{-2/(r-1)}\}$.*

For α to be in the $\mathcal{N}'''$ of Corollary 5, the Fourier coefficients b_j must vanish faster than $\mu_j^{1/(r-1)}$ as j grows. In particular, a sufficient condition for $\alpha \in \mathcal{N}'''$ is that $|b_j| < (\mu_j/L)^{1/(r-1)} c_j$ for any positive sequence c_j with $\sum_{j=1}^{\infty} c_j^2 = 1$. These bounds on b_j correspond to a hyperrectangle, while the $\mathcal{N}'''$ in Corollary 5 corresponds to an ellipsoid. The bounds on b_j shrink as L increases, corresponding to a smaller local identification set when more nonlinearity is allowed. Also, it is well known that, at least in certain environments, imposing bounds on Fourier coefficients corresponds to imposing smoothness conditions, like existence of derivatives; see, for example, Kress (1999, Chapter 8). In that sense, the identification set in Corollary 5 imposes smoothness conditions on the deviations of α from the truth α_0.

The bound imposed in $\mathcal{N}'''$ of Corollary 5 is a "source condition" under Assumption 3(b) and is similar to conditions used by Florens, Johannes, and Van Bellegem (2011) and others. Under Assumption 3(a), it is similar to norms in generalized Hilbert scales; for example, see Engl, Hanke, and Neubauer (1996) and Chen and Reiß (2011). Our Assumption 3(a) or 3(b) is imposed on deviations $\alpha - \alpha_0$, while the above references all impose on true function α_0 itself as well as on the parameter space, hence on the deviations.

3.3. *A Quantile IV Example*

To illustrate the results of this section, we consider an endogenous quantile example where $0 < \tau < 1$ is a scalar,

$$\rho(Y, X, \alpha) = 1\big(Y \le \alpha(X)\big) - \tau,$$

$\mathcal{A} = \{\alpha(\cdot): \mathrm{E}[\alpha(X)^2] < \infty\}$, and $\mathcal{B} = \{a(\cdot): \mathrm{E}[a(W)^2] < \infty\}$, with the usual Hilbert spaces of mean squared integrable random variables. Here, we have

$$m(\alpha) = \mathrm{E}\big[1\big(Y \leq \alpha(X)\big)|W\big] - \tau.$$

Let $f_Y(y|X, W)$ denote the conditional probability density function (p.d.f.) of Y given X and W, $f_X(x|W)$ the conditional p.d.f. of X given W, and $f(x)$ the marginal p.d.f. of X.

THEOREM 6: *If $f_Y(y|X, W)$ is continuously differentiable in y with $|df_Y(y|X, W)/dy| \leq L_1$, $f_X(x|W) \leq L_2 f(x)$, and $m'h = \mathrm{E}[f_Y(\alpha_0(X)|X, W)h(X)|W]$ satisfies Assumption 3, then α_0 is locally identified on*

$$\mathcal{N} = \left\{\alpha = \alpha_0 + \sum_{j=1}^{\infty} b_j\phi_j \in \mathcal{A}: \sum_{j=1}^{\infty} b_j^2/\mu_j^2 < (L_1L_2)^{-2}\right\}.$$

This result gives a precise link between a neighborhood on which α_0 is locally identified and the bounds L_1 and L_2. Assumption 3(b) will hold under primitive conditions for m' to be complete, that are given by Chernozhukov, Imbens, and Newey (2007). Theorem 6 corrects Theorem 3.2 of Chernozhukov, Imbens, and Newey (2007) by adding the bound on $\sum_{j=1}^{\infty} b_j^2/\mu_j^2$. It also gives primitive conditions for local identification for general X, while Chernozhukov and Hansen (2005) only gave primitive conditions for identification when X is discrete. Horowitz and Lee (2007) imposed analogous conditions in their paper on convergence rates of nonparametric endogenous quantile estimators, but assumed identification.

4. SEMIPARAMETRIC MODELS

In this section, we consider local identification in possibly nonlinear semiparametric models, where α can be decomposed into a $p \times 1$ dimensional parameter vector β and nonparametric component g, so that $\alpha = (\beta, g)$. Let $|\cdot|$ denote the Euclidean norm for β and assume $g \in \mathcal{G}$, where $\mathcal{G}$ is a Banach space with norm $\|\cdot\|_{\mathcal{G}}$, such as a Hilbert space. We focus here on a conditional moment restriction model

$$\mathrm{E}\big[\rho(Y, X, \beta_0, g_0)|W\big] = 0,$$

where $\rho(y, x, \beta, g)$ is a $J \times 1$ vector of residuals. Here, $m(\alpha) = \mathrm{E}[\rho(Y, X, \beta, g)|W]$ will be considered as an element of the Hilbert space $\mathcal{B}$ of $J \times 1$ random vectors with inner product

$$\langle a, b\rangle = \mathrm{E}\big[a(W)^T b(W)\big].$$

The differential $m'(\alpha - \alpha_0)$ can be expressed as

$$m'(\alpha - \alpha_0) = m'_\beta(\beta - \beta_0) + m'_g(g - g_0),$$

where m'_β is the derivative of $m(\beta, g_0) = \mathrm{E}[\rho(Y, X, \beta, g_0)|W]$ with respect to β at β_0 and m'_g is the Gâteaux derivative of $m(\beta_0, g)$ with respect to g at g_0. To give conditions for local identification of β_0 in the presence of the nonparametric component g, it is helpful to partial out g. Let $\overline{\mathcal{M}}$ be the closure of the linear span $\mathcal{M}$ of $m'_g(g - g_0)$ for $g \in \mathcal{N}'_g$, where $\mathcal{N}'_g$ will be specified below. In general, $\overline{\mathcal{M}} \neq \mathcal{M}$ because the linear operator m'_g need not have closed range (like m' onto, a closed range would also imply a continuous inverse, by the Banach inverse theorem). For the kth unit vector e_k $(k = 1, \dots, p)$, let

$$\zeta_k^* = \arg\min_{\zeta \in \overline{\mathcal{M}}} \mathrm{E}\big[\big\{m'_\beta(W)e_k - \zeta(W)\big\}^T \big\{m'_\beta(W)e_k - \zeta(W)\big\}\big],$$

which exists and is unique by standard Hilbert space results; for example, see Luenberger (1969). Define Π to be the $p \times p$ matrix with

$$\begin{aligned}\Pi_{jk} &:= \mathrm{E}\big[\big\{m'_\beta(W)e_j - \zeta_j^*(W)\big\}^T \big\{m'_\beta(W)e_k - \zeta_k^*(W)\big\}\big] \\ &\quad (j, k = 1, \dots, p).\end{aligned}$$

The following condition is important for local identification of β_0.

ASSUMPTION 4: *$m' : \mathbb{R}^p \times \mathcal{N}'_g \longrightarrow \mathcal{B}$ is linear and bounded, and Π is nonsingular.*

This assumption is similar to those first used by Chamberlain (1992) to establish the possibility of estimating parametric components at root-n rate in semiparametric moment condition problems; see also Ai and Chen (2003) and Chen and Pouzo (2009). In the local identification analysis considered here, it leads to local identification of β_0 without identification of g when $m(\beta_0, g)$ is linear in g. It allows us to separate conditions for identification of β_0 from conditions for identification of g. Note that the parameter β may be identified even when Π is singular, but that case is more complicated, as discussed at the end of Section 2, and we do not analyze this case.

The following condition controls the behavior of the derivative with respect to β:

ASSUMPTION 5: *For every $\varepsilon > 0$, there is a neighborhood B of β_0 and a set $\mathcal{N}_g^\beta$ such that, for all $g \in \mathcal{N}_g^\beta$ with probability* 1, $\mathrm{E}[\rho(Y, X, \beta, g)|W]$ *is continuously* *differentiable in* β *on* B *and*

$$\sup_{g\in\mathcal{N}_g^\beta}\Big(\mathrm{E}\Big[\sup_{\beta\in B}\big|\partial\mathrm{E}\big[\rho(Y,X,\beta,g)|W\big]/\partial\beta - \partial\mathrm{E}\big[\rho(Y,X,\beta_0,g_0)|W\big]/\partial\beta\big|^2\Big]\Big)^{1/2}<\varepsilon.$$

It turns out that Assumptions 4 and 5 will be sufficient for local identification of β_0 when $m(\beta_0, g)$ is linear in g, that is, for $m(\beta, g) = 0$ to imply $\beta = \beta_0$ when (β, g) is in some neighborhood of (β_0, g_0). This works because Assumption 4 partials out the effect of unknown g on local identification of β_0.

THEOREM 7: *If Assumptions* 4 *and* 5 *are satisfied and* $m(\beta_0, g)$ *is linear in* g, *then there is an* $\varepsilon > 0$ *such that, for* B *and* $\mathcal{N}_g^\beta$ *from Assumption* 5 *and* $\mathcal{N}_g'$ *from Assumption* 4, β_0 *is locally identified for* $\mathcal{N} = B\times(\mathcal{N}_g'\cap\mathcal{N}_g^\beta)$. *If, in addition, Assumption* 1 *is satisfied for* m_g' *and* $\mathcal{N}_g'\cap\mathcal{N}_g^\beta$ *replacing* m' *and* $\mathcal{N}'$, *then* $\alpha_0 = (\beta_0, g_0)$ *is locally identified for* $\mathcal{N}$.

This result is more general than Florens, Johannes, and Van Bellegem (2012) and Santos (2011) since it allows for nonlinearities in β, and dependence on g of the partial derivatives $\partial\mathrm{E}[\rho(Y,X,\beta,g)|W]/\partial\beta$. When the partial derivatives $\partial\mathrm{E}[\rho(Y,X,\beta,g)|W]/\partial\beta$ do not depend on g, then Assumption 5 could be satisfied with $\mathcal{N}_g' = \mathcal{G}$, and Theorem 7 could then imply local identification of β_0 in some neighborhood of β_0 only.

For semiparametric models that are nonlinear in g, we can give local identification results based on Theorem 2 or the more specific conditions of Theorem 4 and Corollary 5. For brevity, we give just a result based on Theorem 2.

THEOREM 8: *If Assumptions* 4 *and* 5 *are satisfied and* $m(\beta_0, g)$ *satisfies Assumption* 2 *with* $\mathcal{N}'' = \mathcal{N}_g''$, *then there is an* $\varepsilon > 0$ *such that, for* B *and* $\mathcal{N}_g^\beta$ *from Assumption* 5, $\mathcal{N}_g'$ *from Assumption* 4, *and*

$$\mathcal{N}_g''' = \big\{g : \big\|m_g'(g-g_0)\big\|_{\mathcal{B}} > \varepsilon^{-1}L\|g-g_0\|_{\mathcal{A}}^r\big\},$$

it is the case that $\alpha_0 = (\beta_0, g_0)$ *is locally identified for* $\mathcal{N} = B\times(\mathcal{N}_g^\beta\cap\mathcal{N}_g'\cap\mathcal{N}_g''\cap\mathcal{N}_g''')$.

One example that highlights the role of semiparametric models in reducing the need for numbers of instruments is a single index example where $Y = g_0(X_1 + X_2^T\beta_0) + U$ with $\mathrm{E}[U|W] = 0$. This example is considered in the Supplemental Material (Chen et al. (2014)). There we find that two instruments can suffice for local identification of β_0 and g_0 even when there are more than two endogenous variables.

5. SEMIPARAMETRIC CCAPM

Consumption capital asset pricing models (CCAPM) provide interesting examples of nonparametric and semiparametric moment restrictions; see Gallant and Tauchen (1989), Newey and Powell (1988), Hansen, Heaton, Lee, and Roussanov (2007), Chen and Ludvigson (2009), and others. In this section, we apply our general theorems to develop new results on identification of a particular semiparametric specification of marginal utility of consumption. Our results could easily be extended to other specifications, and so are of independent interest.

To describe the model, let C_t denote consumption level at time t and $c_t \equiv C_t/C_{t-1}$ be consumption growth. Suppose that the marginal utility of consumption at time t is given by

$$\mathrm{MU}_t = C_t^{-\gamma_0} g_0(C_t/C_{t-1}) = C_t^{-\gamma_0} g_0(c_t),$$

where $g_0(c)$ is an unknown positive function. For this model, the intertemporal marginal rate of substitution is

$$\delta_0 \mathrm{MU}_{t+1}/\mathrm{MU}_t = \delta_0 c_{t+1}^{-\gamma_0} g_0(c_{t+1})/g_0(c_t),$$

where $0 < \delta_0 \leq 1$ is the rate of time preference. Let $R_{t+1} = (R_{t+1,1}, \ldots, R_{t+1,J})^T$ be a $J \times 1$ vector of gross asset returns. A semiparametric CCAPM equation is then given by

$$\mathrm{E}\left[R_{t+1}\delta_0 c_{t+1}^{-\gamma_0}\{g_0(c_{t+1})/g_0(c_t)\}|W_t\right] = e, \tag{5.1}$$

where e is a $J \times 1$ vector of ones, and $W_t \equiv (Z_t^T, c_t)^T$ is a vector of random variables observed by the agent at time t, with Z_t not a measurable function of c_t. This corresponds to an external habit formation model with only one lag, a special case of Chen and Ludvigson (2009). As emphasized in Cochrane (2005), habit formation models can help explain the high risk premia embedded in asset prices. We focus here on consumption growth $c_t = C_t/C_{t-1}$ to circumvent the potential nonstationarity of the level of consumption (see Hall (1978)), as has long been done in this literature, for example, Hansen and Singleton (1982).

From economic theory it is known that, under complete markets, there is a unique intertemporal marginal rate of substitution that solves equation (5.1), when R_t is allowed to vary over all possible vectors of asset returns. Of course, that does not guarantee a unique solution for a fixed vector of returns R_t. Note, though, that the semiparametric model does impose restrictions on the marginal rate of substitution that should be helpful for identification. We show how these restrictions lead to local identification of this model via the results of Section 4.

This model can be formulated as a semiparametric conditional moment restriction by letting $Y = (R_{t+1}^T, c_{t+1}, c_t)^T$, $\beta = (\delta, \gamma)^T$, $W = W_t = (Z_t^T, c_t)^T$, and

$$\rho(Y, \beta, g) = R_{t+1}\delta c_{t+1}^{-\gamma} g(c_{t+1}) - g(c_t)e. \tag{5.2}$$

Then, multiplying equation (5.1) through by $g_0(c_t)$ gives the conditional moment restriction $\mathrm{E}[\rho(Y, \beta_0, g_0)|W] = 0$. Let $A_t = R_{t+1}\delta_0 c_{t+1}^{-\gamma_0}$. The nonparametric rank condition (Assumption 1 for g) will be uniqueness, up to scale, of the solution g_0 of

$$\mathrm{E}\big[A_t g(c_{t+1})|W_t\big] = g(c_t)e. \tag{5.3}$$

This equation differs from the linear nonparametric IV restriction where the function $g_0(X)$ would solve $\mathrm{E}[Y|W] = \mathrm{E}[g(X)|W]$. That equation is an integral equation of the first kind, while equation (5.3) is a homogeneous integral equation of the second kind. The rank condition for this second kind equation is that the null space of the operator $\mathrm{E}[A_t g(c_{t+1})|W_t] - g(c_t)e$ is one-dimensional, which is different than the completeness condition for first kind equations. This example illustrates that the rank condition of Assumption 1 need not be equivalent to completeness of a conditional expectation. Escanciano and Hoderlein (2010) and Lewbel, Linton, and Srisuma (2012) have previously shown how homogeneous integral equations of the second kind arise in CCAPM models, though their models and identification results are different than those given here, as further discussed below.

Let $X_t = (1/\delta_0, -\ln(c_{t+1}))^T$. Then, differentiating inside the integral, as allowed under regularity conditions given below, and applying the Gateaux derivative calculation gives

$$m'_\beta(W) = \mathrm{E}\big[A_t g_0(c_{t+1}) X_t^T|W_t\big], \quad m'_g g = \mathrm{E}\big[A_t g(c_{t+1})|W_t\big] - g(c_t)e.$$

When $\mathrm{E}[A_t g(c_{t+1})|W_t]$ is a compact operator, as holds under conditions described below, it follows from the theory of integral equations of the second kind (e.g., Kress (1999, Theorem 3.2)) that the set of nonparametric directions $\mathcal{M}$ will be closed, that is,

$$\overline{\mathcal{M}} = \mathcal{M} = \big\{\mathrm{E}\big[A_t g(c_{t+1})|W_t\big] - g(c_t)e : \|g\|_{\mathcal{G}} < \infty\big\},$$

where we will specify $\|g\|_{\mathcal{G}}$ below. Let Π be the two-dimensional second moment matrix Π of the residuals from the projection of each column of m'_β on $\overline{\mathcal{M}}$, as described in Section 4. Then nonsingularity of Π leads to local identification of β_0 via Theorem 7.

To give a precise result, let Δ be any finite positive number,

$$D_t = 1 + \big(1 + |R_{t+1}|\big)\big[2 + \big|\ln(c_{t+1})\big|^2\big] \sup_{\gamma \in [\gamma_0 - \Delta, \gamma_0 + \Delta]} c_{t+1}^{-\gamma},$$

$$\mathcal{G} = \Big\{g : \|g\|_{\mathcal{G}} \equiv \sqrt{\mathrm{E}\big[\mathrm{E}\big[D_t^2 | W_t\big] g(c_{t+1})^2\big]} < \infty\Big\}.$$

The following assumption imposes some regularity conditions.

ASSUMPTION 6: *(R_t^T, c_t, Z_t^T) is strictly stationary*, $\mathrm{E}[D_t^2] < \infty$; $0 < \delta_0 \leq 1$, $\|g_0\|_{\mathcal{G}} < \infty$.

The following result applies Theorem 7 to this CCAPM.

THEOREM 9: *Consider equation* (5.3). *Suppose that Assumption 6 is satisfied. Then the linear mapping $m' : \mathbb{R}^2 \times \mathcal{G} \longrightarrow \mathcal{B}$ is bounded, and if Π is nonsingular, there is a neighborhood B of β_0 and $\varepsilon > 0$ such that, for $\mathcal{N}_g^\beta = \{g : \|g - g_0\|_{\mathcal{G}} < \varepsilon\}$, β_0 is locally identified for $\mathcal{N} = B \times \mathcal{N}_g^\beta$. If, in addition, $m'_g(g - g_0) \neq 0$ for all $g \neq g_0$ and $g \in \mathcal{N}_g^\beta$, then (β_0, g_0) is locally identified for $\mathcal{N} = B \times \mathcal{N}_g^\beta$.*

Primitive conditions for nonsingularity of Π and for $m'_g(g - g_0) \neq 0$ when $g \neq g_0$ are needed to make this result interesting. It turns out that some completeness conditions suffice, as shown by the following result. Let $\tilde{W}_t = (w(Z_t), c_t)$ for some measurable function $w(Z_t)$ of Z_t, and $f_{c,\tilde{W}}(c, \tilde{w})$ denote the joint p.d.f. of $(c_{t+1}, \tilde{W}_t)$, $f_c(c)$ and $f_{\tilde{W}}(\tilde{w})$ the marginal p.d.f.s of c_{t+1} and $\tilde{W}_t$, respectively.

THEOREM 10: *Consider equation* (5.3). *Suppose that Assumption 6 is satisfied*, $\Pr(g_0(c_t) = 0) = 0$, *for some $w(Z_t)$ and $\tilde{W}_t = (w(Z_t), c_t)$, $(c_{t+1}, \tilde{W}_t)$ is continuously distributed and there is some j with $A_{tj} = \delta_0 R_{t+1,j} c_{t+1}^{-\gamma_0}$ satisfying*

$$\mathrm{E}\big[A_{tj}^2 f_c(c_{t+1})^{-1} f_{\tilde{W}}(\tilde{W}_t)^{-1} f_{c,\tilde{W}}(c_{t+1}, \tilde{W}_t)\big] < \infty. \tag{5.4}$$

Then (a) *if* $\mathrm{E}[A_{tj}\tilde{h}(c_{t+1}, c_t)|\tilde{W}_t] = 0$ *implies $\tilde{h}(c_{t+1}, c_t) = 0$ a.s. and $a(c_{t+1}) + b(c_t) = 0$ for $c_t \in \mathcal{C}$ with* $\Pr(\mathcal{C}) > 0$ *implies $a(c_{t+1})$ is constant, then Π is nonsingular*; (b) *if $g_0 \in \mathcal{G}_{\bar{c}} \equiv \{g \in \mathcal{G} : g(\bar{c}) \neq 0\}$ for some $\bar{c}$ and* $\mathrm{E}[A_{tj}h(c_{t+1})|w(Z_t), c_t = \bar{c}] = 0$ *with $h \in \mathcal{G}_{\bar{c}}$ implies $h(c_{t+1}) = 0$ a.s., then g_0 is the unique solution to* $\mathrm{E}[A_t g(c_{t+1})|W_t] = g(c_t)$ *up to scale*.

Equation (5.4) implies $\mathrm{E}[A_{tj} g(c_{t+1})|\tilde{W}_t]$ is a Hilbert–Schmidt integral operator and hence compact. Analogous conditions could be imposed to ensure that $\mathcal{M}$ is closed. The sufficient conditions for nonsingularity of Π involve completeness of the conditional expectation $\mathrm{E}[A_{tj} h(c_{t+1}, c_t)|\tilde{W}_t]$ and a stronger version of a measurably separable condition from Florens, Mouchart, and Rolin (1990). As previously noted, sufficient conditions for completeness can be found in Newey and Powell (2003) and Andrews (2011) and completeness is generic in the sense of Andrews (2011) and Lemma 3. A simple sufficient condition for the measurably separable hypothesis is that the support of (c_{t+1}, c_t) is $\Re_+^2$, where $\Re_+ = [0, \infty)$.

Condition (b) is weaker than condition (a). Condition (b) turns out to imply global identification of β_0 and g_0 (up to scale) if $g_0(c)$ is bounded, and bounded away from zero. Because we focus on applying the results of Section 4, we reserve this result to Theorem S5 in the Supplemental Material (Chen et al. (2014)). Even with global identification, the result of Theorem 10(a) is of interest, because nonsingularity of Π will be necessary for γ_0 to be estimable at a root-n rate. The identification result for γ_0 in Theorem S5 involves large and small values of consumption growth, and so amounts to identification at infinity, which may not lead to root-n consistent estimation; for example, see Chamberlain (1986).

A different approach to the nonparametric rank condition, that does not require any instrument $w(Z_t)$ in addition to c_t, can be based on positivity of $g_0(c)$. The linear operator $\mathrm{E}[A_{tj}g(c_{t+1})|c_t]$ and $g(c)$ will be infinite dimensional (functional) analogs of a positive matrix and a positive eigenvector, respectively, by equation (5.3). The Perron–Frobenius theorem says that there is a unique positive eigenvalue and eigenvector (up to scale) pair for a positive matrix. A functional analog, based on Krein and Rutman (1950), gives uniqueness of $g_0(c)$, as well as of the discount factor δ_0. To describe this result, let $r(c, s) = \mathrm{E}[R_{t+1,j}|c_{t+1} = s, c_t = c]$, $f(s, c)$ be the joint p.d.f. of (c_{t+1}, c_t), $f(c)$ the marginal p.d.f. of c_t at c, and $K(c, s) = r(c, s)s^{-\gamma_0}f(s, c)/[f(s)f(c)]$. Then the equation $\mathrm{E}[A_{tj}g(c_{t+1})|c_t] = g(c_t)$ can be written as

$$\delta \int K(c, s)g(s)f(s)\,ds = g(c), \tag{5.5}$$

for $\delta = \delta_0 \in (0, 1]$. Here, the matrix analogy is clear, with $K(c, s)f(s)$ being like a positive matrix, $g(c)$ an eigenvector, and δ^{-1} an eigenvalue.

THEOREM 11: *Suppose that* $(R_{t,j}, c_t)$ *is strictly stationary,* $f(s, c) > 0$ *and* $r(c, s) > 0$ *almost everywhere, and* $\int\int K(c, s)^2 f(c)f(s)\,dc\,ds < \infty$. *Then equation* (5.5) *has a unique positive solution* (δ_0, g_0) *in the sense that* $\delta_0 > 0$, $g_0 > 0$ *almost everywhere and* $\mathrm{E}[g_0(c_t)^2] = 1$.

The conditions of this result include $r(c, s) > 0$, which will hold if $R_{t+1,j}$ is a positive risk free rate. Under square-integrability of K, we obtain global identification of the pair (δ_0, g_0). The uniqueness of $g_0(c)$ in the conclusion of this result implies the nonparametric rank condition. Note that by iterated expectation and inclusion of $R_{t+1,j}$ in R_{t+1}, any solution to equation (5.3) must also satisfy equation (5.5). Thus Theorem 11 implies that g_0 is the unique solution to (5.3). Theorem 11 actually gives more, identification of the discount factor given identification of γ_0.

Previously, Escanciano and Hoderlein (2010) and Lewbel, Linton, and Srisuma (2012) considered nonparametric identification of marginal utility in consumption level, $\mathrm{MU}(C_t)$, by solving the homogeneous integral equation of the second kind:

$$\mathrm{E}\big[R_{t+1,j}\delta_0 \mathrm{MU}(C_{t+1})|C_t\big] = \mathrm{MU}(C_t).$$

In particular, Escanciano and Hoderlein (2010) gave an insightful identification result for the discount factor δ and the marginal utility MU based on the positivity of marginal utility, but assuming that $\mathrm{MU}(c)$ is bounded and continuous, including at zero and infinity if the support of C_t is $[0, \infty)$. Lewbel, Linton, and Srisuma (2012) used a genericity argument for identification of $\mathrm{MU}(C_t)$. While we also use the positivity of the function g, we identify the "correction term" g in the habit-formation model, rather than the marginal utility MU. Furthermore, we base our result on a Krein and Rutman (1950) theorem, a particular functional analog of Perron–Frobenius, that allows us to avoid restricting $g(c)$ to be bounded or continuous. Functional Perron–Frobenius theory has been extensively used by Hansen and Scheinkman (2009, 2012) and Hansen (2012) in their research on the long-run risk and dynamic valuations in a general Markov environment in which their valuation operators may not be compact. Recently, a revised version of Escanciano and Hoderlein (2010) was made available, that uses Krein–Rutman (1950) type results to consider cases without $\mathrm{MU}(c)$ being continuous and bounded.[2] Independently, Christensen (2013) applied a version of Krein and Rutman (1950) result for identification of the first principal eigenfunction of a class of operators closely related to those in Hansen (2012).

The models considered here will generally be highly overidentified. We have analyzed identification using only a single asset return $R_{t+1,j}$. The presence of more asset returns in equation (5.1) provides overidentifying restrictions. Also, in Theorem 10 we only use a function $w(Z_t)$ of the available instrumental variables Z_t in addition to c_t. The additional information in Z_t may provide overidentifying restrictions. These sources of overidentification are familiar in CCAPM models. See, for example, Hansen and Singleton (1982) and Chen and Ludvigson (2009).

[2]A revised version of Escanciano and Hoderlein (2010) became available to us and online in June, 2013. A first revised version of our paper including Theorem 11 was submitted in September, 2012.

6. CONCLUSION

We provide sufficient conditions for local identification for a general class of semiparametric and nonparametric conditional moment restriction models. We give new identification conditions for several important models that illustrate the usefulness of our general results. In particular, we provide primitive conditions for local identification in nonseparable quantile IV models, single-index IV models, and semiparametric consumption-based asset pricing models.

APPENDIX A: PROOFS FOR SECTION 2

A.1. *Proof of Parametric Result*

By $\operatorname{rank}(m') = p$, the nonnegative square root η of the smallest eigenvalue η^2 of $(m')^T m'$ is positive and $|m'h| \geq \eta|h|$ for $h \in \mathbb{R}^p$. Also, by the definition of the derivative, there is $\varepsilon > 0$ such that $|m(\alpha) - m(\alpha_0) - m'(\alpha - \alpha_0)|/|\alpha - \alpha_0| < \eta$ for all $|\alpha - \alpha_0| < \varepsilon$ with $\alpha \neq \alpha_0$. Then

$$\begin{aligned}
\text{(A.1)} \qquad & \frac{|m(\alpha) - m'(\alpha - \alpha_0)|}{|m'(\alpha - \alpha_0)|} \\
&= \frac{|m(\alpha) - m(\alpha_0) - m'(\alpha - \alpha_0)|}{|\alpha - \alpha_0|} \frac{|\alpha - \alpha_0|}{|m'(\alpha - \alpha_0)|} \\
&< \frac{\eta}{\eta} = 1.
\end{aligned}$$

This inequality implies $m(\alpha) \neq 0$, so α_0 is locally identified on $\{\alpha : |\alpha - \alpha_0| < \varepsilon\}$. *Q.E.D.*

A.2. *Proof of Theorem 1*

If $m'h = m'\tilde{h}$ for $h \neq \tilde{h}$, then, for any $\lambda > 0$, we have $m'\bar{h} = 0$ for $\bar{h} = \lambda(h - \tilde{h}) \neq 0$. For λ small enough, $\bar{h}$ would be in any open ball around zero. Therefore, Assumption 1 holding on an open ball containing α_0 implies that m' is invertible. By m' onto and the Banach Inverse Theorem (Luenberger (1969, p. 149)), it follows that $(m')^{-1}$ is continuous. Since any continuous linear map is bounded, it follows that there exists $\eta > 0$ such that $\|m'(\alpha - \alpha_0)\|_{\mathcal{B}} \geq \eta\|\alpha - \alpha_0\|_{\mathcal{A}}$ for all $\alpha \in \mathcal{A}$.

Next, by Fréchet differentiability at α_0, there exists an open ball $\mathcal{N}_\varepsilon$ centered at α_0 such that, for all $\alpha \in \mathcal{N}_\varepsilon$, $\alpha \neq \alpha_0$,

$$\text{(A.2)} \qquad \frac{\|m(\alpha) - m(\alpha_0) - m'(\alpha - \alpha_0)\|_{\mathcal{B}}}{\|\alpha - \alpha_0\|_{\mathcal{A}}} < \eta.$$

Therefore, at all such $\alpha \neq \alpha_0$,

$$\begin{aligned}
&\frac{\|m(\alpha) - m'(\alpha - \alpha_0)\|_{\mathcal{B}}}{\|m'(\alpha - \alpha_0)\|_{\mathcal{B}}} \\
&\quad = \frac{\|m(\alpha) - m(\alpha_0) - m'(\alpha - \alpha_0)\|_{\mathcal{B}}}{\|\alpha - \alpha_0\|_{\mathcal{A}}} \frac{\|\alpha - \alpha_0\|_{\mathcal{A}}}{\|m'(\alpha - \alpha_0)\|_{\mathcal{B}}} \\
&\quad < \eta/\eta = 1.
\end{aligned}$$

Therefore, as in the proof of the parametric result above, $m(\alpha) \neq 0$ for all $\alpha \in \mathcal{N}_\varepsilon$ with $\alpha \neq \alpha_0$. *Q.E.D.*

A.3. *Proof of Theorem 2*

Consider $\alpha \in \mathcal{N}$ with $\alpha \neq \alpha_0$. Then

$$\begin{aligned}
\frac{\|m(\alpha) - m'(\alpha - \alpha_0)\|_{\mathcal{B}}}{\|m'(\alpha - \alpha_0)\|_{\mathcal{B}}} &= \frac{\|m(\alpha) - m(\alpha_0) - m'(\alpha - \alpha_0)\|_{\mathcal{B}}}{\|m'(\alpha - \alpha_0)\|_{\mathcal{B}}} \\
&\leq \frac{L\|\alpha - \alpha_0\|_{\mathcal{A}}^r}{\|m'(\alpha - \alpha_0)\|_{\mathcal{B}}} < 1.
\end{aligned}$$

The conclusion follows as in the proof of Theorem 1. *Q.E.D.*

APPENDIX B: PROOFS FOR SECTION 3

B.1. *Proof of Lemma 3*

By assumptions, there exists a compact, injective operator $\mathcal{K}: \mathcal{A} \mapsto \mathcal{B}$. By Theorem 15.16 in Kress (1999), $\mathcal{K}$ admits a singular value decomposition:

$$\mathcal{K}\delta = \sum_{j=0}^{N} \mu_j \langle \phi_j, \delta \rangle \varphi_j,$$

where $\{\phi_j\}$ is an orthonormal subset of $\mathcal{A}$, either finite or countably infinite, with cardinality $N \leq \infty$, $\{\varphi_j\}$ is an orthonormal subset of $\mathcal{B}$ of equal cardinality, and $(\mu_j)_{j=1}^{\infty}$ is bounded. Since $\|\mathcal{K}\delta\|_{\mathcal{B}}^2 = \sum_{j=0}^{N} \mu_j^2 \langle \phi_j, \delta \rangle^2$, injectivity of $\mathcal{K}$ requires that $\{\phi_j\}$ must be an orthonormal basis in $\mathcal{A}$ and $\mu_j \neq 0$ for all j. Therefore, step 1 is always feasible by using these $\{\phi_j\}$ and $\{\varphi_j\}$ in the construction. The order of eigenvectors in these sets need not be preserved and could be arbitrary. Step 2 is also feasible by using a product of Lebesgue-dominated measures on a bounded subset of $\mathbb{R}$ to define a measure over $\mathbb{R}^N$, or, more generally, using any construction of measure on $\mathbb{R}^N$ from finite dimensional measures obeying Kolmogorov's consistency conditions (e.g., Dudley

(1989)) and the additional condition that $\eta\{\lambda_{j_1} \in A, \lambda_{j_2} \in \mathbb{R}, \ldots, \lambda_{j_k} \in \mathbb{R}\} = 0$ if $\mathrm{Leb}(A) = 0$, for any finite subset $\{j_1, \ldots, j_k\} \subset \{0, \ldots, N\}$. This verifies claim 1.

To verify claim 2, by Bessel's inequality we have that

$$\|m'\delta\|_{\mathcal{B}} \geq \sum_{j=0}^{N} \lambda_j^2 \kappa^2 \langle \phi_j, \delta\rangle^2.$$

m' is not injective iff $\lambda_j^2\kappa^2 = 0$ for some j. By countable additivity and by $\mathrm{Leb}(\{0\}) = 0 \Rightarrow \eta(\{\lambda_j = 0\}) = 0$ holding by assumption,

$$\Pr_\eta\big(\exists j \in \{0, 1, \ldots, N\} : \lambda_j = 0\big) \leq \sum_{j=0}^{N} \eta\big(\{\lambda_j = 0\}\big) = 0.$$

The final claim follows from the penultimate display. *Q.E.D.*

B.2. *Proof of Theorem 4*

By Assumption 3, for any $\alpha \neq \alpha_0$ and $\alpha \in \mathcal{N}'''$ with Fourier coefficients b_j, we have

$$\|m'(\alpha - \alpha_0)\|_{\mathcal{B}} \geq \left(\sum_j \mu_j^2 b_j^2\right)^{1/2} > L\left(\sum_j b_j^2\right)^{r/2} = L\|\alpha - \alpha_0\|_{\mathcal{A}}^r,$$

so the conclusion follows from Theorem 2. *Q.E.D.*

B.3. *Proof of Corollary 5*

Consider $\alpha \in \mathcal{N}'''$. Then

$$\sum_j \mu_j^{-2/(r-1)} b_j^2 < L^{-2/(r-1)}. \tag{B.1}$$

For $b_j = \langle \alpha - \alpha_0, \phi_j\rangle$, note that $\|\alpha - \alpha_0\|_{\mathcal{A}} = (\sum_j b_j^2)^{1/2}$ by $\phi_1, \phi_2, \ldots$ being an orthonormal basis. Then

$$\begin{aligned}\left(\sum_j b_j^2\right)^{1/2} &= \left(\sum_j \mu_j^{-2/r}\mu_j^{2/r} b_j^2\right)^{1/2} \\ &\leq \left(\sum_j \mu_j^{-2/(r-1)} b_j^2\right)^{(r-1)/2r} \left(\sum_j \mu_j^2 b_j^2\right)^{1/2r} \\ &< L^{-1/r}\left(\sum_j \mu_j^2 b_j^2\right)^{1/2r} \leq L^{-1/r}\big(\|m'(\alpha - \alpha_0)\|_{\mathcal{B}}\big)^{1/r},\end{aligned}$$

where the first inequality holds by the Hölder inequality, the second by equation (B.1), and the third by Assumption 3. Raising both sides to the rth power and multiplying through by L gives

$$L\|\alpha-\alpha_0\|_{\mathcal{A}}^r < \left\|m'(\alpha-\alpha_0)\right\|_{\mathcal{B}}. \tag{B.2}$$

The conclusion then follows from Theorem 4. *Q.E.D.*

B.4. *Proof of Theorem 6*

Let $F(y|X,W)=\Pr(Y\le y|X,W)$, $m(\alpha)=\mathrm{E}[1(Y\le\alpha(X))|W]-\tau$, and

$$m'h=\mathrm{E}\big[f_Y\big(\alpha_0(X)|X,W\big)h(X)|W\big],$$

so that by iterated expectations,

$$m(\alpha)=\mathrm{E}\big[F\big(\alpha(X)|X,W\big)|W\big]-\tau.$$

Then, by a pathwise mean value expansion, and by $f_Y(y|X,W)$ continuously differentiable,

$$\begin{aligned}
&\big|F\big(\alpha(X)|X,W\big)-F\big(\alpha_0(X)|X,W\big)\\
&\quad -f_Y\big(\alpha_0(X)|X,W\big)\big(\alpha(X)-\alpha_0(X)\big)\big|\\
&\ =\big|\big[f_Y\big(\bar\alpha(X)|X,W\big)-f_Y\big(\alpha_0(X)|X,W\big)\big]\big[\alpha(X)-\alpha_0(X)\big]\big|\\
&\ \le L_1\big[\alpha(X)-\alpha_0(X)\big]^2,
\end{aligned}$$

where $\bar\alpha(X)$ is the mean value of a pathwise Taylor expansion that lies between $\alpha(X)$ and $\alpha_0(X)$. Then, for $L_1L_2=L$,

$$\begin{aligned}
&\big|m(\alpha)(W)-m(\alpha_0)(W)-m'(\alpha-\alpha_0)(W)\big|\\
&\ \le L_1\mathrm{E}\big[\big\{\alpha(X)-\alpha_0(X)\big\}^2|W\big]\\
&\ \le L\mathrm{E}\big[\big\{\alpha(X)-\alpha_0(X)\big\}^2\big]=L\|\alpha-\alpha_0\|_{\mathcal{A}}^2.
\end{aligned}$$

Therefore,

$$\big\|m(\alpha)-m(\alpha_0)-m'(\alpha-\alpha_0)\big\|_{\mathcal{B}}\le L\|\alpha-\alpha_0\|_{\mathcal{A}}^2,$$

so that Assumption 2 is satisfied with $r=2$ and $\mathcal{N}''=\mathcal{A}$. The conclusion then follows from Corollary 5. *Q.E.D.*

REFERENCES

AI, C., AND X. CHEN (2003): "Efficient Estimation of Models With Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71, 1795–1844. [796]

ANDREWS, D. W. K. (2011): "Examples of L2-Complete and Boundedly-Complete Distributions," Cowles Foundation Discussion Paper 1801, Yale University. [788,789,793,801]

BLUNDELL, R., X. CHEN, AND D. KRISTENSEN (2007): "Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves," *Econometrica*, 75, 1613–1669. [785,789]

CARRASCO, M., J. P. FLORENS, AND E. RENAULT (2007): "Linear Inverse Problems in Structural Econometrics: Estimation Based on Spectral Decomposition and Regularization," in *Handbook of Econometrics*, Vol. 6B, ed. by J. J. Heckman and E. E. Leamer. Amsterdam: North-Holland. [789,792]

CHAMBERLAIN, G. (1986): "Asymptotic Efficiency in Semiparametric Models With Censoring," *Journal of Econometrics*, 32, 189–218. [801]

——— (1987): "Asymptotic Efficiency in Estimation With Conditional Moment Restrictions," *Journal of Econometrics*, 34, 305–334. [785]

——— (1992): "Efficiency Bounds for Semiparametric Regression," *Econometrica*, 60, 567–596. [796]

CHEN, X., AND S. LUDVIGSON (2009): "Land of Addicts? An Empirical Investigation of Habit-Based Asset Pricing Models," *Journal of Applied Econometrics*, 24, 1057–1093. [798,802]

CHEN, X., AND D. POUZO (2009): "Efficient Estimation of Semiparametric Conditional Moment Models With Possibly Nonsmooth Residuals," *Journal of Econometrics*, 152, 46–60. [796]

——— (2012): "Estimation of Nonparametric Conditional Moment Models With Possibly Nonsmooth Generalized Residuals," *Econometrica*, 80, 277–321. [793]

CHEN, X., AND M. REISS (2011): "On Rate Optimality for Ill-Posed Inverse Problems in Econometrics," *Econometric Theory*, 27, 497–521. [793,794]

CHEN, X., V. CHERNOZHUKOV, S. LEE, AND W. K. NEWEY (2014): "Supplement to 'Local Identification of Nonparametric and Semiparametric Models'," *Econometrica Supplemental Material*, 82, http://www.econometricsociety.org/ecta/supmat/9988_proofs.pdf. [786,787,792,793, 797,801]

CHERNOZHUKOV, V., AND C. HANSEN (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245–261. [785,789,795]

CHERNOZHUKOV, V., G. W. IMBENS, AND W. K. NEWEY (2007): "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics*, 139, 4–14. [786,787,795]

CHRISTENSEN, T. (2013): "Nonparametric Identification of Positive Eigenfunctions," Working Paper, Yale. [802]

COCHRANE, J. (2005): *Asset Pricing*. Princeton: Princeton University Press. [798]

DAROLLES, S., Y. FAN, J.-P. FLORENS, AND E. RENAULT (2011): "Nonparametric Instrumental Regression," *Econometrica*, 79, 1541–1565. [785]

D'HAULTFOEUILLE, X. (2011): "On the Completeness Condition in Nonparametric Instrumental Regression," *Econometric Theory*, 27, 460–471. [789]

DUDLEY, R. (1989): *Real Analysis and Probability*. Pacific Grove, CA: Wadsworth & Brooks/Cole. [804,805]

ENGL, H. W., M. HANKE, AND A. NEUBAUER (1996): *Regularization of Inverse Problems*. Dordrecht: Kluwer. [794]

ESCANCIANO, J. C., AND S. HODERLEIN (2010): "Nonparametric Identification of Euler Equations," Working Paper. [799,802]

FISHER, F. (1966): *The Identification Problem in Econometrics*. New York: McGraw-Hill. [785,786]

FLORENS, J.-P., AND E. SBAI (2010): "Local Identification in Empirical Games of Incomplete Information," *Econometric Theory*, 26, 1638–1662. [786]

FLORENS, J.-P., J. JOHANNES, AND S. VAN BELLEGEM (2011): "Identification and Estimation by Penalization in Nonparametric Instrumental Regression," *Econometric Theory*, 27, 472–496. [794]

——— (2012): "Instrumental Regression in Partially Linear Models," *The Econometrics Journal*, 15, 304–324. [797]

FLORENS, J.-P., M. MOUCHART, AND J. M. ROLIN (1990): *Elements of Bayesian Statistics*. New York: Dekker. [801]

GALLANT, A. R., AND G. TAUCHEN (1989): "Semiparametric Estimation of Conditional Constrained Heterogenous Processes: Asset Pricing Applications," *Econometrica*, 57, 1091–1120. [798]

HALL, P., AND J. HOROWITZ (2005): "Nonparametric Methods for Inference in the Presence of Instrumental Variables," *The Annals of Statistics*, 33, 2904–2929. [785]

HALL, R. E. (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 971–987. [798]

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054. [785]

——— (2012): "Dynamic Valuation Decomposition Within Stochastic Economies," *Econometrica*, 80, 911–967. Fisher-Schultz Lecture at the European Meetings of the Econometric Society. [802]

HANSEN, L. P., AND J. SCHEINKMAN (2009): "Long-Term Risk: An Operator Approach," *Econometrica*, 77, 117–234. [802]

——— (2012): "Recursive Utility in a Markov Environment With Stochastic Growth," *Proceedings of the National Academy of Sciences*, 109, 11967–11972. [802]

HANSEN, L. P., AND K. J. SINGLETON (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286. [798,802]

HANSEN, L. P., J. HEATON, J. LEE, AND N. ROUSSANOV (2007): "Intertemporal Substitution and Risk Aversion," in *Handbook of Econometrics*, Vol. 6, ed. by J. J. Heckman and E. E. Leamer. Amsterdam: Elsevier, 3967–4056. [798]

HOROWITZ, J. L., AND S. LEE (2007): "Nonparametric Instrumental Variables Estimation of a Quantile Regression Model," *Econometrica*, 75, 1191–1208. [795]

KREIN, M. G., AND M. A. RUTMAN (1950): *Linear Operators Leaving Invariant a Cone in a Banach Space*. Amer. Math. Soc. Translation 1950, Vol. 26. Providence: Amer. Math. Soc. [801,802]

KRESS, R. (1999): *Linear Integral Equations* (Second Ed.). New York: Springer-Verlag. [792,794, 799,804]

LEWBEL, A., O. LINTON, AND S. SRISUMA (2012): "Nonparametric Euler Equation Identification and Estimation," Working Paper. [799,802]

LUENBERGER, D. (1969): *Optimization by Vector Space Methods*. New York: Wiley. [789,796,803]

MATTNER, L. (1993): "Some Incomplete But Boundedly Complete Location Families," *The Annals of Statistics*, 21, 2158–2162. [789]

NEWEY, W. K., AND J. L. POWELL (1988): "Instrumental Variables Estimation for Nonparametric Models," Working Paper Version of Published Paper, MIT. [798]

——— (2003): "Instrumental Variables Estimation for Nonparametric Models," *Econometrica*, 71, 1565–1578. [785,786,788,789,801]

ROTHENBERG, T. J. (1971): "Identification in Parametric Models," *Econometrica*, 39, 577–591. [785,786]

SANTOS, A. (2011): "Instrumental Variables Methods for Recovering Continuous Linear Functionals," *Journal of Econometrics*, 161, 129–146. [797]

SARGAN, D. (1983): "Identification and Lack of Identification," *Econometrica*, 51, 1605–1634. [792]

ZIMMER, R. (1990): *Essential Results of Functional Analysis*. Chicago Lectures in Mathematics. Chicago: University of Chicago Press. [792]

Dept. of Economics, Yale University, Box 208281, New Haven, CT 06520, U.S.A.; xiaohong.chen@yale.edu, *Dept. of Economics, MIT, Cambridge, MA 02142, U.S.A.; vchern@mit.edu,*

Dept. of Economics, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 151-742, Respublic of Korea; sokbae@snu.ac.kr,

and

Dept. of Economics, MIT, Cambridge, MA 02142, U.S.A.; wnewey@mit.edu.

Manuscript received April, 2011; final revision received September, 2013.