An Expository Note on the Existence of Moments of Fuller and HFUL Estimators*

John C. Chao, Department of Economics, University of Maryland, chao@econ.umd.edu.

Jerry A Hausman, Department of Economics, MIT, jhausman@mit.edu

Whitney K. Newey, Department of Economics, MIT, wnewey@mit.edu.

Norman R. Swanson, Department of Economics, Rutgers University, nswanson@econ.rutgers.edu.

Tiemen Woutersen, Department of Economics, University of Arizona, woutersen@email.arizona.edu.

April 2012

JEL classification: C13, C31.

Keywords: endogeneity, instrumental variables, jackknife estimation, many moments, existence of moments.

----

*Prepared for Advances in Econometrics, Volume 29, Essays in Honor of Jerry Hausman. We all are indebted to Jerry for his advice, insights, wisdom, and wit.

Proposed Running Head: The Existence of Moments of Fuller and HFUL Estimators

Corresponding Author:

Whitney K. Newey

Department of Economics

MIT, E52-262D

Cambridge, MA 02142-1347

Abstract

In a recent paper, Hausman et al. (2012) propose a new estimator, HFUL (Heteroscedasticity robust Fuller), for the linear model with endogeneity. This estimator is consistent and asymptotically normally distributed in the many instruments and many weak instruments asymptotics. Moreover, this estimator has moments, just like the estimator by Fuller (1977). The purpose of this note is to discuss at greater length the existence of moments result given in Hausman et al. (2012). In particular, we intend to answer the following questions: Why does $LIML$ not have moments? Why does the Fuller modification lead to estimators with moments? Is normality required for the Fuller estimator to have moments? Why do we need a condition such as Hausman et al. (2012), Assumption 9? Why do we have the adjustment formula?

I. Introduction

The linear model with endogeneity is one of the most popular models in economics and there exist several estimators for this model. The Two Stage Least Squares estimator is inconsistent if there are many moments, see Kunitomo (1980) and Bekker (1994). Another estimator, the LIML (Limited Information Maximum Likelihood) estimator does not have moments. As a result, the estimates of the last estimator are dispersed when simulated data are used, see for example Hahn, Hausman, Kuersteiner (2005). These authors suggested the Fuller (1977) estimator as a solution to this problem for LIML. However, the Fuller estimator is inconsistent in the many instrument asymptotic if the data is heteroscedastic. In a recent paper, Hausman et al. (2012) propose a new estimator, HFUL (Heteroscedasticity robust Fuller), for the linear model with endogeneity. In that paper, we show that HFUL is consistent and asymptotically normally distributed in the many instruments and many weak instruments asymptotics, even in the presence of heteroscedasticity. Moreover, we also show that HFUL has moments, just like the estimator by Fuller (1977). The purpose of this note is to expound the existence of moments results given in Hausman et al. (2012). Thus, in this note, we intend to answer the following questions:

Q1: Why does $LIML$ not have moments?
Q2: Why does the Fuller modification lead to estimators with moments?
Q3: Is normality required for the Fuller estimator to have moments?
Q4: Why do we need a condition such as Hausman et al. (2012) Assumption 9?
Q5: Why do we have the adjustment formula, equation (2.1), $\widehat{\alpha} = [\widetilde{\alpha} - (1-\widetilde{\alpha})\, C/n]\, [1-(1-\widetilde{\alpha})\, C/n]^{-1}$ in $HFUL$, and what are the effects of C on the asymptotic properties of $HFUL$?

To keep our discussion as intuitive as possible, we adopt the simplest possible setup: a Gaussian, exactly identified IV regression with one endogenous regressor, orthonormal instrument, and a canonical error structure; i.e.,

$$y = x\delta_0 + \varepsilon, \tag{1}$$

$$x = z\pi_0 + v, \tag{2}$$

where $z'z/n = 1$. The reduced form representation y is easily seen as

$$y = z\phi_0 + \zeta,$$

where $\phi_0 = \pi_0 \delta_0$ and $\zeta = \varepsilon + v\delta_0$. To keep notation simple, we also assume that the IV regression is in what has called the canonical form, so that

$$\begin{pmatrix} \zeta_i \\ v_i \end{pmatrix} \sim i.i.d.N\left(0, I_2\right), \tag{3}$$

where ζ_i and v_i are the i^{th} element of the random vectors $\zeta = (\zeta_1, ..., \zeta_n)'$ and $v = (v_1, ..., v_n)'$, respectively. With this simple, stripped-down setup, we can present the essential ideas behind our results while avoiding some of the technical difficulties and tedious calculations associated with having non-normality, heteroskedasticity, and many and/or weak instruments.

In this simple setting, it is easily seen that the OLS estimators of the reduced form parameters ϕ and π have the following joint normal distribution

$$\begin{pmatrix} \widehat{\phi}_n \\ \widehat{\pi}_n \end{pmatrix} = \begin{pmatrix} z'y/n \\ z'x/n \end{pmatrix} \sim N\left(\begin{pmatrix} \phi_0 \\ \pi_0 \end{pmatrix}, n^{-1} I_2\right), \tag{4}$$

so that, in particular, note that $\widehat{\phi}_n$ and $\widehat{\pi}_n$ are independent in this case.

Given the simplicity of the setup here, the existence and non-existence of moment results given below are not new but are presented here so as to illustrate some of the issues involved. In fact, Fuller (1977) has already established the existence of moments of his estimator for a IV regression model under homoskedastic, Gaussian error assumptions and a fixed number of instruments. However, here, we provide some intuitive explanation for why the Fuller modification works based on certain geometric properties of the high-dimensional sphere. Similar discussion does not appear in Fuller (1977) and does not seem to appear elsewhere in the literature. In addition, the existence of moments result which we give in the paper is new, as it generalizes the Fuller (1977) result to IV regression models with heteroskedasticity, non-Gaussian error distributions, and possibly many weak instruments, and it establishes such a result for a new estimator $HFUL$.

In the remainder, we answer each of the questions posed above in turn.

II. Q1: Why does $LIML$ not have moments?

Now, to address **Q1**, we note that, under exact identification, we have

$$\widehat{\delta}_{LIML} = \widehat{\delta}_{2SLS} = \frac{\widehat{\pi}_n \left(z'z\right) \widehat{\phi}_n}{\widehat{\pi}_n \left(z'z\right) \widehat{\pi}_n} = \frac{\widehat{\phi}_n}{\widehat{\pi}_n}.$$

The non-existence of finite sample moments for this estimator is easily established by the following calculations

$$\begin{aligned}
& E\left|\widehat{\delta}_{LIML}\right|^p \\
= \; & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|\frac{\widehat{\phi}}{\widehat{\pi}}\right|^p \frac{n}{(2\pi)}\exp\left\{-\frac{1}{2}\left[n\left(\widehat{\pi}-\pi_0\right)^2 + n\left(\widehat{\phi}-\phi_0\right)^2\right]\right\} d\widehat{\pi}\, d\widehat{\phi} \\
= \; & \int_{-\infty}^{\infty}\left|\widehat{\phi}\right|^p \sqrt{\frac{n}{2\pi}}\exp\left\{-\frac{n}{2}\left(\widehat{\phi}-\phi_0\right)^2\right\} d\widehat{\phi} \\
& \times \int_{-\infty}^{\infty}\left|\widehat{\pi}\right|^{-p}\sqrt{\frac{n}{2\pi}}\exp\left\{-\frac{n}{2}\left(\widehat{\pi}-\pi_0\right)^2\right\} d\widehat{\pi} \\
\geq \; & \int_{-\infty}^{\infty}\left|\widehat{\phi}\right|^p \sqrt{\frac{n}{2\pi}}\exp\left\{-\frac{n}{2}\left(\widehat{\phi}-\phi_0\right)^2\right\} d\widehat{\phi} \\
& \times \int_{-|\pi_0|}^{|\pi_0|}\left|\widehat{\pi}\right|^{-p}\sqrt{\frac{n}{2\pi}}\exp\left\{-\frac{n}{2}\left(\widehat{\pi}-\pi_0\right)^2\right\} d\widehat{\pi} \\
\geq \; & \int_{-\infty}^{\infty}\left|\widehat{\phi}\right|^p \sqrt{\frac{n}{2\pi}}\exp\left\{-\frac{n}{2}\left(\widehat{\phi}-\phi_0\right)^2\right\} d\widehat{\phi} \\
& \times \sqrt{\frac{n}{2\pi}}\exp\left\{-2n\left|\pi_0\right|^2\right\}\int_{-|\pi_0|}^{|\pi_0|}\left|\widehat{\pi}\right|^{-p} d\widehat{\pi} \\
= \; & +\infty
\end{aligned} \tag{5}$$

for all p such that $1 \leq p < \infty$ and for each finite n, since

$$\int_{-|\pi_0|}^{|\pi_0|}\left|\widehat{\pi}\right|^{-p} d\widehat{\pi} = +\infty.$$

Note that problem here is that part of the integrand (i.e., $\left|\widehat{\pi}\right|^{-p}$) has a pole at $\widehat{\pi} = 0$, so that if there is sufficient probability mass in the neighborhood of $\widehat{\pi} = 0$, then the integral does not exist. We will provide more discussion and intuition when we contrast this case with the case where the estimator has been modified in the sense of Fuller (1977). Please see remark in section III below.

III. Q2: Why does the Fuller modification lead to estimators with moments?

To address this question, note first that, under the current setup, Fuller estimator can be written as

$$\widehat{\delta}_{FULL} = \frac{\widehat{\pi}\,(z'z)\,\widehat{\phi} + (C/n)\,x'My}{\widehat{\pi}\,(z'z)\,\widehat{\pi} + (C/n)\,x'Mx} = \frac{n\widehat{\pi}\widehat{\phi} + (C/n)\,v'M\zeta}{n\widehat{\pi}^2 + (C/n)\,v'Mv} = \frac{\widehat{\pi}\widehat{\phi} + (C/n)\,(v'M\zeta/n)}{\widehat{\pi}^2 + (C/n)\,(v'Mv/n)} \tag{6}$$

where $M = I_n - z\,(z'z)^{-1}\,z' = I_n - zz'/n$. To understand why this estimator solves the moment problems it may be helpful to draw an analogy with ridge-regression. In particular, the ridge version of the least squares estimator has its denominator perturbed by an extra term which ensures that the denominator is nonzero. Similarly, in this case the Fuller modification modifies the denominator of 2SLS/LIML by adding an extra term.

To show that this added term is effective in ensuring the existence of moments, first partition $z = \left(\underset{1\times 1}{z_1}, \underset{1\times (n-1)}{z_{2\cdot}'} \right)'$ and consider the decomposition

$$M = H^{\perp}H^{\perp\prime},$$

where

$$\underset{n\times(n-1)}{H^{\perp}} = \begin{pmatrix} -z_{2\cdot}'/z_1 \\ I_{n-1} \end{pmatrix} \left[I_{n-1} + \frac{z_{2\cdot}z_{2\cdot}'}{z_1^2} \right]^{-1/2} \in V_{n-1,n}$$

and where

$$V_{n-1,n} = \left\{ \underset{n\times(n-1)}{X} : X'X = I_{n-1} \right\}$$

denotes the Stiefel manifold. Consider the transformation $v_* = (n-1)^{-1/2}\,H^{\perp\prime}v$ and $\zeta_* = (n-1)^{-1/2}\,H^{\perp\prime}\zeta$, and it is easily verified that in the present case

$$\begin{pmatrix} \zeta_{*,i} \\ v_{*,i} \end{pmatrix} \equiv i.i.d.N\left(0, (n-1)^{-1}\,I_2\right),$$

where $v_{*,i}$ and $\zeta_{*,i}$ are the i^{th} element of v_* and ζ_*, respectively. Moreover, v_* and ζ_* are independent of $\widehat{\pi}$ and $\widehat{\phi}$ in this case. Using this change of variables, we can rewrite the Fuller estimator in the representation

$$\widehat{\delta}_{FULL} = \frac{\widehat{\pi}\widehat{\phi} + (1-1/n)\,(C/n)\,(v_*'\zeta_*)}{\widehat{\pi}^2 + (1-1/n)\,(C/n)\,(v_*'v_*)}.$$

Next, define

$$\xi_1 = \frac{1}{\sqrt{n}}z'\zeta,\ \ \xi_2 = \frac{1}{\sqrt{n}}z'v,$$

so that

$$\widehat{\phi}=\phi_0+\frac{1}{\sqrt{n}}\xi_1,\ \widehat{\pi}=\pi_0+\frac{1}{\sqrt{n}}\xi_2,$$

and we can further represent the Fuller estimator as

$$\begin{aligned}\widehat{\delta}_{FULL} &= \frac{\widehat{\pi}\widehat{\phi}+(1-1/n)(C/n)(v_*'\zeta_*)}{\widehat{\pi}^2+(1-1/n)(C/n)(v_*'v_*)} \\ &= \frac{\pi_0^2\delta_0+\pi_0\left(n^{-1/2}\xi_1\right)+\pi_0\delta_0\left(n^{-1/2}\xi_2\right)+n^{-1}\xi_1\xi_2+(1-1/n)(C/n)(v_*'\zeta_*)}{\pi_0^2+2\pi_0\left(n^{-1/2}\xi_2\right)+n^{-1}\xi_2^2+(1-1/n)(C/n)(v_*'v_*)} \qquad (7)\end{aligned}$$

Note that (7) makes clear that the Fuller estimator can be written as a function of several random components, some of which are linear in the error vectors such as $n^{-1/2}\xi_1$ and $n^{-1/2}\xi_2$ while others are bilinear such as $v_*'\zeta_*$ and $v_*'v_*$. To show the existence of moments for the Fuller estimator, we divide the domain of integration into a region where all of these random components are in some small neighborhood of their asymptotic limit (denoted by the event $\mathcal{A}$ below) and the complement of this region (denoted by $\mathcal{A}^C$). More precisely, let

$$\begin{aligned}\mathcal{A}_1 &= \left\{\left|v_*'\zeta_*\right|<\eta_1\right\},\mathcal{A}_2=\left\{\left|v_*'v_*-1\right|<\eta_2\right\},\mathcal{A}_3=\left\{\left|n^{-1/2}\xi_1\right|<\eta_3\right\},\mathcal{A}_4=\left\{\left|n^{-1/2}\xi_2\right|<\eta_4\right\}, \\ \mathcal{A} &= \mathcal{A}_1\cap\mathcal{A}_2\cap\mathcal{A}_3\cap\mathcal{A}_4\end{aligned}$$

for constants $\eta_1,\eta_2,\eta_3,\eta_4>0$ and $\eta_4<|\pi_0|/2$. Now,

$$\begin{aligned}&\left|\widehat{\delta}_{FULL}\right|\mathbb{I}_{\mathcal{A}} \\ =& \left|\frac{\pi_0^2\delta_0+\pi_0\left(n^{-1/2}\xi_1\right)+\pi_0\delta_0\left(n^{-1/2}\xi_2\right)+n^{-1}\xi_1\xi_2+(1-1/n)(C/n)(v_*'\zeta_*)}{\left(\pi_0+n^{-1/2}\xi_2\right)^2+(1-1/n)(C/n)(v_*'v_*)}\right|\mathbb{I}_{\mathcal{A}} \\ \leq& \frac{\pi_0^2|\delta_0|+|\pi_0|\left|n^{-1/2}\xi_1\right|+|\pi_0||\delta_0|\left|n^{-1/2}\xi_2\right|+\left|n^{-1/2}\xi_1\right|\left|n^{-1/2}\xi_2\right|+C\left|v_*'\zeta_*\right|}{\pi_0^2-2|\pi_0|\left|n^{-1/2}\xi_2\right|} \\ \leq& \frac{\pi_0^2|\delta_0|+|\pi_0|\eta_3+|\pi_0||\delta_0|\eta_4+\eta_3\eta_4+C\eta_1}{\pi_0^2-2|\pi_0|\eta_4}.\end{aligned}$$

It follows that for any fixed $p>0$ and any true parameter value (δ_0,π_0) there exists a positive constant C_1, possibly depending on (δ_0,π_0) and p, such that

$$E\left[\left|\widehat{\delta}_{FULL}\right|^p\mathbb{I}_{\mathcal{A}}\right]\leq C_1<\infty \qquad (8)$$

Moreover, suppose that the parameter space of (δ,π) is some bounded set $\mathcal{D}\subset\mathbb{R}^2$, then (8) holds under some constant not depending on the true value (δ_0,π_0)

Next, consider what happens under the event $\mathcal{A}^C$. In this case, we first form the upper bound

$$\left|\widehat{\delta}_{FULL}\right| = \left|\frac{\widehat{\pi}\widehat{\phi} + (C/n)(v_*'\zeta_*/n)}{\widehat{\pi}^2 + (C/n)(v_*'v_*/n)}\right| \leq \frac{\left|\widehat{\pi}\widehat{\phi}\right| + (1-1/n)(C/n)(|v_*'\zeta_*|)}{(1-1/n)(C/n)(v_*'v_*)} = \left(\frac{n}{n-1}\right)\frac{n\left|\widehat{\pi}\widehat{\phi}\right|}{Cv_*'v_*} + \frac{|v_*'\zeta_*|}{v_*'v_*}.$$

By Loève's inequality, we have for any fixed $p > 0$

$$E\left[\left|\widehat{\delta}_{FULL}\right|^p \mathbb{I}_{\mathcal{A}^C}\right] \leq c_p\left\{\left(\frac{n}{n-1}\right)^p E\left|\frac{n\left|\widehat{\pi}\widehat{\phi}\right|\mathbb{I}_{\mathcal{A}^C}}{Cv_*'v_*}\right|^p + E\left[\frac{|v_*'\zeta_*|}{v_*'v_*}\mathbb{I}_{\mathcal{A}^C}\right]^p\right\} \tag{9}$$

To analyze the first term in (9), note that, by the Cauchy-Schwarz (CS) inequality, we have

$$E\left|\frac{n\left|\widehat{\pi}\widehat{\phi}\right|\mathbb{I}_{\mathcal{A}^C}}{Cv_*'v_*}\right|^p \leq \frac{n^p}{C^p}\sqrt{\Pr\left\{\mathcal{A}^C\right\}}\sqrt{E\left|\frac{\left|\widehat{\pi}\widehat{\phi}\right|}{v_*'v_*}\right|^{2p}}$$

Moreover,

$$\begin{aligned}
& \Pr\left\{\mathcal{A}^C\right\} \\
= \; & \Pr\left\{\mathcal{A}_1^C \cup \mathcal{A}_2^C \cup \mathcal{A}_3^C \cup \mathcal{A}_4^C\right\} \\
\leq \; & \Pr\left\{\mathcal{A}_1^C\right\} + \Pr\left\{\mathcal{A}_2^C\right\} + \Pr\left\{\mathcal{A}_3^C\right\} + \Pr\left\{\mathcal{A}_4^C\right\} \\
= \; & \Pr\left\{\left|v_*'\zeta_*\right| \geq \eta_1\right\} + \Pr\left\{\left|v_*'v_* - 1\right| \geq \eta_2\right\} + \Pr\left\{\left|n^{-1/2}\xi_1\right| \geq \eta_3\right\} \\
& + \Pr\left\{\left|n^{-1/2}\xi_2\right| \geq \eta_4\right\} \\
\leq \; & \frac{E\left|v_*'\zeta_*\right|^{4p}}{\eta_1^{4p}} + \frac{E\left|v_*'v_* - 1\right|^{4p}}{\eta_2^{4p}} + \frac{E\left[n^{-1/2}\xi_1\right]^{4p}}{\eta_3^{4p}} + \frac{E\left[n^{-1/2}\xi_2\right]^{4p}}{\eta_4^{4p}} \\
& \text{(by Markov's inequality)} \\
< \; & \frac{C_p}{n^{2p}\eta_1^{4p}} + \frac{C_p}{n^{2p}\eta_2^{4p}} + \frac{C_p}{n^{2p}\eta_3^{4p}} + \frac{C_p}{n^{2p}\eta_4^{4p}} \\
& \text{(by Lemma 1 below and by the fact that both } \xi_1 \text{ and } \xi_2 \text{ are } N(0,1) \text{ random variables)} \\
< \; & \frac{\overline{C}}{n^{2p}}
\end{aligned} \tag{10}$$

Turning our attention now to the expectation $E\left|\left|\widehat{\pi}\widehat{\phi}\right|/(v_*'v_*)\right|^{2p}$, note that

$$E\left|\frac{\left|\widehat{\pi}\widehat{\phi}\right|}{v_*'v_*}\right|^{2p} = E\left|\widehat{\pi}\widehat{\phi}\right|^{2p} E\left|v_*'v_*\right|^{-2p}$$

which the equality above follows from the fact that v_* is independent of $\widehat{\pi}$ and $\widehat{\phi}$, as noted above. Given the joint normality of the OLS estimators as noted in (4) and given that $\widehat{\pi}$ and $\widehat{\phi}$ are independent, it is trivial to show (since all moments of the normal distribution exist) that there exists a constant C_p such that

$$E\left|\widehat{\pi}\widehat{\phi}\right|^{2p} = E\left|\widehat{\pi}\right|^{2p} E\left|\widehat{\phi}\right|^{2p} \leq C_p < \infty \tag{11}$$

for any fixed $p > 0$. Next, we consider $E\left|v_*'v_*\right|^{-2p}$. To evaluate this expectation, it is helpful to change into (generalized) polar coordinates, viz

$$v_* = hr \quad \text{with } h = v_*\left(v_*'v_*\right)^{-1/2} \text{ and } r = \left(v_*'v_*\right)^{1/2}, \tag{12}$$

where $h \in V_{1,n-1}$, so that $h'h = I_{n-1}$. The Jacobian of this transformation is given by

$$(dv_*) = c_n r^{(n-2)} dr\,[dh] \tag{13}$$

where $[dh]$ denotes exterior differential form of the normalized invariant measure on the Stiefel manifold $V_{1,n-1}$ and where

$$c_n = \frac{2\pi^{(n-1)/2}}{\Gamma\left[\frac{1}{2}(n-1)\right]}. \tag{14}$$

(See, for example, Lemma 1.5.2 of Chikuse, 2003). We show in Lemma 2 below that

$$E\left|v_*'v_*\right|^{-2p} = 1 + O\left(n^{-1}\right), \tag{15}$$

so that the first term in (9) is bounded for all n sufficiently large, i.e.,

$$\begin{aligned}
E\left|\frac{n\left|\widehat{\pi}\widehat{\phi}\right|\mathbb{I}_{\mathcal{A}^C}}{Cv_*'v_*}\right|^p &\leq \frac{n^p}{C^p}\sqrt{\Pr\left\{\mathcal{A}^C\right\}}\sqrt{E\left|\frac{\left|\widehat{\pi}\widehat{\phi}\right|}{v_*'v_*}\right|^{2p}} \\
&= \frac{n^p}{C^p}\sqrt{\frac{\overline{C}}{n^{2p}}}\sqrt{E\left|\widehat{\pi}\widehat{\phi}\right|^{2p} E\left|v_*'v_*\right|^{-2p}} \\
&\leq \frac{n^p}{C^p}\sqrt{\frac{\overline{C}}{n^{2p}}}\sqrt{C_p\left(1 + O\left(n^{-1}\right)\right)} \\
&= O(1).
\end{aligned} \tag{16}$$

Moreover, it is easy to show that the second term in (9) is also bounded for all n sufficiently large, i.e., $E\left[\left\{\left|v_*'\zeta_*\right|/\left(v_*'v_*\right)\right\}\mathbb{I}_{\mathcal{A}^C}\right]^p = O(1)$. In particular, by making use of various forms of the CS inequality, we have that

$$E\left[\frac{\left|v_*'\zeta_*\right|}{v_*'v_*}\mathbb{I}_{\mathcal{A}^C}\right]^p \leq \sqrt{\Pr\left(\mathcal{A}^C\right)}\sqrt{E\left|\frac{\left|v_*'\zeta_*\right|}{v_*'v_*}\right|^{2p}} \leq \sqrt{E\left[\left|v_*'v_*\right|^{-p}\left|\zeta_*'\zeta_*\right|^p\right]} \leq \left(E\left|v_*'v_*\right|^{-2p}\right)^{1/4}\left(E\left|\zeta_*'\zeta_*\right|^{2p}\right)^{1/4}.$$

Now, since ζ_* has the same multivariate normal distribution as v_*, we can apply Lemma 2 (given below) for the case $q = -p < 0$ to obtain

$$E\left|\zeta_*'\zeta_*\right|^{2p} = 1 + O\left(n^{-1}\right). \tag{17}$$

Using this result in conjunction with (15), we have that

$$E\left[\frac{|v_*'\zeta_*|}{v_*'v_*}\mathbb{I}_{\mathcal{A}^C}\right]^p \leq \sqrt{1 + O\left(n^{-1}\right)}\sqrt{1 + O\left(n^{-1}\right)} = O\left(1\right) \tag{18}$$

It follows from (8), (16) and (18) that

$$\begin{aligned} E\left[\left|\widehat{\delta}_{FULL}\right|^p\right] &= E\left[\left|\widehat{\delta}_{FULL}\right|^p\mathbb{I}_{\mathcal{A}}\right] + E\left[\left|\widehat{\delta}_{FULL}\right|^p\mathbb{I}_{\mathcal{A}^C}\right] \\ &\leq E\left[\left|\widehat{\delta}_{FULL}\right|^p\mathbb{I}_{\mathcal{A}}\right] + c_p\left\{\left(\frac{n}{n-1}\right)^p E\left|\frac{n\left|\widehat{\pi}\widehat{\phi}\right|\mathbb{I}_{\mathcal{A}^C}}{Cv_*'v_*}\right|^p + E\left[\frac{|v_*'\zeta_*|}{v_*'v_*}\mathbb{I}_{\mathcal{A}^C}\right]^p\right\} = O\left(1\right), \end{aligned}$$

which yields the desired existence of moments result for the Fuller estimator.

Remark:

Looking back at expression (14), note that

$$c_n = \frac{2\pi^{(n-1)/2}}{\Gamma\left[\frac{1}{2}\left(n-1\right)\right]}$$

is the normalization factor for the invariant (uniform) measure on the sphere, and it gives the surface area of the unit sphere (i.e., a sphere with unit radius) in $\mathbb{R}^{n-1}$. By Stirling's approximation one can show that

$$c_n = \frac{2\pi^{(n-1)/2}}{\Gamma\left[\frac{1}{2}\left(n-1\right)\right]} \sim \left(\frac{2e\pi}{n}\right)^{(n-2)/2},$$

so that this surface area is very small in high dimension; i.e., when n is large. Moreover, the volume of the unit sphere is proportional to $\left(n-1\right)^{-1}c_n$, so that it is also small. It follows that in high dimension, the probability of landing within a fixed neighborhood of the origin under the invariant uniform measure is very small. This basic fact of high-dimensional convex geometry provides an intuition for why the Fuller modification is effective in solving the moment problem. In particular, in our simple setting here, the denominator of the Fuller estimator is given by

$$\widehat{\pi}^2 + \left(1 - 1/n\right)\left(C/n\right)v_*'v_* \sim \widehat{\pi}^2 + \left(C/n\right)\left\|v_*\right\|^2$$

Moreover, the denominator of the $2SLS$ estimator is simply a special case of the expression above with $C = 0$. Hence, the $2SLS$ has denominator that is the square of an one-dimensional random variable $\widehat{\pi}$, so that, in particular, its probability of being zero is relatively high given its low dimensionality. This, in turn, leads to the non-existence of moments of the $2SLS$, as we have already shown in (5) above by explicit calculations. On the other hand, as noted above, the Fuller modification amounts to a ridge-regression-type perturbation of the denominator of the $2SLS$, so that the denominator of the modified estimator is now dependent on the square of the norm of a high-dimensional vector v_*, in addition to $\widehat{\pi}^2$. In consequence, the probability of the denominator being zero is now very small for finite n sufficiently large, leading to the existence of moments.

IV. Q3: Is normality required for the Fuller estimator to have moments?

Note that although the results here are shown under a Gaussian error assumption as in Fuller (1977), the existence of moments of the Fuller estimator is not hinged upon such an assumption. In particular, such an assumption is not needed to establish the boundedness of the inverse moment $E\left|v_*'v_*\right|^{-2p}$ even though in proving Lemma 2 below, we have evaluated an integral of the form

$$\int_0^\infty (2\pi)^{-1/2}(n-1)^{1/2} r^{(n-4p-2)} \exp\left\{-\frac{(n-1)}{2}r^2\right\} dr,$$

which may suggest that we need distributions with exponentially decaying tails in order to ensure the existence of all moments as n becomes large. Note, however, that we can easily modify the proof of Lemma 2 by a truncation argument since the existence of $E\left|v_*'v_*\right|^{-2p}$ depends only on the behavior of the high-dimensional random vector v_* near the origin. More specifically, note that

$$\begin{aligned} E\left|v_*'v_*\right|^{-2p} &= E\left[\left|v_*'v_*\right|^{-2p}\mathbb{I}\left\{\left|v_*'v_*\right|<1\right\}\right]+E\left[\left|v_*'v_*\right|^{-2p}\mathbb{I}\left\{\left|v_*'v_*\right|\geq 1\right\}\right] \\ &\leq E\left[\left|v_*'v_*\right|^{-2p}\mathbb{I}\left\{\left|v_*'v_*\right|<1\right\}\right]+1. \end{aligned}$$

Hence, changing into polar coordinates, i.e., $v_* = hr$ with $h = v_*\left(v_*'v_*\right)^{-1/2}$ and $r = \left(v_*'v_*\right)^{1/2}$, we have

$$E\left|v_*'v_*\right|^{-2p} \leq E\left[r^{-4p}\mathbb{I}\left\{r^2<1\right\}\right]+1,$$

so that we really only need to specify conditions on the density of r near zero in order to ensure the existence of $E\left|v_*'v_*\right|^{-2p}$.

On the other hand, to show the existence of $E\left[\left|\widehat{\delta}_{FULL}\right|^p\right]$ for some $p>0$ (possibly large) we do of course need the error distributions of the IV regression model to have enough (positive order) moments. Assumption of normality ensures that this is not a problem for any p. However, if we are only interested in the existence of $E\left[\left|\widehat{\delta}_{FULL}\right|^{p}\right]$ for some fixed p, it is sufficient to specify weaker moment conditions on the error distributions, as we do in the paper.

V. Q4: Why do we need a condition such as Hausman et al. (2012), Assumption 9?

Assumption 9 ensures the existence of certain inverse moments[1]. Under Gaussian error assumption, we show explicitly in Lemma 2 below that inverse moments of the form $E\left|v_{*}'v_{*}\right|^{-2p}$ (for $p > 0$) do exist. However, in general under possibly non-normal distributions, it is not true that inverse moments will necessarily exist.

To show that a condition such as Assumption 9 is not superfluous, we give an example of a pathological density under which inverse moments do not exist. To proceed, let w be an $(n-1)\times 1$ random vector, and consider the family of densities

$$f_n(w) = \frac{\Gamma\left[\frac{1}{2}(n-1)\right]}{2^{1/2}\pi^{n/2}} \frac{1}{(w'w)^{(n-2)/2}} \exp\left\{-\frac{1}{2}w'w\right\}, \quad \text{for } w \in \mathbb{R}^{n-1}$$

First, we show that $f_n(w)$ is indeed a density. To proceed, we again change into polar coordinates, viz

$$\begin{aligned} w &= hr, \text{ where } h = w\left(w'w\right)^{-1/2} \text{ and } r = \left(w'w\right)^{1/2}; \\ (dw) &= c_n r^{(n-2)} dr\,[dh]\,, \text{ where } c_n = \frac{2\pi^{(n-1)/2}}{\Gamma\left[\frac{1}{2}(n-1)\right]}. \end{aligned}$$

[1] In the revised version of the paper, we have changed Assumption 9 a bit. However, the current Assumption 9 and the Assumption 9 given in the previous version of this paper serve a similar purpose as both are designed to ensure the existence of inverse moments.

It follows that

$$
\begin{aligned}
& \int_{\mathbb{R}^{n-K}} \frac{\Gamma\left[\frac{1}{2}(n-1)\right]}{2^{3/2}\pi^{n/2}} \frac{1}{(w'w)^{(n-2)/2}} \exp\left\{-\frac{1}{2}w'w\right\}(dw) \\
= \;& \frac{\Gamma\left[\frac{1}{2}(n-1)\right]}{2^{1/2}\pi^{n/2}} \int_{V_{1,n-1}} \int_0^\infty \frac{1}{r^{(n-2)}} \exp\left\{-\frac{1}{2}r^2\right\} c_n r^{(n-2)} dr\,[dh] \\
= \;& \frac{\Gamma\left[\frac{1}{2}(n-1)\right]}{2^{1/2}\pi^{n/2}} \frac{2\pi^{(n-1)/2}}{\Gamma\left[\frac{1}{2}(n-1)\right]} \int_{V_{1,n-1}} \int_0^\infty \exp\left\{-\frac{1}{2}r^2\right\} dr\,[dh] \\
= \;& \sqrt{\frac{2}{\pi}} \int_{V_{1,n-1}} \int_0^\infty \exp\left\{-\frac{1}{2}r^2\right\} dr\,[dh] \\
= \;& 2\int_{V_{1,n-1}} \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}r^2\right\} dr\,[dh] \\
= \;& \int_{V_{1,n-1}} [dh] \quad (\text{since } \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}r^2\right\} dr = \frac{1}{2}) \\
= \;& \int_{V_{1,n-1}} [dh] = 1 \;\; (\text{since } [dh] \text{ defines the differential form for the normalized Haar measure})
\end{aligned}
$$

Moreover, it is easy to see that the inverse moment $E\left[(w'w)^{-1}\right]$ does not exist in this case, since by previous calculations

$$
\begin{aligned}
& E\left[\left(w'w\right)^{-1}\right] \\
= \;& \int_{\mathbb{R}^{n-K}} \frac{\Gamma\left[\frac{1}{2}(n-1)\right]}{2^{3/2}\pi^{n/2}} \frac{1}{(w'w)^{n/2}} \exp\left\{-\frac{1}{2}w'w\right\}(dw) \\
= \;& \frac{\Gamma\left[\frac{1}{2}(n-1)\right]}{2^{1/2}\pi^{n/2}} \int_{V_{1,n-1}} \int_0^\infty \frac{1}{r^n} \exp\left\{-\frac{1}{2}r^2\right\} c_n r^{(n-2)} dr\,[dh] \\
= \;& \frac{\Gamma\left[\frac{1}{2}(n-1)\right]}{2^{1/2}\pi^{n/2}} \frac{2\pi^{(n-1)/2}}{\Gamma\left[\frac{1}{2}(n-1)\right]} \int_{V_{1,n-1}} \int_0^\infty \frac{1}{r^2} \exp\left\{-\frac{1}{2}r^2\right\} dr\,[dh] \\
= \;& \sqrt{\frac{2}{\pi}} \int_0^\infty \frac{1}{r^2} \exp\left\{-\frac{1}{2}r^2\right\} dr \int_{V_{1,n-1}} [dh] \\
\geq \;& \sqrt{\frac{2}{\pi}} \int_0^\varepsilon \frac{1}{r^2} \exp\left\{-\frac{1}{2}r^2\right\} dr \;\; \text{for some } \varepsilon \text{ such that } 0 < \varepsilon < 1 \\
\geq \;& \sqrt{\frac{2}{\pi}} \exp\left\{-\frac{1}{2}\varepsilon^2\right\} \int_0^\varepsilon \frac{1}{r^2} dr = +\infty,
\end{aligned}
$$

so that without ruling out pathological cases such as this one, we will not in general be able to establish the existence of moments of the Fuller estimator.

V. Q5: Why do we have the adjustment formula $\widehat{\alpha} = [\widetilde{\alpha} - (1-\widetilde{\alpha})\, C/n]\; [1-(1-\widetilde{\alpha})\, C/n]^{-1}$ in $HFUL$, and what are the effects of C on the asymptotic properties of $HFUL$?

Consider here the more general IV regression model given in Hausman et al. (2012, section 2) with G endogenous regressors and K instruments. In this case, the $HFUL$ estimator has the form

$$\widehat{\delta} = \left(X'\left[P - D_P\right]X - \widehat{\alpha}X'X\right)^{-1}\left(X'\left[P - D_P\right]y - \widehat{\alpha}X'y\right) \tag{19}$$

where $P = Z(Z'Z)^{-1}Z'$, $D_P = diag\left(P_{11}, ..., P_{nn}\right)$ with P_{ii} being the i^{th} diagonal element of P, and

$$\widehat{\alpha} = [\widetilde{\alpha} - (1-\widetilde{\alpha})\, C/n]\,[1-(1-\widetilde{\alpha})\, C/n]^{-1}, \tag{20}$$

with $\widetilde{\alpha}$ being the α value which appears in $HLIM$. To better understand the adjustment formula (20), note that in Lemma B1 of a supplementary technical appendix to this paper (available on the web at http://econweb.umd.edu/˜chao/Research/research.html), we show that the $HFUL$ estimator (19) has an equivalent representation of the form

$$\widehat{\delta} = \left(X'\left[P - D_P\right]X - \left\{\widetilde{\kappa} - \frac{C}{n}\right\}X'\left[M + D_P\right]X\right)^{-1}\left(X'\left[P - D_P\right]y - \left\{\widetilde{\kappa} - \frac{C}{n}\right\}X'\left[M + D_P\right]y\right), \tag{21}$$

where $M = I_n - P$ and $\widetilde{\kappa}$ is the smallest root of the determinantal equation

$$\det\left\{\overline{X}'\left[P - D_P\right]\overline{X} - \kappa\overline{X}'\left[M + D_P\right]\overline{X}\right\} = 0,$$

with $\overline{X} = \begin{bmatrix} y & X \end{bmatrix}$. From (21), it is apparent that the adjustment factor $\widehat{\alpha}$ allows $HFUL$ to be rewritten in a form where the "denominator" also contains a ridge-regression-type perturbation term (i.e., the term $(C/n)\, X'\left[M + D_P\right]X$) analogous to that of the Fuller estimator (6). As we have shown earlier, this additional term leads to the existence of moments.

C does not affect the first-order asymptotic properties of $HFUL$ as can be seen from Theorem 2 of the paper and the surrounding discussion. It will have higher-order effects, but that is beyond the scope of this paper.

VII. Lemmas and Proofs:

Lemma 1: Suppose that $v_* \sim N\left(0, (n-1)^{-1}I_{n-1}\right)$ and $\zeta_* \sim N\left(0, (n-1)^{-1}I_{n-1}\right)$; and v_* and ζ_* are independent. Then, for any positive integer p, the following results hold:

(a) $E\left[v_*'\zeta_*\right]^{4p} \leq C_p/n^{2p}$;

(b) $E\left[v_*'v_* - 1\right]^{4p} \leq C_p/n^{2p}$;

where C_p is a constant which may be different in parts (a) and (b) above and which may depend on p.

Proof:

Define $\widetilde{v} = (n-1)^{1/2}\,v_*$ and $\widetilde{\zeta} = (n-1)^{1/2}\,\zeta_*$, so that $\widetilde{v} \sim N\left(0, I_{n-1}\right)$ and $\widetilde{\zeta} \sim N\left(0, I_{n-1}\right)$. Now, to show part (a), note first that

$$\begin{aligned}
& E\left[v_*'\zeta_*\right]^{4p} \\
= \;& (n-1)^{-4p}\,E\left[\widetilde{v}'\widetilde{\zeta}\right]^{4p} \\
= \;& \frac{1}{(n-1)^{4p}} E\left[\sum_{i=1}^{n-1}\widetilde{v}_i\widetilde{\zeta}_i\right]^{4p} \\
\leq \;& \frac{1}{(n-1)^{4p}} E\left[\sum_{i=1}^{n-1}\widetilde{v}_i\widetilde{\zeta}_i\right]^{4p} \\
= \;& \frac{1}{(n-1)^{4p}} \sum_{1\leq\, i_1,\ldots,i_{4p}\leq\, n-1} E\left[\prod\nolimits_{j=1}^{4p}\left(\widetilde{v}_{i_j}\widetilde{\zeta}_{i_j}\right)\right].
\end{aligned} \tag{22}$$

Observe that, $E\left[\widetilde{v}_{i_j}\widetilde{\zeta}_{i_\ell}\right] = 0$ for all $j \neq \ell$, since $E\left[\widetilde{v}_{i_j}\right] = E\left[\widetilde{\zeta}_{i_j}\right] = 0$ for all i_j and $\widetilde{v}_{i_j}$ is independent of $\widetilde{\zeta}_{i_\ell}$ for all $j \neq \ell$. In consequence, each $\widetilde{v}_{i_j}\widetilde{\zeta}_{i_j}$ must appear at least twice in the product $\prod_{j=1}^{4p}\left(\widetilde{v}_{i_j}\widetilde{\zeta}_{i_j}\right)$; otherwise, the expectation of this product is equal to zero. Hence, at most $2p$ distinct factors of the form $\widetilde{v}_{i_j}\widetilde{\zeta}_{i_j}$ will appear in the product $\prod_{j=1}^{4p}\left(\widetilde{v}_{i_j}\widetilde{\zeta}_{i_j}\right)$ if its expectation is nonzero. To account for the different cases, consider a product with $2p-r$ distinct factors, where $r = 0, 1, ..., 2p-1$; and let N_r be the number of such products, i.e., N_r is the number of ways that one can assign $i_1, ..., i_{4p}$ from the set $\{1, ..., n\}$ such that each i_j appears at least twice and such that exactly $2p-r$ different integers are selected. Moreover, note that since $\widetilde{v}_{i_j}$ and $\widetilde{\zeta}_{i_j}$ are normally distributed, so that, in particular, moments of all order exist. From these facts, we deduce that there exists some positive constant $\overline{C}_p$ (depending on p) such that

$$\begin{aligned} & E\left[v_*'\zeta_*\right]^{4p} \\ \leq\ & \frac{1}{(n-1)^{4p}} \sum_{1\leq i_1,\ldots,i_{4p}\leq n} E\left[\prod_{j=1}^{4p}\left(\widetilde{v}_{i_j}\widetilde{\zeta}_{i_j}\right)\right] \\ \leq\ & \frac{\overline{C}_p}{(n-1)^{4p}} \sum_{r=0}^{2p-1} N_r. \end{aligned}$$

Now, a crude upper bound for N_r is given by

$$N_r \leq \binom{n}{2p-r}(2p-r)^{4p} \leq \frac{n^{2p-r}}{(2p-r)!}(2p-r)^{4p} \leq (en)^{2p-r}(2p-r)^{2p+r} \leq (en)^{2p-r}(2p)^{2p+r},$$

where the third inequality above has made use of the inequality $(2p-r)! \geq (2p-r)^{2p-r}e^{-(2p-r)}$. Applying this upper estimate of N_r, we obtain

$$\begin{aligned} & \frac{1}{(n-1)^{4p}}E\left[\widetilde{v}'\widetilde{\zeta}\right]^{4p} \\ \leq\ & \overline{C}_p\frac{1}{(n-1)^{4p}}\sum_{r=0}^{2p-1}(en)^{2p-r}(2p)^{2p+r} \\ =\ & \overline{C}_p\frac{(2pen)^{2p}}{(n-1)^{4p}}\sum_{r=0}^{2p-1}\left(\frac{2p}{en}\right)^r \\ =\ & \overline{C}_p\left(\frac{n}{n-1}\right)^{4p}\left(\frac{2pe}{n}\right)^{2p}\sum_{r=0}^{2p-1}\left(\frac{2p}{en}\right)^r \\ \leq\ & \overline{C}_p\left(\frac{n}{n-1}\right)^{4p}\left(\frac{2pe}{n}\right)^{2p}\left(1-\frac{2p}{en}\right)^{-1} \quad \text{(for } n \text{ sufficiently large so that } \frac{2p}{n}<1\text{)} \\ <\ & 2\overline{C}_p\left(\frac{n}{n-1}\right)^{4p}\left(\frac{2pe}{n}\right)^{2p} \\ \leq\ & C_p n^{-2p}. \end{aligned} \tag{23}$$

To show part (b), note that

$$E\left[v_*'v_*-1\right]^{4p} = \frac{1}{(n-1)^{4p}}E\left[\sum_{i=1}^{n-1}\left(\widetilde{v}_i^2-1\right)\right]^{4p} = \frac{1}{(n-1)^{4p}}\sum_{1\leq i_1,\ldots,i_{4p}\leq n}E\left[\prod_{j=1}^{4p}\left(\widetilde{v}_{i_j}^2-1\right)\right].$$

Moreover, observe that $E\left[\left(\widetilde{v}_{i_j}^2-1\right)\left(\widetilde{v}_{i_\ell}^2-1\right)\right]=0$ for all $j\neq\ell$ by mutual independence of the elements of the subsequence $\left\{\widetilde{v}_{i_j}^2\right\}$ and by the fact that $E\left[\widetilde{v}_{i_j}^2\right]=1$ for all i_j. It follows again that each $\left(\widetilde{v}_{i_j}^2-1\right)$ must appear at least twice in the product $\prod_{j=1}^{4p}\left(\widetilde{v}_{i_j}^2-1\right)$ for this product to have a nonzero expectation. Hence, we can bound $E\left[v_*'v_*-1\right]^{4p}$ in a way similar to done in the proof of part (a) to obtain

$$E\left[v_*'v_*-1\right]^{4p}\leq\frac{\overline{C}_p}{(n-1)^{4p}}\sum_{r=0}^{2p-1}N_r\leq C_pn^{-2p}\text{ for }n\text{ sufficiently large. }\square$$

Lemma 2: Suppose that $v_*\sim N\left(0,(n-1)^{-1}I_{n-1}\right)$. Then, for any real constant q,

$$E\left|v_*'v_*\right|^{-2q}=1+O\left(n^{-1}\right).$$

Proof: Making use of the change-of-variable formulae (12)-(14), we can evaluate $E\left|v_*'v_*\right|^{-2q}$ as follows:

$$\begin{aligned}
& E\left|v_*'v_*\right|^{-2q}\\
={}& \int_{\mathbb{R}^{n-1}}\left|v_*'v_*\right|^{-2q}\frac{(n-1)^{\frac{1}{2}(n-1)}}{(2\pi)^{(n-1)/2}}\exp\left\{-\frac{(n-1)}{2}v_*'v_*\right\}(dv_*)\\
={}& \int_{V_{1,n-1}}\int_0^\infty\left|rh'hr\right|^{-2q}\frac{(n-1)^{\frac{1}{2}(n-1)}}{(2\pi)^{(n-1)/2}}\exp\left\{-\frac{(n-1)}{2}rh'hr\right\}c_nr^{(n-2)}dr\,[dh]\\
={}& \int_{V_{1,n-1}}\int_0^\infty r^{-4q}\frac{(n-1)^{\frac{1}{2}(n-1)}}{(2\pi)^{(n-1)/2}}\exp\left\{-\frac{(n-1)}{2}r^2\right\}\frac{2\pi^{(n-1)/2}}{\Gamma\left[\frac{1}{2}(n-1)\right]}r^{(n-2)}dr\,[dh]\\
={}& \int_{V_{1,n-1}}\frac{2^{-(n-4)/2}\pi^{1/2}(n-1)^{\frac{1}{2}(n-2)}}{\Gamma\left[\frac{1}{2}(n-1)\right]}\int_0^\infty(2\pi)^{-1/2}r^{(n-4q-2)}(n-1)^{1/2}\exp\left\{-\frac{(n-1)}{2}r^2\right\}dr\,[dh]
\end{aligned}$$

Now, the integral

$$\begin{aligned}
& \int_0^\infty(2\pi)^{-1/2}(n-1)^{1/2}r^{(n-4q-2)}\exp\left\{-\frac{(n-1)}{2}r^2\right\}dr\\
={}& \frac{1}{2}\int_{-\infty}^\infty(2\pi)^{-1/2}(n-1)^{1/2}|r|^{(n-4q-2)}\exp\left\{-\frac{(n-1)}{2}r^2\right\}dr\\
={}& \frac{1}{2}(n-1)^{-(n-4q-2)/2}\frac{2^{(n-4q-2)/2}\Gamma\left[\frac{1}{2}(n-4q-1)\right]}{\sqrt{\pi}}\\
={}& \frac{2^{(n-4q-4)/2}(n-1)^{-(n-4q-2)/2}\Gamma\left[\frac{1}{2}(n-4q-1)\right]}{\sqrt{\pi}}
\end{aligned}$$

for n sufficiently large. Plugging this into the original multiple integral we have

$$
\begin{aligned}
& E\left|v_*' v_*\right|^{-2q} \\
= & \int_{V_{1,n-1}} \frac{2^{-(n-4)/2}\pi^{1/2}\left(n-1\right)^{\frac{1}{2}(n-2)}}{\Gamma\left[\frac{1}{2}\left(n-1\right)\right]} \int_0^{\infty} \left(2\pi\right)^{-1/2} r^{(n-4q-2)} \left(n-1\right)^{1/2} \exp\left\{-\frac{(n-1)}{2} r^2\right\} dr \, [dh] \\
= & \frac{2^{-(n-4)/2}\pi^{1/2}\left(n-1\right)^{\frac{1}{2}(n-2)}}{\Gamma\left[\frac{1}{2}\left(n-1\right)\right]} \frac{2^{(n-4q-4)/2}\left(n-1\right)^{-(n-4q-2)/2}\Gamma\left[\frac{1}{2}\left(n-4q-1\right)\right]}{\sqrt{\pi}} \int_{V_{1,n-1}} [dh] \\
= & 2^{-2q}\left(n-1\right)^{2q} \frac{\Gamma\left[\frac{1}{2}\left(n-4q-1\right)\right]}{\Gamma\left[\frac{1}{2}\left(n-1\right)\right]}
\end{aligned}
$$

where the last equality follows from the fact that

$$\int_{V_{1,n-1}} [dh] = 1.$$

Next, note that by the Stirling approximation

$$
\begin{aligned}
& \Gamma\left[\frac{1}{2}\left(n-4q-1\right)\right] \\
= & \left(\frac{4\pi}{n-4q-1}\right)^{1/2}\left(\frac{n-4q-1}{2e}\right)^{(n-4q-1)/2}\left(1+O\left(n^{-1}\right)\right) \\
= & \left(4\pi\right)^{1/2} n^{(n-4q-2)/2}\left(2e\right)^{-(n-4q-1)/2}\left(1-\frac{4q+1}{n}\right)^{n/2}\left(1+O\left(n^{-1}\right)\right) \\
= & \left(4\pi\right)^{1/2} n^{(n-4q-2)/2}\left(2e\right)^{-(n-4q-1)/2} e^{-(4q+1)/2}\left(1+O\left(n^{-1}\right)\right) \\
= & \left(\frac{n}{2e}\right)^{(n-4q-2)/2}\left(\frac{2\pi}{e^{2(2q+1)}}\right)^{1/2}\left(1+O\left(n^{-1}\right)\right)
\end{aligned}
$$

Similarly,

$$\Gamma\left[\frac{1}{2}\left(n-1\right)\right] = \left(\frac{n}{2e}\right)^{(n-2)/2}\left(\frac{2\pi}{e^2}\right)^{1/2}\left(1+O\left(n^{-1}\right)\right).$$

Putting things together, we have

$$
\begin{aligned}
& E\left|v_*'v_*\right|^{-2q} \\
= \; & \int_{\mathbb{R}^{n-1}} \left|v_*'v_*\right|^{-2q} \frac{(n-1)^{\frac{1}{2}(n-1)}}{(2\pi)^{(n-1)/2}} \exp\left\{-\frac{(n-1)}{2} v_*'v_*\right\} (dv_*) \\
= \; & 2^{-2q} (n-1)^{2q} \frac{\Gamma\left[\frac{1}{2}(n-4q-1)\right]}{\Gamma\left[\frac{1}{2}(n-1)\right]} \\
= \; & 2^{-2q} (n-1)^{2q} \left(\frac{n}{2e}\right)^{(n-4q-2)/2} \left(\frac{2\pi}{e^{2(2q+1)}}\right)^{1/2} \left(\frac{n}{2e}\right)^{-(n-2)/2} \left(\frac{2\pi}{e^2}\right)^{-1/2} \left(1+O\left(n^{-1}\right)\right) \\
= \; & 2^{-2q} (n-1)^{2q} \left(\frac{n}{2e}\right)^{-2q} e^{-2q} \left(1+O\left(n^{-1}\right)\right) \\
= \; & \left(1-\frac{1}{n}\right)^{2q} \left(1+O\left(n^{-1}\right)\right). \\
= \; & 1+O\left(n^{-1}\right). \;\square
\end{aligned}
$$

1 References

Bekker, P.A. (1994): Alternative approximations to the distributions of instrumental variable estimators, *Econometrica*, 62, 657-681.

Fuller, W.A. (1977): Some properties of a modification of the limited information estimator, *Econometrica*, 45, 939-954.

Hahn, J., J.A. Hausman, and G.M. Kuersteiner (2004): Estimation with Weak Instruments: Accuracy of Higher-order Bias and MSE Approximations, *Econometrics Journal, Volume 7.*

Hausman, J.A., W.K. Newey, T. Woutersen, J. Chao, and N.R. Swanson (2012): IV estimation with heteroskedasticity and many instruments, *Quantitative Economics*, forthcoming.

Kunitomo, N. (1980): Asymptotic expansions of distributions of estimators in a linear functional relationship and simultaneous equations, *Journal of the American Statistical Association,* 75, 693-700.