Appendix B: Supplemental Material

The first two parts of this supplementary appendix illustrates how to implement the testing procedures proposed in the paper and cluster-robust GMM bootstrap procedure in Hall and Horowitz (1996) in the context of linear dynamic panel model under clustered dependence. Next, we presents some further numerical results, which are referenced in the paper. The last one applies our proposed test procedures to make inference in empirical data to revisit the study of Emran and Hou (2013).

B.1 Application to Linear Dynamic Panel Model

Consider

$$y_{it} = \gamma y_{it-1} + x_{it}'\beta + \eta_i + u_{it}, \tag{B.1}$$

for $i = 1, ..., n$, $t = 1, ..., T$, where $x_{it} = (x_{it}^1, ..., x_{it}^{d-1})' \in \mathbb{R}^{d-1}$. The unknown parameter vector is $\theta = (\gamma, \beta')' \in \mathbb{R}^d$. We assume that the vector of regressors $w_{it} = (y_{it-1}, x_{it}')'$ is correlated with η_i and is predetermined with respect to u_{it}, i.e., $E(w_{it}u_{it+s}) = 0$ for $s = 0, ..., T - t$.

When T is small, popular panel estimators such as the fixed-effects estimator or first-differenced estimator suffer from the Nickel bias (Nickel, 1981). Anderson and Hsiao (1981) consider the first-differenced equation

$$\Delta y_{it} = \Delta w_{it}'\theta + \Delta u_{it},\ t = 2, ..., T$$

and propose a consistent IV estimator that employs the lagged w_{it} as the instrument. Building upon Anderson and Hsiao (1981), Arellano and Bond (1991, AB hereafter) examine the problem in a GMM framework and find $dT(T-1)/2$ sequential instruments:

$$\underset{(T-1)\times d(T-1)T/2}{Z_i} = diag(z_{i2}', ..., z_{iT}'),$$

where $z_{it} = (y_{i0}, ..., y_{it-2}, x_{i1}', ..., x_{it-1}')', 2 \leq t \leq T$. The moment conditions are then given by

$$E\left(Z_i'\Delta u_i\right) = 0,$$

where Δu_i is the $(T-1)$ vector $(\Delta u_{i2}, ..., \Delta u_{iT})'$. The original AB method assumes away cross-sectional dependence, but clustered dependence can be easily accommodated. Here we assume that the moment vector $\{Z_i'\Delta u_i\}_{i=1}^n$ can be partitioned into independent clusters. That is, $\{Z_i'\Delta u_i\}_{i=1}^n = \cup_{g=1}^G \cup_{k=1}^L \{Z_k^{g\prime}\Delta u_k^g\}$ with $Z_k^{g\prime}\Delta u_k^g$ and $Z_l^{h\prime}\Delta u_l^h$ being independent for all $g \neq h$.

The first-step GMM estimator with initial weighting matrix W_n^{-1} is given by

$$\hat{\theta}_1 = \left(\Delta w' Z W_n^{-1} Z'\Delta w\right)^{-1} \Delta w' Z W_n^{-1} Z'\Delta y,$$

where Z' is the $dT(T-1)/2 \times n(T-1)$ matrix $(Z_1', Z_2', ..., Z_n')$, Δw_i is the $(T-1) \times d$ matrix $(\Delta w_{i2,} ..., \Delta w_{iT})'$, Δy_i is the $(\Delta y_{i2,} ..., \Delta y_{iT})'$, Δw and Δy are $(\Delta w_1', ..., \Delta w_n')'$ and $(\Delta y_1', ..., \Delta y_n')'$, respectively. The examples of W_n's can be $Z'Z/n$ for 2SLS and $n^{-1}\sum_{i=1}^n Z_i'HZ_i$ where H is a matrix that consists with 2's on the main diagonal, with -1's on the main diagonal, and zeros elsewhere.

The Wald statistic[5] for testing $H_0 : R\theta_0 = r$ vs $H_1 : R\theta_0 \neq r$ is given by

$$F(\hat{\theta}_1) := \frac{1}{p}(R\hat{\theta}_1 - r)' \left\{R\widehat{var}(\hat{\theta}_1)R'\right\}^{-1} (R\hat{\theta}_1 - r),$$

[5] The formula for the t statistic, which is omitted here, can be similarly constructed.

where

$$\widehat{var}(\hat{\theta}_1) = n\left(\Delta w' Z W_n^{-1} Z' \Delta w\right)^{-1}\left(\Delta w' Z W_n^{-1} \hat{\Omega}(\hat{\theta}_1) W_n^{-1} Z' \Delta w\right)\left(\Delta w' Z W_n^{-1} Z' \Delta w\right)^{-1}.$$

Let $Z_{(g)}$ be the $L(T-1) \times dT(T-1)/2$ matrix obtained by stacking all Z_i's belonging to cluster g. Similarly, let $\Delta \hat{u}_{(g)}$ be the $\sqrt{L}(T-1)$ stacked vector of the estimated first-differenced errors $\Delta \hat{u}_i = \Delta y_i - \Delta w_i' \hat{\theta}_1$. Then, in the presence of clustered dependence, the CCE and centered CCE are constructed as

$$\hat{\Omega}(\hat{\theta}_1) = \frac{1}{G} \sum_{g=1}^{G} \left(\frac{Z_{(g)}' \Delta \hat{u}_{(g)}}{\sqrt{L}} \right) \left(\frac{Z_{(g)}' \Delta \hat{u}_{(g)}}{\sqrt{L}} \right)'$$

and

$$\hat{\Omega}^c(\hat{\theta}_1) = \frac{1}{G} \sum_{g=1}^{G} \left(\frac{Z_{(g)}' \Delta \hat{u}_{(g)}}{\sqrt{L}} - \frac{1}{G} \sum_{\tilde{g}=1}^{G} \frac{Z_{(\tilde{g})}' \Delta \hat{u}_{(\tilde{g})}}{\sqrt{L}} \right) \left(\frac{Z_{(g)}' \Delta \hat{u}_{(g)}}{\sqrt{L}} - \frac{1}{G} \sum_{\tilde{g}=1}^{G} \frac{Z_{(\tilde{g})}' \Delta \hat{u}_{(\tilde{g})}}{\sqrt{L}} \right)'.$$

Using the centered CCE $\hat{\Omega}^c(\hat{\theta}_1)$ as the weighting matrix, the two-step GMM estimator $\hat{\theta}_2^c$ is

$$\hat{\theta}_2^c = \left(\Delta w' Z \left[\hat{\Omega}^c(\hat{\theta}_1) \right]^{-1} Z' \Delta w \right)^{-1} \Delta w' Z \left[\hat{\Omega}^c(\hat{\theta}_1) \right]^{-1} Z' \Delta y,$$

and the t and Wald statistics for $\hat{\theta}_2^c$ is

$$F_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c) = \frac{1}{p}(R\hat{\theta}_2^c - r)' \{ R \widehat{var}_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2) R' \}^{-1} (R\hat{\theta}_2^c - r),$$

$$\widehat{var}_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c) = n \left\{ \Delta w' Z \left[\hat{\Omega}^c(\hat{\theta}_1) \right]^{-1} Z' \Delta w \right\}^{-1}.$$

Under the conventional large-G asymptotics, both $F(\hat{\theta}_1)$ and $F_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c)$ are asymptotically χ_p^2/p. Under our small-G asymptotics, we have

$$F(\hat{\theta}_1) \xrightarrow{d} \frac{G}{G-p} F_{p,G-p} \text{ and}$$

$$F_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c) \xrightarrow{d} \frac{G}{G-p-q} F_{p,G-p-q}\left(\|\Delta\|^2 \right). \tag{B.2}$$

In addition to utilizing these new approximations, we suggest a variance correction in order to capture the higher order effect of $\hat{\theta}_1$ on $\hat{\Omega}^c(\hat{\theta}_1)$. The finite sample corrected variance is

$$\widehat{var}^{\text{adj}}_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c) = \widehat{var}_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c) + \widehat{\mathcal{E}}_n \widehat{var}_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c) + \widehat{var}_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c) \widehat{\mathcal{E}}_n' + \widehat{\mathcal{E}}_n \widehat{var}(\hat{\theta}_1) \widehat{\mathcal{E}}_n', \tag{B.3}$$

where the j-th column is given by

$$\widehat{\mathcal{E}}_n[., j] = -\left(\Delta w' Z \left[\hat{\Omega}^c(\hat{\theta}_1) \right]^{-1} Z' \Delta w \right)^{-1} \Delta w' Z \left[\hat{\Omega}^c(\hat{\theta}_1) \right]^{-1} \times$$
$$\left. \frac{\partial \hat{\Omega}^c(\theta)}{\partial \theta_j} \right|_{\theta = \hat{\theta}_1} \left[\hat{\Omega}^c(\hat{\theta}_1) \right]^{-1} Z' \Delta \hat{u}_2,$$
$$\Delta \hat{u}_2 = \Delta y - \Delta w \hat{\theta}_2^c,$$

and

$$\left.\frac{\partial\hat{\Omega}^c(\theta)}{\partial\theta_j}\right|_{\theta=\hat{\theta}_1} = \Upsilon_j(\hat{\theta}_1) + \Upsilon_j'(\hat{\theta}_1),$$

$$\Upsilon_j(\hat{\theta}_1) = -\frac{1}{G}\sum_{g=1}^{G}\left(\frac{Z_{(g)}'\Delta w_{j,(g)}}{\sqrt{L}} - \frac{1}{G}\sum_{\tilde{g}=1}^{G}\frac{Z_{(\tilde{g})}'\Delta w_{j,(\tilde{g})}}{\sqrt{L}}\right)\left(\frac{Z_{(g)}'\Delta\hat{u}_{(g)}}{\sqrt{L}} - \frac{1}{G}\sum_{\tilde{g}=1}^{G}\frac{Z_{(\tilde{g})}'\Delta\hat{u}_{(\tilde{g})}}{\sqrt{L}}\right)',$$

$$\underset{(T-1)\times d}{\Delta w_i} = (\Delta w_{1,i},...,\Delta w_{d,i}) \text{ and } \underset{l(T-1)\times d}{\Delta w_{(g)}} = (\Delta w_{1,(g)},...,\Delta w_{d,(g)})$$

for each $j=1,...,d$. Here, $\Delta w_{(g)}$ is a $L(T-1)\times d$ matrix that stacks $\{\Delta w_i\}_{i=1}^n$ belonging to the group g.

Based on the finite sample corrected variance estimator in (B.3) and the usual J statistic, we construct the modified Wald and t statistics

$$\tilde{F}^{\text{adj}}_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c) = \frac{G-p-q}{G}\frac{F^{\text{adj}}_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c)}{1+\frac{1}{G}J(\hat{\theta}_2^c)}, \tag{B.4}$$

where

$$\begin{aligned} J(\hat{\theta}_2^c) &= n\cdot g_n(\hat{\theta}_2^c)'\left(\hat{\Omega}^c(\hat{\theta}_2^c)\right)^{-1}(\theta)g_n(\hat{\theta}_2^c) \\ &= \left(\frac{Z'\Delta u(\hat{\theta}_2^c)}{\sqrt{n}}\right)'\left[\hat{\Omega}^c(\hat{\theta}_2^c)\right]^{-1}\left(\frac{Z'\Delta u(\hat{\theta}_2^c)}{\sqrt{n}}\right). \end{aligned}$$

From the t and F limit theory in Section 4 and the small-G approximation of the finite sample corrected variance in Section 6, we have

$$\tilde{F}^{\text{adj}}_{\hat{\Omega}^c(\hat{\theta}_1)}(\hat{\theta}_2^c) \xrightarrow{d} F_{p,G-p-q} \tag{B.5}$$

and

$$\frac{G-q}{Gq}J(\hat{\theta}_2^c) \xrightarrow{d} F_{q,G-q}.$$

The critical values from the t and F distributions are readily available from statistical tables.

B.2 Procedure for cluster-robust Hall and Horowitz (1996) two-step GMM-bootstrap

Conditioning on the original sample $\{Z_{(g)}, \Delta u_{(g)}(\theta), \Delta y_{(g)}, \Delta w_{(g)}\}_{g=1}^G$, let $\{Z_i^*, \Delta y_i^*, \Delta u_i^*(\theta), \Delta w_i^*\}_{i=1}^n = \cup_{g=1}^G\{Z_{(g)}^*, \Delta y_{(g)}^*, \Delta u_{(g)}^*(\theta), \Delta w_{(g)}^*\}$ be a cluster-wise bootstrap sample. Given the b-th resampled data $\cup_{g=1}^G\{Z_{(g)}^*, \Delta y_{(g)}^*, \Delta u_{(g)}^*(\theta), \Delta w_{(g)}^*\}$, we can implement the GMM bootstrap procedure in Hall and Horowitz (1996) as follows.

Step 1: With the recentered moment function $(Z_i^*)'\Delta u_i^*(\theta) - E^*[(Z_i^*)'\Delta u_i^*(\hat{\theta}_1)]$, obtain a bootstrap version of the initial estimator:

$$\hat{\theta}_{1,(b)}^* = \left[(\Delta w^*)'Z^*[W_n^*]^{-1}(Z^*)'\Delta w^*\right]^{-1}(\Delta w^*)'Z^*[W_n^*]^{-1}\left((Z^*)'\Delta y^* - Z'\Delta u(\hat{\theta}_1)\right).$$

, where $W_n^* = n^{-1}\sum_{i=1}^n (Z_i^*)' Z_i^*$.

Step 2: Obtain a bootstrap version of the two-step GMM weighting matrix $\hat{\Omega}^{c*}(\hat{\theta}^*_{1,(b)})$ with the recentered moment process $(Z_i^*)'\Delta u_i^*(\theta) - E^*[(Z_i^*)'\Delta u_i^*(\hat{\theta}_2)]$:

$$\hat{\Omega}^{c*}(\hat{\theta}^*_{1,(b)};\hat{\theta}_2) = \frac{1}{G}\sum_{g=1}^{G}\left[\left(\frac{(Z^*_{(g)})'\Delta u^*_{(g)}(\hat{\theta}^*_{1,(b)})}{\sqrt{L}} - \frac{1}{G}\sum_{h=1}^{G}\frac{Z'_{(h)}\Delta u_{(h)}(\hat{\theta}_2)}{\sqrt{L}}\right)\times \left(\frac{(Z^*_{(g)})'\Delta u^*_{(g)}(\hat{\theta}^*_{1,(b)})}{\sqrt{L}} - \frac{1}{G}\sum_{h=1}^{G}\frac{Z'_{(h)}\Delta u_{(h)}(\hat{\theta}_2)}{\sqrt{L}}\right)'\right].$$

, and corresponding two-step GMM estimator:

$$\hat{\theta}^*_{2,(b)} = \left[(\Delta w^*)' Z^* \left[\hat{\Omega}^{c*}(\hat{\theta}^*_{1,(b)})\right]^{-1}(Z^*)'\Delta w^*\right]^{-1}(\Delta w^*)' Z^*\left[\hat{\Omega}^{c*}(\hat{\theta}^*_{1,(b)})\right]^{-1}\times \left[(Z^*)'\Delta y^* - Z'\Delta u(\hat{\theta}_1)\right].$$

Step 3: Construct the b-th bootstrap version of t statistic:

$$t^*(\hat{\theta}^*_{2,(b)}) = \frac{R(\hat{\theta}^*_{2,(b)} - \hat{\theta}_2)}{\sqrt{R\widehat{var}_{\hat{\Omega}^{c*}(\hat{\theta}^*_{1,(b)};\hat{\theta}_2)}(\hat{\theta}^*_{2,(b)})R'}},$$

where $\widehat{var}_{\hat{\Omega}^{c*}(\hat{\theta}^*_{1,(b)})}(\hat{\theta}^*_{2,(b)}) = n\left((\Delta w^*)' Z^*\left[\hat{\Omega}^{c*}(\hat{\theta}^*_{1,(b)};\hat{\theta}_2)\right]^{-1}(Z^*)'\Delta w^*\right)^{-1}$, and J statistic:

$$J^*(\hat{\theta}^*_{2,(b)};\hat{\theta}_2) = \left[\frac{Z^{*\prime}\Delta u^*(\hat{\theta}^*_{2,(b)}) - Z'\Delta u(\hat{\theta}_2)}{\sqrt{n}}\right]'\left[\hat{\Omega}^{c*}(\hat{\theta}^*_{2,(b)};\hat{\theta}_2)\right]^{-1}\times \left[\frac{Z^{*\prime}\Delta u^*(\hat{\theta}^*_{2,(b)}) - Z'\Delta u(\hat{\theta}_2)}{\sqrt{n}}\right].$$

Step 4: After repeating 1∼3 steps B-times, compute the two-side bootstrap p-values for t and J statistics with

$$\hat{p}_{t,\text{HH}} = \frac{1}{B}\sum_{b=1}^{B} 1\left(\left|t^*(\hat{\theta}^*_{2,(b)})\right| > \left|t(\hat{\theta}_2)\right|\right) \text{ and } \hat{p}_{J,\text{HH}} = \frac{1}{B}\sum_{b=1}^{B} 1\left(J^*(\hat{\theta}^*_{2,(b)}) > J(\hat{\theta}_2)\right).$$

Then, we reject Reject the null hypothesis $H_0 : R\theta = r$ iff

$$\hat{p}_{t,HH} \leqq \alpha$$

, and reject the null hypothesis of J test $H_0 : E\left(Z_i'\Delta u_i\right) = 0$ iff

$$\hat{p}_{J,HH} \leqq \alpha.$$

B.3 Further simulation results when L=100

Table B.1: Empirical size of GMM tests based on the centered CCE when $L = 100$, $G = 35, 50$, $p = 1 \sim 3$, and $m = 12, 24$, with $T = 4$.

$G = 35$	Test statistic	Critical values	$m = 24$			$m = 12$		
			$p = 1$	$p = 2$	$p = 3$	$p = 1$	$p = 2$	$p = 3$
First-step	F_1	$\chi_p^{1-\alpha}/p$	0.204	0.214	0.217	0.169	0.171	0.181
	$\frac{G-p}{G}F_1$	$F_{p,G-p}^{1-\alpha}$	0.177	0.163	0.148	0.147	0.127	0.119
Two-step	F_2	$\chi_p^{1-\alpha}/p$	0.444	0.648	0.782	0.169	0.235	0.290
	$\tilde{F}_2$	$F_{p,G-p-q}^{1-\alpha}$	0.049	0.049	0.047	0.061	0.054	0.055
	$\tilde{F}_2^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.043	0.039	0.036	0.053	0.050	0.050
	F_2	HH-Bootstrap	0.000	0.000	0.000	0.037	0.025	0.021
CU-type	$F_{\text{CU-GEE}}$	$\chi_p^{1-\alpha}/p$	0.457	0.646	0.783	0.180	0.242	0.295
	$\tilde{F}_{\text{CU-GEE}}$	$F_{p,G-p-q}^{1-\alpha}$	0.060	0.059	0.060	0.065	0.058	0.058
	$\tilde{F}_{\text{CU-GEE}}^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.061	0.062	0.059	0.056	0.051	0.052
	$F_{\text{CU-GMM}}$	$\chi_p^{1-\alpha}/p$	0.472	0.668	0.793	0.177	0.238	0.287
	$\tilde{F}_{\text{CU-GMM}}$	$F_{p,G-p-q}^{1-\alpha}$	0.069	0.064	0.061	0.065	0.058	0.057
	$\tilde{F}_{\text{CU-GMM}}^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.072	0.063	0.066	0.054	0.049	0.050
J test	J	$\chi_q^{1-\alpha}$	–	0.002	–	–	0.037	–
	$\frac{1}{G}J$	$B_{q/2,(G-q)/2}^{1-\alpha}$	–	0.112	–	–	0.062	–
	J^c	$\chi_q^{1-\alpha}$	–	0.800	–	–	0.206	–
	$\frac{G-q}{Gq}J^c$	$F_{q,G-q}^{1-\alpha}$	–	0.058	–	–	0.057	–
	J^c	HH-Bootstrap	–	0.734	–	–	0.125	–

$G = 50$	Test statistic	Critical values	$m = 24$			$m = 12$		
First-step	F_1	$\chi_p^{1-\alpha}/p$	0.166	0.168	0.169	0.149	0.144	0.145
	$\frac{G-p}{G}F_1$	$F_{p,G-p}^{1-\alpha}$	0.148	0.139	0.121	0.135	0.116	0.107
Two-step	F_2	$\chi_p^{1-\alpha}/p$	0.282	0.401	0.507	0.126	0.159	0.193
	$\tilde{F}_2$	$F_{p,G-p-q}^{1-\alpha}$	0.060	0.058	0.051	0.059	0.059	0.055
	$\tilde{F}_2^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.059	0.054	0.048	0.055	0.055	0.052
	F_2	HH-Bootstrap	0.016	0.004	0.001	0.045	0.042	0.036
CU-type	$F_{\text{CU-GEE}}$	$\chi_p^{1-\alpha}/p$	0.286	0.408	0.510	0.126	0.161	0.194
	$\tilde{F}_{\text{CU-GEE}}$	$F_{p,G-p-q}^{1-\alpha}$	0.062	0.061	0.056	0.062	0.060	0.056
	$\tilde{F}_{\text{CU-GEE}}^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.061	0.058	0.049	0.056	0.053	0.053
	$F_{\text{CU-GMM}}$	$\chi_p^{1-\alpha}/p$	0.294	0.411	0.516	0.121	0.154	0.193
	$\tilde{F}_{\text{CU-GMM}}$	$F_{p,G-p-q}^{1-\alpha}$	0.066	0.062	0.056	0.055	0.056	0.056
	$\tilde{F}_{\text{CU-GMM}}^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.061	0.056	0.049	0.052	0.051	0.051
J test	J	$\chi_q^{1-\alpha}/q$	–	0.016	–	–	0.036	–
	$\frac{1}{G}J$	$B_{q/2,(G-q)/2}^{1-\alpha}$	–	0.067	–	–	0.053	–
	J^c	$\chi_q^{1-\alpha}/q$	–	0.562	–	–	0.151	–
	$\frac{G-q}{Gq}J^c$	$F_{q,G-q}^{1-\alpha}$	–	0.051	–	–	0.049	–
	J^c	HH-Bootstrap	–	0.284	–	–	0.121	–

See footnote to Table 1.

Table B.2: Empirical size of GMM tests based on the centered CCE when $L = 100$, number of clusters $G = 70, 100$, number of joint hypothesis $p = 1 \sim 3$ and number of moment conditions $m = 12, 24$, with $T = 4$.

$G = 70$	Test statistic	Critical values	$m = 12$			$m = 24$		
			$p = 1$	$p = 2$	$p = 3$	$p = 1$	$p = 2$	$p = 3$
First-step	F_1	$\chi_p^{1-\alpha}/p$	0.131	0.131	0.131	0.120	0.122	0.118
	$\frac{G-p}{G}F_1$	$F_{p,G-p}^{1-\alpha}$	0.119	0.112	0.101	0.108	0.104	0.092
Two-step	F_2	$\chi_p^{1-\alpha}/p$	0.189	0.264	0.319	0.092	0.114	0.130
	$\tilde{F}_2$	$F_{p,G-p-q}^{1-\alpha}$	0.057	0.060	0.057	0.051	0.054	0.053
	$\tilde{F}_2^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.053	0.057	0.053	0.048	0.051	0.051
	F_2	HH-Bootstrap	0.038	0.027	0.018	0.045	0.042	0.040
CU-type	$F_{\text{CU-GEE}}$	$\chi_p^{1-\alpha}/p$	0.190	0.265	0.324	0.093	0.116	0.130
	$\tilde{F}_{\text{CU-GEE}}$	$F_{p,G-p-q}^{1-\alpha}$	0.057	0.061	0.059	0.052	0.056	0.054
	$\tilde{F}_{\text{CU-GEE}}^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.053	0.057	0.056	0.048	0.050	0.051
	$F_{\text{CU-GMM}}$	$\chi_p^{1-\alpha}/p$	0.191	0.261	0.324	0.090	0.113	0.126
	$\tilde{F}_{\text{CU-GMM}}$	$F_{p,G-p-q}^{1-\alpha}$	0.059	0.060	0.057	0.050	0.049	0.051
	$\tilde{F}_{\text{CU-GMM}}^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.052	0.055	0.051	0.048	0.047	0.048
J test	J	$\chi_q^{1-\alpha}/q$	–	0.023	–	–	0.035	–
	$\frac{1}{G}J$	$B_{q/2,(G-q)/2}^{1-\alpha}$	–	0.052	–	–	0.045	–
	J^c	$\chi_q^{1-\alpha}/q$	–	0.367	–	–	0.102	–
	$\frac{G-q}{Gq}J^c$	$F_{q,G-q}^{1-\alpha}$	–	0.048	–	–	0.045	–
	J^c	HH-Bootstrap	–	0.240	–	–	0.093	–

$G = 100$	Test statistic	Critical values	$m = 12$			$m = 24$		
			$p = 1$	$p = 2$	$p = 3$	$p = 1$	$p = 2$	$p = 3$
First-step	F_1	$\chi_p^{1-\alpha}/p$	0.118	0.117	0.114	0.104	0.098	0.099
	$\frac{G-p}{G}F_1$	$F_{p,G-p}^{1-\alpha}$	0.111	0.103	0.097	0.100	0.085	0.086
Two-step	F_2	$\chi_p^{1-\alpha}/p$	0.142	0.181	0.218	0.087	0.093	0.107
	$\tilde{F}_2$	$F_{p,G-p-q}^{1-\alpha}$	0.058	0.062	0.059	0.059	0.051	0.050
	$\tilde{F}_2^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.055	0.061	0.058	0.056	0.049	0.049
	F_2	HH-Bootstrap	0.042	0.042	0.037	0.052	0.046	0.043
CU-type	$F_{\text{CU-GEE}}$	$\chi_p^{1-\alpha}/p$	0.143	0.184	0.221	0.088	0.093	0.107
	$\tilde{F}_{\text{CU-GEE}}$	$F_{p,G-p-q}^{1-\alpha}$	0.058	0.064	0.060	0.060	0.052	0.051
	$\tilde{F}_{\text{CU-GEE}}^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.057	0.060	0.058	0.056	0.049	0.049
	$F_{\text{CU-GMM}}$	$\chi_p^{1-\alpha}/p$	0.136	0.173	0.214	0.080	0.090	0.100
	$\tilde{F}_{\text{CU-GMM}}$	$F_{p,G-p-q}^{1-\alpha}$	0.056	0.057	0.057	0.054	0.048	0.046
	$\tilde{F}_{\text{CU-GMM}}^{\text{adj}}$	$F_{p,G-p-q}^{1-\alpha}$	0.053	0.055	0.055	0.053	0.047	0.046
J test	J	$\chi_q^{1-\alpha}/q$	–	0.034	–	–	0.044	–
	$\frac{1}{G}J$	$B_{q/2,(G-q)/2}^{1-\alpha}$	–	0.058	–	–	0.053	–
	J^c	$\chi_q^{1-\alpha}/q$	–	0.236	–	–	0.100	–
	$\frac{G-q}{Gq}J^c$	$F_{q,G-q}^{1-\alpha}$	–	0.054	–	–	0.053	–
	J^c	HH-Bootstrap	–	0.187	–	–	0.096	–

See footnote to Table 1.

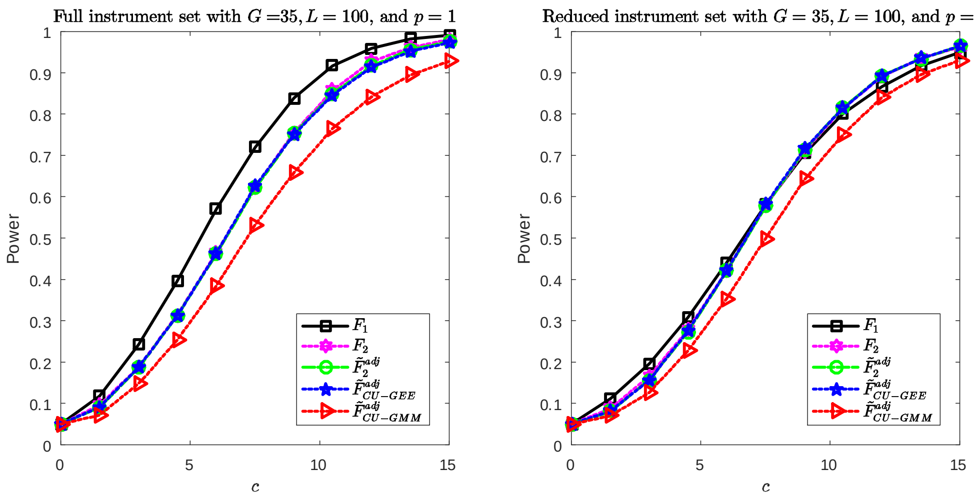

Figure B.1: Size-adjusted power of the first-step (2SLS) and two-step tests with $G = 35$ and $L = 100$.

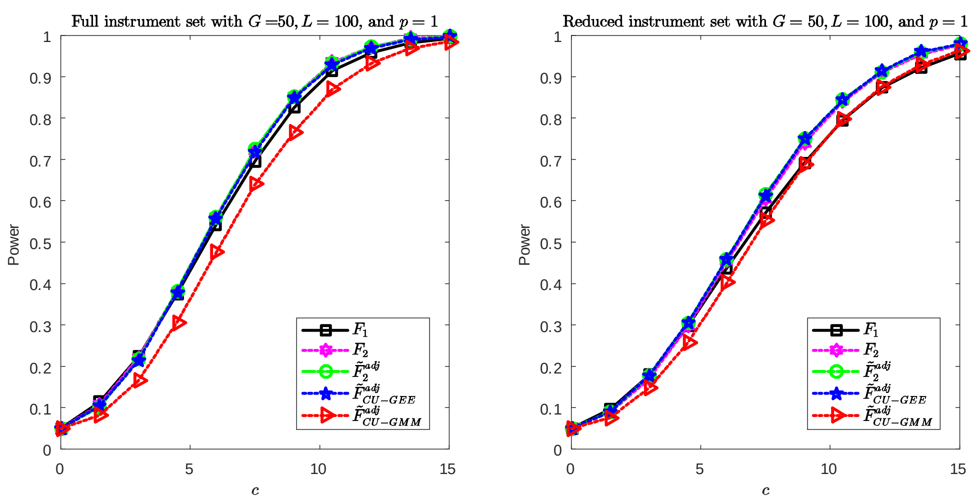

Figure B.2: Size-adjusted power of the first-step (2SLS) and two-step tests with $G = 50$ and $L = 100$.

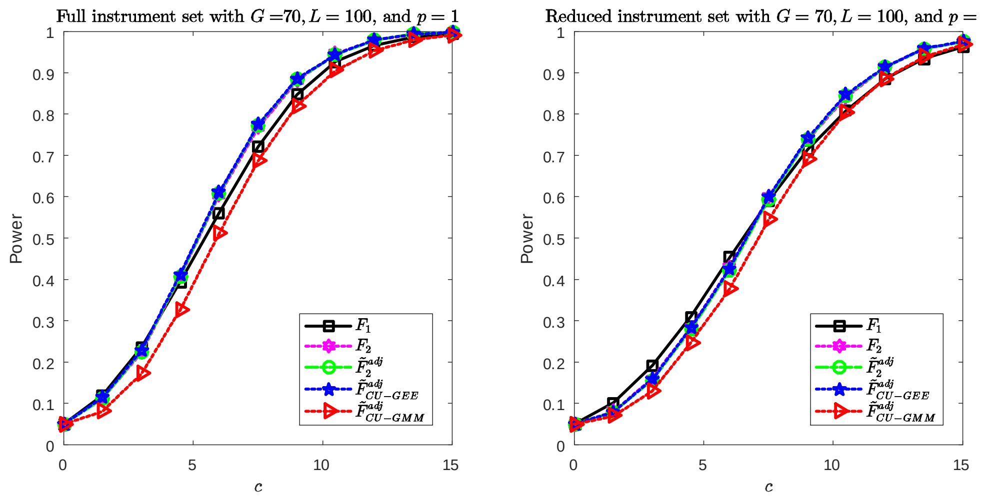

Figure B.3: Size-adjusted power of the first-step (2SLS) and two-step tests with $G = 70$ and $L = 100$.

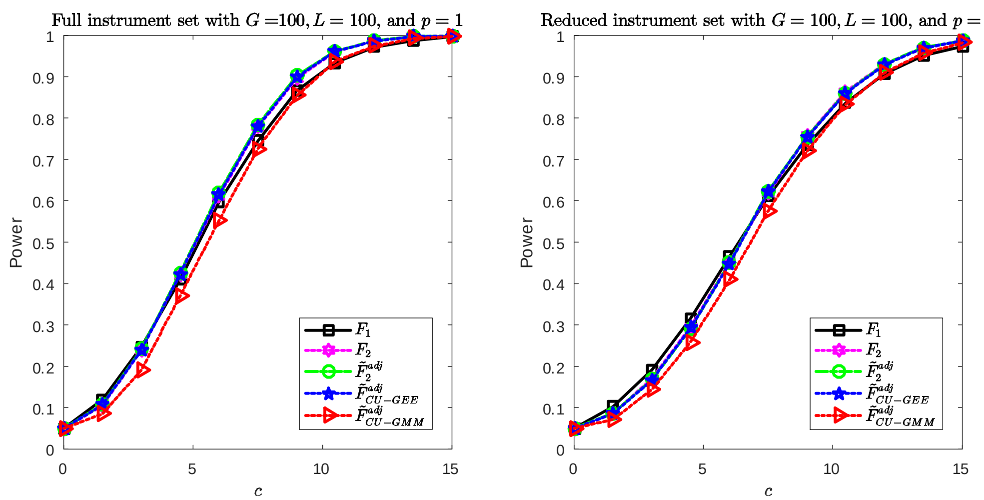

Figure B.4: Size-adjusted power of the first-step (2SLS) and two-step tests with $G = 100$ and $L = 100$.

B.4 Empirical Application

In this section we employ the proposed procedures to revisit the study of Emran and Hou (2013), which investigates the casual effects of access to domestic and international markets on household consumption in rural China. They use a survey data of 7998 rural households across 19 provinces in China. The survey data comes from Chinese Household Income Project (ICPSR 3012) in 1995.[6]

The regression equation for per capita consumption for household i, C_i, in 1995 (yuan) is specified as

$$\begin{aligned} C_i &= \alpha_0 + \alpha_p + \beta_d A_i^d + \beta_s A_i^s + \beta_{ds}(A_i^d \times A_i^s) + X_i' \beta_h \\ &\quad + X_i^{v\prime} \beta_v + X_i^{c'} \beta_c + \epsilon_i, \end{aligned} \tag{B.6}$$

where A_i^d and A_i^s are the log distances of access to domestic (km) and international markets (km), respectively. X_i is the vector of household characteristics that may affect consumption choice, and X_i^v, X_i^c are village, county level controls, respectively, which capture the heterogeneity in economic environments across different regions, and α_p is the province level fixed effect.

Among the unknown parameters in vector $\theta = (\alpha_0, \alpha_p, \beta_m', \beta_h', \beta_v', \beta_c')'$, our focus of interest is $\beta_m = (\beta_d, \beta_s, \beta_{ds})'$ which measures the casual effect of access to domestic and international markets on household consumption in the rural areas. To identify these parameters, Emran and Hou (2013) employs geographic instrumental variables that capture exogenous variations in access to markets, e.g., straight-line distances to the nearest navigable river and coastline, along with the topographic and agroclimatic features of the counties.[7] There are 21 instrument variables and 31 control variables, including province dummy variables so that the number of moment conditions m is 52. The number of estimated parameters d is 34, and the degree of over-identification q is 18. Because of the close economic and cultural ties within the same county in rural Chinese areas, the study clusters the data by the county level and estimates the model using 2SLS and two-step GMM with uncentered cluster-robust weighting matrix. The data set consists of 7462 observations divided into 86 clusters where the number of households vary across from a low of 49 to a high of 270. Statistical inferences in Emran and Hou (2013) are conducted using the large-G asymptotics only. We apply our more accurate asymptotics to Emran and Hou's study. The inference methods we use here are described in Tables B.3 and B.4 which present the test statistics, the reference distribution, and the standard error formula (finite sample corrected or not) for each method. Here we view all corrections, including the degree-of-freedom correction, the J correction, and the finite sample variance corrections as corrections to the variance estimator.

Table B.5 shows the point estimation results, standard error estimates, and associated confidence intervals (CIs) for each of 2SLS and the uncentered and centered two-step GMM estimators. Similar to Emran and Hou (2013), our results show that the better access to domestic and international markets has a substantial positive effect on household consumption, and that the domestic market effect is significantly higher. For the 2SLS method, there are no much differences in confidence interval and standard error between the large-G and small-G results. This is well expected because the number of clusters $G = 86$ is large enough so that the large-G and small-G approximations are close to each other.

The uncentered two-step GMM estimate of the effect of access to domestic market is $\beta_d = -2722.22$. The reported standard error 400.5 is about 40% smaller than that of 2SLS. However,

[6] The data set is downloadable from the *Review of Economics and Statistics* website.

[7] For the detailed description of the control variables and instrument variables, see the appendix in Emran and Hou (2013).

the plain two-step standard error estimate might be downward biased because the variation of the cluster-robust weighting matrix is not considered. The centered two-step GMM estimator gives a smaller effect of market access $\beta_d = -2670.0$ with the modified standard error of 519.2, which is 25% larger than the plain two-step standard error. However, the modified standard error is still smaller than that based on the 2SLS method. So the two-step estimator still enjoys the benefit of using the cluster-robust weighting matrix. The results for other parameters deliver similar qualitative messages. Table B.5 also provides the finite sample corrected standard error estimates of two-step estimators that capture the extra variation of feasible CCE, leading to slightly larger standard errors and wider CIs than the uncorrected ones. Overall, our results suggest that the effect of access to markets may be lower than the previous finding after we take into account the randomness of the estimated optimal GMM weighting matrix.

Table B.3: Summary of the difference between the conventional large-G asympototics and alternative fixed-G asympotics for the first-step (2SLS) and two-step GMM methods.

			2SLS				
Asymptotics	Weight	Test statistic			Reference distribution		
·		Wald	t	·	Wald	t	·
large-G	$Z'Z/n$	$F(\hat{\theta}_1)$	$t(\hat{\theta}_1)$	·	χ^2_p/p	$N(0,1)$	·
small-G	$Z'Z/n$	$\frac{G-p}{G}F(\hat{\theta}_1)$	$\sqrt{\frac{G-1}{G}}t(\hat{\theta}_1)$	·	$\mathcal{F}_{p,G-p}$	t_{G-1}	·
			Two-step GMM				
Asymptotics	Weight	Test statistic			Reference distribution		
·		Wald	t	J	Wald	t	J
large-G	$\hat{\Omega}$	$F(\hat{\theta}_2)$	$t(\hat{\theta}_2)$	$J(\hat{\theta}_2)$	χ^2_p/p	$N(0,1)$	χ^2_q
small-G	$\hat{\Omega}^c$	$\tilde{F}(\hat{\theta}^c_2)$	$\tilde{t}(\hat{\theta}^c_2)$	$\frac{G-q}{Gq}J(\hat{\theta}^c_2)$	$\mathcal{F}_{p,G-p-q}$	t_{G-1-q}	$\mathcal{F}_{q,G-q}$
	$\hat{\Omega}^c$	$\tilde{F}^{\text{adj}}(\hat{\theta}^c_2)$	$\tilde{t}^{\text{adj}}(\hat{\theta}^c_2)$	·	$\mathcal{F}_{p,G-p-q}$	t_{G-1-q}	·

Table B.4: Summary of standard error formula when $p = 1$

		2SLS
	Large-G asymptotics	Small-G asymptotics
standard error	$\left[R\widehat{var}(\hat{\theta}_1)R'\right]^{1/2}$	$\left[\frac{G-1}{G}R\widehat{var}_{\Omega(\hat{\theta}_1)}(\hat{\theta}_1)R'\right]^{1/2}$
		Two-step GMM
	Large-G asymptotics	Small-G asymptotics
standard error	$\left[R\widehat{var}(\hat{\theta}_1)R'\right]^{1/2}$	$\left[\frac{G}{G-1-q}R\widehat{var}_{\hat{\Omega}(\hat{\theta}_1)}(\hat{\theta}^c_2)R'\left(1+(q/G)J(\hat{\theta}^c_2)\right)\right]^{1/2}$
corrected standard error	·	$\left[\frac{G}{G-1-q}R\widehat{var}^{\text{adj}}_{\hat{\Omega}(\hat{\theta}_1)}(\hat{\theta})R'\left(1+(q/G)J(\hat{\theta}^c_2)\right)\right]^{1/2}$

Table B.5: Results for Emran and Hou (2013) data

	2SLS	
Variables	Large-G asymptotics	Small-G asymptotics
Domestic market (A_i^d)	-2713.2 (712.1)	-2713.2 (716.8)
	$[-4109.9, -1316.4]$	$[-4138.0, -1288.0]$
International market (A_i^s)	-1993.5 (514.8)	-1993.5 (517.9)
	$[-3002.5, -984.4]$	$[-3023.10, -963.8]$
Interaction $(A_i^d \times A_i^s)$	345.8 (105.0)	345.8 (105.6)
	$[140.0, 551.7]$	$[135.8, 555.9]$
$H_0 : \beta_d = \beta_s$	-2.3218 (2.02%)	-2.771 (2.26%)
	Two-step GMM	
Variables	Large-G asymptotics	Small-G asymptotics
Domestic market (A_i^d)	-2722.8 (400.5)	-2670.0 (519.2)
		(520.7)
	$[-3507.7, -1937.9]$	$[-3706.2, -1633.8]$
		$[-3709.2, -1630.7]$
International market (A_i^s)	-2000.2 (344.3)	-1981.3 (446.4)
		(447.7)
	$[-2675.0, -1325.5]$	$[-2872.3, -1090.3]$
		$[-2874.9, -1087.7]$
Interaction $(A_i^d \times A_i^s)$	362.7 (68.7)	364.1 (89.1)
		(89.4)
	$[228.0, 497.3]$	$[186.2, 541.9]$
		$[187.5, 542.4]$
$H_0 : \beta_d = \beta_s$	-5.239 (0%)	-3.3318 (0%)
		-3.3217 (0%)
J statistic ($q = 18$)	1.1708 (99.8%)	0.3096 (45.83%)

Notes: Standard errors for 2SLS and the weighting matrix for (centered) two-step GMM estimators are clustered at the county level. Numbers in parentheses are standard errors and intervals are 95% confidence intervals. For hypothesis testing, the numbers in parentheses are p-values.

References

[1] Anderson, T. W. and Hsiao, C. (1981): "Estimation of dynamic models with error components." *Journal of the American statistical Association*, 76(375), 598-606.

[2] Arellano, M. and Bond, S. (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The Review of Economic Studies*, 58(2), 277-297.

[3] Emran, M. S., and Hou, Z. (2013): "Access to markets and rural poverty: evidence from household consumption in China." *Review of Economics and Statistics*, 95(2), 682-697.

[4] Nickell, S. (1981): "Biases in dynamic models with fixed effects." *Econometrica*: Journal of the Econometric Society, 1417-1426.

[5] Hall, P., and Horowitz, J. L. (1996): "Bootstrap critical values for tests based on generalized-method-of-moments estimators." *Econometrica: Journal of the Econometric Society*, 891-916.

the plain two-step standard error estimate might be downward biased because the variation of the cluster-robust weighting matrix is not considered. The centered two-step GMM estimator gives a smaller effect of market access $\beta_d = -2670.0$ with the modified standard error of 519.2, which is 25% larger than the plain two-step standard error. However, the modified standard error is still smaller than that based on the 2SLS method. So the two-step estimator still enjoys the benefit of using the cluster-robust weighting matrix. The results for other parameters deliver similar qualitative messages. Table B.5 also provides the finite sample corrected standard error estimates of two-step estimators that capture the extra variation of feasible CCE, leading to slightly larger standard errors and wider CIs than the uncorrected ones. Overall, our results suggest that the effect of access to markets may be lower than the previous finding after we take into account the randomness of the estimated optimal GMM weighting matrix.

Table B.3: Summary of the difference between the conventional large-G asympototics and alternative fixed-G asympotics for the first-step (2SLS) and two-step GMM methods.

			2SLS				
Asymptotics	Weight	Test statistic			Reference distribution		
·		Wald	t	·	Wald	t	·
large-G	$Z'Z/n$	$F(\hat{\theta}_1)$	$t(\hat{\theta}_1)$	·	χ^2_p/p	$N(0,1)$	·
small-G	$Z'Z/n$	$\frac{G-p}{G}F(\hat{\theta}_1)$	$\sqrt{\frac{G-1}{G}}t(\hat{\theta}_1)$	·	$\mathcal{F}_{p,G-p}$	t_{G-1}	·
			Two-step GMM				
Asymptotics	Weight	Test statistic			Reference distribution		
·		Wald	t	J	Wald	t	J
large-G	$\hat{\Omega}$	$F(\hat{\theta}_2)$	$t(\hat{\theta}_2)$	$J(\hat{\theta}_2)$	χ^2_p/p	$N(0,1)$	χ^2_q
small-G	$\hat{\Omega}^c$	$\tilde{F}(\hat{\theta}^c_2)$	$\tilde{t}(\hat{\theta}^c_2)$	$\frac{G-q}{Gq}J(\hat{\theta}^c_2)$	$\mathcal{F}_{p,G-p-q}$	t_{G-1-q}	$\mathcal{F}_{q,G-q}$
	$\hat{\Omega}^c$	$\tilde{F}^{\text{adj}}(\hat{\theta}^c_2)$	$\tilde{t}^{\text{adj}}(\hat{\theta}^c_2)$	·	$\mathcal{F}_{p,G-p-q}$	t_{G-1-q}	·

Table B.4: Summary of standard error formula when $p = 1$

	2SLS	
	Large-G asymptotics	Small-G asymptotics
standard error	$\left[R\widehat{var}(\hat{\theta}_1)R'\right]^{1/2}$	$\left[\frac{G-1}{G}R\widehat{var}_{\Omega(\hat{\theta}_1)}(\hat{\theta}_1)R'\right]^{1/2}$
	Two-step GMM	
	Large-G asymptotics	Small-G asymptotics
standard error	$\left[R\widehat{var}(\hat{\theta}_1)R'\right]^{1/2}$	$\left[\frac{G}{G-1-q}R\widehat{var}_{\hat{\Omega}(\hat{\theta}_1)}(\hat{\theta}^c_2)R'\left(1+(q/G)J(\hat{\theta}^c_2)\right)\right]^{1/2}$
corrected standard error	·	$\left[\frac{G}{G-1-q}R\widehat{var}^{\text{adj}}_{\hat{\Omega}(\hat{\theta}_1)}(\hat{\theta})R'\left(1+(q/G)J(\hat{\theta}^c_2)\right)\right]^{1/2}$

Table B.5: Results for Emran and Hou (2013) data

2SLS		
Variables	Large-G asymptotics	Small-G asymptotics
Domestic market (A_i^d)	-2713.2 (712.1)	-2713.2 (716.8)
	$[-4109.9, -1316.4]$	$[-4138.0, -1288.0]$
International market (A_i^s)	-1993.5 (514.8)	-1993.5 (517.9)
	$[-3002.5, -984.4]$	$[-3023.10, -963.8]$
Interaction $(A_i^d \times A_i^s)$	345.8 (105.0)	345.8 (105.6)
	$[140.0, 551.7]$	$[135.8, 555.9]$
$H_0 : \beta_d = \beta_s$	-2.3218 (2.02%)	-2.771 (2.26%)
Two-step GMM		
Variables	Large-G asymptotics	Small-G asymptotics
Domestic market (A_i^d)	-2722.8 (400.5)	-2670.0 (519.2)
		(520.7)
	$[-3507.7, -1937.9]$	$[-3706.2, -1633.8]$
		$[-3709.2, -1630.7]$
International market (A_i^s)	-2000.2 (344.3)	-1981.3 (446.4)
		(447.7)
	$[-2675.0, -1325.5]$	$[-2872.3, -1090.3]$
		$[-2874.9, -1087.7]$
Interaction $(A_i^d \times A_i^s)$	362.7 (68.7)	364.1 (89.1)
		(89.4)
	$[228.0, 497.3]$	$[186.2, 541.9]$
		$[187.5, 542.4]$
$H_0 : \beta_d = \beta_s$	-5.239 (0%)	-3.3318 (0%)
		-3.3217 (0%)
J statistic ($q = 18$)	1.1708 (99.8%)	0.3096 (45.83%)

Notes: Standard errors for 2SLS and the weighting matrix for (centered) two-step GMM estimators are clustered at the county level. Numbers in parentheses are standard errors and intervals are 95% confidence intervals. For hypothesis testing, the numbers in parentheses are p-values.

References

[1] Anderson, T. W. and Hsiao, C. (1981): "Estimation of dynamic models with error components." *Journal of the American statistical Association*, 76(375), 598-606.

[2] Arellano, M. and Bond, S. (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The Review of Economic Studies*, 58(2), 277-297.

[3] Emran, M. S., and Hou, Z. (2013): "Access to markets and rural poverty: evidence from household consumption in China." *Review of Economics and Statistics*, 95(2), 682-697.

[4] Nickell, S. (1981): "Biases in dynamic models with fixed effects." *Econometrica*: Journal of the Econometric Society, 1417-1426.

[5] Hall, P., and Horowitz, J. L. (1996): "Bootstrap critical values for tests based on generalized-method-of-moments estimators." *Econometrica: Journal of the Econometric Society*, 891-916.