Supplement to “Panel data models with nonadditive unobserved heterogeneity: Estimation and inference”

(*Quantitative Economics,* Vol. 4, No. 3, November 2013, 453–481)

Iván Fernández-Val
Department of Economics, Boston University

Joonhwah Lee
Department of Economics, MIT

This supplement provides additional numerical examples and the proofs of the main results. It is organized in seven appendices. Appendix A contains a Monte Carlo simulation calibrated to the empirical example of the paper. Appendix B gives the proofs of the consistency of the one-step and two-step FE-GMM estimators. Appendix C includes the derivations of the asymptotic distribution of one-step and two-step FE-GMM estimators. Appendix D provides the derivations of the asymptotic distribution of bias corrected FE-GMM estimators. Appendix E and Appendix F contain the characterization of the stochastic expansions for the estimators of the individual effects and the scores. Appendix G includes the expressions for the scores and their derivatives.

Throughout the appendices O_{uP} and o_{uP} will denote uniform orders in probability. For example, for a sequence of random variables $\{\xi_i : 1 \leq i \leq n\}$, $\xi_i = O_{uP}(1)$ means $\sup_{1\leq i\leq n} \xi_i = O_P(1)$ as $n \to \infty$, and $\xi_i = o_{uP}(1)$ means $\sup_{1\leq i\leq n} \xi_i = o_P(1)$ as $n \to \infty$. It can be shown that the usual algebraic properties for O_P and o_P orders apply to the uniform orders O_{uP} and o_{uP}. Let e_j denote a $1 \times d_g$ unitary vector with a one in position j. For a matrix A, $|A|$ denotes Euclidean norm, that is $|A|^2 = \text{trace}[AA']$. HK refers to Hahn and Kuersteiner (2011).

Appendix A: Numerical example

We design a Monte Carlo experiment to closely match the cigarette demand empirical example in the paper. In particular, we consider the linear model with common and individual-specific parameters,

$$
\begin{aligned}
C_{it} &= \alpha_{0i} + \alpha_{1i} P_{it} + \theta_1 C_{i,t-1} + \theta_2 C_{i,t+1} + \psi \varepsilon_{it}, \\
P_{it} &= \eta_{0i} + \eta_{1i} \text{Tax}_{it} + u_{it} \quad (i = 1, 2, \dots, n, t = 1, 2, \dots, T),
\end{aligned}
$$

Iván Fernández-Val: ivanf@bu.edu
Joonhwah Lee: jhlee82@mit.edu

DOI: 10.3982/QE75

where $\{(\alpha_{ji}, \eta_{ji}) : 1 \leq i \leq n\}$ is independent and identically distributed (i.i.d.) bivariate normal with mean (μ_j, μ_{η_j}), variances $(\sigma_j^2, \sigma_{\eta_j}^2)$, and correlation ρ_j for $j \in \{0, 1\}$, independent across j; $\{u_{it} : 1 \leq t \leq T, 1 \leq i \leq n\}$ is i.i.d. $N(0, \sigma_u^2)$; and $\{\varepsilon_{it} : 1 \leq t \leq T, 1 \leq i \leq n\}$ is i.i.d. standard normal. We fix the values of Tax_{it} to the values in the data set. All the parameters other than ρ_1 and ψ are calibrated to the data set. Since the panel is balanced for only 1972–1994, we set $T = 23$ and generate balanced panels for the simulations. Specifically, we consider

$$
\begin{aligned}
&n = 51, \qquad T = 23; \qquad \mu_0 = 72.86, \qquad \mu_1 = -31.26, \qquad \mu_{\eta_0} = 0.81,\\
&\mu_{\eta_1} = 0.13, \qquad \sigma_0 = 18.54, \qquad \sigma_1 = 10.60, \qquad \sigma_{\eta_0} = 0.14, \qquad \sigma_{\eta_1} = 2.05,\\
&\sigma_u = 0.15, \qquad \theta_1 = 0.45, \qquad \theta_2 = 0.27, \qquad \rho_0 = -0.17,\\
&\rho_1 \in \{0, 0.3, 0.6, 0.9\}, \qquad \psi \in \{2, 4, 6\}.
\end{aligned}
$$

In the empirical example, the estimated values of ρ_1 and ψ are close to 0.3 and 5, respectively.

Since the model is dynamic with leads and lags of the dependent variable on the right hand side, we construct the series of C_{it} by solving the difference equation following BGM. The stationary part of the solution is

$$
C_{it} = \frac{1}{\theta_1 \phi_1 (\phi_2 - \phi_1)} \sum_{s=1}^{\infty} \phi_1^s h_i(t+s) + \frac{1}{\theta_1 \phi_2 (\phi_2 - \phi_1)} \sum_{s=0}^{\infty} \phi_2^{-s} h_i(t-s),
$$

where

$$
\begin{aligned}
&h_i(t) = \alpha_{0i} + \alpha_{1i} P_{i,t-1} + \psi \varepsilon_{i,t-1},\\
&\phi_1 = \frac{1 - (1 - 4\theta_1\theta_2)^{1/2}}{2\theta_1}, \qquad \phi_2 = \frac{1 + (1 - 4\theta_1\theta_2)^{1/2}}{2\theta_1}.
\end{aligned}
$$

In our specification, these values are $\phi_1 = 0.31$ and $\phi_2 = 1.91$. The parameters that we vary across the experiments are ρ_1 and ψ. The parameter ρ_1 controls the degree of correlation between α_{1i} and P_{it}, and determines the bias caused by using fixed-coefficient estimators. The parameter ψ controls the degree of endogeneity in $C_{i,t-1}$ and $C_{i,t+1}$, which determines the bias of OLS and the incidental parameter bias of random-coefficient IV estimators. Although ψ is not an ideal experimental parameter because it is the variance of the error, it is the only free parameter that affects the endogeneity of $C_{i,t-1}$ and $C_{i,t+1}$. In this design, we cannot fully remove the endogeneity of $C_{i,t-1}$ and $C_{i,t+1}$ because of the dynamics.

In each simulation, we estimate the parameters with standard fixed-coefficient OLS and IV with additive individual effects (FC), and estimate the FE-GMM OLS and IV estimators with the individual-specific coefficients (RC). For IV, we use the same set of instruments as in the empirical example. We report results only for the common coefficient θ_2, and the mean and standard deviation of the individual-specific coefficient α_{1i}. Throughout the tables, “Bias” refers to the mean of the bias across simulations, SD refers to the standard deviation of the estimates, SE/SD denotes the ratio of the average standard error to the standard deviation, and p; 0.05 is the rejection frequency of a two-sided test with nominal level of 0.05 that the parameter is equal to its true value. For bias-corrected RC estimators, the standard errors are calculated using bias-corrected estimates of the common parameter and individual effects.

Table A.1 reports the results for the estimators of θ_2. We find significant biases in all the OLS estimators relative to the standard deviations of these estimators. The bias of OLS grows with ψ. The IV-RC estimator has bias unless $\rho_1 = 0$, that is unless there is no correlation between α_{1i} and P_{it}, and its test shows size distortions due to the bias and underestimation in the standard errors. IV-RC estimators have no bias in every configuration and their tests display much smaller size distortions than for the other estimators. The bias corrections preserve the bias and inference properties of the RC-IV estimator.

Table A.2 reports similar results for the estimators of the mean of the individual-specific coefficient $\mu_1 = \bar{E}[\alpha_{1i}]$. We find substantial biases for OLS and IV-FC estimators. RC-IV displays some bias, which is removed by the corrections in some configurations. The bias corrections provide significant improvements in the estimation of standard errors. IV-RC standard errors overestimate the dispersion by more than 15% when ψ is greater than 2, whereas IV-BC or IV-IBC estimators have SE/SD ratios close to 1. As a result, bias-corrected estimators show smaller size distortions. This improvement comes from the bias correction in the estimates of the dispersion of α_{1i} that we use to construct the standard errors. The bias of the estimator of the dispersion is generally large and is effectively removed by the correction. We can see more evidence of this phenomenon in Table A.3.

Table A.3 shows the results for the estimators of the standard deviation of the individual-specific coefficient $\sigma_1 = \bar{E}[(\alpha_{1i} - \mu_1)^2]^{1/2}$. As noted above, the bias corrections are relevant in this case. As ψ increases, the bias grows in orders of ψ. Most of the bias is removed by the correction, even when ψ is large. For example, when $\psi = 6$, the bias of the IV-RC estimator is about 4, which is larger than two times its standard deviation. The correction reduces the bias to about 0.5, which is small relative to the standard deviation. Moreover, despite the overestimation in the standard errors, there are important size distortions for IV-RC estimators for tests on σ_1 when ψ is large. The bias corrections bring the rejection frequencies close to their nominal levels.

Overall, the calibrated Monte Carlo experiment confirms that the IV-RC estimator with bias correction provides improved estimation and inference for all the parameters of interest for the model considered in the empirical example.

Appendix B: Consistency of one-step and two-step FE-GMM estimators

Lemma 3. *Suppose that Conditions* 1 *and* 2 *hold. Then, for every* $\eta > 0$,

$$\Pr\left\{ \sup_{1\le i\le n} \sup_{(\theta,\alpha)\in\Upsilon} \left|\widehat{Q}_i^W(\theta,\alpha) - Q_i^W(\theta,\alpha)\right| \ge \eta \right\} = o\left(T^{-1}\right)$$

TABLE A.1. Common parameter θ_2.

Estimator	$\rho_1 = 0$				$\rho_1 = 0.3$				$\rho_1 = 0.6$				$\rho_1 = 0.9$			
	Bias	SD	SE/SD	p; 0.05	Bias	SD	SE/SD	p; 0.05	Bias	SD	SE/SD	p; 0.05	Bias	SD	SE/SD	p; 0.05
								$\psi = 2$								
OLS-FC	0.06	0.01	0.84	1.00	0.06	0.01	0.83	1.00	0.06	0.01	0.84	1.00	0.07	0.01	0.71	1.00
IV-FC	0.00	0.01	0.90	0.08	−0.01	0.02	0.84	0.11	−0.01	0.02	0.78	0.18	−0.01	0.02	0.63	0.28
OLS-RC	0.04	0.01	0.97	1.00	0.04	0.01	0.99	1.00	0.04	0.01	1.02	1.00	0.04	0.01	0.96	1.00
BC-OLS	0.04	0.01	0.97	1.00	0.04	0.01	0.99	1.00	0.04	0.01	1.02	1.00	0.04	0.01	0.96	1.00
IBC-OLS	0.04	0.01	0.97	1.00	0.04	0.01	0.99	1.00	0.04	0.01	1.02	1.00	0.04	0.01	0.96	1.00
IV-RC	0.00	0.01	1.00	0.06	0.00	0.01	1.01	0.05	0.00	0.01	1.00	0.05	0.00	0.01	1.01	0.05
BC-IV	0.00	0.01	0.99	0.06	0.00	0.01	1.01	0.05	0.00	0.01	1.00	0.05	0.00	0.01	1.00	0.05
IBC-IV	0.00	0.01	0.99	0.06	0.00	0.01	1.01	0.05	0.00	0.01	1.00	0.05	0.00	0.01	1.00	0.05
								$\psi = 4$								
OLS-FC	0.12	0.01	1.09	1.00	0.12	0.01	1.03	1.00	0.12	0.01	1.10	1.00	0.12	0.01	1.07	1.00
IV-FC	0.00	0.02	0.94	0.07	−0.01	0.02	0.89	0.08	−0.01	0.02	0.92	0.09	−0.01	0.03	0.79	0.15
OLS-RC	0.10	0.01	1.06	1.00	0.10	0.01	1.05	1.00	0.11	0.01	1.08	1.00	0.11	0.01	1.07	1.00
BC-OLS	0.10	0.01	1.06	1.00	0.10	0.01	1.05	1.00	0.11	0.01	1.08	1.00	0.11	0.01	1.07	1.00
IBC-OLS	0.10	0.01	1.06	1.00	0.10	0.01	1.05	1.00	0.11	0.01	1.08	1.00	0.11	0.01	1.07	1.00
IV-RC	0.00	0.02	0.98	0.06	0.00	0.02	0.96	0.06	0.00	0.02	1.01	0.05	0.00	0.02	1.00	0.06
BC-IV	0.00	0.02	0.97	0.05	0.00	0.02	0.95	0.06	0.00	0.02	1.00	0.05	0.00	0.02	0.99	0.06
IBC-IV	0.00	0.02	0.97	0.05	0.00	0.02	0.95	0.06	0.00	0.02	1.00	0.05	0.00	0.02	0.99	0.06
								$\psi = 6$								
OLS-FC	0.16	0.01	1.27	1.00	0.16	0.01	1.22	1.00	0.16	0.01	1.25	1.00	0.16	0.01	1.34	1.00
IV-FC	0.00	0.03	0.95	0.06	0.00	0.03	0.94	0.06	−0.01	0.03	0.92	0.08	−0.01	0.04	0.92	0.08
OLS-RC	0.15	0.01	1.20	1.00	0.15	0.01	1.21	1.00	0.15	0.01	1.21	1.00	0.15	0.01	1.26	1.00
BC-OLS	0.15	0.01	1.20	1.00	0.15	0.01	1.21	1.00	0.15	0.01	1.21	1.00	0.15	0.01	1.26	1.00
IBC-OLS	0.15	0.01	1.20	1.00	0.15	0.01	1.21	1.00	0.15	0.01	1.21	1.00	0.15	0.01	1.26	1.00
IV-RC	0.00	0.03	0.98	0.06	0.00	0.03	1.00	0.04	0.00	0.03	1.01	0.05	0.00	0.03	1.05	0.04
BC-IV	0.00	0.03	0.95	0.06	0.00	0.03	0.97	0.04	0.00	0.03	0.98	0.05	0.00	0.03	1.02	0.04
IBC-IV	0.00	0.03	0.95	0.06	0.00	0.03	0.97	0.04	0.00	0.03	0.98	0.05	0.00	0.03	1.02	0.04

Note: RC and FC refer to random and fixed coefficient models. BC and IBC refer to bias-corrected and iterated bias-corrected estimates. 1,000 repetitions.

TABLE A.2. Mean of individual specific parameter $\mu_1 = \bar{E}[\alpha_{1i}]$.

	$\rho_1 = 0$				$\rho_1 = 0.3$				$\rho_1 = 0.6$				$\rho_1 = 0.9$			
Estimator	Bias	SD	SE/SD	p; 0.05	Bias	SD	SE/SD	p; 0.05	Bias	SD	SE/SD	p; 0.05	Bias	SD	SE/SD	p; 0.05
									$\psi = 2$							
OLS-FC	2.33	1.65	0.35	0.78	2.58	1.91	0.31	0.79	3.01	1.92	0.30	0.84	3.68	2.20	0.27	0.89
IV-FC	0.08	1.59	0.40	0.44	0.16	1.72	0.37	0.47	0.46	1.66	0.39	0.46	0.96	1.80	0.37	0.53
OLS-RC	1.16	1.53	1.02	0.12	1.15	1.65	0.96	0.12	1.19	1.59	0.99	0.12	1.25	1.62	0.97	0.13
BC-OLS	1.16	1.53	0.97	0.14	1.15	1.65	0.92	0.14	1.19	1.59	0.95	0.14	1.25	1.62	0.93	0.15
IBC-OLS	1.16	1.53	0.97	0.14	1.15	1.65	0.92	0.14	1.19	1.59	0.95	0.14	1.25	1.62	0.93	0.15
IV-RC	0.01	1.51	1.07	0.04	−0.01	1.62	1.00	0.05	0.02	1.56	1.04	0.05	0.08	1.59	1.03	0.05
BC-IV	−0.01	1.51	1.02	0.04	−0.02	1.62	0.96	0.06	0.00	1.56	1.00	0.06	0.06	1.59	0.98	0.06
IBC-IV	−0.01	1.51	1.02	0.04	−0.03	1.62	0.96	0.06	0.00	1.56	1.00	0.06	0.06	1.59	0.98	0.06
									$\psi = 4$							
OLS-FC	4.15	1.84	0.52	0.90	4.43	1.95	0.49	0.90	4.90	2.08	0.46	0.93	5.45	2.22	0.43	0.95
IV-FC	0.09	1.85	0.59	0.25	0.21	1.92	0.57	0.27	0.56	1.89	0.58	0.27	1.03	1.94	0.57	0.32
OLS-RC	3.19	1.76	1.06	0.41	3.12	1.81	1.04	0.38	3.12	1.76	1.07	0.38	3.18	1.78	1.06	0.38
BC-OLS	3.19	1.76	0.93	0.50	3.12	1.81	0.91	0.48	3.12	1.76	0.94	0.47	3.18	1.78	0.93	0.47
IBC-OLS	3.19	1.76	0.93	0.50	3.12	1.81	0.91	0.48	3.12	1.76	0.94	0.47	3.18	1.78	0.93	0.47
IV-RC	0.06	1.78	1.15	0.03	−0.01	1.86	1.10	0.03	0.03	1.78	1.15	0.03	0.10	1.78	1.15	0.03
BC-IV	0.00	1.78	1.02	0.05	−0.08	1.86	0.98	0.05	−0.04	1.78	1.02	0.05	0.03	1.78	1.02	0.05
IBC-IV	−0.01	1.78	1.02	0.05	−0.08	1.86	0.98	0.05	−0.04	1.78	1.02	0.05	0.03	1.78	1.02	0.05
									$\psi = 6$							
OLS-FC	5.62	2.13	0.62	0.93	5.87	2.25	0.58	0.92	6.19	2.31	0.57	0.93	6.35	2.28	0.57	0.95
IV-FC	0.14	2.29	0.69	0.17	0.26	2.34	0.68	0.19	0.53	2.31	0.69	0.19	0.80	2.26	0.70	0.20
OLS-RC	4.69	2.10	1.08	0.53	4.59	2.14	1.07	0.51	4.52	2.11	1.09	0.50	4.30	2.01	1.15	0.46
BC-OLS	4.69	2.10	0.88	0.69	4.59	2.14	0.88	0.67	4.52	2.11	0.89	0.64	4.30	2.01	0.94	0.61
IBC-OLS	4.69	2.10	0.88	0.69	4.59	2.14	0.88	0.67	4.52	2.11	0.89	0.64	4.30	2.01	0.94	0.61
IV-RC	0.09	2.30	1.12	0.04	0.05	2.33	1.11	0.03	0.03	2.23	1.16	0.02	−0.10	2.18	1.19	0.02
BC-IV	−0.05	2.30	0.95	0.06	−0.10	2.33	0.94	0.06	−0.12	2.23	0.97	0.06	−0.26	2.18	1.00	0.05
IBC-IV	−0.06	2.30	0.95	0.06	−0.10	2.32	0.94	0.06	−0.13	2.23	0.97	0.06	−0.26	2.18	1.00	0.05

Note: RC/FC refers to random/fixed coefficient model. BC/IBC refers to bias corrected/iterated bias corrected estimates. 1,000 repetitions.

TABLE A.3. Standard deviation of the individual specific parameter $\sigma_1 = \bar{E}[(\alpha_{1i} - \mu_1)^2]^{1/2}$.

	$\rho_1 = 0$				$\rho_1 = 0.3$				$\rho_1 = 0.6$				$\rho_1 = 0.9$			
Estimator	Bias	SD	SE/SD	p; 0.05	Bias	SD	SE/SD	p; 0.05	Bias	SD	SE/SD	p; 0.05	Bias	SD	SE/SD	p; 0.05
								$\psi = 2$								
OLS-RC	0.01	1.06	1.02	0.05	0.15	1.06	1.02	0.05	0.11	1.08	0.99	0.06	0.17	1.06	0.99	0.06
BC-OLS	−0.63	1.10	1.04	0.10	−0.48	1.11	1.04	0.09	−0.52	1.12	1.01	0.10	−0.46	1.11	1.01	0.09
IBC-OLS	−0.63	1.10	1.04	0.10	−0.48	1.11	1.04	0.09	−0.52	1.12	1.01	0.10	−0.46	1.11	1.01	0.09
IV-RC	0.38	1.08	1.03	0.05	0.47	1.10	1.02	0.06	0.41	1.13	0.98	0.06	0.46	1.11	0.99	0.06
BC-IV	−0.25	1.13	1.05	0.06	−0.16	1.14	1.04	0.06	−0.22	1.18	1.00	0.07	−0.17	1.16	1.00	0.06
IBC-IV	−0.25	1.13	1.05	0.06	−0.16	1.14	1.04	0.06	−0.22	1.18	1.00	0.07	−0.17	1.16	1.00	0.06
								$\psi = 4$								
OLS-RC	0.89	1.21	1.17	0.04	0.98	1.20	1.17	0.05	1.08	1.16	1.19	0.06	1.09	1.05	1.23	0.08
BC-OLS	−1.24	1.46	1.19	0.08	−1.13	1.44	1.20	0.08	−1.02	1.41	1.20	0.05	−0.98	1.23	1.24	0.03
IBC-OLS	−1.24	1.46	1.19	0.08	−1.13	1.44	1.20	0.08	−1.02	1.41	1.20	0.05	−0.98	1.24	1.22	0.03
IV-RC	1.84	1.28	1.17	0.17	1.83	1.29	1.16	0.16	1.85	1.26	1.17	0.18	1.87	1.18	1.17	0.20
BC-IV	−0.25	1.52	1.20	0.03	−0.26	1.52	1.19	0.02	−0.26	1.51	1.18	0.03	−0.21	1.37	1.18	0.02
IBC-IV	−0.25	1.52	1.20	0.03	−0.26	1.52	1.19	0.02	−0.26	1.51	1.18	0.03	−0.21	1.38	1.16	0.02
								$\psi = 6$								
OLS-RC	2.35	1.33	1.38	0.14	2.60	1.40	1.30	0.21	2.57	1.37	1.31	0.21	2.69	1.31	1.28	0.38
BC-OLS	−2.06	2.04	1.41	0.00	−1.71	2.14	1.30	0.01	−1.75	2.06	1.35	0.00	−1.54	1.78	1.28	0.00
IBC-OLS	−2.06	2.04	1.41	0.00	−1.71	2.14	1.30	0.01	−1.75	2.06	1.35	0.00	−1.54	1.80	1.26	0.00
IV-RC	3.79	1.52	1.31	0.46	3.87	1.55	1.28	0.49	3.78	1.50	1.30	0.47	3.87	1.48	1.23	0.60
BC-IV	−0.49	2.13	1.37	0.00	−0.42	2.23	1.29	0.01	−0.55	2.14	1.35	0.00	−0.40	1.96	1.24	0.01
IBC-IV	−0.49	2.13	1.37	0.00	−0.41	2.23	1.29	0.01	−0.55	2.14	1.35	0.00	−0.39	1.97	1.22	0.02

Note: RC/FC refers to random/fixed coefficient model. BC/IBC refers to bias corrected/iterated bias corrected estimates. 1,000 repetitions.

and

$$\sup_{\alpha}\left|Q_i^W(\theta,\alpha)-Q_i^W(\theta',\alpha)\right|\le C\cdot E\left[M(z_{it})\right]^2\left|\theta-\theta'\right|$$

for some constant $C>0$.

Proof. First, note that

$$\begin{aligned}
&\left|\widehat{Q}_i^W(\theta,\alpha)-Q_i^W(\theta,\alpha)\right| \\
&\quad\le\left|\widehat{g}_i(\theta,\alpha)'W_i^{-1}\widehat{g}_i(\theta,\alpha)-g_i(\theta,\alpha)'W_i^{-1}g_i(\theta,\alpha)\right| \\
&\qquad+\left|\widehat{g}_i(\theta,\alpha)'\left(\widehat{W}_i^{-1}-W_i^{-1}\right)\widehat{g}_i(\theta,\alpha)\right| \\
&\quad\le\left|\left[\widehat{g}_i(\theta,\alpha)-g_i(\theta,\alpha)\right]'W_i^{-1}\left[\widehat{g}_i(\theta,\alpha)-g_i(\theta,\alpha)\right]\right| \\
&\qquad+2\cdot\left|g_i(\theta,\alpha)'W_i^{-1}\left[\widehat{g}_i(\theta,\alpha)-g_i(\theta,\alpha)\right]\right| \\
&\qquad+\left|\left[\widehat{g}_i(\theta,\alpha)-g_i(\theta,\alpha)\right]'\left(\widehat{W}_i^{-1}-W_i^{-1}\right)\left[\widehat{g}_i(\theta,\alpha)-g_i(\theta,\alpha)\right]\right| \\
&\qquad+2\left|\left[\widehat{g}_i(\theta,\alpha)-g_i(\theta,\alpha)\right]'\left(\widehat{W}_i^{-1}-W_i^{-1}\right)g_i(\theta,\alpha)\right| \\
&\qquad+\left|g_i(\theta,\alpha)'\left(\widehat{W}_i^{-1}-W_i^{-1}\right)g_i(\theta,\alpha)\right| \\
&\quad\le d_g^2\max_{1\le k\le d_g}\left|\widehat{g}_{k,i}(\theta,\alpha)-g_{k,i}(\theta,\alpha)\right|^2|W_i|^{-1} \\
&\qquad+2d_g^2\sup_{1\le i\le n}E\left[M(z_{it})\right]|W_i|^{-1}\max_{1\le k\le d_g}\left|\widehat{g}_{k,i}(\theta,\alpha)-g_{k,i}(\theta,\alpha)\right| \\
&\qquad+o_P\left(\max_{1\le k\le d_g}\left|\widehat{g}_{k,i}(\theta,\alpha)-g_{k,i}(\theta,\alpha)\right|\right),
\end{aligned}$$

where we use that $\sup_{1\le i\le n}|\widehat{W}_i-W_i|=o_P(1)$. Then, by Condition 2, we can apply Lemma 4 of HK to $|\widehat{g}_{k,i}(\theta,\alpha)-g_{k,i}(\theta,\alpha)|$ to obtain the first part.

The second part follows from

$$\begin{aligned}
&\left|Q_i^W(\theta,\alpha)-Q_i^W(\theta',\alpha)\right| \\
&\quad\le\left|g_i(\theta,\alpha)'W_i^{-1}\left[g_i(\theta,\alpha)-g_i(\theta',\alpha)\right]\right|+\left|\left[g_i(\theta,\alpha)-g_i(\theta',\alpha)\right]'W_i^{-1}g_i(\theta',\alpha)\right| \\
&\quad\le 2\cdot d_g^2E\left[M(z_{it})\right]^2|W_i|^{-1}\left|\theta-\theta'\right|.
\end{aligned}$$

□

B.1 *Proof of Theorem 1*

Proof. *Part I: Consistency of* $\tilde{\theta}$. For any $\eta>0$, let

$$\varepsilon:=\inf_i\left[Q_i^W(\theta_0,\alpha_{i0})-\sup_{\{(\theta,\alpha):|(\theta,\alpha)-(\theta_0,\alpha_{i0})|>\eta\}}Q_i^W(\theta,\alpha)\right]>0$$

as defined in Condition 2. Using the standard argument for consistency of extremum estimators, as in Newey and McFadden (1994), with probability $1 - o(T^{-1})$,

$$\max_{|\theta-\theta_0|>\eta,\alpha_1,\dots,\alpha_n} n^{-1}\sum_{i=1}^{n}\widehat{Q}_i^W(\theta,\alpha_i) < n^{-1}\sum_{i=1}^{n}\widehat{Q}_i^W(\theta_0,\alpha_{i0}) - \frac{1}{3}\varepsilon$$

by definition of ε and Lemma 3. Thus, by continuity of $\widehat{Q}_i^W$ and the definition on the left hand side above, we conclude that $\Pr[|\tilde{\theta}-\theta_0| \geq \eta] = o(T^{-1})$.

Part II: Consistency of $\tilde{\alpha}_i$. By Part I and Lemma 3,

$$\Pr\Big[\sup_{1\leq i\leq n}\sup_{\alpha}\big|\widehat{Q}_i^W(\tilde{\theta},\alpha) - Q_i^W(\theta_0,\alpha)\big| \geq \eta\Big] = o\big(T^{-1}\big) \tag{B.1}$$

for any $\eta > 0$. Let

$$\varepsilon := \inf_i\Big[Q_i^W(\theta_0,\alpha_{i0}) - \sup_{\{\alpha_i:|\alpha_i-\alpha_{i0}|>\eta\}} Q_i^W(\theta_0,\alpha_i)\Big] > 0.$$

Condition on the event

$$\Big\{\sup_{1\leq i\leq n}\sup_{\alpha}\big|\widehat{Q}_i^W(\tilde{\theta},\alpha) - Q_i^W(\theta_0,\alpha)\big| \leq \frac{1}{3}\varepsilon\Big\},$$

which has a probability equal to $1 - o(T^{-1})$ by (B.1). Then

$$\begin{aligned}\max_{|\alpha_i-\alpha_{i0}|>\eta}\widehat{Q}_i^W(\tilde{\theta},\alpha_i) &< \max_{|\alpha_i-\alpha_{i0}|>\eta} Q_i^W(\theta_0,\alpha_i) + \frac{1}{3}\varepsilon < Q_i^W(\theta_0,\alpha_{i0}) - \frac{2}{3}\varepsilon \\ &< \widehat{Q}_i^W(\tilde{\theta},\alpha_{i0}) - \frac{1}{3}\varepsilon.\end{aligned}$$

This is inconsistent with $\widehat{Q}_i^W(\tilde{\theta},\tilde{\alpha}_i) \geq \widehat{Q}_i^W(\tilde{\theta},\alpha_{i0})$ and, therefore, $|\tilde{\alpha}_i - \alpha_{i0}| \leq \eta$ with probability $1 - o(T^{-1})$ for every i.

Part III: Consistency of $\tilde{\lambda}_i$. First, note that

$$\begin{aligned}|\tilde{\lambda}_i| &= \big|\widehat{W}_i^{-1}\widehat{g}_i(\tilde{\theta},\tilde{\alpha}_i)\big| \leq d_g|\widehat{W}_i|^{-1}\max_{1\leq k\leq d_g}\big(\big|\widehat{g}_{k,i}(\tilde{\theta},\tilde{\alpha}_i) - g_{k,i}(\tilde{\theta},\tilde{\alpha}_i)\big| + \big|g_{k,i}(\tilde{\theta},\tilde{\alpha}_i)\big|\big) \\ &\leq d_g|\widehat{W}_i|^{-1}\max_{1\leq k\leq d_g}\sup_{(\theta,\alpha_i)\in\Upsilon}\big|\widehat{g}_{k,i}(\theta,\alpha_i) - g_{k,i}(\theta,\alpha_i)\big| \\ &\quad + d_g|\widehat{W}_i|^{-1}M(z_{it})|\tilde{\theta}-\theta_0| + d_g|\widehat{W}_i|^{-1}M(z_{it})|\tilde{\alpha}_i - \alpha_{i0}|.\end{aligned}$$

Then the result follows because $\sup_{1\leq i\leq n}|\widehat{W}_i - W_i| = o_P(1)$ and $\{W_i : 1 \leq i \leq n\}$ are positive definite by Condition 2, $\max_{1\leq k\leq d_g}\sup_{(\theta,\alpha_i)\in\Upsilon}|\widehat{g}_{k,i}(\theta,\alpha_i) - g_{k,i}(\theta,\alpha_i)| = o_P(1)$ by Lemma 4 in HK, and $|\tilde{\theta}-\theta_0| = o_P(1)$ and $\sup_{1\leq i\leq n}|\tilde{\alpha}_i - \alpha_{i0}| = o_P(1)$ by Parts I and II. □

B.2 *Proof of Theorem 3*

Proof. First, assume that Conditions 1, 2, 3, and 4 hold. The proof is exactly the same as that of Theorem 1 using the uniform convergence of the criterion function.

To establish the uniform convergence of the criterion function as in Lemma 3, we need

$$\sup_{1\le i\le n}\left|\widehat{\Omega}_i(\tilde{\theta},\tilde{\alpha}_i)-\Omega_i(\theta_0,\alpha_{i0})\right| = o_P(1)$$

along with an extended version of the continuous mapping theorem for o_{uP}. This can be shown by noting that

$$\begin{aligned}
&\left|\widehat{\Omega}_i(\tilde{\theta},\tilde{\alpha}_i)-\Omega_i(\theta_0,\alpha_{i0})\right| \\
&\quad\le \left|\widehat{\Omega}_i(\tilde{\theta},\tilde{\alpha}_i)-\Omega_i(\tilde{\theta},\tilde{\alpha}_i)\right| + \left|\Omega_i(\tilde{\theta},\tilde{\alpha}_i)-\Omega_i(\theta_0,\alpha_{i0})\right| \\
&\quad\le \left|\widehat{\Omega}_i(\tilde{\theta},\tilde{\alpha}_i)-\Omega_i(\tilde{\theta},\tilde{\alpha}_i)\right| + d_g^2 E\left[M(z_{it})^2\right]\left|(\tilde{\theta},\tilde{\alpha}_i)-(\theta_0,\alpha_{i0})\right|.
\end{aligned}$$

The convergence follows by the consistency of $\tilde{\theta}$ and $\tilde{\alpha}_i$'s, and the application of Lemma 2 of HK to $g_k(z_{it};\theta,\alpha_i)g_l(z_{it};\theta,\alpha_i)$ using that $|g_k(z_{it};\theta,\alpha_i)g_l(z_{it};\theta,\alpha_i)| \le M(z_{it})^2$. □

Appendix C: Asymptotic distribution of one-step and two-step FE-GMM estimators

C.1 *Some lemmas*

Lemma 4. *Assume that Condition 1 holds. Let $h(z_{it};\theta,\alpha_i)$ be a function such that* (i) *$h(z_{it};\theta,\alpha_i)$ is continuously differentiable in $(\theta,\alpha_i)\in\Upsilon\subset\mathbb{R}^{d_\theta+d_\alpha}$,* (ii) *$\Upsilon$ is convex, and* (iii) *there exists a function $M(z_{it})$ such that $|h(z_{it};\theta,\alpha_i)|\le M(z_{it})$ and $|\partial h(z_{it};\theta,\alpha_i)/\partial(\theta,\alpha_i)|\le M(z_{it})$ with $E[M(z_{it})^{5(d_\theta+d_\alpha+6)/(1-10v)+\delta}]<\infty$ for some $\delta>0$ and $0<v<1/10$. Define $\widehat{H}_i(\theta,\alpha_i):=T^{-1}\sum_{t=1}^T h(z_{it};\theta,\alpha_i)$ and $H_i(\theta,\alpha_i):=E[\widehat{H}_i(\theta,\alpha_i)]$. Let*

$$\alpha_i^* = \arg\max_{\alpha_i}\widehat{Q}_i^W\left(\theta^*,\alpha_i\right)$$

such that $\alpha_i^*-\alpha_{i0}=o_{uP}(T^{a_\alpha})$ and $\theta^*-\theta_0=o_P(T^{a_\theta})$, with $-2/5\le a\le 0$ for $a=$ $\max(a_\alpha,a_\theta)$. *Then, for any $\overline{\theta}$ between θ^* and θ_0, and any $\overline{\alpha}_i$ between α_i^* and α_{i0},*

$$\sqrt{T}\left[\widehat{H}_i(\overline{\theta},\overline{\alpha}_i)-H_i(\overline{\theta},\overline{\alpha}_i)\right]=o_{uP}\left(T^{1/10}\right),\qquad \widehat{H}_i(\overline{\theta},\overline{\alpha}_i)-H_i(\theta_0,\alpha_{i0})=o_{uP}\left(T^{a}\right).$$

Proof. The first statement follows from Lemma 2 in HK. The second statement follows by the first statement and the conditions of the lemma by a mean value expansion since

$$\begin{aligned}
\left|\widehat{H}_i(\overline{\theta},\overline{\alpha}_i)-H_i(\theta_0,\alpha_{i0})\right| &\le \underbrace{|\overline{\theta}-\theta_0|}_{=o_{uP}(T^a)}\underbrace{\left|\frac{1}{T}\sum_{t=1}^T M(z_{it})\right|}_{=O_{uP}(1)} + \underbrace{|\overline{\alpha}_i-\alpha_{i0}|}_{=o_{uP}(T^a)}\underbrace{\left|\frac{1}{T}\sum_{t=1}^T M(z_{it})\right|}_{=O_{uP}(1)} \\
&\quad + \underbrace{\left|\widehat{H}_i(\theta_0,\alpha_{i0})-H_i(\theta_0,\alpha_{i0})\right|}_{=o_{uP}(T^{-2/5})} \\
&= o_{uP}\left(T^a\right).
\end{aligned}$$

□

Lemma 5. *Assume that Conditions* 1–4 *hold. Let* $\hat{t}_i^W(\theta, \gamma_i)$ *denote the first stage GMM score of the fixed effects, that is,*

$$\hat{t}_i^W(\theta, \gamma_i) = -\begin{pmatrix} \widehat{G}_{\alpha_i}(\theta, \alpha_i)'\lambda_i \\ \hat{g}_i(\theta, \alpha_i) + \widehat{W}_i\lambda_i \end{pmatrix},$$

where $\gamma_i = (\alpha_i', \lambda_i')'$, $\hat{s}_i^W(\theta, \gamma_i)$ *denotes the one-step GMM score for the common parameter, that is,*

$$\hat{s}_i^W(\theta, \gamma_i) = -\widehat{G}_{\theta i}(\theta, \alpha_i)'\lambda_i,$$

and $\tilde{\gamma}_i(\theta)$ *is such that* $\hat{t}_i^W(\theta, \tilde{\gamma}_i(\theta)) = 0$.

Let $\widehat{T}_{i,j}^W(\theta, \gamma_i)$ *denote* $\partial\hat{t}_i^W(\theta, \gamma_i)/\partial\gamma_i'\partial\gamma_{i,j}$ *and let* $\widehat{M}_{i,j}^W(\theta, \gamma_i)$ *denote* $\partial\hat{s}_i^W(\theta, \gamma_i)/\partial\gamma_i'\partial\gamma_{i,j}$ *for some* $0 \le j \le d_g + d_\alpha$, *where* $\gamma_{i,j}$ *is the* j*th element of* γ_i *and* $j = 0$ *denotes no second derivative. Let* $\widehat{N}_i^W(\theta, \gamma_i)$ *denote* $\partial\hat{t}_i^W(\theta, \gamma_i)/\partial\theta'$ *and let* $\widehat{S}_i^W(\theta, \gamma_i)$ *denote* $\partial\hat{s}_i^W(\theta, \gamma_i)/\partial\theta'$. *Let* $(\tilde{\theta}, \tilde{\gamma}_1, \dots, \tilde{\gamma}_n)$ *be the one-step GMM estimator. Then, for any* $\overline{\theta}$ *between* $\tilde{\theta}$ *and* θ_0, *and any* $\overline{\gamma}_i$ *between* $\tilde{\gamma}_i$ *and* γ_{i0},

$$\widehat{T}_{i,j}^W(\overline{\theta}, \overline{\gamma}_i) - T_{i,j}^W = o_{uP}(1), \qquad \widehat{M}_{i,j}^W(\overline{\theta}, \overline{\gamma}_i) - M_{i,j}^W = o_{uP}(1),$$
$$\widehat{N}_i^W(\overline{\theta}, \overline{\gamma}_i) - N_i^W = o_{uP}(1), \qquad \widehat{S}_i^W(\overline{\theta}, \overline{\gamma}_i) - S_i^W = o_{uP}(1).$$

Also, for any $\overline{\gamma}_{i0}$ *between* γ_{i0} *and* $\tilde{\gamma}_{i0} = \tilde{\gamma}_i(\theta_0)$,

$$\sqrt{T}\hat{t}_i^W(\theta_0, \overline{\gamma}_{i0}) = o_{uP}(T^{1/10}), \qquad \sqrt{T}\big(\widehat{T}_{i,j}^W(\theta_0, \overline{\gamma}_{i0}) - T_{i,j}^W\big) = o_{uP}(T^{1/10}),$$
$$\sqrt{T}\big(\widehat{M}_{i,j}^W(\theta_0, \overline{\gamma}_{i0}) - M_{i,j}^W\big) = o_{uP}(T^{1/10}).$$

Proof. The first set of results follows by inspection of the scores and their derivatives (the expressions are given in Appendix G), uniform consistency of $\tilde{\gamma}_i$ by Theorem 1, and application of the first part of Lemma 4 to $\theta^* = \tilde{\theta}$ and $\alpha_i^* = \tilde{\alpha}_i$ with $a = 0$.

The following steps are used to prove the second set of result. By Lemma 4,

$$\sqrt{T}\hat{t}_i^W = o_{uP}(T^{1/10}), \qquad \widehat{T}_i^W(\theta_0, \overline{\gamma}_{i0}) - T_i^W = o_{uP}(1),$$

where $\overline{\gamma}_{i0}$ is between $\tilde{\gamma}_{i0}$ and γ_{i0}. Then a mean value expansion of the FOC of $\tilde{\gamma}_{i0}$, $\hat{t}_i^W(\theta_0, \tilde{\gamma}_{i0}) = 0$, around $\tilde{\gamma}_{i0} = \gamma_{i0}$ gives

$$\begin{aligned}
&\sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0}) \\
&= -\underbrace{(T_i^W)^{-1}}_{=O_u(1)} \underbrace{\sqrt{T}\hat{t}_i^W}_{=o_{uP}(T^{1/10})} - \underbrace{(T_i^W)^{-1}}_{=O_u(1)} \underbrace{\big(\widehat{T}_i^W(\theta_0, \overline{\gamma}_{i0}) - T_i^W\big)}_{=o_{uP}(1)} \sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0}) \\
&= o_{uP}(T^{1/10}) + o_{uP}\big(\sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0})\big)
\end{aligned}$$

by Condition 3 and the previous result. Therefore,

$$\big(1 + o_{uP}(1)\big)\sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0}) = o_{uP}(T^{1/10}) \quad \Rightarrow \quad \sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0}) = o_{uP}(T^{1/10}).$$

Given this uniform rate for $\tilde{\gamma}_{i0}$, the desired result can be obtained by applying the second part of Lemma 4 to $\theta^* = \theta_0$ and $\alpha_i^* = \tilde{\alpha}_{i0}$ with $a = -2/5$. □

C.2 *Proof of Theorem 2*

PROOF. By a mean value expansion of the FOC for $\tilde{\theta}$ around $\tilde{\theta} = \theta_0$,

$$0 = \widehat{s}^W(\tilde{\theta}) = \widehat{s}^W(\theta_0) + \frac{d\widehat{s}^W(\overline{\theta})}{d\theta'}(\tilde{\theta} - \theta_0),$$

where $\overline{\theta}$ lies between $\tilde{\theta}$ and θ_0.

Part I: Asymptotic limit of $d\widehat{s}^W(\overline{\theta})/d\theta'$. Note that

$$\begin{aligned} \frac{d\widehat{s}^W(\overline{\theta})}{d\theta'} &= \frac{1}{n}\sum_{i=1}^{n} \frac{d\widehat{s}_i^W(\overline{\theta}, \tilde{\gamma}_i(\overline{\theta}))}{d\theta'}, \\ \frac{d\widehat{s}_i^W(\overline{\theta}, \tilde{\gamma}_i(\overline{\theta}))}{d\theta'} &= \frac{\partial\widehat{s}_i^W(\overline{\theta}, \tilde{\gamma}_i(\overline{\theta}))}{\partial\theta'} + \frac{\partial\widehat{s}_i^W(\overline{\theta}, \tilde{\gamma}_i(\overline{\theta}))}{\partial\gamma_i'}\frac{\partial\tilde{\gamma}_i(\overline{\theta})}{\theta'}. \end{aligned} \tag{C.1}$$

By Lemma 5,

$$\frac{\partial\widehat{s}_i^W(\overline{\theta}, \tilde{\gamma}_i(\overline{\theta}))}{\partial\theta'} = S_i^W + o_{uP}(1), \qquad \frac{\partial\widehat{s}_i^W(\overline{\theta}, \tilde{\gamma}_i(\overline{\theta}))}{\partial\gamma_i'} = M_i^W + o_{uP}(1).$$

Then differentiation of the FOC for $\tilde{\gamma}_i(\overline{\theta})$, $\widehat{t}_i^W(\overline{\theta}, \tilde{\gamma}_i(\overline{\theta})) = 0$, with respect to θ and $\tilde{\gamma}_i$ gives

$$\widehat{T}_i^W\big(\overline{\theta}, \tilde{\gamma}_i(\overline{\theta})\big)\frac{\partial\tilde{\gamma}_i(\overline{\theta})}{\partial\theta'} + \widehat{N}_i^W\big(\overline{\theta}, \tilde{\gamma}_i(\overline{\theta})\big) = 0.$$

By repeated application of Lemma 5 and Condition 3,

$$\frac{\partial\tilde{\gamma}_i(\overline{\theta})}{\partial\theta'} = -\big(T_i^W\big)^{-1}N_i^W + o_{uP}(1).$$

Finally, replacing the expressions for the components in (C.1) and using the formulae for the derivatives, which are provided in the Appendix G, we get

$$\frac{d\widehat{s}^W(\overline{\theta})}{d\theta'} = \frac{1}{n}\sum_{i=1}^{n} G_{\theta_i}'P_{\alpha_i}^W G_{\theta_i} + o_P(1) = J_s^W + o_P(1), \quad J_s^W = \bar{E}\big[G_{\theta_i}'P_{\alpha_i}^W G_{\theta_i}\big]. \tag{C.2}$$

Part II: Asymptotic expansion for $\tilde{\theta} - \theta_0$. By (C.2) and Lemma 22, which states the stochastic expansion of $\sqrt{nT}\widehat{s}^W(\theta_0)$, we can write

$$0 = \underbrace{\sqrt{nT}\widehat{s}^W(\theta_0)}_{O_P(1)} + \underbrace{\frac{d\widehat{s}^W(\overline{\theta})}{d\theta'}}_{O_P(1)}\sqrt{nT}(\tilde{\theta} - \theta_0).$$

Therefore, $\sqrt{nT}(\tilde{\theta} - \theta_0) = O_P(1)$, and by Part I, Lemma 22, and Condition 3, we obtain

$$\sqrt{nT}(\tilde{\theta} - \theta_0) \xrightarrow{d} -\big(J_s^W\big)^{-1}N\big(\kappa B_s^W, V_s^W\big).$$

□

C.3 *Proof of Theorem 4*

Proof. Applying Lemma 4 with a minor modification, along with Condition 4, we can prove an exact counterpart to Lemma 5 for the two-step GMM score for the fixed effects

$$\widehat{t}_i(\theta, \gamma_i) = \widehat{t}_i^{\Omega}(\theta, \gamma_i) + \widehat{t}_i^{R}(\theta, \gamma_i),$$

where the expressions of $\widehat{t}_i^{\Omega}$ and $\widehat{t}_i^{R}$ are given in the Appendix G, and for the two-step score of the common parameter

$$\widehat{s}_i\big(\theta, \widehat{\gamma}_i(\theta)\big) = -\widehat{G}_{\theta i}\big(\theta, \widehat{\alpha}_i(\theta)\big)'\widehat{\lambda}_i(\theta).$$

The only difference arises due to the term $\widehat{t}_i^{R}(\theta, \gamma_i)$, which involves $\widehat{\Omega}_i(\tilde{\theta}, \tilde{\alpha}_i) - \Omega_i$. Lemma 8 shows that $\sqrt{T}(\widehat{\Omega}_i(\tilde{\theta}, \tilde{\alpha}_i) - \Omega_i) = o_{uP}(T^{1/10})$, so that a result similar to Lemma 5 holds for the two-step scores.

Thus, we can make the same argument as in the proof of Theorem 2, using the stochastic expansion of $\sqrt{nT}\widehat{s}(\theta_0)$ given in Lemma 23. $\square$

Appendix D: Asymptotic distribution of the bias-corrected two-step GMM estimator

D.1 *Some lemmas*

Lemma 6. *Assume that Conditions 1–4 hold. Let $\widehat{t}_i(\theta, \gamma_i)$ denote the two-step GMM score for the fixed effects, let $\widehat{s}_i(\theta, \gamma_i)$ denote the two-step GMM score for the common parameter, and let $\widehat{\gamma}_i(\theta)$ be such that $\widehat{t}_i(\theta, \widehat{\gamma}_i(\theta)) = 0$. Let $\widehat{T}_{i,j}(\theta, \gamma_i)$ denote $\partial \widehat{t}_i(\theta, \gamma_i)/\partial\gamma_i'\partial\gamma_{i,j}$ for some $0 \le j \le d_g + d_\alpha$, where $\gamma_{i,j}$ is the jth component of γ_i and $j = 0$ denotes no second derivative. Let $\widehat{N}_i(\theta, \gamma_i)$ denote $\partial \widehat{t}_i(\theta, \gamma_i)/\partial\theta'$. Let $\widehat{M}_{i,j}(\theta, \gamma_i)$ denote $\partial \widehat{s}_i(\theta, \gamma_i)/\partial\gamma_i'\partial\gamma_{i,j}$ for some $0 \le j \le d_g + d_\alpha$. Let $\widehat{S}_i(\theta, \gamma_i)$ denote $\partial \widehat{s}_i(\theta, \gamma_i)/\partial\theta'$. Let $(\widehat{\theta}, \{\widehat{\gamma}_i\}_{i=1}^n)$ be the two-step GMM estimators.*

Then, for any $\overline{\theta}$ between $\widehat{\theta}$ and θ_0, and any $\overline{\gamma}_i$ between $\widehat{\gamma}_i$ and γ_{i0},

$$\sqrt{T}\big(\widehat{T}_{i,d}(\overline{\theta}, \overline{\gamma}_i) - T_{i,d}\big) = o_{uP}\big(T^{1/10}\big), \qquad \sqrt{T}\big(\widehat{M}_{i,j}(\overline{\theta}, \overline{\gamma}_i) - M_{i,j}\big) = o_{uP}\big(T^{1/10}\big),$$
$$\sqrt{T}\big(\widehat{N}_i(\overline{\theta}, \overline{\gamma}_i) - N_i\big) = o_{uP}\big(T^{1/10}\big), \qquad \sqrt{T}\big(\widehat{S}_i(\overline{\theta}, \overline{\gamma}_i) - S_i\big) = o_{uP}\big(T^{1/10}\big).$$

Proof. Let $\widehat{\gamma}_i = \widehat{\gamma}_i(\widehat{\theta})$ and $\widehat{\gamma}_{i0} = \widehat{\gamma}_i(\theta_0)$. First, note that

$$\begin{aligned}\sqrt{T}(\widehat{\gamma}_i - \widehat{\gamma}_{i0}) &= \frac{\partial \widehat{\gamma}_i(\overline{\theta})}{\partial\theta'}\sqrt{T}(\widehat{\theta} - \theta_0)\\ &= -\underbrace{\big(T_i^{\Omega}\big)^{-1}N_i}_{=O_u(1)}\underbrace{\sqrt{T}(\widehat{\theta} - \theta_0)}_{=O_P(n^{-1/2})} + o_{uP}\big(\sqrt{T}(\widehat{\theta} - \theta_0)\big) = O_{uP}\big(n^{-1/2}\big),\end{aligned}$$

where the second equality follows from the proof of Theorems 2 and 4. Thus, by the same argument used in the proof of Lemma 5,

$$\sqrt{T}(\widehat{\gamma}_i - \gamma_{i0}) = \sqrt{T}(\widehat{\gamma}_i - \widehat{\gamma}_{i0}) + \sqrt{T}(\widehat{\gamma}_{i0} - \gamma_{i0}) = o_{uP}\big(T^{1/10}\big).$$

Given this result, and inspection of the scores and their derivatives (see the Appendix G), the proof is similar to the proof of the second part of Lemma 5. □

LEMMA 7. *Assume that Condition* 1 *holds. Let* $h_j(z_{it};\theta,\alpha_i)$, $j=1,2$, *be two functions such that* (i) $h_j(z_{it};\theta,\alpha_i)$ *is continuously differentiable in* $(\theta,\alpha_i)\in\Upsilon\subset\mathbb{R}^{d_\theta+d_\alpha}$, (ii) Υ *is convex, and* (iii) *there exists a function* $M(z_{it})$ *such that* $|h_j(z_{it};\theta,\alpha_i)|\le M(z_{it})$ *and* $|\partial h_j(z_{it};\theta,\alpha_i)/\partial(\theta,\alpha_i)|\le M(z_{it})$ *with* $E[M(z_{it})^{10(d_\theta+d_\alpha+6)/(1-10v)+\delta}]<\infty$ *for some* $\delta>0$ *and* $0<v<1/10$. *Define* $\widehat{F}_i(\theta,\alpha_i):=T^{-1}\sum_{t=1}^{T}h_1(z_{it};\theta,\alpha_i)h_2(z_{it};\theta,\alpha_i)$ *and* $F_i(\theta,\alpha_i):=E[\widehat{F}_i(\theta,\alpha_i)]$. *Let*

$$\alpha_i^*=\arg\sup_{\alpha}\widehat{Q}_i^W(\theta^*,\alpha)$$

such that $\alpha_i^*-\alpha_{i0}=o_{uP}(T^{a_\alpha})$ *and* $\theta^*-\theta_0=o_P(T^{a_\theta})$, *with* $-2/5\le a\le 0$ *for* $a=\max(a_\alpha,a_\theta)$. *Then, for any* $\overline{\theta}$ *between* θ^* *and* θ_0, *and any* $\overline{\alpha}_i$ *between* α_i^* *and* α_{i0},

$$\widehat{F}_i(\overline{\theta},\overline{\alpha}_i)-F_i(\theta_0,\alpha_{i0})=o_{uP}(T^a),\qquad \sqrt{T}\big[\widehat{F}_i(\overline{\theta},\overline{\alpha}_i)-F_i(\overline{\theta},\overline{\alpha}_i)\big]=o_{uP}(T^{1/10}).$$

The proof is the same as for Lemma 4, replacing H_i by F_i, and $M(z_{it})$ by $M(z_{it})^2$.

LEMMA 8. *Assume that Conditions* 1–6 *hold. Let* $\widehat{\Omega}_i(\overline{\theta},\overline{\alpha}_i)=T^{-1}\sum_{t=1}^{T}g(z_{it};\overline{\theta},\overline{\alpha}_i)g(z_{it};\overline{\theta},\overline{\alpha}_i)'$ *be an estimator of the covariance function* $\Omega_i=E[g(z_{it})g(z_{it})']$, *where* $\overline{\theta}=\theta_0+o_P(T^{-2/5})$ *and* $\overline{\alpha}_i=\alpha_{i0}+o_{uP}(T^{-2/5})$. *Let* $\widehat{\Omega}_{\alpha_i^{d_1}\theta_i^{d_2}}(\overline{\theta},\overline{\alpha}_i)=\partial^{d_1+d_2}\widehat{\Omega}_i(\overline{\theta},\overline{\alpha}_i)/\partial^{d_1}\alpha_i\,\partial^{d_2}\theta$ *for* $0\le d_1+d_2\le 2$. *Then*

$$\sqrt{T}\big(\widehat{\Omega}_{\alpha_i^{d_1}\theta_i^{d_2}}(\overline{\theta},\overline{\alpha}_i)-\Omega_{\alpha_i^{d_1}\theta_i^{d_2}}\big)=o_{up}\big(T^{1/10}\big).$$

PROOF. Note that

$$\begin{aligned}&\big|g(z_{it};\overline{\theta},\overline{\alpha}_i)g(z_{it};\overline{\theta},\overline{\alpha}_i)'-E\big[g(z_{it};\overline{\theta},\overline{\alpha}_i)g(z_{it};\overline{\theta},\overline{\alpha}_i)'\big]\big|\\&\quad\le d_g^2\max_{1\le k\le l\le d_g}\big|g_k(z_{it};\overline{\theta},\overline{\alpha}_i)g_l(z_{it};\overline{\theta},\overline{\alpha}_i)'-E\big[g_k(z_{it};\overline{\theta},\overline{\alpha}_i)g_l(z_{it};\overline{\theta},\overline{\alpha}_i)'\big]\big|.\end{aligned}$$

Then we can apply Lemma 7 to $h_1=g_k$ and $h_2=g_l$ with $a=-2/5$. A similar argument applies to the derivatives, since they are sums of products of elements that satisfy the assumption of Lemma 7. □

LEMMA 9. *Assume that Conditions* 1–6 *hold and that* $\ell\to\infty$ *such that* $\ell/T\to 0$ *as* $T\to\infty$. *For any* $\overline{\theta}$ *between* $\widehat{\theta}$ *and* θ_0, *let* $\widehat{\Sigma}_{\alpha_i}(\overline{\theta})=[\widehat{G}_{\alpha_i}(\overline{\theta})'\widehat{\Omega}_i^{-1}\widehat{G}_{\alpha_i}(\overline{\theta})]^{-1}$, $\widehat{H}_{\alpha_i}(\overline{\theta})=\widehat{\Sigma}_{\alpha_i}(\overline{\theta})\widehat{G}_{\alpha_i}(\overline{\theta})'\widehat{\Omega}_i^{-1}$, $\widehat{P}_{\alpha_i}(\overline{\theta})=\widehat{\Omega}_i^{-1}-\widehat{\Omega}_i^{-1}\widehat{G}_{\alpha_i}(\overline{\theta})\widehat{H}_{\alpha_i}(\overline{\theta})$, $\widehat{\Sigma}_{\alpha_i}^W(\overline{\theta})=[\widehat{G}_{\alpha_i}(\overline{\theta})'W_i^{-1}\widehat{G}_{\alpha_i}(\overline{\theta})]^{-1}$, $\widehat{H}_{\alpha_i}^W(\overline{\theta})=\widehat{\Sigma}_{\alpha_i}^W(\overline{\theta})\widehat{G}_{\alpha_i}(\overline{\theta})'W_i^{-1}$, $\widehat{J}_{si}(\overline{\theta})=\widehat{G}_{\theta_i}(\overline{\theta})'\widehat{P}_{\alpha_i}(\overline{\theta})\widehat{G}_{\theta_i}(\overline{\theta})$, $\widehat{B}_{si}^C(\bar{\theta})=T^{-1}\times\sum_{j=0}^{\ell}\sum_{t=j+1}^{T}\widehat{G}_{\theta_{it}}(\overline{\theta})'\widehat{P}_{\alpha_i}(\overline{\theta})\widehat{g}_{i,t-j}(\bar{\theta})$, *and* $\widehat{B}_{si}^B(\bar{\theta})=-\widehat{G}_{\theta_i}(\overline{\theta})'[\widehat{B}_{\lambda_i}^I(\overline{\theta})+\widehat{B}_{\lambda_i}^G(\overline{\theta})+\widehat{B}_{\lambda_i}^\Omega(\overline{\theta})+$

$\widehat{B}^W_{\lambda_i}(\bar{\theta})]$, *where*

$$\widehat{B}^I_{\lambda_i}(\theta) = -\widehat{P}_{\alpha_i}(\theta)\sum_{j=1}^{d_\alpha}\widehat{G}_{\alpha\alpha_{i,j}}(\theta)\widehat{\Sigma}_{\alpha_i}(\theta)/2$$
$$+\widehat{P}_{\alpha_i}(\theta)\sum_{j=0}^{\ell}T^{-1}\sum_{t=j+1}^{T}\widehat{G}_{\alpha_{it}}(\theta)\widehat{H}_{\alpha_i}(\theta)\widehat{g}_{i,t-j}(\theta),$$

$$\widehat{B}^G_{\lambda_i}(\theta) = \widehat{H}_{\alpha_i}(\theta)'\sum_{j=0}^{\infty}T^{-1}\sum_{t=j+1}^{T}\widehat{G}_{\alpha_{it}}(\theta)'\widehat{P}_{\alpha_i}(\theta)\widehat{g}_{i,t-j}(\theta),$$

$$\widehat{B}^\Omega_{\lambda_i}(\theta) = \widehat{P}_{\alpha_i}(\theta)\sum_{j=0}^{\ell}T^{-1}\sum_{t=j+1}^{T}\widehat{g}_{it}(\theta)\widehat{g}_{it}(\theta)'\widehat{P}_{\alpha_i}(\theta)\widehat{g}_{i,t-j}(\theta),$$

$$\widehat{B}^W_{\lambda_i}(\theta) = \widehat{P}_{\alpha_i}(\theta)\sum_{j=1}^{d_\alpha}\widehat{\Omega}_{\alpha_{i,j}}\left[\widehat{H}^{W'}_{\alpha_{i,j}}(\theta)-\widehat{H}'_{\alpha_{i,j}}(\theta)\right]$$

are estimators of Σ_{α_i}, H_{α_i}, P_{α_i}, $\Sigma^W_{\alpha_i}$, $H^W_{\alpha_i}$, J_{si}, B^C_{si}, *and* B^B_{si}. *Let* $\widehat{F}_{\alpha^{d_1}\theta^{d_2}i}(\theta,\widehat{\alpha}_i(\theta))$ *and* $F_{\alpha^{d_1}\theta^{d_2}i}(\theta,\alpha_i)$, *with* $F\in\{\Sigma,H,P,\Sigma^W,H^W,J_{si},B^C_{si},B^B_{si}\}$, *denote their derivatives for* $0\le d_1+d_2\le 1$. *Then*

$$\sqrt{T}\left(\widehat{F}_{\alpha^{d_1}\theta^{d_2}i}\left(\overline{\theta},\widehat{\alpha}_i(\overline{\theta})\right)-F_{\alpha^{d_1}\theta^{d_2}i}\right)=o_{uP}\left(T^{1/10}\right),$$

where $F_{\alpha^{d_1}\theta^{d_2}i} := F$ *if* $d_1+d_2=0$.

The results follow by Theorem 3 and Lemma 6, using the algebraic properties of the o_{uP} orders and Lemma 12 of HK to show the properties of the estimators of the spectral expectations.

Lemma 10. *Assume that Conditions* 1–6 *hold. Then, for any* $\overline{\theta}$ *between* $\widehat{\theta}$ *and* θ_0,

$$\widehat{J}_s(\overline{\theta}) = J_s + o_P\left(T^{-2/5}\right).$$

Proof. Note that

$$\sqrt{T}\left[\widehat{G}_{\theta_i}(\overline{\theta})'\widehat{P}_{\alpha_i}(\overline{\theta})\widehat{G}_{\theta_i}(\overline{\theta})-G'_{\theta_i}P_{\alpha_i}G_{\theta_i}\right]=o_{uP}\left(T^{1/10}\right)$$

by Theorem 3 and Lemmas 6 and 9, using the algebraic properties of the o_{uP} orders. The result then follows by a central limit theorem (CLT) for independent sequences since

$$\widehat{J}_s(\overline{\theta})-J_s=\widehat{\bar{E}}\left[\widehat{G}_{\theta_i}(\overline{\theta})'\widehat{P}_{\alpha_i}(\overline{\theta})\widehat{G}_{\theta_i}(\overline{\theta})\right]-\bar{E}\left[G'_{\theta_i}P_{\alpha_i}G_{\theta_i}\right]$$
$$=n^{-1}\sum_{i=1}^{n}\left(G'_{\theta_i}P_{\alpha_i}G_{\theta_i}-\bar{E}\left[G'_{\theta_i}P_{\alpha_i}G_{\theta_i}\right]\right)+o_{uP}\left(T^{-2/5}\right).$$

□

Lemma 11. *Assume that Conditions* 1–6 *hold. Then, for any* $\overline{\theta}$ *between* $\widehat{\theta}$ *and* θ_0,

$$\widehat{B}_s(\tilde{\theta}) = B_s + o_P\big(T^{-2/5}\big).$$

The proof is analogous to the proof of Lemma 10, replacing J_s by B_s.

Lemma 12. *Assume that Conditions* 1–6 *hold. Then, for any* $\overline{\theta}$ *between* $\widehat{\theta}$ *and* θ_0, *and* $\mathcal{B} = -J_s^{-1}B_s$,

$$\widehat{\mathcal{B}}(\widehat{\theta}) = -\widehat{J}_s(\widehat{\theta})^{-1}\widehat{B}_s(\widehat{\theta}) = \mathcal{B} + o_P\big(T^{-2/5}\big).$$

The result follows from Lemmas 10 and 11, using a Taylor expansion argument.

D.2 *Proof of Theorem 5*

Proof. *Case I*: $C = \mathrm{BC}$. By Lemmas 10 and 25,

$$\begin{aligned}\sqrt{nT}(\widehat{\theta} - \theta_0) &= -\widehat{J}_s(\overline{\theta})^{-1}\widehat{s}(\theta_0) = -J_s^{-1}\widehat{s}(\theta_0) + o_P\big(T^{-2/5}\big)O_P\left(\sqrt{\frac{n}{T}}\right) \\ &= -J_s^{-1}\widehat{s}(\theta_0) + o_P(1).\end{aligned}$$

Then, by Lemmas 12 and 25,

$$\begin{aligned}\sqrt{nT}\big(\widehat{\theta}^{\mathrm{BC}} - \theta_0\big) &= \sqrt{nT}(\widehat{\theta} - \theta_0) - \sqrt{nT}\frac{1}{T}\widehat{\mathcal{B}}(\widehat{\theta}) = -J_s^{-1}\widehat{s}(\theta_0) + \sqrt{\frac{n}{T}}J_s^{-1}B_s + o_P(1) \\ &= -J_s^{-1}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_{si} + \sqrt{\frac{n}{T}}B_s - \sqrt{\frac{n}{T}}B_s\right] + o_P(1) \xrightarrow{d} N\big(0, J_s^{-1}\big).\end{aligned}$$

Case II: $C = \mathrm{SBC}$. First, note that $\widehat{\theta}^{\mathrm{SBC}} - \widehat{\theta} = O_P(T^{-1})$ because the correction of the score is of order $O_P(T^{-1})$. Then, by a Taylor expansion of the corrected FOC around $\widehat{\theta}^{\mathrm{SBC}} = \theta_0$,

$$0 = \widehat{s}\big(\widehat{\theta}^{\mathrm{SBC}}\big) - T^{-1}\widehat{B}_s\big(\widehat{\theta}^{\mathrm{SBC}}\big) = \widehat{s}(\theta_0) + \widehat{J}_s(\overline{\theta})\big(\widehat{\theta}^{\mathrm{SBC}} - \theta_0\big) - T^{-1}B_s + o_P\big(T^{-2}\big),$$

where $\overline{\theta}$ lies between $\widehat{\theta}^{\mathrm{SBC}}$ and θ_0. Then by Lemma 25,

$$\begin{aligned}&\sqrt{nT}\big(\widehat{\theta}^{\mathrm{SBC}} - \theta_0\big) \\ &\quad= -\widehat{J}_s(\overline{\theta})^{-1}\big[\sqrt{nT}\widehat{s}(\theta_0) - n^{1/2}T^{-1/2}B_s\big] + o_P(1) \\ &\quad= -\widehat{J}_s(\overline{\theta})^{-1}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_{si} + \sqrt{\frac{n}{T}}B_s - \sqrt{\frac{n}{T}}B_s\right] + o_P(1) \xrightarrow{d} N\big(0, J_s^{-1}\big).\end{aligned}$$

Case III: $C = \mathrm{IBC}$. A similar argument applies to the estimating equation (5.2), since $\widehat{\theta}^{\mathrm{IBC}}$ is in an $O(T^{-1})$ neighborhood of θ_0. □

APPENDIX E: STOCHASTIC EXPANSION FOR $\tilde{\gamma}_{i0} = \tilde{\gamma}_i(\theta_0)$ AND $\widehat{\gamma}_{i0} = \widehat{\gamma}_i(\theta_0)$

We characterize the stochastic expansions up to second order for one-step and two-step estimators of the individual effects at the true common parameter. We only provide detailed proofs of the results for the two-step estimator $\widehat{\gamma}_{i0}$, because the proofs the one-step estimator $\tilde{\gamma}_{i0}$ follow by similar arguments. Lemmas 1 and 2 in the main text are corollaries of these expansions. The expressions for the scores and their derivatives in the components of the expansions are given in Appendix G.

LEMMA 13. *Suppose that Conditions* 1–4 *hold. Then*

$$\sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0}) = \tilde{\psi}_i^W + T^{-1/2} R_{1i}^W \xrightarrow{d} N(0, V_i^W),$$

where

$$\tilde{\psi}_i^W = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \psi_{it}^W = -(T_i^W)^{-1} \sqrt{T} \hat{t}_i^W = o_{uP}(T^{1/10}),$$

$$R_{1i}^W = o_{uP}(T^{1/5}), \qquad V_i^W = E[\tilde{\psi}_i^W \tilde{\psi}_i^{W'}].$$

Also

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{\psi}_i^W = O_P(1).$$

PROOF. We just show the part of the remainder term because the rest of the proof is similar to the proof of Lemma 16. By the proof of Lemma 5, $\sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0}) = o_{uP}(T^{1/10})$ and

$$R_{1i}^W = -\underbrace{(T_i^W)^{-1}}_{=O_u(1)} \underbrace{(\widehat{T}_i^W(\theta_0, \overline{\gamma}_{i0}) - T_i^W)}_{=o_{uP}(T^{1/10})} \underbrace{\sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0})}_{=o_{uP}(T^{1/10})} = o_{uP}(T^{1/5}).$$

□

LEMMA 14. *Suppose that Conditions* 1–4 *hold. Then*

$$\sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0}) = \tilde{\psi}_i^W + T^{-1/2} Q_{1i}^W + T^{-1} R_{2i}^W,$$

where

$$Q_{1i}^W = -(T_i^W)^{-1} \left[\tilde{A}_i^W \tilde{\psi}_i^W + \frac{1}{2} \sum_{j=1}^{d_g + d_\alpha} \tilde{\psi}_{i,j}^W T_{i,j}^W \tilde{\psi}_i^W \right] = o_{uP}(T^{1/5}),$$

$$\tilde{A}_i^W = \sqrt{T}(\widehat{T}_i^W - T_i^W) = o_{uP}(T^{1/10}), \qquad R_{2i}^W = o_{uP}(T^{3/10}).$$

Also,

$$\frac{1}{n} \sum_{i=1}^{n} Q_{1i}^W = O_P(1).$$

The proof is similar to the proof of Lemma 18.

LEMMA 15. *Suppose that Conditions* 1–4 *hold. Then*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_i^W \xrightarrow{d} N\big(0,\bar{E}[V_i^W]\big), \qquad \frac{1}{n}\sum_{i=1}^{n} Q_{1i}^W \xrightarrow{p} \bar{E}\big[B_{\gamma_i}^{W,I}+B_{\gamma_i}^{W,G}+B_{\gamma_i}^{W,1S}\big] =: B_\gamma^W,$$

where

$$V_i^W = \begin{pmatrix} H_{\alpha_i}^W \\ P_{\alpha_i}^W \end{pmatrix} \Omega_i\big(H_{\alpha_i}^{W'}, P_{\alpha_i}^W\big),$$

$$B_{\gamma_i}^{W,I} = \begin{pmatrix} B_{\alpha_i}^{W,I} \\ B_{\lambda_i}^{W,I} \end{pmatrix} = \begin{pmatrix} H_{\alpha_i}^W \\ P_{\alpha_i}^W \end{pmatrix}\left(\sum_{j=-\infty}^{\infty} E\big[G_{\alpha_i}(z_{it})H_{\alpha_i}^W g(z_{i,t-j})\big] - \sum_{j=1}^{d_\alpha} G_{\alpha\alpha_{i,j}} H_{\alpha_i}^W \Omega_i H_{\alpha_i}^{W'}/2\right),$$

$$B_{\gamma_i}^{W,G} = \begin{pmatrix} B_{\alpha_i}^{W,G} \\ B_{\lambda_i}^{W,G} \end{pmatrix} = \begin{pmatrix} -\Sigma_{\alpha_i}^W \\ H_{\alpha_i}^{W'} \end{pmatrix} \sum_{j=-\infty}^{\infty} E\big[G_{\alpha_i}(z_{it})' P_{\alpha_i}^W g(z_{i,t-j})\big],$$

$$B_{\gamma_i}^{W,1S} = \begin{pmatrix} B_{\alpha_i}^{W,1S} \\ B_{\lambda_i}^{W,1S} \end{pmatrix} = \begin{pmatrix} \Sigma_{\alpha_i}^W \\ -H_{\alpha_i}^{W'} \end{pmatrix}\left(\sum_{j=1}^{d_\alpha} G'_{\alpha\alpha_{i,j}} P_{\alpha_i}^W \Omega_i H_{\alpha_i}^{W'}/2 + \sum_{j=1}^{d_g} G'_{\alpha\alpha_i}(I_{d_\alpha}\otimes e_j) H_{\alpha_i}^W \Omega_i P_{\alpha_i,j}^W/2\right)$$
$$+ \begin{pmatrix} H_{\alpha_i}^W \\ P_{\alpha_i}^W \end{pmatrix} \sum_{j=-\infty}^{\infty} E\big[\xi_i(z_{it}) P_{\alpha_i}^W g(z_{i,t-j})\big],$$

for $\Sigma_{\alpha_i}^W = (G'_{\alpha_i} W_i^{-1} G_{\alpha_i})^{-1}$, $H_{\alpha_i}^W = \Sigma_{\alpha_i}^W G'_{\alpha_i} W_i^{-1}$, *and* $P_{\alpha_i}^W = W_i^{-1} - W_i^{-1} G_{\alpha_i} H_{\alpha_i}^W$.

PROOF. The results follow from Lemmas 13 and 14, noting that

$$(T_i^W)^{-1} = -\begin{pmatrix} -\Sigma_{\alpha_i}^W & H_{\alpha_i}^W \\ H_{\alpha_i}^{W'} & P_{\alpha_i}^W \end{pmatrix}, \qquad \psi_{it}^W = -\begin{pmatrix} H_{\alpha_i}^W \\ P_{\alpha_i}^W \end{pmatrix} g(z_{it}),$$

$$E\big[\tilde{\psi}_i^W \tilde{\psi}_i^{W'}\big] = \begin{pmatrix} H_{\alpha_i}^W \\ P_{\alpha_i}^W \end{pmatrix} \Omega_i\big(H_{\alpha_i}^{W'}, P_{\alpha_i}^W\big),$$

$$E\big[\tilde{A}_i^W \tilde{\psi}_i^W\big] = \sum_{j=-\infty}^{\infty} \begin{pmatrix} E[G_{\alpha_i}(z_{it})' P_{\alpha_i}^W g(z_{i,t-j})] \\ E[G_{\alpha_i}(z_{it})' H_{\alpha_i}^W g(z_{i,t-j})] + E\big[\xi_i(z_{it}) P_{\alpha_i}^W g(z_{i,t-j})\big] \end{pmatrix},$$

$$E\big[\tilde{\psi}_{i,j}^W T_{i,j}^W \tilde{\psi}_i^W\big] = \begin{cases} -\begin{pmatrix} G'_{\alpha\alpha_{i,j}} P_{\alpha_i}^W \Omega_i H_{\alpha_i}^{W'} \\ G'_{\alpha\alpha_{i,j}} H_{\alpha_i}^W \Omega_i H_{\alpha_i}^{W'} \end{pmatrix}, & \text{if } j \le d_\alpha, \\ G'_{\alpha\alpha_i}(I_{d_\alpha}\otimes e_{j-d_\alpha}) H_{\alpha_i}^W \Omega_i P_{\alpha_i,j}^W, & \text{if } j > d_\alpha. \end{cases}$$

□

Lemma 16. *Suppose that Conditions* 1–6 *hold. Then*

$$\sqrt{T}(\widehat{\gamma}_{i0}-\gamma_{i0}) = \tilde{\psi}_i + T^{-1/2}R_{1i} \xrightarrow{d} N(0, V_i),$$

where

$$\tilde{\psi}_i = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\psi_{it} = -\big(T_i^{\Omega}\big)^{-1}\sqrt{T}\hat{t}_i^{\Omega} = o_{uP}\big(T^{1/10}\big),$$
$$R_{1i} = o_{uP}\big(T^{1/5}\big), \qquad V_i = E\big[\tilde{\psi}_i\tilde{\psi}_i'\big].$$

Also

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_i = O_P(1).$$

Proof. The statements about $\tilde{\psi}_i$ follow by the proof of Lemma 5 applied to the second stage and by the CLT in Lemma 3 of HK. From a similar argument to the proof of Lemma 5,

$$\begin{aligned} R_{1i} &= -\underbrace{\big(T_i^{\Omega}\big)^{-1}}_{=O_u(1)}\underbrace{\sqrt{T}\big(\widehat{T}_i^{\Omega}(\theta_0,\overline{\gamma}_i)-T_i^{\Omega}\big)}_{=o_{uP}(T^{1/10})}\underbrace{\sqrt{T}(\widehat{\gamma}_{i0}-\gamma_{i0})}_{=o_{uP}(T^{1/10})} \\ &\quad -\underbrace{\big(T_i^{\Omega}\big)^{-1}}_{=O_u(1)}\underbrace{\sqrt{T}\big(\widehat{T}_i^{R}(\theta_0,\overline{\gamma}_i)-T_i^{R}\big)}_{=o_{uP}(T^{1/10})}\underbrace{\sqrt{T}(\widehat{\gamma}_{i0}-\gamma_{i0})}_{=o_{uP}(T^{1/10})} \\ &= o_{uP}\big(T^{1/5}\big) \end{aligned}$$

by Conditions 3 and 4. □

Lemma 17. *Assume that Conditions* 1–4 *hold. Then,*

$$\widehat{\Omega}_i(\tilde{\theta},\tilde{\alpha}_i) = \Omega_i + T^{-1/2}\tilde{\psi}_{\Omega i}^{W} + T^{-1}R_{1\Omega i}^{W},$$

where

$$\tilde{\psi}_{\Omega i}^{W} = \sqrt{T}(\widehat{\Omega}_i - \Omega_i) + \sum_{j=1}^{d_\alpha}\Omega_{\alpha_{i,j}}\tilde{\psi}_{i,j}^{W} = o_{up}\big(T^{1/10}\big), \qquad R_{1\Omega i}^{W} = o_{up}\big(T^{1/5}\big)$$

and $\tilde{\psi}_{i,j}^{W}$ is the jth element of $\tilde{\psi}_{i,j}^{W}$.

Proof. By a mean value expansion around (θ_0, α_{i0}),

$$\widehat{\Omega}_i(\tilde{\theta},\tilde{\alpha}_i) = \widehat{\Omega}_i + \sum_{j=1}^{d_\alpha}\widehat{\Omega}_{\alpha_{i,j}}(\overline{\theta},\overline{\alpha}_i)(\tilde{\alpha}_{i,j}-\alpha_{i0,j}) + \sum_{j=1}^{d_\theta}\widehat{\Omega}_{\theta_j}(\overline{\theta},\overline{\alpha}_i)(\tilde{\theta}_j-\theta_{0,j}),$$

where $(\overline{\theta},\overline{\alpha}_i)$ lies between $(\tilde{\theta},\tilde{\alpha}_i)$ and (θ_0,α_{i0}). The expressions for $\tilde{\psi}_{\Omega i}^{W}$ can be obtained using the expansions for $\tilde{\gamma}_{i0}$ in Lemma 13 since $\tilde{\gamma}_i - \tilde{\gamma}_{i0} = o_{uP}(T^{-3/10})$. The order of this term follows from Lemma 13 and the CLT for independent sequences. The remainder term is

$$R_{1\Omega i}^{W}=\sum_{j=1}^{d_\alpha}\big[\Omega_{\alpha_{i,j}}R_{1i,j}^{W}+\sqrt{T}\big(\widehat{\Omega}_{\alpha_{i,j}}(\overline{\theta},\overline{\alpha}_i)-\Omega_{\alpha_{i,j}}\big)\sqrt{T}(\tilde{\alpha}_{i,j}-\alpha_{i0,j})\big]$$
$$+\sum_{j=1}^{d_\theta}\widehat{\Omega}_{\theta_j}(\overline{\theta},\overline{\alpha}_i)T(\tilde{\theta}_j-\theta_{0,j}).$$

The uniform rate of convergence then follows by Lemmas 8 and 13, and Theorem 1. □

LEMMA 18. *Suppose that Conditions* 1–4 *hold. Then*

$$\sqrt{T}(\widehat{\gamma}_{i0}-\gamma_{i0})=\tilde{\psi}_i+T^{-1/2}Q_{1i}+T^{-1}R_{2i}, \tag{E.1}$$

where

$$Q_{1i}(\tilde{\psi}_i,\tilde{a}_i)=-\big(T_i^{\Omega}\big)^{-1}\left[\tilde{A}_i^{\Omega}\tilde{\psi}_i+\frac{1}{2}\sum_{j=1}^{d_g+d_\alpha}\tilde{\psi}_{i,j}T_{i,j}^{\Omega}\tilde{\psi}_i+\operatorname{diag}\big[0,\tilde{\psi}_{\Omega_i}^{W}\big]\tilde{\psi}_i\right]$$
$$=o_{uP}\big(T^{1/5}\big),$$
$$\tilde{A}_i^{\Omega}=\sqrt{T}\big(\widehat{T}_i^{\Omega}-T_i^{\Omega}\big)=o_{uP}\big(T^{1/10}\big),\qquad R_{2i}=o_{uP}\big(T^{3/10}\big).$$

Also,

$$\frac{1}{n}\sum_{i=1}^{n}Q_{1i}=O_P(1).$$

PROOF. By a second order Taylor expansion of the FOC for $\widehat{\gamma}_{i0}$, we have

$$0=\widehat{t}_i(\theta_0,\widehat{\gamma}_{i0})=\widehat{t}_i^{\Omega}+\widehat{T}_i(\widehat{\gamma}_{i0}-\gamma_{i0})+\frac{1}{2}\sum_{j=1}^{d_g+d_\alpha}(\widehat{\gamma}_{i0,j}-\gamma_{i0,j})\widehat{T}_{i,j}(\theta_0,\overline{\gamma}_i)(\widehat{\gamma}_{i0}-\gamma_{i0}),$$

where $\overline{\gamma}_i$ is between $\widehat{\gamma}_{i0}$ and γ_{i0}. The expression for Q_{1i} can be obtained in a similar fashion as in Lemma A4 in Newey and Smith (2004). The rest of the properties for Q_{1i} follow by Lemma 5 applied to the second stage, Lemma 16, and an argument similar to the proof of Theorem 1 in HK that uses Corollary A.2 of Hall and Heide (1980, p. 278) and Lemma 1 of Andrews (1991). The remainder term is

$$R_{2i}=-\big(T_i^{\Omega}\big)^{-1}\left[\tilde{A}_i^{\Omega}R_{1i}+\sum_{j=1}^{d_g+d_\alpha}\big[R_{1i,j}T_{i,j}^{\Omega}\sqrt{T}(\widehat{\gamma}_{i0}-\gamma_{i0})+\tilde{\psi}_{i,j}T_{i,j}^{\Omega}R_{1i}\big]/2\right]$$
$$-\big(T_i^{\Omega}\big)^{-1}\sum_{j=1}^{d_g+d_\alpha}\sqrt{T}(\widehat{\gamma}_{i0,j}-\gamma_{i0,j})\sqrt{T}\big(\widehat{T}_{i,j}^{\Omega}(\theta_0,\overline{\gamma}_i)-T_{i,j}^{\Omega}\big)\sqrt{T}(\widehat{\gamma}_{i0}-\gamma_{i0})/2$$
$$-\big(T_i^{\Omega}\big)^{-1}\big[\operatorname{diag}\big[0,R_{1\Omega_i}^{W}\big]\sqrt{T}(\widehat{\gamma}_{i0}-\gamma_{i0})+\operatorname{diag}\big[0,\tilde{\psi}_{\Omega_i}^{W}\big]R_{1i}\big].$$

The uniform rate of convergence then follows by Lemmas 5 and 16 and Conditions 3 and 4. □

LEMMA 19. *Suppose that Conditions* 1–6 *hold. Then*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_i \xrightarrow{d} N\big(0,\bar{E}[V_i]\big), \qquad \frac{1}{n}\sum_{i=1}^{n}Q_{1i} \xrightarrow{p} \bar{E}\big[B^I_{\gamma_i}+B^G_{\gamma_i}+B^\Omega_{\gamma_i}+B^W_{\gamma_i}\big] =: B_\gamma,$$

where

$$\begin{aligned}
V_i &= \mathrm{diag}(\Sigma_{\alpha_i}, P_{\alpha_i}),\\
B^I_{\gamma_i} &= \begin{pmatrix} B^I_{\alpha_i}\\ B^I_{\lambda_i}\end{pmatrix} = \begin{pmatrix} H_{\alpha_i}\\ P_{\alpha_i}\end{pmatrix}\left(-\sum_{j=1}^{d_\alpha} G_{\alpha\alpha_{i,j}}\Sigma_{\alpha_i}/2 + E\big[G_{\alpha_i}(z_{it})H_{\alpha_i}g(z_{i,t-j})\big]\right),\\
B^G_{\gamma_i} &= \begin{pmatrix} B^G_{\alpha_i}\\ B^G_{\lambda_i}\end{pmatrix} = \begin{pmatrix} -\Sigma_{\alpha_i}\\ H'_{\alpha_i}\end{pmatrix}\sum_{j=0}^{\infty} E\big[G_{\alpha_i}(z_{it})'P_{\alpha_i}g(z_{i,t-j})\big],\\
B^\Omega_{\gamma_i} &= \begin{pmatrix} B^\Omega_{\alpha_i}\\ B^\Omega_{\lambda_i}\end{pmatrix} = \begin{pmatrix} H_{\alpha_i}\\ P_{\alpha_i}\end{pmatrix}\sum_{j=0}^{\infty} E\big[g(z_{it})g(z_{it})'P_{\alpha_i}g(z_{i,t-j})\big],\\
B^W_{\gamma_i} &= \begin{pmatrix} B^W_{\alpha_i}\\ B^W_{\lambda_i}\end{pmatrix} = \begin{pmatrix} H_{\alpha_i}\\ P_{\alpha_i}\end{pmatrix}\sum_{j=1}^{d_\alpha} \Omega_{\alpha_{i,j}}(H^{W'}_{\alpha_{i,j}} - H'_{\alpha_{i,j}})
\end{aligned}$$

for $\Sigma_{\alpha_i} = (G'_{\alpha_i}\Omega_i^{-1}G_{\alpha_i})^{-1}$, $H_{\alpha_i} = \Sigma_{\alpha_i}G'_{\alpha_i}\Omega_i^{-1}$, *and* $P_{\alpha_i} = \Omega_i^{-1} - \Omega_i^{-1}G_{\alpha_i}H_{\alpha_i}$.

PROOF. The results follow by Lemmas 16 and 18, noting that

$$\begin{aligned}
(T_i^\Omega)^{-1} &= -\begin{pmatrix} -\Sigma_{\alpha_i} & H_{\alpha_i}\\ H'_{\alpha_i} & P_{\alpha_i}\end{pmatrix}, \qquad \psi_{it} = -\begin{pmatrix} H_{\alpha_i}\\ P_{\alpha_i}\end{pmatrix} g(z_{it}),\\
E[\tilde{\psi}_i\tilde{\psi}_i'] &= \begin{pmatrix} \Sigma_{\alpha_i} & 0\\ 0 & P_{\alpha_i}\end{pmatrix}, \qquad E[\tilde{A}_i^\Omega\tilde{\psi}_i] = \sum_{j=0}^{\infty}\begin{pmatrix} E[G_{\alpha_i}(z_{it})'P_{\alpha_i}g(z_{i,t-j})]\\ E[G_{\alpha_i}(z_{it})'H_{\alpha_i}g(z_{i,t-j})]\end{pmatrix},\\
E\big[\tilde{\psi}_{i,j}T^\Omega_{i,j}\tilde{\psi}_i\big] &= \begin{cases} -\begin{pmatrix} 0\\ G'_{\alpha\alpha_{i,j}}\Sigma_{\alpha_i}\end{pmatrix}, & \text{if } j \le d_\alpha,\\ 0, & \text{if } j > d_\alpha,\end{cases}\\
E\big[&\mathrm{diag}[0,\tilde{\psi}^W_{\Omega_i}]\tilde{\psi}_i\big]\\
&= \begin{pmatrix} 0\\ \sum_{j=0}^{\infty} E[g(z_{it})g(z_{it})'P_{\alpha_i}g(z_{i,t-j})] + \sum_{j=1}^{d_\alpha}\Omega_{\alpha_{i,j}}(H^{W'}_{\alpha_{i,j}} - H_{\alpha_{i,j}})\end{pmatrix}.
\end{aligned}$$

□

APPENDIX F: STOCHASTIC EXPANSION FOR $\widehat{s}_i^W(\theta_0, \tilde{\gamma}_{i0})$ AND $\widehat{s}_i(\theta_0, \widehat{\gamma}_{i0})$

We characterize stochastic expansions up to second order for one-step and two-step profile scores of the common parameter evaluated at the true value of the common parameter. The expressions for the scores and their derivatives in the components of the expansions are given in Appendix G.

LEMMA 20. *Suppose that Conditions* 1–4 *hold. Then*

$$\widehat{s}_i^W(\theta_0, \tilde{\gamma}_{i0}) = T^{-1/2}\tilde{\psi}_{si}^W + T^{-1}Q_{1si}^W + T^{-3/2}R_{2si}^W,$$

where

$$\tilde{\psi}_{si}^W = M_i^W \tilde{\psi}_i^W = o_{uP}(T^{1/10}),$$

$$Q_{1si}^W = M_i^W Q_{1i}^W + \tilde{C}_i^W \tilde{\psi}_i^W + \frac{1}{2}\sum_{j=1}^{d_g+d_\alpha} \tilde{\psi}_{i,j}^W M_{i,j}^W \tilde{\psi}_i^W = o_{uP}(T^{1/5}),$$

$$\tilde{C}_i^W = \sqrt{T}(\widehat{M}_i^W - M_i^W) = o_{uP}(T^{1/10}), \qquad R_{2si}^W = o_{uP}(T^{2/5}).$$

Also,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{\psi}_{si}^W = O_P(1), \qquad \frac{1}{n}\sum_{i=1}^n Q_{1si}^W = O_P(1).$$

PROOF. By a second order Taylor expansion of $\widehat{s}_i^W(\theta_0, \tilde{\gamma}_{i0})$ around $\tilde{\gamma}_{i0} = \gamma_{i0}$,

$$\widehat{s}_i^W(\theta_0, \tilde{\gamma}_{i0}) = \widehat{s}_i^W + \widehat{M}_i^W(\tilde{\gamma}_{i0} - \gamma_{i0}) + \frac{1}{2}\sum_{j=1}^{d_g+d_\alpha} (\tilde{\gamma}_{i0,j} - \gamma_{i0,j})\widehat{M}_{i,j}^W(\theta_0, \overline{\gamma}_i)(\tilde{\gamma}_{i0} - \gamma_{i0}),$$

where $\overline{\gamma}_i$ is between $\tilde{\gamma}_{i0}$ and γ_{i0}. Noting that $\widehat{s}_i^W(\theta_0, \gamma_{i0}) = 0$ and using the expansion for $\tilde{\gamma}_{i0}$ in Lemma 14, we can obtain the expressions for $\tilde{\psi}_{si}^W$ and Q_{1si}^W, after some algebra. The rest of the properties for these terms follow by the properties of $\tilde{\psi}_i^W$ and Q_{1i}^W. The remainder term is

$$R_{2si}^W = M_i^W R_{2i}^W + \tilde{C}_i^W R_{1i}^W + \frac{1}{2}\sum_{j=1}^{d_g+d_\alpha} [R_{1i,j}^W M_{i,j}^W \sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0}) + \tilde{\psi}_{i,j}^W M_{i,j}^W R_{1i}^W]$$
$$+ \frac{1}{2}\sum_{j=1}^{d_g+d_\alpha} \sqrt{T}(\tilde{\gamma}_{i0,j} - \gamma_{i0,j})\sqrt{T}\big(\widehat{M}_{i,j}^W(\theta_0, \overline{\gamma}_i) - M_{i,j}^W\big)\sqrt{T}(\tilde{\gamma}_{i0} - \gamma_{i0}).$$

The uniform order of R_{2si}^W follows by the properties of the components in the expansion of $\tilde{\gamma}_{i0}$, Lemma 5, and Conditions 3 and 4. □

LEMMA 21. *Suppose that Conditions* 1–4 *hold. We then have*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_{si}^{W}\xrightarrow{d}N\big(0,V_s^W\big),\qquad V_s^W=\bar{E}\big[G_{\theta_i}'P_{\alpha_i}^{W}\Omega_iP_{\alpha_i}^{W}G_{\theta_i}\big],$$

$$\frac{1}{n}\sum_{i=1}^{n}Q_{1si}^{W}\xrightarrow{p}\bar{E}E\big[Q_{1si}^{W}\big]=\bar{E}\big[B_{si}^{W,B}+B_{si}^{W,C}+B_{si}^{W,V}\big]=:B_s^W,$$

where $B_{si}^{W,B}=-G_{\theta_i}'B_{\lambda_i}^{W}=-G_{\theta_i}'(B_{\lambda_i}^{W,I}+B_{\lambda_i}^{W,G}+B_{\lambda_i}^{W,1S})$, $B_{si}^{W,C}=\sum_{j=-\infty}^{\infty}E[G_{\theta_i}(z_{it})'P_{\alpha_i}^{W}\times g(z_{i,t-j})]$, $B_{si}^{W,V}=-\sum_{j=1}^{d_\alpha}G_{\theta\alpha_{i,j}}'P_{\alpha_i}^{W}\Omega_iH_{\alpha_i}^{W'}/2-\sum_{j=1}^{d_g}G_{\theta\alpha_i}'(I_{d_\alpha}\otimes e_j)H_{\alpha_i}^{W}\Omega_iP_{\alpha_i,j}^{W}/2$, $H_{\alpha_i}^{W}=\Sigma_{\alpha_i}^{W}G_{\alpha_i}'W_i^{-1}$, $\Sigma_{\alpha_i}^{W}=(G_{\alpha_i}'W_i^{-1}G_{\alpha_i})^{-1}$, *and* $P_{\alpha_i}^{W}=W_i^{-1}-W_i^{-1}G_{\alpha_i}H_{\alpha_i}^{W}$.

PROOF. The results follow by Lemmas 20 and 15, noting that

$$E\big[\tilde{\psi}_{si}^{W}\tilde{\psi}_{si}^{W'}\big]=M_i^W\begin{pmatrix}H_{\alpha_i}^{W}\Omega_iH_{\alpha_i}^{W'} & H_{\alpha_i}^{W}\Omega_iP_{\alpha_i}^{W}\\ P_{\alpha_i}^{W}\Omega_iH_{\alpha_i}^{W'} & P_{\alpha_i}^{W}\Omega_iP_{\alpha_i}^{W}\end{pmatrix}M_i^{W'},$$

$$E\big[\tilde{C}_i^W\tilde{\psi}_i^W\big]=\sum_{j=-\infty}^{\infty}E\big[G_{\theta_i}(z_{it})'P_{\alpha_i}^{W}g(z_{i,t-j})\big],$$

$$E\big[\tilde{\psi}_{i,j}^{W}M_{i,j}^{W}\tilde{\psi}_i^W\big]=\begin{cases}-G_{\theta\alpha_{i,j}}'P_{\alpha_i}^{W}\Omega_iH_{\alpha_i}^{W'}, & \text{if } j\le d_\alpha,\\ -G_{\theta\alpha_i}'(I_{d_\alpha}\otimes e_{j-d_\alpha})H_{\alpha_i}^{W}\Omega_iP_{\alpha_i,j}^{W}, & \text{if } j>d_\alpha.\end{cases}\qquad\square$$

LEMMA 22. *Suppose that Conditions* 1–4 *hold. Then, for* $\widehat{s}^W(\theta_0)=n^{-1}\sum_{i=1}^{n}\widehat{s}_i^W(\theta_0,\tilde{\gamma}_{i0})$,

$$\sqrt{nT}\widehat{s}^W(\theta_0)\xrightarrow{d}N\big(\kappa B_s^W,V_s^W\big),$$

where B_s^W *and* V_s^W *are defined in Lemma* 21.

PROOF. By Lemma 20,

$$\sqrt{nT}\widehat{s}^W(\theta_0)=\underbrace{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_{si}^{W}}_{=O_P(1)}+\underbrace{\sqrt{\frac{n}{T}}\frac{1}{n}\sum_{i=1}^{n}Q_{1si}^{W}}_{=O_P(1)}+\underbrace{\sqrt{\frac{n}{T^2}}\frac{1}{n}\sum_{i=1}^{n}R_{2si}^{W}}_{=o_P(1)}$$
$$=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_{si}^{W}+\sqrt{\frac{n}{T}}\frac{1}{n}\sum_{i=1}^{n}Q_{1si}^{W}+o_P(1).$$

Then the result follows by Lemma 21. $\square$

LEMMA 23. *Suppose that Conditions* 1–6 *hold. Then*

$$\widehat{s}_i(\theta_0,\widehat{\gamma}_{i0})=T^{-1/2}\tilde{\psi}_{si}+T^{-1}Q_{1si}+T^{-3/2}R_{2si},$$

where all the terms are identical to that of Lemma 20 *after replacing* W *by* Ω. *Also, the properties of all the terms of the expansion are analogous to those of Lemma* 20.

The proof is similar to the proof of Lemma 20.

LEMMA 24. *Suppose that Conditions* 1–6 *hold. Then*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_{si}\xrightarrow{d}N(0,J_s),\qquad J_s=\bar{E}\big[G'_{\theta_i}P_{\alpha_i}G_{\theta_i}\big],$$

$$\frac{1}{n}\sum_{i=1}^{n}Q_{1si}\xrightarrow{p}\bar{E}E[Q_{1si}]=\bar{E}\big[B_{si}^{B}+B_{si}^{C}\big]=:B_s,$$

where $B_{si}^{B}=-G'_{\theta_i}(B_{\lambda_i}^{I}+B_{\lambda_i}^{G}+B_{\lambda_i}^{\Omega}+B_{\lambda_i}^{W})$, $B_{si}^{C}=\sum_{j=0}^{\infty}E[G_{\theta_i}(z_{it})'P_{\alpha_i}g(z_{i,t-j})]$, $P_{\alpha_i}=\Omega_i^{-1}-\Omega_i^{-1}G_{\alpha_i}H_{\alpha_i}$, $H_{\alpha_i}=\Sigma_{\alpha_i}G'_{\alpha_i}\Omega_i^{-1}$, *and* $\Sigma_{\alpha_i}=(G'_{\alpha_i}\Omega_i^{-1}G_{\alpha_i})^{-1}$.

PROOF. The results follow by Lemmas 16, 18, 19, and 23, noting that

$$E\big[\tilde{\psi}_{si}\tilde{\psi}'_{si}\big]=M_i^{\Omega}\begin{pmatrix}\Sigma_{\alpha_i} & 0\\ 0 & P_{\alpha_i}\end{pmatrix}M_i^{\Omega'},$$

$$E\big[\tilde{C}_i^{\Omega}\tilde{\psi}_i\big]=\sum_{j=0}^{\infty}E\big[G_{\theta_i}(z_{it})'P_{\alpha_i}g(z_{i,t-j})\big],\qquad E\big[\tilde{\psi}_{i,j}M_{i,j}^{\Omega}\tilde{\psi}_i\big]=0.$$

□

LEMMA 25. *Suppose that Conditions* 1–4 *hold. Then for* $\widehat{s}(\theta_0)=n^{-1}\sum_{i=1}^{n}\widehat{s}_i(\theta_0,\widehat{\gamma}_{i0})$,

$$\sqrt{nT}\widehat{s}(\theta_0)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\psi}_{si}+\sqrt{\frac{n}{T}}B_s+o_P(1)\xrightarrow{d}N(\kappa B_s,J_s),$$

where $\tilde{\psi}_{si}$ *and* B_s *are defined in Lemmas* 23 *and* 24, *respectively.*

Using the expansion form obtained in Lemma 23, we can get the result by examining each term with Lemma 24.

APPENDIX G: SCORES AND DERIVATIVES

G.1 *One-step score and derivatives: Individual effects*

We denote dimensions of $g(z_{it})$, α_i, and θ by d_g, d_α, and d_θ. The symbol $\otimes$ denotes Kronecker product of matrices, and I_{d_α} denotes a d_α-order identity matrix. Let $G_{\alpha\alpha_i}(z_{it};\theta,\alpha_i):=(G_{\alpha\alpha_{i,1}}(z_{it};\theta,\alpha_i)',\dots,G_{\alpha\alpha_{i,d_\alpha}}(z_{it};\theta,\alpha_i)')'$, where

$$G_{\alpha\alpha_{i,j}}(z_{it};\theta,\alpha_i)=\frac{\partial G_{\alpha_i}(z_{it};\theta,\alpha_i)}{\partial\alpha_{i,j}}.$$

We denote derivatives of $G_{\alpha\alpha_i}(z_{it};\theta,\alpha_i)$ with respect to $\alpha_{i,j}$ by $G_{\alpha\alpha,\alpha_{i,j}}(z_{it};\theta,\alpha_i)$, and use additional subscripts for higher-order derivatives.

G.1.1 *Score:*

$$\widehat{t}_i^W(\theta,\gamma_i) = -\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} G_{\alpha_i}(z_{it};\theta,\alpha_i)'\lambda_i \\ g(z_{it};\theta,\alpha_i)+\widehat{W}_i\lambda_i \end{pmatrix} = -\begin{pmatrix} \widehat{G}_{\alpha_i}(\theta,\alpha_i)'\lambda_i \\ \widehat{g}_i(\theta,\alpha_i)+\widehat{W}_i\lambda_i \end{pmatrix}.$$

G.1.2 *Derivatives with respect to the fixed effects*

First derivatives:

$$\widehat{T}_i^W(\theta,\gamma_i) = \frac{\partial \widehat{t}_i^W(\gamma_i,\theta)}{\partial \gamma_i'} = -\begin{pmatrix} \widehat{G}_{\alpha\alpha_i}(\theta,\alpha_i)'(I_{d_\alpha}\otimes\lambda_i) & \widehat{G}_{\alpha_i}(\theta,\alpha_i)' \\ \widehat{G}_{\alpha_i}(\theta,\alpha_i) & \widehat{W}_i \end{pmatrix},$$

$$T_i^W = E[\widehat{T}_i^W] = -\begin{pmatrix} 0 & G_{\alpha_i}' \\ G_{\alpha_i} & W_i \end{pmatrix},$$

$$(T_i^W)^{-1} = -\begin{pmatrix} -\Sigma_{\alpha_i}^W & H_{\alpha_i}^W \\ H_{\alpha_i}^{W'} & P_{\alpha_i}^W \end{pmatrix}.$$

Second derivatives:

$$\widehat{T}_{i,j}^W(\theta,\gamma_i) = \frac{\partial^2 \widehat{t}_i^W(\theta,\gamma_i)}{\partial\gamma_{i,j}\,\partial\gamma_i'}$$
$$= \begin{cases} -\begin{pmatrix} \widehat{G}_{\alpha\alpha,\alpha_{i,j}}(\theta,\alpha_i)'(I_{d_\alpha}\otimes\lambda_i) & \widehat{G}_{\alpha\alpha_{i,j}}(\theta,\alpha_i)' \\ \widehat{G}_{\alpha\alpha_{i,j}}(\theta,\alpha_i) & 0 \end{pmatrix}, & \text{if } j \le d_\alpha, \\ -\begin{pmatrix} \widehat{G}_{\alpha\alpha_i}(\theta,\alpha_i)'(I_{d_\alpha}\otimes e_{j-d_\alpha}) & 0 \\ 0 & 0 \end{pmatrix}, & \text{if } j > d_\alpha, \end{cases}$$

$$T_{i,j}^W = E[\widehat{T}_{i,j}^W(\gamma_{i0};\theta_0)] = \begin{cases} -\begin{pmatrix} 0 & G_{\alpha\alpha_{i,j}}' \\ G_{\alpha\alpha_{i,j}} & 0 \end{pmatrix}, & \text{if } j \le d_\alpha, \\ -\begin{pmatrix} G_{\alpha\alpha_i}'(I_{d_\alpha}\otimes e_{j-d_\alpha}) & 0 \\ 0 & 0 \end{pmatrix}, & \text{if } j > d_\alpha. \end{cases}$$

Third derivatives:

$$\widehat{T}_{i,jk}^W(\theta,\gamma_i) = \frac{\partial^3 \widehat{t}_i^W(\theta,\gamma_i)}{\partial\gamma_{i,k}\,\partial\gamma_{i,j}\,\partial\gamma_i'}$$

$$= \begin{cases} -\begin{pmatrix} \widehat{G}_{\alpha\alpha,\alpha\alpha_{i,jk}}(\theta,\alpha_i)'(I_{d_\alpha}\otimes\lambda_i) & \widehat{G}_{\alpha\alpha\alpha_{i,jk}}(\theta,\alpha_i)' \\ \widehat{G}_{\alpha\alpha\alpha_{i,jk}}(\theta,\alpha_i) & 0 \end{pmatrix}, \\ \qquad \text{if } j \le d_\alpha, k \le d_\alpha, \\ -\begin{pmatrix} \widehat{G}_{\alpha\alpha,\alpha_{i,j}}(\theta,\alpha_i)'(I_{d_\alpha}\otimes e_{k-d_\alpha}) & 0 \\ 0 & 0 \end{pmatrix}, \\ \qquad \text{if } j \le d_\alpha, k > d_\alpha, \\ -\begin{pmatrix} \widehat{G}_{\alpha\alpha,\alpha_{i,k}}(\theta,\alpha_i)'(I_{d_\alpha}\otimes e_{j-d_\alpha}) & 0 \\ 0 & 0 \end{pmatrix}, \\ \qquad \text{if } j > d_\alpha, k \le d_\alpha, \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{if } j > d_\alpha, k > d_\alpha, \end{cases}$$

$$T^W_{i,jk} = E\big[\widehat{T}^W_{i,jk}\big] = \begin{cases} -\begin{pmatrix} 0 & G'_{\alpha\alpha\alpha_{i,jk}} \\ G_{\alpha\alpha\alpha_{i,jk}} & 0 \end{pmatrix}, & \text{if } j \le d_\alpha, k \le d_\alpha, \\ -\begin{pmatrix} G'_{\alpha\alpha,\alpha_{i,j}}(I_{d_\alpha}\otimes e_{k-d_\alpha}) & 0 \\ 0 & 0 \end{pmatrix}, & \text{if } j \le d_\alpha, k > d_\alpha, \\ -\begin{pmatrix} G'_{\alpha\alpha,\alpha_{i,k}}(I_{d_\alpha}\otimes e_{j-d_\alpha}) & 0 \\ 0 & 0 \end{pmatrix}, & \text{if } j > d_\alpha, k \le d_\alpha, \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, & \text{if } j > d_\alpha, k > d_\alpha. \end{cases}$$

G.1.3 *Derivatives with respect to the common parameter*

First derivatives:

$$\widehat{N}^W_{i,j}(\theta,\gamma_i) = \frac{\partial \widehat{t}^W_i(\gamma_i,\theta)}{\partial\theta_j} = -\begin{pmatrix} \widehat{G}_{\theta_j\alpha_i}(\theta,\alpha_i)'\lambda_i \\ \widehat{G}_{\theta_{i,j}}(\theta,\alpha_i) \end{pmatrix},$$

$$N^W_{i,j} = E\big[\widehat{N}^W_{i,j}\big] = -\begin{pmatrix} 0 \\ G_{\theta_{i,j}} \end{pmatrix}.$$

G.2 *One-step score and derivatives: Common parameters*

Let $G_{\theta\alpha_i}(z_{it};\theta,\alpha_i) := (G_{\theta\alpha_{i,1}}(z_{it};\theta,\alpha_i)',\dots,G_{\theta\alpha_{i,d_\alpha}}(z_{it};\theta,\alpha_i)')'$, where

$$G_{\theta\alpha_{i,j}}(z_{it};\theta,\alpha_i) = \frac{\partial G_\theta(z_{it};\theta,\alpha_i)}{\partial\alpha_{i,j}}.$$

We denote the derivatives of $G_{\theta\alpha_i}(z_{it};\theta,\alpha_i)$ with respect to $\alpha_{i,j}$ by $G_{\theta\alpha,\alpha_{i,j}}(z_{it};\theta,\alpha_i)$, and use additional subscripts for higher-order derivatives.

G.2.1 *Score:*

$$\widehat{s}^W_i(\theta,\gamma_i) = -\frac{1}{T}\sum_{t=1}^T G_\theta(z_{it};\theta,\alpha_i)'\lambda_i = -\widehat{G}_{\theta_i}(\theta,\alpha_i)'\lambda_i.$$

G.2.2 *Derivatives with respect to the fixed effects*
First derivatives:

$$\widehat{M}_i^W(\theta,\gamma_i) = \frac{\partial \widehat{s}_i^W(\theta,\gamma_i)}{\partial \gamma_i'} = -\big(\widehat{G}_{\theta\alpha_i}(\theta,\alpha_i)'(I_{d_\alpha}\otimes\lambda_i) \quad \widehat{G}_{\theta_i}(\theta,\alpha_i)'\big),$$
$$M_i^W = E\big[\widehat{M}_i^W\big] = -\big(0 \quad G_{\theta_i}'\big).$$

Second derivatives:

$$\widehat{M}_{i,j}^W(\theta,\gamma_i) = \frac{\partial^2 \widehat{s}_i^W(\theta,\gamma_i)}{\partial \gamma_{i,j}\,\partial \gamma_i'}$$
$$= \begin{cases} -\big(\widehat{G}_{\theta\alpha,\alpha_{i,j}}(\theta,\alpha_i)'(I_{d_\alpha}\otimes\lambda_i) \quad \widehat{G}_{\theta\alpha_{i,j}}(\theta,\alpha_i)'\big), & \text{if } j\le d_\alpha, \\ -\big(\widehat{G}_{\theta\alpha_i}(\theta,\alpha_i)'(I_{d_\alpha}\otimes e_{j-d_\alpha}) \quad 0\big), & \text{if } j> d_\alpha, \end{cases}$$
$$M_{i,j}^W = E\big[\widehat{M}_{i,j}^W(\theta_0,\gamma_{i0})\big] = \begin{cases} -\big(0 \quad G_{\theta\alpha_{i,j}}'\big), & \text{if } j\le d_\alpha, \\ -\big(G_{\theta\alpha_i}'(I_{d_\alpha}\otimes e_{j-d_\alpha}) \quad 0\big), & \text{if } j> d_\alpha. \end{cases}$$

Third derivatives:

$$\widehat{M}_{i,jk}^W(\theta,\gamma_i) = \frac{\partial^3 \widehat{s}_i^W(\theta,\gamma_i)}{\partial \gamma_{i,k}\,\partial \gamma_{i,j}\,\partial \gamma_i'}$$
$$= \begin{cases} -\big(\widehat{G}_{\theta\alpha,\alpha\alpha_{i,jk}}(\theta,\alpha_i)'(I_{d_\alpha}\otimes\lambda_i) \quad \widehat{G}_{\theta\alpha\alpha_{i,jk}}(\theta,\alpha_i)'\big), \\ \qquad \text{if } j\le d_\alpha, k\le d_\alpha, \\ -\big(\widehat{G}_{\theta\alpha,\alpha_{i,j}}(\theta,\alpha_i)'(I_{d_\alpha}\otimes e_{k-d_\alpha}) \quad 0\big), \\ \qquad \text{if } j\le d_\alpha, k> d_\alpha, \\ -\big(\widehat{G}_{\theta\alpha,\alpha_{i,k}}(\theta,\alpha_i)'(I_{d_\alpha}\otimes e_{j-d_\alpha}) \quad 0\big), \\ \qquad \text{if } j> d_\alpha, k\le d_\alpha, \\ -(0 \quad 0), \quad \text{if } j> d_\alpha, k> d_\alpha, \end{cases}$$
$$M_{i,jk}^W = E\big[\widehat{M}_{i,jk}^W\big] = \begin{cases} -\big(0 \quad G_{\theta\alpha\alpha_{i,jk}}'\big), & \text{if } j\le d_\alpha, k\le d_\alpha, \\ -\big(G_{\theta\alpha,\alpha_{i,j}}'(I_{d_\alpha}\otimes e_{k-d_\alpha}) \quad 0\big), & \text{if } j\le d_\alpha, k> d_\alpha, \\ -\big(G_{\theta\alpha,\alpha_{i,k}}'(I_{d_\alpha}\otimes e_{j-d_\alpha}) \quad 0\big), & \text{if } j> d_\alpha, k\le d_\alpha, \\ -(0 \quad 0), & \text{if } j> d_\alpha, k> d_\alpha. \end{cases}$$

G.2.3 *Derivatives with respect to the common parameters*
First derivatives:

$$\widehat{S}_{i,j}^W(\theta,\gamma_i) = \frac{\partial \widehat{s}_i^W(\theta,\gamma_i)}{\partial \theta_j} = -\widehat{G}_{\theta\theta_{i,j}}(\theta,\alpha_i)'\lambda_i,$$
$$S_{i,j}^W = E\big[\widehat{S}_{i,j}^W\big] = 0.$$

G.3 *Two-step score and derivatives: Fixed effects*

G.3.1 *Score:*

$$\begin{aligned}\widehat{t}_i(\theta,\gamma_i) &= -\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} G_{\alpha_i}(z_{it};\theta,\alpha_i)'\lambda_i \\ g(z_{it};\theta,\alpha_i)+\widehat{\Omega}_i(\tilde{\theta},\tilde{\alpha}_i)\lambda_i\end{pmatrix} \\ &= -\begin{pmatrix}\widehat{G}_{\alpha_i}(\theta,\alpha_i)'\lambda_i \\ \widehat{g}_i(\theta,\alpha_i)+\Omega_i\lambda_i\end{pmatrix} - \begin{pmatrix} 0 \\ (\widehat{\Omega}_i-\Omega_i)\lambda_i\end{pmatrix} \\ &= \widehat{t}_i^{\Omega}(\theta,\gamma_i)+\widehat{t}_i^{R}(\theta,\gamma_i).\end{aligned}$$

Note that the formulae for the derivatives of Appendix G.1 apply for $\widehat{t}_i^{\Omega}$, replacing $\widehat{W}$ by Ω. Hence, we only need to obtain the derivatives for $\widehat{t}_i^{R}$.

G.3.2 *Derivatives with respect to the fixed effects*

First derivatives:

$$\widehat{T}_i^{R}(\theta,\gamma_i) = \frac{\partial \widehat{t}_i^{R}(\theta,\gamma_i)}{\partial \gamma_i'} = -\begin{pmatrix} 0 & 0 \\ 0 & \widehat{\Omega}_i(\tilde{\theta},\tilde{\alpha}_i)-\Omega_i\end{pmatrix},$$

$$T_i^{R} = E\big[\widehat{T}_i^{R}\big] = -\begin{pmatrix} 0 & 0 \\ 0 & E[\widehat{\Omega}_i-\Omega_i]\end{pmatrix}.$$

Second and third derivatives:

Since $\widehat{T}_i^{R}(\gamma_i,\theta)$ does not depend on γ_i, the derivatives (and its expectation) of order greater than 1 are 0.

G.3.3 *Derivatives with respect to the common parameters*

First derivatives:

$$\widehat{N}_i^{R}(\theta,\gamma_i) = \frac{\partial \widehat{t}_i^{R}(\theta,\gamma_i)}{\partial \theta'} = 0.$$

G.4 *Two-step score and derivatives: Common parameters*

G.4.1 *Score:*

$$\widehat{s}_i(\theta,\gamma_i) = -\frac{1}{T}\sum_{t=1}^{T} G_{\theta}(z_{it};\theta,\alpha_i)'\lambda_i = -\widehat{G}_{\theta_i}(\theta,\alpha_i)'\lambda_i.$$

Since this score does not depend explicitly on $\widehat{\Omega}_i(\tilde{\theta},\tilde{\alpha}_i)$, the formulae for the derivatives are the same as in Appendix G.2.

References

Andrews, D. W. K. (1991), "Heteroskedasticity and autocorrelation consistent covariance matrix estimation." *Econometrica*, 59, 817–858. [19]

Hahn, J. and G. Kuersteiner (2011), "Bias reduction for dynamic nonlinear panel models with fixed effects." *Econometric Theory*, 27, 1152–1191. [1]

Hall, P. and C. Heide (1980), *Martingale Limit Theory and Applications*. Academic Press, San Diego. [19]

Newey, W. K. and D. McFadden (1994), "Large sample estimation and hypothesis testing." In *Handbook of Econometrics*, Vol. 4 (R. F. Engle and D. L. McFadden, eds.), North-Holland, Amsterdam. [8]

Newey, W. K. and R. Smith (2004), "Higher order properties of GMM and generalized empirical likelihood estimators." *Econometrica*, 72, 219–255. [19]

Submitted April, 2010. Final version accepted January, 2013.