Supplement to "Irregular N2SLS and LASSO estimation of the matrix exponential spatial specification model"

Fei Jin[a,b] and Lung-fei Lee[c*]

[a]School of Economics, Shanghai University of Finance and Economics, Shanghai 200433, China

[b]Key Laboratory of Mathematical Economics (SUFE), Ministry of Education, Shanghai 200433, China

[c]Department of Economics, The Ohio State University, Columbus, OH 43210, USA

June 2, 2016

1 Additional Monte Carlo results

Tables 1.1–1.5 present Monte Carlo results in addition to those in the main text.

2 Lemma and Proofs

The proofs frequently use the following lemma on the uniform convergence of a quadratic form, which we prove first.

Lemma 1. *Suppose that A_n, B_n and C_n are $n \times n$ nonstochastic matrices bounded in both row and column sum norms, $u_{n1} = (u_{n1,1}, \ldots, u_{n1,n})'$ and $u_{n2} = (u_{n2,1}, \ldots, u_{n2,n})'$ are $n \times 1$ vectors, $(u_{n1,i}, u_{n2,i})$'s are i.i.d., $\mathrm{E}(u_{n1,i}^4) < \infty$ and $\mathrm{E}(u_{n2,i}^4) < \infty$. Then $\frac{1}{n} u_{n1}' A_n e^{\alpha B_n} C_n u_{n2} - \frac{1}{n}\mathrm{E}(u_{n1}' A_n e^{\alpha B_n} C_n u_{n2}) = o_p(1)$ uniformly in $\alpha \in [-\eta, \eta]$ and $\frac{1}{n}\mathrm{E}(u_{n1}' A_n e^{\alpha B_n} C_n u_{n2})$ is uniformly equicontinuous.*

Proof. Since B_n is bounded in row sum norm, $\|e^{\alpha B_n}\|_\infty = \|\sum_{i=0}^\infty \frac{\alpha^i B_n^i}{i!}\|_\infty \leq \sum_{i=0}^\infty \frac{|\alpha|^i \|B_n\|_\infty^i}{i!} = e^{|\alpha| \|B_n\|_\infty} \leq e^{|\eta| \|B_n\|_\infty}$. Thus, $e^{\alpha B_n}$ is bounded in row sum norm. Similarly, $e^{\alpha B_n}$ is bounded in column sum norm. Then $A_n e^{\alpha B_n} C_n$ is bounded in both row and column sum norms. Thus, for each $\alpha \in [-\eta, \eta]$, $\frac{1}{n} u_{n1}' A_n e^{\alpha B_n} C_n u_{n2} - \frac{1}{n}\mathrm{E}(u_{n1}' A_n e^{\alpha B_n} C_n u_{n2}) = o_p(1)$ by Lemma 1 in Qu and Lee (2012). By the mean value theorem, for α_1 and α_2 in $[-\eta, \eta]$,

$$[\frac{1}{n} u_{n1}' A_n e^{\alpha_1 B_n} C_n u_{n2} - \frac{1}{n}\mathrm{E}(u_{n1}' A_n e^{\alpha_1 B_n} C_n u_{n2})] - [\frac{1}{n} u_{n1}' A_n e^{\alpha_2 B_n} C_n u_{n2} - \frac{1}{n}\mathrm{E}(u_{n1}' A_n e^{\alpha_2 B_n} C_n u_{n2})]$$

*Corresponding author. Tel.: +1 614 292 5508; fax: +1 614 292 4192. E-mail addresses: jin.fei@sufe.edu.cn (F. Jin), lee.1777@osu.edu (L.-F. Lee).

Table 1.1: Ratios of the RMSE when $n = 144$ to that when $n = 400$

	α	β_1	β_2	β_3	β_4
$\zeta_0 = 0$					
queen, $R^2 = 0.2$, $\alpha_0 = -0.2$	1.274[1.639]1.724	1.738[1.711]1.801	1.016[1.573]1.642	1.260[—]—	1.795[—]—
queen, $R^2 = 0.2$, $\alpha_0 = -1$	1.179[1.569]1.678	1.664[1.660]1.707	1.001[1.531]1.455	1.196[—]—	1.702[—]—
rook, $R^2 = 0.2$, $\alpha_0 = -0.2$	1.155[1.650]1.686	1.550[1.645]1.677	0.961[1.775]1.706	1.224[—]—	1.689[—]—
rook, $R^2 = 0.2$, $\alpha_0 = -1$	1.144[1.607]1.767	1.617[1.710]1.756	1.075[1.794]2.029	1.242[—]—	1.668[—]—
queen, $R^2 = 0.8$, $\alpha_0 = -0.2$	1.323[1.638]2.281	1.672[1.670]1.753	1.233[1.634]2.388	1.276[—]—	1.683[—]—
queen, $R^2 = 0.8$, $\alpha_0 = -1$	1.337[1.663]2.354	1.749[1.681]1.781	1.586[1.703]2.400	1.340[—]—	1.750[—]—
rook, $R^2 = 0.8$, $\alpha_0 = -0.2$	1.306[1.584]2.498	1.716[1.705]1.768	1.379[1.606]2.664	1.325[—]—	1.699[—]—
rook, $R^2 = 0.8$, $\alpha_0 = -1$	1.356[1.649]2.956	1.708[1.701]1.763	1.488[1.643]3.269	1.334[—]—	1.765[—]—
$\zeta_0 = (1, 1)$					
queen, $R^2 = 0.2$, $\alpha_0 = -0.2$	1.763[1.581]1.763	1.782[1.696]1.781	2.395[2.493]2.395	1.679[—]1.679	1.676[—]1.676
queen, $R^2 = 0.2$, $\alpha_0 = -1$	1.705[1.550]1.705	1.748[1.745]1.747	1.827[1.824]1.827	1.678[—]1.677	1.700[—]1.700
rook, $R^2 = 0.2$, $\alpha_0 = -0.2$	1.691[1.720]1.691	1.683[1.751]1.683	1.789[2.132]1.789	1.626[—]1.626	1.782[—]1.782
rook, $R^2 = 0.2$, $\alpha_0 = -1$	1.928[1.711]1.928	1.725[1.729]1.725	1.944[2.173]1.944	1.720[—]1.720	1.721[—]1.721
queen, $R^2 = 0.8$, $\alpha_0 = -0.2$	1.750[1.035]1.750	1.697[1.470]1.697	1.797[1.008]1.797	1.671[—]1.671	1.777[—]1.777
queen, $R^2 = 0.8$, $\alpha_0 = -1$	1.753[1.054]1.753	1.690[1.424]1.690	1.730[1.026]1.730	1.674[—]1.674	1.694[—]1.694
rook, $R^2 = 0.8$, $\alpha_0 = -0.2$	1.723[1.072]1.749	1.719[1.262]1.719	1.715[1.029]1.721	1.740[—]1.765	1.730[—]1.731
rook, $R^2 = 0.8$, $\alpha_0 = -1$	1.775[1.044]1.775	1.741[1.260]1.741	1.817[1.018]1.817	1.744[—]1.744	1.712[—]1.712
$\zeta_0 = (0, 1)$					
queen, $R^2 = 0.2$, $\alpha_0 = -0.2$	1.777[1.801]1.778	1.725[1.694]1.725	2.168[2.519]2.170	1.643[—]1.644	1.746[—]1.746
queen, $R^2 = 0.2$, $\alpha_0 = -1$	1.672[1.642]1.672	1.733[1.707]1.733	1.933[2.649]1.932	1.709[—]1.709	1.730[—]1.730
rook, $R^2 = 0.2$, $\alpha_0 = -0.2$	1.759[1.732]1.759	1.639[1.609]1.639	1.819[1.979]1.819	1.670[—]1.670	1.669[—]1.669
rook, $R^2 = 0.2$, $\alpha_0 = -1$	1.764[1.683]1.764	1.659[1.648]1.659	1.757[1.768]1.757	1.674[—]1.674	1.739[—]1.739
queen, $R^2 = 0.8$, $\alpha_0 = -0.2$	1.813[1.670]1.813	1.809[1.699]1.808	2.167[1.747]2.166	1.851[—]1.851	1.763[—]1.763
queen, $R^2 = 0.8$, $\alpha_0 = -1$	1.697[1.689]1.696	1.721[1.681]1.721	2.400[1.826]2.400	1.693[—]1.693	1.692[—]1.692
rook, $R^2 = 0.8$, $\alpha_0 = -0.2$	1.811[1.665]1.811	1.724[1.659]1.724	1.856[1.715]1.856	1.808[—]1.808	1.770[—]1.770
rook, $R^2 = 0.8$, $\alpha_0 = -1$	1.919[1.690]1.919	1.764[1.691]1.764	2.087[1.686]2.087	1.850[—]1.850	1.793[—]1.793

The numbers show the ratios of the RMSE when $n = 144$ to that when $n = 400$ in each case. The three numbers in each cell correspond to: N2SLS[N2SLS-r]AGLASSO. $\beta_{10} = 1$ and $\beta_{20} = 1$.

Table 1.2: Biases, SEs and CPs when $\zeta_0 = 0$ and $n = 400$

	α	β_1	β_2	β_3	β_4
queen, $R^2 = 0.2$, $\alpha_0 = -0.2$					
N2SLS	-0.419[0.705]0.968	0.036[0.117]0.993	-0.147[0.784]0.810	-0.320[0.688]0.989	0.000[0.049]0.992
N2SLS-r	-0.043[0.288]0.989	-0.005[0.105]0.986	-0.000[0.341]0.963	0.000[0.000]—	0.000[0.000]—
AGLASSO	-0.065[0.381]0.973	-0.007[0.110]0.984	0.005[0.404]0.934	-0.038[0.213]—	0.000[0.013]—
queen, $R^2 = 0.2$, $\alpha_0 = -1$					
N2SLS	-0.319[0.795]0.960	0.042[0.119]0.994	0.043[1.240]0.825	-0.203[0.797]0.975	0.001[0.050]0.994
N2SLS-r	-0.052[0.293]0.985	-0.001[0.105]0.988	-0.007[0.345]0.951	0.000[0.000]—	0.000[0.000]—
AGLASSO	-0.077[0.375]0.974	-0.004[0.110]0.988	-0.010[0.399]0.924	-0.027[0.190]—	0.000[0.013]—
rook, $R^2 = 0.2$, $\alpha_0 = -0.2$					
N2SLS	-0.041[0.544]0.969	0.037[0.124]0.991	0.127[0.894]0.916	-0.052[0.562]0.995	0.000[0.050]0.991
N2SLS-r	-0.014[0.190]0.994	-0.002[0.109]0.985	0.005[0.230]0.982	0.000[0.000]—	0.000[0.000]—
AGLASSO	-0.017[0.252]0.968	-0.002[0.112]0.981	0.015[0.303]0.961	-0.011[0.112]—	0.000[0.012]—
rook, $R^2 = 0.2$, $\alpha_0 = -1$					
N2SLS	0.075[0.592]0.961	0.039[0.123]0.994	0.332[1.227]0.940	0.058[0.581]0.992	-0.001[0.051]0.989
N2SLS-r	-0.010[0.192]0.994	-0.003[0.103]0.985	0.010[0.236]0.977	0.000[0.000]—	0.000[0.000]—
AGLASSO	-0.036[0.235]0.981	-0.001[0.106]0.987	-0.007[0.266]0.955	-0.012[0.123]—	0.001[0.014]—
queen, $R^2 = 0.8$, $\alpha_0 = -0.2$					
N2SLS	-0.046[0.404]0.954	0.011[0.030]0.987	0.040[0.457]0.938	-0.019[0.416]0.952	0.002[0.052]0.990
N2SLS-r	-0.002[0.076]0.977	0.000[0.026]0.984	0.002[0.083]0.976	0.000[0.000]—	0.000[0.000]—
AGLASSO	-0.019[0.139]0.955	0.001[0.027]0.984	-0.010[0.120]0.957	-0.013[0.102]—	0.001[0.015]—
queen, $R^2 = 0.8$, $\alpha_0 = -1$					
N2SLS	0.040[0.441]0.926	0.012[0.031]0.991	0.151[0.572]0.958	0.068[0.450]0.922	0.003[0.050]0.984
N2SLS-r	-0.002[0.076]0.980	-0.000[0.026]0.986	-0.000[0.083]0.983	0.000[0.000]—	0.000[0.000]—
AGLASSO	-0.016[0.117]0.970	0.001[0.026]0.986	-0.011[0.116]0.965	-0.006[0.083]—	0.001[0.014]—
rook, $R^2 = 0.8$, $\alpha_0 = -0.2$					
N2SLS	0.090[0.275]0.965	0.009[0.029]0.990	0.137[0.330]0.988	0.087[0.266]0.982	0.002[0.050]0.988
N2SLS-r	-0.002[0.051]0.986	-0.001[0.026]0.986	-0.001[0.060]0.980	0.000[0.000]—	0.000[0.000]—
AGLASSO	-0.009[0.070]0.978	0.000[0.026]0.986	-0.007[0.078]0.970	-0.002[0.047]—	0.000[0.014]—
rook, $R^2 = 0.8$, $\alpha_0 = -1$					
N2SLS	0.104[0.272]0.958	0.010[0.029]0.990	0.154[0.333]0.991	0.099[0.261]0.981	0.001[0.049]0.989
N2SLS-r	-0.001[0.050]0.991	0.000[0.026]0.985	0.002[0.059]0.985	0.000[0.000]—	0.000[0.000]—
AGLASSO	-0.006[0.059]0.986	0.001[0.026]0.985	-0.003[0.067]0.980	-0.001[0.031]—	0.001[0.012]—

"N2SLS" denotes the unrestricted N2SLS estimator and "N2SLS-r" denotes the restricted N2SLS estimator with the restriction $\zeta = 0$ imposed. The three numbers in each cell are: bias[SE]CP. $\beta_{10} = 1$ and $\beta_{20} = 1$.

Table 1.3: Biases, SEs and CPs when $\zeta_0 = (1,1)'$ and $n = 400$

	α	β_1	β_2	β_3	β_4
queen, $R^2 = 0.2$, $\alpha_0 = -0.2$					
N2SLS	0.006[0.160]0.987	0.003[0.113]0.989	0.017[0.224]0.983	0.012[0.388]0.985	-0.003[0.053]0.989
N2SLS-r	-0.125[0.284]1.000	-0.023[0.230]0.986	-0.081[0.377]0.984	-1.000[0.000]—	-1.000[0.000]—
AGLASSO	0.006[0.160]0.987	0.003[0.113]0.989	0.017[0.224]0.983	0.011[0.388]0.985	-0.003[0.053]0.989
queen, $R^2 = 0.2$, $\alpha_0 = -1$					
N2SLS	0.004[0.155]0.989	-0.001[0.113]0.988	0.017[0.219]0.979	-0.007[0.374]0.987	-0.002[0.053]0.986
N2SLS-r	-0.128[0.276]1.000	-0.021[0.225]0.982	-0.085[0.399]0.979	-1.000[0.000]—	-1.000[0.000]—
AGLASSO	0.003[0.155]0.989	-0.001[0.113]0.988	0.017[0.219]0.979	-0.007[0.374]0.987	-0.002[0.053]0.986
rook, $R^2 = 0.2$, $\alpha_0 = -0.2$					
N2SLS	-0.001[0.103]0.993	0.004[0.120]0.986	0.006[0.183]0.984	0.000[0.243]0.991	-0.004[0.051]0.993
N2SLS-r	-0.089[0.201]1.000	-0.021[0.237]0.984	-0.065[0.312]0.987	-1.000[0.000]—	-1.000[0.000]—
AGLASSO	-0.001[0.103]0.993	0.004[0.120]0.986	0.006[0.183]0.984	0.000[0.243]0.992	-0.004[0.051]0.993
rook, $R^2 = 0.2$, $\alpha_0 = -1$					
N2SLS	-0.001[0.099]0.994	0.002[0.117]0.986	0.005[0.173]0.989	-0.009[0.242]0.990	-0.003[0.053]0.984
N2SLS-r	-0.087[0.199]1.000	-0.023[0.237]0.981	-0.064[0.305]0.989	-1.000[0.000]—	-1.000[0.000]—
AGLASSO	-0.001[0.099]0.994	0.002[0.117]0.986	0.005[0.173]0.989	-0.009[0.242]0.990	-0.003[0.053]0.984
queen, $R^2 = 0.8$, $\alpha_0 = -0.2$					
N2SLS	-0.003[0.118]0.990	0.001[0.032]0.988	0.004[0.125]0.984	-0.000[0.176]0.983	0.000[0.051]0.990
N2SLS-r	-0.624[0.121]0.034	-0.049[0.058]0.949	-0.461[0.074]0.028	-1.000[0.000]—	-1.000[0.000]—
AGLASSO	-0.003[0.118]0.990	0.001[0.032]0.988	0.004[0.125]0.984	-0.000[0.176]0.983	0.000[0.051]0.990
queen, $R^2 = 0.8$, $\alpha_0 = -1$					
N2SLS	-0.003[0.119]0.994	-0.000[0.032]0.990	0.004[0.127]0.988	-0.001[0.176]0.990	0.000[0.053]0.986
N2SLS-r	-0.621[0.125]0.041	-0.052[0.062]0.933	-0.458[0.076]0.035	-1.000[0.000]—	-1.000[0.000]—
AGLASSO	-0.004[0.119]0.994	-0.000[0.032]0.990	0.004[0.127]0.988	-0.001[0.176]0.990	0.000[0.053]0.986
rook, $R^2 = 0.8$, $\alpha_0 = -0.2$					
N2SLS	-0.001[0.072]0.989	-0.000[0.034]0.988	0.003[0.082]0.987	-0.003[0.083]0.993	0.000[0.051]0.993
N2SLS-r	-0.547[0.127]0.099	-0.098[0.069]0.803	-0.417[0.086]0.091	-1.000[0.000]—	-1.000[0.000]—
AGLASSO	-0.001[0.072]0.989	-0.000[0.034]0.988	0.002[0.082]0.987	-0.003[0.083]0.993	0.000[0.051]0.993
rook, $R^2 = 0.8$, $\alpha_0 = -1$					
N2SLS	-0.002[0.071]0.993	0.000[0.033]0.991	0.002[0.079]0.989	0.000[0.083]0.993	-0.001[0.050]0.995
N2SLS-r	-0.552[0.123]0.080	-0.096[0.069]0.822	-0.420[0.081]0.077	-1.000[0.000]—	-1.000[0.000]—
AGLASSO	-0.002[0.071]0.993	0.000[0.033]0.991	0.002[0.079]0.989	0.000[0.083]0.993	-0.001[0.050]0.995

"N2SLS" denotes the unrestricted N2SLS estimator and "N2SLS-r" denotes the restricted N2SLS estimator with the restriction $\zeta = 0$ imposed. The three numbers in each cell are: bias[SE]CP. $\beta_{10} = 1$ and $\beta_{20} = 1$.

Table 1.4: Biases, SEs and CPs when $\zeta_0 = (0,1)'$ and $n = 400$

	α	β_1	β_2	β_3	β_4
queen, $R^2 = 0.2$, $\alpha_0 = -0.2$					
N2SLS	0.002[0.157]0.994	0.006[0.103]0.987	0.017[0.226]0.980	0.001[0.356]0.986	-0.003[0.052]0.988
N2SLS-r	0.010[0.283]1.000	0.003[0.213]0.990	0.061[0.408]0.995	0.000[0.000]—	-1.000[0.000]—
AGLASSO	0.002[0.157]0.994	0.006[0.103]0.987	0.017[0.225]0.980	0.001[0.356]0.986	-0.003[0.052]0.988
queen, $R^2 = 0.2$, $\alpha_0 = -1$					
N2SLS	0.001[0.161]0.987	0.001[0.102]0.993	0.014[0.218]0.973	0.014[0.346]0.987	-0.002[0.052]0.990
N2SLS-r	0.010[0.301]1.000	-0.000[0.210]0.983	0.057[0.425]0.995	0.000[0.000]—	-1.000[0.000]—
AGLASSO	0.001[0.161]0.987	0.001[0.102]0.993	0.014[0.218]0.973	0.014[0.346]0.986	-0.002[0.052]0.989
rook, $R^2 = 0.2$, $\alpha_0 = -0.2$					
N2SLS	0.001[0.103]0.997	0.002[0.107]0.984	0.006[0.166]0.989	-0.005[0.225]0.987	-0.002[0.053]0.988
N2SLS-r	0.023[0.201]1.000	0.000[0.219]0.977	0.040[0.330]0.996	0.000[0.000]—	-1.000[0.000]—
AGLASSO	0.001[0.103]0.997	0.002[0.107]0.984	0.006[0.166]0.989	-0.005[0.225]0.987	-0.002[0.053]0.988
rook, $R^2 = 0.2$, $\alpha_0 = -1$					
N2SLS	0.001[0.107]0.993	-0.001[0.105]0.989	0.005[0.169]0.984	0.000[0.225]0.989	-0.001[0.052]0.988
N2SLS-r	0.020[0.211]1.000	-0.006[0.218]0.980	0.046[0.346]0.997	0.000[0.000]—	-1.000[0.000]—
AGLASSO	0.001[0.107]0.993	-0.001[0.105]0.989	0.005[0.169]0.984	0.000[0.225]0.988	-0.001[0.052]0.987
queen, $R^2 = 0.8$, $\alpha_0 = -0.2$					
N2SLS	0.003[0.151]0.992	0.001[0.025]0.991	0.014[0.167]0.984	0.003[0.168]0.985	-0.003[0.052]0.988
N2SLS-r	0.008[0.139]0.989	-0.001[0.053]0.984	0.017[0.156]0.980	0.000[0.000]—	-1.000[0.000]—
AGLASSO	0.003[0.151]0.992	0.001[0.025]0.991	0.014[0.167]0.984	0.003[0.168]0.984	-0.003[0.052]0.987
queen, $R^2 = 0.8$, $\alpha_0 = -1$					
N2SLS	0.006[0.162]0.985	0.000[0.027]0.993	0.019[0.180]0.980	0.009[0.179]0.983	-0.005[0.053]0.990
N2SLS-r	0.005[0.137]0.991	-0.003[0.055]0.985	0.012[0.154]0.986	0.000[0.000]—	-1.000[0.000]—
AGLASSO	0.006[0.162]0.985	0.000[0.027]0.993	0.019[0.180]0.980	0.009[0.179]0.982	-0.005[0.053]0.989
rook, $R^2 = 0.8$, $\alpha_0 = -0.2$					
N2SLS	0.004[0.104]0.987	0.000[0.026]0.992	0.009[0.111]0.985	0.002[0.113]0.989	-0.003[0.052]0.987
N2SLS-r	0.007[0.097]0.991	-0.002[0.054]0.984	0.011[0.114]0.986	0.000[0.000]—	-1.000[0.000]—
AGLASSO	0.004[0.104]0.987	0.000[0.026]0.992	0.009[0.111]0.985	0.002[0.113]0.988	-0.003[0.052]0.986
rook, $R^2 = 0.8$, $\alpha_0 = -1$					
N2SLS	-0.001[0.101]0.991	0.003[0.026]0.986	0.005[0.108]0.992	-0.003[0.111]0.987	-0.002[0.052]0.987
N2SLS-r	0.008[0.097]0.989	0.001[0.053]0.985	0.013[0.114]0.990	0.000[0.000]—	-1.000[0.000]—
AGLASSO	-0.001[0.101]0.991	0.003[0.026]0.986	0.005[0.108]0.992	-0.003[0.111]0.987	-0.002[0.052]0.987

"N2SLS" denotes the unrestricted N2SLS estimator and "N2SLS-r" denotes the restricted N2SLS estimator with the restriction $\zeta = 0$ imposed. The three numbers in each cell are: bias[SE]CP. $\beta_{10} = 1$ and $\beta_{20} = 1$.

Table 1.5: Size and power of the distance difference and gradient tests in the homoskedastic case

	distance difference test							gradient test						
	size	power						size	power					
		(1)	(2)	(3)	(4)	(5)	(6)		(1)	(2)	(3)	(4)	(5)	(6)
W_n, R^2, α_0	n=144													
queen, 0.2, −0.2	0.062	0.977	0.999	1.000	1.000	1.000	1.000	0.045	0.986	0.999	0.998	0.999	0.997	0.995
queen, 0.2, −1	0.063	0.974	1.000	1.000	1.000	1.000	1.000	0.054	0.980	1.000	0.997	0.997	0.998	0.998
rook, 0.2, −0.2	0.048	0.971	0.999	1.000	1.000	1.000	1.000	0.041	0.973	0.998	0.997	0.995	0.997	0.996
rook, 0.2, −1	0.048	0.973	0.999	1.000	1.000	1.000	1.000	0.032	0.974	0.999	0.999	0.997	0.997	0.992
queen, 0.8, −0.2	0.057	0.991	1.000	1.000	1.000	1.000	1.000	0.054	0.992	1.000	1.000	1.000	1.000	1.000
queen, 0.8, −1	0.051	0.987	1.000	1.000	1.000	1.000	1.000	0.046	0.988	1.000	1.000	1.000	1.000	1.000
rook, 0.8, −0.2	0.064	0.986	1.000	1.000	1.000	1.000	1.000	0.059	0.960	1.000	1.000	1.000	1.000	1.000
rook, 0.8, −1	0.059	0.983	1.000	1.000	1.000	1.000	1.000	0.055	0.961	1.000	1.000	1.000	1.000	1.000
	n=400													
queen, 0.2, −0.2	0.065	1.000	1.000	1.000	1.000	1.000	1.000	0.046	1.000	1.000	1.000	1.000	1.000	0.999
queen, 0.2, −1	0.061	1.000	1.000	1.000	1.000	1.000	1.000	0.041	1.000	1.000	1.000	1.000	0.999	0.997
rook, 0.2, −0.2	0.048	1.000	1.000	1.000	1.000	1.000	1.000	0.044	1.000	1.000	1.000	1.000	0.999	0.998
rook, 0.2, −1	0.053	1.000	1.000	1.000	1.000	1.000	1.000	0.042	1.000	1.000	1.000	0.999	0.999	0.997
queen, 0.8, −0.2	0.058	1.000	1.000	1.000	1.000	1.000	1.000	0.039	1.000	1.000	1.000	1.000	1.000	1.000
queen, 0.8, −1	0.047	1.000	1.000	1.000	1.000	1.000	1.000	0.050	1.000	1.000	1.000	1.000	1.000	1.000
rook, 0.8, −0.2	0.055	1.000	1.000	1.000	1.000	1.000	1.000	0.053	1.000	1.000	1.000	1.000	1.000	1.000
rook, 0.8, −1	0.048	1.000	1.000	1.000	1.000	1.000	1.000	0.048	1.000	1.000	1.000	1.000	1.000	1.000

For the power, (1), (2), (3), (4), (5) and (6) in the table mean that in the DGP $\zeta_0 = (1, 0.5)'$, $\zeta_0 = (1, 1)'$, $\zeta_0 = (1, 1.5)'$, $\zeta_0 = (1, 2)'$, $\zeta_0 = (1, 2.5)'$ and $\zeta_0 = (1, 3)'$, respectively. $\beta_{10} = 1$ and $\beta_{20} = 1$.

$$= \frac{1}{n} u_{n1}' A_n e^{\alpha_3 B_n} B_n C_n u_{n2} (\alpha_1 - \alpha_2) - \frac{1}{n} \operatorname{tr}[A_n e^{\alpha_3 B_n} B_n C_n \operatorname{E}(u_{n2} u_{n1}')](\alpha_1 - \alpha_2),$$

where α_3 lies between α_1 and α_2. By the Cauchy-Schwarz inequality, $|\frac{1}{n} u_{n1}' A_n e^{\alpha_3 B_n} B_n C_n u_{n2}| \le \sqrt{\frac{1}{n} u_{n1}' E_n u_{n1}} \sqrt{\frac{1}{n} u_{n2}' u_{n2}}$, where $E_n = A_n e^{\alpha_3 B_n} B_n C_n C_n' B_n' e^{\alpha_3 B_n'} A_n'$, and $\frac{1}{n} u_{n2}' u_{n2} = O_p(1)$ by Markov's inequality. Let the eigendecomposition of E_n be $E_n = E_{n1} E_{n2} E_{n1}'$, where E_{n1} is an orthogonal matrix and E_{n2} is a diagonal matrix with its diagonal elements being the eigenvalues of E_n. Then by the spectral radius theorem, $\frac{1}{n} u_{n1}' E_n u_{n1} = \frac{1}{n} u_{n1}' E_{n1} E_{n2} E_{n1}' u_{n1} \le \|E_n\|_\infty \frac{1}{n} u_{n1}' E_{n1} E_{n1}' u_{n1} = \|E_n\|_\infty \frac{1}{n} u_{n1}' u_{n1}$, where $\|E_n\|_\infty$ is smaller than a finite constant not depending on α_3 and $\frac{1}{n} u_{n1}' u_{n1} = O_p(1)$. In addition, $|\frac{1}{n} \operatorname{tr}[A_n e^{\alpha_3 B_n} B_n C_n \operatorname{E}(u_{n2} u_{n1}')]| \le \|A_n e^{\alpha_3 B_n} B_n C_n \operatorname{E}(u_{n2} u_{n1}')\|_\infty < c$ for some finite c not depending on α_3 and n. Thus, $\frac{1}{n} \operatorname{E}(u_{n1}' A_n e^{\alpha B_n} C_n u_{n2})$ is uniformly equicontinuous, and $|\big[\frac{1}{n} u_{n1}' A_n e^{\alpha_1 B_n} C_n u_{n2} - \frac{1}{n} \operatorname{E}(u_{n1}' A_n e^{\alpha_1 B_n} C_n u_{n2})\big] - \big[\frac{1}{n} u_{n1}' A_n e^{\alpha_2 B_n} C_n u_{n2} - \frac{1}{n} \operatorname{E}(u_{n1}' A_n e^{\alpha_2 B_n} C_n u_{n2})\big]| \le r_n |\alpha_1 - \alpha_2|$, where $r_n = O_p(1)$ does not depend on α_1 or α_2. It follows that $\frac{1}{n} u_{n1}' A_n e^{\alpha B_n} C_n u_{n2} - \frac{1}{n} \operatorname{E}(u_{n1}' A_n e^{\alpha B_n} C_n u_{n2})$ is stochastically equicontinuous (Davidson, 1994, p. 339, Theorem 21.10). The pointwise convergence and stochastic equicontinuity of $\frac{1}{n} u_{n1}' A_n e^{\alpha B_n} C_n u_{n2} - \frac{1}{n} \operatorname{E}(u_{n1}' A_n e^{\alpha B_n} C_n u_{n2})$ implies that it is $o_p(1)$ uniformly in $\alpha \in [-\eta, \eta]$ (Davidson, 1994, p. 337, Theorem 21.9). □

Proof of Proposition 2.1. For a given α, the N2SLS estimator for β is $\check{\beta}_n(\alpha) = (D_n' H_n D_n)^{-1} D_n' H_n e^{\alpha W_n} Y_n$, where $H_n = F_n \Pi_n^{-1} F_n'$. Substituting $\check{\beta}_n(\alpha)$ into the N2SLS criterion function yields the function

$$Q_n(\alpha) = Y_n' e^{\alpha W_n'} [H_n - H_n D_n (D_n' H_n D_n)^{-1} D_n' H_n] e^{\alpha W_n} Y_n.$$

Let $\bar{Q}_n(\alpha) = \operatorname{E}(Y_n' e^{\alpha W_n'} F_n) A_n \operatorname{E}(F_n' e^{\alpha W_n} Y_n)$, where $A_n = \bar{\Pi}_n^{-1} - \bar{\Pi}_n^{-1} \operatorname{E}(F_n' D_n)[\operatorname{E}(D_n' F_n) \bar{\Pi}_n^{-1} \operatorname{E}(F_n' D_n)]^{-1} \operatorname{E}(D_n' F_n) \bar{\Pi}_n^{-1}$, with $\bar{\Pi}_n = \operatorname{E}(\Pi_n)$. Note that $\operatorname{E}(F_n' e^{\alpha W_n} Y_n) = \operatorname{E}[F_n' e^{(\alpha - \alpha_0) W_n} (D_n \beta_0 + V_n)] = \operatorname{E}(F_n' e^{(\alpha - \alpha_0) W_n} D_n) \beta_0$. By Lemma 1, under Assumption 3, $\frac{1}{n} F_n' e^{\alpha W_n} Y_n - \frac{1}{n} \operatorname{E}(F_n' e^{\alpha W_n} Y_n) = \frac{1}{n} [F_n' e^{(\alpha - \alpha_0) W_n} D_n - \operatorname{E}(F_n' e^{(\alpha - \alpha_0) W_n} D_n)] \beta_0 + \frac{1}{n} F_n' e^{(\alpha - \alpha_0) W_n} V_n = o_p(1)$ uniformly in $\alpha \in [-\eta, \eta]$, $\frac{1}{n} \Pi_n - \frac{1}{n} \bar{\Pi}_n = o_p(1)$ and $\frac{1}{n} F_n' D_n - \frac{1}{n} \operatorname{E}(F_n' D_n) = o_p(1)$. Thus, $\frac{1}{n}[Q_n(\alpha) - \bar{Q}_n(\alpha)] = o_p(1)$ uniformly in $\alpha \in [-\eta, \eta]$. Notice that $A_n = \bar{\Pi}_n^{-1/2} B_n \bar{\Pi}_n^{-1/2}$, where B_n is the projection matrix $I_n - \bar{\Pi}_n^{-1/2} \operatorname{E}(F_n' D_n)[\operatorname{E}(D_n' F_n) \bar{\Pi}_n^{-1} \operatorname{E}(F_n' D_n)]^{-1} \operatorname{E}(D_n' F_n) \bar{\Pi}_n^{-1/2}$. Then by the partitioned matrix formula, Assumption 5 implies that $\frac{1}{n} \bar{Q}_n(\alpha)$ is uniquely zero at α_0 for large enough n. In addition, $\frac{1}{n} \bar{Q}_n(\alpha)$ is uniformly equicontinuous by Lemma 1. The identification condition and the uniform equicontinuity of $\frac{1}{n} \bar{Q}_n(\alpha)$ imply that the identification uniqueness condition for $\frac{1}{n} \bar{Q}_n(\alpha)$ holds. The consistency of $\check{\alpha}_n$ follows from the uniform convergence and identification uniqueness conditions (White, 1994, p. 28, Theorem 3.4). The consistency of $\check{\beta}_n$ can be seen by applying the mean value theorem to $\check{\beta}_n = \check{\beta}_n(\check{\alpha}_n) = (D_n' H_n D_n)^{-1} D_n' H_n e^{\check{\alpha}_n W_n} Y_n$ and stochastic boundedness of $\sup_{\alpha \in [-\eta, \eta]} \frac{1}{n} F_n' W_n e^{\alpha W_n} Y_n$. □

Proof of Proposition 2.2. The first order condition of the N2SLS estimation is $G_n'(\check{\theta}_n) \Pi_n^{-1} g_n(\check{\theta}_n) = 0$. By the mean value theorem, $0 = G_n'(\check{\theta}_n) \Pi_n^{-1} g_n(\check{\theta}_n) = G_n'(\check{\theta}_n) \Pi_n^{-1} [g_n(\theta_0) + G_n(\ddot{\theta}_n)(\check{\theta}_n - \theta_0)]$, where $\ddot{\theta}_n$ is between $\check{\theta}_n$ and θ_0. As $\frac{1}{n} G_n(\theta) = \frac{1}{n} F_n' [W_n e^{(\alpha - \alpha_0) W_n} (D_n \beta_0 + V_n), -D_n]$, $\frac{1}{n} G_n(\check{\theta}_n) = \frac{1}{n} \operatorname{E}[G_n(\theta)]|_{\theta = \check{\theta}_n} + o_p(1) = \frac{1}{n} \operatorname{E}[G_n(\theta_0)] + o_p(1)$, where the first equality follows by Lemma 1 and the second by the mean value theorem. By the central limit theorem in

Lemma 2 of Qu and Lee (2012), $\frac{1}{\sqrt{n}} g_n(\theta_0) = \frac{1}{\sqrt{n}} F_n' V_n \xrightarrow{d} N(0, \lim_{n\to\infty} \frac{1}{n}\bar{\Pi}_n)$. Hence,

$$\begin{aligned}\sqrt{n}(\check{\theta}_n - \theta_0) &= -[\frac{1}{n} G_n'(\check{\theta}_n)\Pi_n^{-1} G_n(\ddot{\theta}_n)]^{-1} \frac{1}{\sqrt{n}} G_n'(\check{\theta}_n)\Pi_n^{-1} g_n(\theta_0) \\ &= -\big(\frac{1}{n}\bar{G}_n'\bar{\Pi}_n^{-1}\bar{G}_n\big)^{-1} \frac{1}{\sqrt{n}} \bar{G}_n'\bar{\Pi}_n^{-1} g_n(\theta_0) + o_p(1) \\ &\xrightarrow{d} N\big(0, \lim_{n\to\infty}\big(\frac{1}{n}\bar{G}_n'\bar{\Pi}_n^{-1}\bar{G}_n\big)^{-1}\big),\end{aligned}$$

under the assumption that $\lim_{n\to\infty} \bar{G}_n$ has full rank. Since $\bar{\Pi}_n = \mathrm{E}(F_n' T_n T_n' F_n)$ and $\bar{G}_n = [\mathrm{E}(F_n' W_n D_n)\beta_0, -\mathrm{E}(F_n' D_n)]$, $\bar{G}_n'\bar{\Pi}_n^{-1}\bar{G}_n \le \mathrm{E}[(T_n^{-1} W_n D_n \beta_0, -T_n^{-1} D_n)'(T_n^{-1} W_n D_n \beta_0, -T_n^{-1} D_n)]$ by the generalized Cauchy-Schwarz inequality. The equality holds when F_n is equal to $(T_n T_n')^{-1}\,\mathrm{E}(W_n D_n \beta_0, -D_n|\mathbb{X}_n)$, or equivalently, the matrix formed by the independent columns of $(T_n T_n')^{-1}\,\mathrm{E}(D_n, W_n D_n|\mathbb{X}_n)$ or more compactly the independent columns of

$$(T_n T_n')^{-1}[X_n^*, W_n X_n, W_n^2 X_n, \mathrm{E}(Z_n, W_n Z_n|\mathbb{X}_n)]. \qquad \square$$

Proof of Proposition 2.3. By taking a third order Taylor expansion of the first order condition $\frac{\partial Q_n^*(\check{\omega}_n)}{\partial \omega} = 0$, because $Q_n^*(\omega)$ is quadratic in ψ, by eliminating higher order derivative terms with zero values as in Appendix A, we have

$$\begin{aligned}0 = \frac{\partial Q_n^*(\check{\omega}_n)}{\partial \phi} &= \frac{\partial Q_n^*(\omega_0)}{\partial \phi} + \frac{\partial^2 Q_n^*(\omega_0)}{\partial \phi^2}(\check{\phi}_n - \phi_0) + \frac{\partial^2 Q_n^*(\omega_0)}{\partial \phi \partial \psi'}(\check{\psi}_n - \psi_0) + \frac{1}{2}\frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^3}(\check{\phi}_n - \phi_0)^2 \\ &+ \frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^2 \partial \psi'}(\check{\psi}_n - \psi_0)(\check{\phi}_n - \phi_0) + \frac{1}{6}\frac{\partial^4 Q_n^*(\omega_0)}{\partial \phi^4}(\check{\phi}_n - \phi_0)^3 + \frac{1}{2}\frac{\partial^4 Q_n^*(\omega_0)}{\partial \phi^3 \partial \psi'}(\check{\psi}_n - \psi_0)(\check{\phi}_n - \phi_0)^2 \\ &+ \frac{1}{24}\frac{\partial^5 Q_n^*(\ddot{\omega}_n)}{\partial \phi^5}(\check{\phi}_n - \phi_0)^4 + \frac{1}{6}\frac{\partial^5 Q_n^*(\ddot{\omega}_n)}{\partial \phi^4 \partial \psi'}(\check{\psi}_n - \psi_0)(\check{\phi}_n - \phi_0)^3,\end{aligned} \tag{2.1}$$

and

$$\begin{aligned}0 = \frac{\partial Q_n^*(\check{\omega}_n)}{\partial \psi} &= \frac{\partial Q_n^*(\omega_0)}{\partial \psi} + \frac{\partial^2 Q_n^*(\omega_0)}{\partial \phi \partial \psi}(\check{\phi}_n - \phi_0) + \frac{\partial^2 Q_n^*(\omega_0)}{\partial \psi \partial \psi'}(\check{\psi}_n - \psi_0) + \frac{1}{2}\frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^2 \partial \psi}(\check{\phi}_n - \phi_0)^2 \\ &+ \frac{1}{6}\frac{\partial^4 Q_n^*(\ddot{\omega}_n)}{\partial \phi^3 \partial \psi}(\check{\phi}_n - \phi_0)^3,\end{aligned} \tag{2.2}$$

where $\ddot{\omega}_n$ lies between ω_0 and $\check{\omega}_n$ elementwise. Since the N2SLS estimator $\check{\theta}_n$ is consistent, so is $\check{\omega}_n$. Let $\bar{o}_p(\cdot)$ denote terms with order smaller than those of some terms on the r.h.s. in the same equation. Using the consistency of $\ddot{\omega}_n$ and $\check{\omega}_n$, and relevant orders of derivatives (Appendix A gives expressions of derivatives and their orders), by keeping only possible leading order terms in (2.1), but dropping terms with surely relatively smaller orders into $\bar{o}_p(\cdot)$, we have

$$\begin{aligned}0 = &\frac{\partial Q_n^*(\omega_0)}{\partial \phi} + \frac{\partial^2 Q_n^*(\omega_0)}{\partial \phi^2}(\check{\phi}_n - \phi_0) + \frac{\partial^2 Q_n^*(\omega_0)}{\partial \phi \partial \psi'}(\check{\psi}_n - \psi_0) + \frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^2 \partial \psi'}(\check{\psi}_n - \psi_0)(\check{\phi}_n - \phi_0) \\ &+ \frac{1}{6}\frac{\partial^4 Q_n^*(\omega_0)}{\partial \phi^4}(\check{\phi}_n - \phi_0)^3 + \bar{o}_p(\cdot),\end{aligned} \tag{2.1$'$}$$

which follows because the fourth term on the r.h.s. of (2.1) is dominated by the second term, the seventh term is dominated by the fifth term, and the last two terms are dominated by the sixth term. Furthermore,

$$0 = \frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial \psi} + \frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial \psi \partial \psi'}\sqrt{n}(\check{\psi}_n - \psi_0) + \frac{1}{2n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^2 \partial \psi}\sqrt{n}(\check{\phi}_n - \phi_0)^2 + \bar{o}_p(\cdot), \tag{2.2$'$}$$

because the second term on the r.h.s. of (2.2) is dominated by the first term, and the fifth term is dominated by the fourth term. Because $(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'})^{-1} = O_p(1)$,

$$\sqrt{n}(\check{\psi}_n - \psi_0) = -(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'})^{-1}\frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial\psi} - (\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'})^{-1}\frac{1}{2n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial\phi^2\partial\psi}\sqrt{n}(\check{\phi}_n - \phi_0)^2 + \bar{o}_p(\cdot). \quad (2.3)$$

Substituting (2.3) into (2.1′), and by multiplying $n^{-1/4}$ on the whole equation, we have, after rearrangement,

$$\begin{aligned} 0 = {} & n^{-1/4}\frac{\partial Q_n^*(\omega_0)}{\partial\phi} - n^{-1/4}\frac{1}{\sqrt{n}}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\phi\partial\psi'}(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'})^{-1}[\frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial\psi} + \frac{1}{2n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial\phi^2\partial\psi}\sqrt{n}(\check{\phi}_n - \phi_0)^2] \\ & + n^{1/4}(\check{\phi}_n - \phi_0)\Big\{\frac{1}{\sqrt{n}}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\phi^2} - \frac{1}{n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial\phi^2\partial\psi'}(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'})^{-1}\frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial\psi} \qquad (2.4) \\ & + \Big[\frac{1}{6n}\frac{\partial^4 Q_n^*(\omega_0)}{\partial\phi^4} - \frac{1}{n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial\phi^2\partial\psi'}(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'})^{-1}\frac{1}{2n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial\phi^2\partial\psi}\Big]\sqrt{n}(\check{\phi}_n - \phi_0)^2\Big\} + \bar{o}_p(\cdot). \end{aligned}$$

Note that:

i) $n^{-1/4}\frac{\partial Q_n^*(\omega_0)}{\partial\phi} - n^{-1/4}\frac{1}{\sqrt{n}}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\phi\partial\psi'}\big(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'}\big)^{-1}\frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial\psi} = 2n^{-1/4}V_n'W_n'\mathbb{M}_D V_n = O_p(n^{-1/4})$,

ii) $n^{-1/4}\frac{1}{\sqrt{n}}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\phi\partial\psi'}\big(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'}\big)^{-1}\frac{1}{2n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial\phi^2\partial\psi} = n^{-3/4}(W_n^2X_n\delta_0)'\mathbb{P}_D W_n V_n + O_p(n^{-3/4}) = O_p(n^{-1/4})$,

iii) $\frac{1}{\sqrt{n}}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\phi^2} - \frac{1}{n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial\phi^2\partial\psi'}(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'})^{-1}\frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial\psi} = R_n + O_p(n^{-1/2})$, where $R_n = \frac{2}{\sqrt{n}}(W_n^2X_n\delta_0)'\mathbb{M}_D V_n = O_p(1)$, and

iv) $\frac{1}{6n}\frac{\partial^4 Q_n^*(\omega_0)}{\partial\phi^4} - \frac{1}{n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial\phi^2\partial\psi'}(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'})^{-1}\frac{1}{2n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial\phi^2\partial\psi} = S_n + O_p(n^{-1/2})$, where $S_n = \frac{1}{n}(W_n^2X_n\delta_0)'\mathbb{M}_D W_n^2 X_n\delta_0 = O(1)$.

Hence, (2.4) implies that

$$0 = O_p(n^{-1/4}) + O_p(n^{-1/4})[n^{1/4}(\check{\phi}_n - \phi_0)]^2 + n^{1/4}(\check{\phi}_n - \phi_0)R_n + [n^{1/4}(\check{\phi}_n - \phi_0)]^3 S_n + \bar{o}_p(\cdot). \quad (2.4')$$

As $S_n > 0$ for large enough n, $n^{1/4}(\check{\phi}_n - \phi_0)$ cannot grow with a rate as an increasing function of n. This is so, otherwise, $S_n[n^{1/4}(\check{\phi}_n - \phi_0)]^3$ would be the dominating term of (2.4′) on the r.h.s., which grows to infinity. Hence, (2.4′) implies $\sqrt{n}(\check{\phi}_n - \phi_0)^2 = O_p(1)$. On the other hand, it follows from (2.3) that $\sqrt{n}(\check{\psi}_n - \psi_0) = O_p(1)$.

When $R_n > 0$, $R_n + S_n\sqrt{n}(\check{\phi}_n - \phi_0)^2 \geq R_n > 0$. Thus, conditional on $R_n > 0$, $R_n + S_n\sqrt{n}(\check{\phi}_n - \phi_0)^2$ cannot converge in distribution to a random variable with an atom of probability at 0 along any subsequence of n. Hence, the equality (2.4′) is possible only if $n^{1/4}(\check{\phi}_n - \phi_0) = o_p(1)$. Therefore, conditional on $R_n > 0$, $\sqrt{n}(\check{\phi}_n - \phi_0)^2 = o_p(1)$. Then, when $R_n > 0$, (2.3) becomes

$$\sqrt{n}(\check{\psi}_n - \psi_0) = -(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial\psi\partial\psi'})^{-1}\frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial\psi} + o_p(1) = L_n + o_p(1),$$

where $L_n = (\frac{1}{n}D_n'H_nD_n)^{-1}\frac{1}{\sqrt{n}}D_n'H_nV_n$. Note that $L_n \xrightarrow{d} L$, where L is the normal random vector

$$N(0, \lim_{n\to\infty}[\frac{1}{n}\,\mathrm{E}(D_n'F_n)\bar{\Pi}_n^{-1}\,\mathrm{E}(F_n'D_n)]^{-1}).$$

Next, when $R_n < 0$, we can prove that $\sqrt{n}(\check{\phi}_n - \phi_0)^2 = J_{1n} + o_p(1)$, where $J_{1n} = -S_n^{-1}R_n$, by showing that there exists no subsequence n' of n such that ${n'}^{1/4}(\check{\phi}_{n'} - \phi_0)$ converges in distribution to a random variable with an atom of probability at 0. That is, if ${n'}^{1/4}(\check{\phi}_{n'} - \phi_0) \xrightarrow{d} U$, then for any $\delta > 0$, there exists an $\epsilon > 0$ such that $\mathrm{P}(|U| > \epsilon) > 1 - \delta$. We show this by contradiction.[1] Suppose that there exists a subsequence n' such that

[1]Such an argument appears in Rotnitzky et al. (2000). We adopt the analysis for our model.

${n'}^{1/4}(\check{\phi}_{n'} - \phi_0) \xrightarrow{d} U$ and $\mathrm{P}(U = 0) = \delta > 0$. By a fourth order Taylor expansion and with the orders of derivatives in Appendix A,

$$Q_n^*(\check{\omega}_n) - Q_n^*(\omega_0) = \Big[\frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial \psi'} + \frac{1}{2n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^2 \partial \psi'}\sqrt{n}(\check{\phi}_n - \phi_0)^2\Big]\sqrt{n}(\check{\psi}_n - \psi_0) + \frac{1}{2\sqrt{n}}\frac{\partial^2 Q_n^*(\omega_0)}{\partial \phi^2}\sqrt{n}(\check{\phi}_n - \phi_0)^2$$
$$+ \sqrt{n}(\check{\psi}_n - \psi_0)'\frac{1}{2n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial \psi \partial \psi'}\sqrt{n}(\check{\psi}_n - \psi_0) + \frac{1}{24n}\frac{\partial^4 Q_n^*(\omega_0)}{\partial \phi^4} n(\check{\phi}_n - \phi_0)^4 + O_p(n^{-1/4}). \tag{2.5}$$

Note that the order $\bar{o}_p(\cdot)$ in (2.3) is $O_p(n^{-1/4})$. Substituting the expression for $\sqrt{n}(\check{\psi}_n - \psi_0)$ in (2.3) into (2.5) yields

$$\begin{aligned}
&Q_n^*(\check{\omega}_n) - Q_n^*(\omega_0) \\
&= -\frac{1}{2\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial \psi'}\Big(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial \psi \partial \psi'}\Big)^{-1}\frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial \psi} \\
&\quad + \sqrt{n}(\check{\phi}_n - \phi_0)^2\Big\{\Big[\frac{1}{2\sqrt{n}}\frac{\partial^2 Q_n^*(\omega_0)}{\partial \phi^2} - \frac{1}{2n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^2 \partial \psi'}\Big(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial \psi \partial \psi'}\Big)^{-1}\frac{1}{\sqrt{n}}\frac{\partial Q_n^*(\omega_0)}{\partial \psi}\Big] \\
&\quad + \Big[\frac{1}{24n}\frac{\partial^4 Q_n^*(\omega_0)}{\partial \phi^4} - \frac{1}{8n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^2 \partial \psi'}\Big(\frac{1}{n}\frac{\partial^2 Q_n^*(\omega_0)}{\partial \psi \partial \psi'}\Big)^{-1}\frac{1}{n}\frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^2 \partial \psi}\Big]\sqrt{n}(\check{\phi}_n - \phi_0)^2\Big\} + O_p(n^{-1/4}) \\
&= -V_n'\mathbb{P}_D V_n + \sqrt{n}(\check{\phi}_n - \phi_0)^2\Big[\frac{1}{2}R_n + \frac{1}{4}S_n\sqrt{n}(\check{\phi}_n - \phi_0)^2\Big] + O_p(n^{-1/4}).
\end{aligned} \tag{2.6}$$

Since ${n'}^{1/4}(\check{\phi}_{n'} - \phi_0) \xrightarrow{d} U$ and $\mathrm{P}(U = 0) = \delta > 0$, there exists a negative constant M such that for all $\epsilon > 0$,

$$Q_{n'}^*(\check{\omega}_{n'}) - Q_{n'}^*(\omega_0) > -V_{n'}'\mathbb{P}_D V_{n'} + \epsilon M$$

with probability converging to a number greater than $\delta/2$ along the subsequence. Note that (2.6) still holds if we replace $\check{\omega}_n$ by any $\bar{\omega}_n$ satisfying $\sqrt{n}(\bar{\phi}_n - \phi_0)^2 = O_p(1)$ and (2.3) with $\check{\omega}_n$ replaced by $\bar{\omega}_n$. In particular, if we let $\sqrt{n}(\bar{\phi}_n - \phi_0)^2 = -S_n^{-1}R_n$ and define $\bar{\psi}_n$ according to (2.3) after $\check{\phi}_n$ is replaced by $\bar{\phi}_n$ in that formula, then

$$Q_{n'}^*(\bar{\omega}_{n'}) - Q_{n'}^*(\omega_0) = -V_{n'}'\mathbb{P}_D V_{n'} - \frac{1}{4}S_{n'}^{-1}R_{n'}^2 + O_p(n^{-1/4}).$$

Hence, by taking ϵ small enough, $Q_{n'}^*(\bar{\omega}_{n'}) - Q_{n'}^*(\check{\omega}_{n'}) < -\epsilon M - \frac{1}{4}S_{n'}^{-1}R_{n'}^2 < -\frac{1}{8}S_{n'}^{-1}R_{n'}^2 < 0$ with probability converging along the subsequence to a strictly positive number. This is a contradiction since $\check{\omega}_n$ is the N2SLS estimator that minimizes $Q_n^*(\omega)$. Therefore, (2.4′) holds only if $\sqrt{n}(\check{\phi}_n - \phi_0)^2 = J_{1n} + o_p(1)$ when $R_n < 0$, where $J_{1n} = -S_n^{-1}R_n$. Then by (2.3), $\sqrt{n}(\check{\psi}_n - \psi_0) = J_{2n} + o_p(1)$ when $R_n < 0$, where

$$J_{2n} = L_n + \Big(\frac{2}{n}D_n'H_nD_n\Big)^{-1}\frac{1}{n}D_n'H_nW_n^2X_n\delta_0J_{1n}.$$

Alternatively, when $R_n < 0$, by (2.1′) and (2.2′), we are essentially solving (8). Thus, the leading order term of $[\sqrt{n}(\check{\phi}_n - \phi_0)^2, \sqrt{n}(\check{\psi}_n - \psi_0)']'$ is $J_n = (J_{1n}, J_{2n}')'$ in (9). By Lemma 1, $J_n = \mathbb{J}_n + o_p(1)$. By Lemma 2 in Qu and Lee (2012), $\mathbb{J}_n \xrightarrow{d} J$, where $J = N(0, \lim_{n\to\infty}\Delta_n)$ with Δ_n being the covariance matrix of $\mathbb{J}_n$.

We next calculate the statistic that asymptotically determines the sign of $n^{1/4}(\check{\phi}_n - \phi_0)$ conditional on $R_n < 0$. By a fifth order Taylor expansion of $Q_n^*(\check{\omega}_n)$ at ω_0, we have with the orders of derivatives in Appendix A that

$$Q_n^*(\check{\omega}_n) = \Xi_n + K_n + O_p(n^{-1/2}),$$

where

$$\Xi_n = V_n' H_n V_n + \frac{\partial Q_n^*(\omega_0)}{\partial \psi'}(\check{\psi}_n - \psi_0) + \frac{1}{2}\frac{\partial^2 Q_n^*(\omega_0)}{\partial \phi^2}(\check{\phi}_n - \phi_0)^2 + \frac{1}{2}(\check{\psi}_n - \psi_0)'\frac{\partial^2 Q_n^*(\omega_0)}{\partial \psi \partial \psi'}(\check{\psi}_n - \psi_0)$$
$$+ \frac{1}{2}\frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^2 \partial \psi'}(\check{\psi}_n - \psi_0)(\check{\phi}_n - \phi_0)^2 + \frac{1}{24}\frac{\partial^4 Q_n^*(\omega_0)}{\partial \phi^4}(\check{\phi}_n - \psi_0)^4$$
$$= O_p(1),$$

and

$$K_n = \frac{\partial Q_n^*(\omega_0)}{\partial \phi}(\check{\phi}_n - \phi_0) + \frac{\partial^2 Q_n^*(\omega_0)}{\partial \phi \partial \psi'}(\check{\psi}_n - \psi_0)(\check{\phi}_n - \phi_0) + \frac{1}{6}\frac{\partial^3 Q_n^*(\omega_0)}{\partial \phi^3}(\check{\phi}_n - \phi_0)^3$$
$$+ \frac{1}{6}\frac{\partial^4 Q_n^*(\omega_0)}{\partial \phi^3 \partial \psi'}(\check{\psi}_n - \psi_0)(\check{\phi}_n - \phi_0)^3 + \frac{1}{120}\frac{\partial^5 Q_n^*(\omega_0)}{\partial \phi^5}(\check{\phi}_n - \phi_0)^5 \quad (2.7)$$
$$= O_p(n^{-1/4}).$$

Since the sign of $n^{1/4}(\check{\phi}_n - \phi_0)$ does not affect the value of Ξ_n, the sign of $n^{1/4}(\check{\phi}_n - \phi_0)$ must be chosen to minimize K_n. Note that $K_n = (\check{\phi}_n - \phi_0)\mathbb{K}_n + o_p(n^{-1/4})$, where

$$\mathbb{K}_n = 2V_n' W_n' H_n V_n + \begin{pmatrix} \sqrt{n}(\check{\phi}_n - \phi_0)^2 \\ \sqrt{n}(\check{\psi}_n - \psi_0) \end{pmatrix}' \Bigg[\begin{pmatrix} \frac{1}{3\sqrt{n}}(X_n\delta_0)'(W_n'^3 H_n + 3W_n'^2 H_n W_n)V_n \\ -\frac{2}{\sqrt{n}} D_n' H_n W_n V_n \end{pmatrix}$$
$$- \begin{pmatrix} \frac{1}{6n}(W_n^3 X_n \delta_0)' H_n(-W_n^2 X_n\delta_0, 2D_n) \\ 0 \end{pmatrix} \begin{pmatrix} \sqrt{n}(\check{\phi}_n - \phi_0)^2 \\ \sqrt{n}(\check{\psi}_n - \psi_0) \end{pmatrix} \Bigg]$$
$$= 2V_n' W_n' H_n V_n + J_n' \Bigg[\begin{pmatrix} \frac{1}{3\sqrt{n}}(X_n\delta_0)'(W_n'^3 H_n + 3W_n'^2 H_n W_n)V_n \\ -\frac{2}{\sqrt{n}} D_n' H_n W_n V_n \end{pmatrix} - \begin{pmatrix} \frac{1}{6n}(W_n^3 X_n \delta_0)' H_n(-W_n^2 X_n\delta_0, 2D_n) \\ 0 \end{pmatrix} J_n \Bigg] + o_p(1).$$

Thus, $\mathrm{P}(n^{1/4}(\check{\phi}_n - \phi_0)\mathbb{K}_n < 0 | R_n < 0) \to 1$ as $n \to \infty$. Equivalently, $\mathrm{P}(I(n^{1/4}(\check{\phi}_n - \phi_0) < 0) = I(\mathbb{K}_n > 0) | R_n < 0) \to 1$ as $n \to \infty$. By Lemma 1 and using U_{jn} and Υ_{jn} for $j = 1, 2$ defined above Proposition 2.3, we have K_n^* in (10). Since U_n is uncorrelated with $\mathbb{J}_n$, by Lemma 2 in Qu and Lee (2012), $(U_n', J_n')' \xrightarrow{d} (U', J')'$, where $U = N(0, \lim_{n\to\infty} \mathrm{E}(U_n U_n'))$ is independent of J. Hence, $K_n^* \xrightarrow{d} K^*$. As $J_{1n} = -S_n^{-1} R_n$, J_{1n} has a sign opposite to that of R_n. Therefore, the asymptotic distribution of $\check{\omega}_n$ in the proposition follows. □

Proof of Proposition 2.4. For the proof of $\frac{1}{n}\hat{\Pi}_n - \frac{1}{n}\Pi_n = o_p(1)$, we first show that that Theorem 2 on the consistency of a HAC estimator for a covariance matrix in Kelejian and Prucha (2007) (KP for short henceforth) still holds with a stochastic IV matrix whose elements have uniformly bounded moments of an order higher than the fourth (q in KP) and are independent of the disturbances (ε_n in KP) and the measure errors for the distances. Then we verify that the assumptions for Theorem 1 in KP, a special case of their theorem 2, are satisfied in our current setting. For the proof of KP's Theorem 2, we just pay attentions to places where the elements of the IV matrix may affect the arguments. All notations are now as in KP and their proof of Theorem 2. We need to prove that $a_{rs,n} = o_p(1)$, $b_{rs,n} = o_p(1)$ and $c_{rs,n} = o_p(1)$. As in KP, with a stochastic IV matrix, we have $|a_{rs,n}| \le A_{rs,n}^{(1)} + A_{rs,n}^{(2)} + A_{rs,n}^{(3)}$, where

$$A_{rs,n}^{(1)} = \|\Delta_n\| n^{-1} \sum_{i=1}^n \sum_{j=1}^n \Big(1 - \prod_{m=1}^M \mathbf{1}_{d_{ij,m,n}^* > d_{m,n}}\Big) |h_{ir,n} h_{js,n} u_{i,n}| \|z_{j,n}\|,$$
$$A_{rs,n}^{(2)} = \|\Delta_n\| n^{-1} \sum_{i=1}^n \sum_{j=1}^n \Big(1 - \prod_{m=1}^M \mathbf{1}_{d_{ij,m,n}^* > d_{m,n}}\Big) |h_{ir,n} h_{js,n} u_{j,n}| \|z_{i,n}\|,$$

$$A^{(3)}_{rs,n} = \|\Delta_n\|^2 n^{-1} \sum_{i=1}^n \sum_{j=1}^n \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}\Big) |h_{ir,n} h_{js,n}| \|z_{i,n}\| \|z_{j,n}\|.$$

Then for the q in KP,

$$\begin{aligned} A^{(1)}_{rs,n} &\leq \|\Delta_n\| n^{-1} \sum_{j=1}^n |h_{js,n}| \|z_{j,n}\| \Big[\sum_{i=1}^n \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}\Big)^{1/(1-1/q)}\Big]^{1-1/q} \Big(\sum_{i=1}^n |h_{ir,n} u_{i,n}|^q\Big)^{1/q} \\ &\leq \|\Delta_n\| n^{-1+1/q} l_n^{1-1/q} \sum_{j=1}^n |h_{js,n}| \|z_{j,n}\| \Big(n^{-1} \sum_{i=1}^n |h_{ir,n} u_{i,n}|^q\Big)^{1/q} \\ &\leq n^{1/2} \|\Delta_n\| n^{-1/2+1/q} l_n^{1-1/q} \Big(n^{-1} \sum_{j=1}^n |h_{js,n}|^2\Big)^{1/2} \Big(n^{-1} \sum_{j=1}^n \|z_{j,n}\|^2\Big)^{1/2} \Big(n^{-1} \sum_{i=1}^n |h_{ir,n} u_{i,n}|^q\Big)^{1/q}, \end{aligned}$$

where the first inequality follows by Hölder's inequality, the second inequality holds because

$$\sum_{i=1}^n \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}\Big)^{1/(1-1/q)} \leq \sum_{i=1}^n \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}\Big) \leq l_n,$$

and the third inequality holds by the Cauchy-Schwarz inequality. Note that $n^{-1} \sum_{j=1}^n |h_{js,n}|^2 = O_p(1)$ by Markov's inequality, and $n^{-1} \sum_{i=1}^n |h_{ir,n} u_{i,n}|^q = O_p(1)$ by Markov's inequality and independence of $h_{ir,n}$ and $u_{i,n}$. It is shown in KP that $n^{-1/2+1/q} l_n^{1-1/q} = o_p(1)$. Then $A^{(1)}_{rs,n} = o_p(1)$ under the maintained assumptions. Similarly $A^{(2)}_{rs,n} = o_p(1)$. By generalized Hölder's inequality,

$$\begin{aligned} A^{(3)}_{rs,n} &\leq \|\Delta_n\|^2 n^{-1} \sum_{j=1}^n |h_{js,n}| \|z_{j,n}\| \Big(\sum_{i=1}^n |h_{ir,n}|^4\Big)^{1/4} \Big[\sum_{i=1}^n \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}\Big)\Big]^{1/4} \Big(\sum_{i=1}^n \|z_{i,n}\|^2\Big)^{1/2} \\ &\leq (n^{1/2} \|\Delta_n\|)^2 n^{-1/4} l_n^{1/4} \Big(n^{-1} \sum_{j=1}^n |h_{js,n}| \|z_{j,n}\|\Big) \Big(n^{-1} \sum_{i=1}^n |h_{ir,n}|^4\Big)^{1/4} \Big(n^{-1} \sum_{i=1}^n \|z_{i,n}\|^2\Big)^{1/2} \\ &\leq (n^{1/2} \|\Delta_n\|)^2 n^{-1/4} l_n^{1/4} \Big(n^{-1} \sum_{j=1}^n |h_{js,n}|^2\Big)^{1/2} \Big(n^{-1} \sum_{j=1}^n \|z_{j,n}\|^2\Big)^{1/2} \Big(n^{-1} \sum_{i=1}^n |h_{ir,n}|^4\Big)^{1/4} \Big(n^{-1} \sum_{i=1}^n \|z_{i,n}\|^2\Big)^{1/2}. \end{aligned}$$

It is shown in KP that $l_n = o_p(n^{1/2})$. Thus $A^{(3)}_{rs,n} = o_p(1)$. It follows that $a_{rs,n} = o_p(1)$.

For $b_{rs,n}$, it is shown in KP that $b_{rs,n} = n^{-1} \sum_{l=1}^n \sum_{k=1}^n \gamma_{lk,n} [\varepsilon_{l,n} \varepsilon_{k,n} - \mathrm{E}(\varepsilon_{l,n} \varepsilon_{k,n})]$, where

$$\gamma_{lk,n} = \sum_{i=1}^n \sum_{j=1}^n h_{ir,n} h_{js,n} r_{il,n} r_{jk,n} K(\min_m \{d^*_{ij,m,n}/d_{m,n}\}).$$

Because V_n and ε_n are independent and so are H_n and ε_n, $\mathrm{E}(b_{rs,n}) = 0$. Then it suffices to show that $\mathrm{var}(b_{rs,n}) = o(1)$. As in KP, $\mathrm{var}(b_{rs,n}|V_n, H_n) = 2n^{-2} \sum_{l=1}^n \sum_{k=1}^n \gamma^2_{lk,n} + n^{-2} \sum_{l=1}^n \gamma^2_{ll} [\mathrm{E}(\varepsilon^4_{l,n}) - 3] \leq (c_E + 6) n^{-2} \sum_{l=1}^n (\sum_{k=1}^n |\gamma_{lk,n}|)^2$. Note that

$$\sum_{k=1}^n |\gamma_{lk,n}| \leq \sum_{i=1}^n |h_{ir,n} r_{il,n}| \sum_{j=1}^n |h_{js,n}| \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}\Big) \sum_{k=1}^n |r_{jk,n}| \leq C_R \sum_{i=1}^n \sum_{j=1}^n |h_{ir,n} h_{js,n} r_{il,n}| \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}\Big).$$

Thus,

$$\begin{aligned} \mathrm{E}\Big(\sum_{k=1}^n |\gamma_{lk,n}|\Big)^2 &\leq C_R^2\, \mathrm{E} \sum_{i=1}^n \sum_{j=1}^n \sum_{i'=1}^n \sum_{j'=1}^n |h_{ir,n} h_{js,n} r_{il,n} h_{i'r,n} h_{j's,n} r_{i'l,n}| \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}\Big) \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{i'j',m,n} > d_{m,n}}\Big) \\ &\leq C_R^2 C_H\, \mathrm{E} \sum_{i=1}^n |r_{il,n}| \sum_{i'=1}^n |r_{i'l,n}| \sum_{j=1}^n \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}\Big) \sum_{j'=1}^n \Big(1 - \prod_{m=1}^M \mathbf{1}_{d^*_{i'j',m,n} > d_{m,n}}\Big) \end{aligned}$$

$$\leq C_R^4 C_H \operatorname{E}(l_n^2),$$

where C_H is a constant such that $\operatorname{E}|h_{ir,n}h_{js,n}h_{i'r,n}h_{j's,n}| < C_H$ for any i, i', j, j', r, s, n, the second inequality follows by the independence of H_n and V_n, and the last inequality holds because $\sum_{j=1}^n (1 - \prod_{m=1}^M \mathbf{1}_{d^*_{ij,m,n} > d_{m,n}}) \leq l_n$. It is shown in KP that $\operatorname{E}(l_n^2) = o(n)$. Hence, $\operatorname{var}(b_{rs,n}) = o(1)$ and $b_{rs,n} = o_p(1)$.

From the first line on p. 151 and the bottom on p. 150 of KP,

$$\begin{aligned}
|c_{rs,n}| &\leq (c_K + 1) d_{1,n}^{-\rho_*} n^{-1} \sum_{i=1}^n \sum_{j=1}^n |\sigma_{ij,n} h_{ir,n} h_{js,n}| (d_{ij,1,n} + c_V)^{\rho_*} \\
&\leq (c_K + 1) d_{1,n}^{-\rho_*} n^{-1} \sum_{i=1}^n \sum_{j=1}^n |\sigma_{ij,n} h_{ir,n} h_{js,n}| (d_{ij,1,n} + c_V + 1)^{\rho_S} \\
&\leq 2^{\rho_S} (c_K + 1) d_{1,n}^{-\rho_*} n^{-1} \sum_{i=1}^n \Big((c_V + 1)^{\rho_S} \sum_{j=1}^n |\sigma_{ij,n} h_{ir,n} h_{js,n}| + \sum_{j=1}^n |\sigma_{ij,n} h_{ir,n} h_{js,n}| d_{ij,1,n}^{\rho_S} \Big),
\end{aligned}$$

where $d_{1,n}^{-\rho_*} = o(1)$ as in KP, $n^{-1} \sum_{i=1}^n \sum_{j=1}^n |\sigma_{ij,n} h_{ir,n} h_{js,n}| = O_p(1)$ and $n^{-1} \sum_{i=1}^n \sum_{j=1}^n |\sigma_{ij,n} h_{ir,n} h_{js,n}| d_{ij,1,n}^{\rho_S} = O_p(1)$ by Markov's inequality. Thus $c_{rs,n} = o_p(1)$. Therefore, Theorem 2 in KP with a stochastic IV matrix still holds.

Next we show that Theorem 1 in KP, a special case of Theorem 2, with a stochastic IV matrix is applicable in our current situation. Now the notations are as in the main text of this paper. All assumptions for that theorem are assumed except for their Assumption 6. By the mean value theorem, $\hat{V}_n - V_n = (W_n e^{\ddot{\alpha}_n W_n} Y_n, -D_n)(\dot{\theta}_n - \theta_0)$, where $\ddot{\alpha}_n$ lies between $\dot{\alpha}_n$ and α_0. Using the reduced form for Y_n and Lemma 1, similar to the proof for $\frac{1}{n} G_n(\check{\theta}_n) - \frac{1}{n} \operatorname{E}[G_n(\theta_0)] = o_p(1)$ in the proof of Proposition 2.2, we can show that $\frac{1}{n}(W_n e^{\ddot{\alpha}_n W_n} Y_n, -D_n)'(W_n e^{\ddot{\alpha}_n W_n} Y_n, -D_n) - \frac{1}{n} \operatorname{E}[(W_n e^{\alpha_0 W_n} Y_n, -D_n)'(W_n e^{\alpha_0 W_n} Y_n, -D_n)] = o_p(1)$, where $\frac{1}{n} \operatorname{E}[(W_n e^{\alpha_0 W_n} Y_n, -D_n)'(W_n e^{\dot{\alpha}_n W_n} Y_n, -D_n)] = O(1)$. Thus, Assumption 6 in KP is also satisfied. Therefore, $\frac{1}{n}\hat{\Pi}_n - \frac{1}{n}\Pi_n = o_p(1)$.

Since $\hat{Q}_n(\theta)$ is quadratic in $F_n'(e^{\alpha W_n} Y_n - D_n \beta)$ and $\frac{1}{n}\hat{\Pi}_n - \frac{1}{n}\Pi_n = o_p(1)$, the arguments in the proof of Proposition 2.1 still hold and the feasible N2SLS estimator $\hat{\theta}_n$ is consistent. Replacing Π_n by $\hat{\Pi}_n$ in $Q_n(\theta)$ and $Q_n^*(\omega)$ does not affect the analyses for Propositions 2.2 and 2.3, because neither the orders of terms nor asymptotic distributions for these propositions will change. Thus, Proposition 2.2 still holds if $\check{\theta}_n$ is replaced by $\hat{\theta}_n$, and Proposition 2.3 still holds if $\check{\omega}_n$ is replaced by $\hat{\omega}_n$. □

Proof of Proposition 3.1. Let $P_{\Pi^{-1/2}F'D} = \Pi_n^{-1/2} F_n' D_n (D_n' H_n D_n)^{-1} D_n' F_n \Pi_n^{-1/2}$ be the orthogonal projector onto the column space of $\Pi_n^{-1/2} F_n' D_n$. Similarly, let

$$P_{\Pi^{-1/2}F'(-WX\delta_0, X)} = \Pi_n^{-1/2} F_n' (-W_n X_n \delta_0, X_n)[(-W_n X_n \delta_0, X_n)' H_n (-W_n X_n \delta_0, X_n)]^{-1} (-W_n X_n \delta_0, X_n)' F_n \Pi_n^{-1/2}$$

and $P_{\Pi^{-1/2}F'(-W^2X\delta_0, D)} = \Pi_n^{-1/2} F_n' (-W_n^2 X_n \delta_0, D_n)[(-W_n^2 X_n \delta_0, D_n)' H_n (-W_n^2 X_n \delta_0, D_n)]^{-1} (-W_n^2 X_n \delta_0, D_n)' F_n \Pi_n^{-1/2}$. Then $\mathbb{P}_D = F_n \Pi_n^{-1/2} P_{\Pi^{-1/2}F'D} \Pi_n^{-1/2} F_n'$, $\mathbb{P}_{(-WX\delta_0, X)} = F_n \Pi_n^{-1/2} P_{\Pi^{-1/2}F'(-WX\delta_0, X)} \Pi_n^{-1/2} F_n'$, and

$$\mathbb{P}_{(-W^2X\delta_0, D)} = F_n \Pi_n^{-1/2} P_{\Pi^{-1/2}F'(-W_n^2 X_n \delta_0, D)} \Pi_n^{-1/2} F_n'.$$

Let the $k_f \times (k_{x^*} + k_x + k_z)$ matrix $[\Pi_n^{-1/2} F_n'(-W_n X_n \delta_0, X_n), A_n]$ be a basis matrix for the column space of $\Pi_n^{-1/2} F_n' D_n$, where A_n is an $k_f \times (k_{x^*} + k_z - 1)$ matrix perpendicular to $\Pi_n^{-1/2} F_n'(-W_n X_n \delta_0, X_n)$. Then, $P_{\Pi^{-1/2}F'D} =$

$P_{\Pi^{-1/2}F'(-WX\delta_0,X)} + P_A$, where $P_A = A_n(A_n'A_n)^{-1}A_n'$. By (3.25) on p. 71 in Ruud (2000), $P_{\Pi^{-1/2}F'(-W^2X\delta_0,D)} = P_{\Pi^{-1/2}F'D} + P_B$, where P_B is the orthogonal projector onto the column space of $B_n = M_{\Pi^{-1/2}F'D}\Pi_n^{-1/2}F_n'W_n^2X_n\delta_0$ with $M_{\Pi^{-1/2}F'D} = I_{k_f} - P_{\Pi^{-1/2}F'D}$, and B_n is perpendicular to A_n. Thus, $P_{\Pi^{-1/2}F'(-W^2X\delta_0,D)} - P_{\Pi^{-1/2}F'(-WX\delta_0,X)} = P_B + P_A$. It follows that (16) becomes

$$\hat{Q}_n(\tilde{\theta}_n) - \hat{Q}_n(\hat{\theta}_n) = I(R_n < 0)V_n'F_n\Pi_n^{-1/2}(P_B + P_A)\Pi_n^{-1/2}F_n'V_n + I(R_n > 0)V_n'F_n\Pi_n^{-1/2}P_A\Pi_n^{-1/2}F_n'V_n + o_p(1).$$

By the central limit theorem in Lemma 2 of Qu and Lee (2012), $\Pi_n^{-1/2}F_n'V_n \xrightarrow{d} N(0, I_{k_f})$. Since $R_n = \frac{2}{\sqrt{n}}B_n'\Pi_n^{-1/2}F_n'V_n$, $\hat{Q}_n(\tilde{\theta}_n) - \hat{Q}_n(\hat{\theta}_n) \xrightarrow{d} T$, where $T = \sum_{i=1}^{k_{x^*}+k_z} r_i^2 I(r_1 < 0) + \sum_{i=2}^{k_{x^*}+k_z} r_i^2 I(r_1 > 0)$, with $r_1, \ldots, r_{k_{x^*}}$ being i.i.d. standard normal random variables. Because the probability density function of r_1 is symmetric, T is a mixture of a $\chi^2(k_{x^*} + k_z - 1)$ variable and a $\chi^2(k_{x^*} + k_z)$ variable with mixing probabilities equal to 1/2. □

Proof of Proposition 3.2. Note that $\mathbb{M}_{(-WX\delta_0,X)} = F_n\Pi_n^{-1/2}M_{\Pi^{-1/2}F'(-WX\delta_0,X)}\Pi_n^{-1/2}F_n'$, where $M_{\Pi^{-1/2}F'(-WX\delta_0,X)} = I_{k_f} - P_{\Pi^{-1/2}F'(-WX\delta_0,X)}$ with $P_{\Pi^{-1/2}F'(-WX\delta_0,X)}$ being defined in the proof of Proposition 3.1. As in the proof of Proposition 3.1, $\Pi_n^{-1/2}F_n \xrightarrow{d} N(0, I_{k_f})$. Thus, the asymptotic distribution of the gradient vector in (17) holds. Furthermore, by the partitioned matrix formula,

$$\begin{aligned}
&\operatorname{rk}\big(\operatorname{plim}_{n\to\infty} \frac{1}{n}(W_nX_n^{**}, Z_n)'F_n\Pi_n^{-1/2}M_{\Pi^{-1/2}F'(-WX\delta_0,X)}\Pi_n^{-1/2}F_n'(W_nX_n^{**}, Z_n)\big)\\
&= \operatorname{rk}\big(\operatorname{plim}_{n\to\infty} \frac{1}{n}(W_nX_n^{**}, Z_n, -W_nX_n\delta_0, X_n)'H_n(W_nX_n^{**}, Z_n, -W_nX_n\delta_0, X_n)\big)\\
&\quad - \operatorname{rk}\big(\operatorname{plim}_{n\to\infty} \frac{1}{n}(-W_nX_n\delta_0, X_n)'H_n(-W_nX_n\delta_0, X_n)\big)\\
&= \operatorname{rk}\big(\operatorname{plim}_{n\to\infty} \frac{1}{n}(W_nX_n^{*}, Z_n, X_n)'H_n(W_nX_n^{*}, Z_n, X_n)\big) - k_x - 1\\
&= k_{x^*} + k_z - 1.
\end{aligned}$$

Similarly, $\operatorname{rk}\big(\frac{1}{n}(W_nX_n^{**}, Z_n)'\tilde{\mathbb{M}}_{(-WX\delta_0,X)}(W_nX_n^{**}, Z_n)\big) = \operatorname{rk}\big(\frac{1}{n}(W_nX_n^{*}, Z_n, X_n)'H_n(W_nX_n^{*}, Z_n, X_n)\big) - k_x - 1$. Thus, w.p.a.1. $\operatorname{rk}\big(\frac{1}{n}(W_nX_n^{**}, Z_n)'\tilde{\mathbb{M}}_{(-WX\delta_0,X)}(W_nX_n^{**}, Z_n)\big) = \operatorname{rk}\big(\operatorname{plim}_{n\to\infty} \frac{1}{n}(W_nX_n^{**}, Z_n)'\mathbb{M}_{(-WX\delta_0,X)}(W_nX_n^{**}, Z_n)\big)$. By Theorem 1 in Andrews (1987), the result in the proposition follows. □

Proof of Proposition 3.3. The proof is similar to that of Proposition 2.3, thus we pay attention to the differences brought by the drift in Assumption 8. Proposition 2.1 shows the consistency of the N2SLS estimator for model (1) when the true parameter vector θ_0 is fixed regardless of the sample size n. Suppose that the true parameter vector θ_{0n} changes with n but still satisfies $\theta_{0n} \to \theta_0$ as $n \to \infty$. Because

$$\lim_{n\to\infty} \frac{1}{n}[\mathrm{E}(F_n'e^{(\alpha-\alpha_0)W_n}D_n)\delta_{0n}, \mathrm{E}(F_n'D_n)] = \lim_{n\to\infty} \frac{1}{n}[\mathrm{E}(F_n'e^{(\alpha-\alpha_0)W_n}D_n)\delta_0, \mathrm{E}(F_n'D_n)],$$

$\frac{1}{n}\bar{Q}_n(\alpha)$ in the proof of Proposition 2.1 is uniquely zero at α_0 for large enough n under Assumption 5. Following almost the same argument as the proof of Proposition 2.1, Theorem 3.4 in White (1994, p. 28) applies under Assumption 8, so the N2SLS estimator $\hat{\theta}_n = \theta_{0n} + o_p(1)$. With the reparameterization ω, we have $\hat{\omega}_n - \omega_{0n} = o_p(1)$. The derivatives of $\hat{Q}_n^*(\omega)$ at ω_{0n} still have the same orders as in the case without the drift. Then $\sqrt{n}(\hat{\phi}_n - \phi_{0n})^2 = O_p(1)$, $\sqrt{n}(\hat{\psi}_n - \psi_{0n}) = O_p(1)$, and (2.1′)–(2.4) with ω_0 replaced by ω_{0n} hold. Note that $\frac{1}{\sqrt{n}}\frac{\partial^2\hat{Q}_n^*(\omega_{0n})}{\partial\phi^2}-\frac{1}{n}\frac{\partial^3\hat{Q}_n^*(\omega_{0n})}{\partial\phi^2\partial\psi'}(\frac{1}{n}\frac{\partial^2\hat{Q}_n^*(\omega_{0n})}{\partial\psi\partial\psi'})^{-1}\frac{1}{\sqrt{n}}\frac{\partial\hat{Q}_n^*(\omega_{0n})}{\partial\psi}=R_n+O_p(n^{-1/2})$, where $R_n=\frac{2}{\sqrt{n}}(W_n^2X_n\delta_0)'\mathbb{M}_DV_n=O_p(1)$, and $\frac{1}{6n}\frac{\partial^4\hat{Q}_n^*(\omega_{0n})}{\partial\phi^4}-\frac{1}{n}\frac{\partial^3\hat{Q}_n^*(\omega_{0n})}{\partial\phi^2\partial\psi'}(\frac{1}{n}\frac{\partial^2\hat{Q}_n^*(\omega_{0n})}{\partial\psi\partial\psi'})^{-1}\frac{1}{2n}\frac{\partial^3\hat{Q}_n^*(\omega_{0n})}{\partial\phi^2\partial\psi}=S_n+O_p(n^{-1/2})$, where

$$S_n=\frac{1}{n}(W_n^2X_n\delta_0)'\mathbb{M}_DW_n^2X_n\delta_0=O(1)$$

and $S_n\geq 0$. Similar to the proof of Proposition 2.3, when $R_n>0$, $\sqrt{n}(\hat{\phi}_n-\phi_{0n})^2=o_p(1)$, and

$$\sqrt{n}(\hat{\psi}_n-\psi_{0n})=-\big(\frac{1}{n}\frac{\partial^2\hat{Q}_n^*(\omega_{0n})}{\partial\psi\partial\psi'}\big)^{-1}\frac{1}{\sqrt{n}}\frac{\partial\hat{Q}_n^*(\omega_{0n})}{\partial\psi}+o_p(1)=L_n+o_p(1),$$

where $L_n=(\frac{1}{n}D_n'H_nD_n)^{-1}\frac{1}{\sqrt{n}}D_n'H_nV_n\xrightarrow{d}L$ with $L=N(0,(\lim_{n\to\infty}\frac{1}{n}\operatorname{E}(D_n'F_n)\bar{\Pi}_n^{-1}\operatorname{E}(F_n'D_n))^{-1})$. For $R_n<0$, $R_n+S_n\sqrt{n}(\hat{\phi}_n-\phi_{0n})^2=O_p(n^{-1/4})$ and thus $\sqrt{n}(\hat{\phi}_n-\phi_{0n})^2=J_{1n}+o_p(1)$, where $J_{1n}=-S_n^{-1}R_n$. Then by (2.3), $\sqrt{n}(\hat{\psi}_n-\psi_{0n})=J_{2n}+o_p(1)$ when $R_n<0$, where

$$J_{2n}=L_n+\big(\frac{2}{n}D_n'H_nD_n\big)^{-1}\frac{1}{n}D_n'H_nW_n^2X_n\delta_0J_{1n}.$$

Alternatively, when $R_n<0$, by (2.1′) and (2.2′), we are essentially solving (8) with ω_0 replaced by ω_{0n}. Then the leading order term of $[\sqrt{n}(\hat{\phi}_n-\phi_{0n})^2,\sqrt{n}(\hat{\psi}_n-\psi_{0n})']'$ is $J_n=\mathbb{J}_n+o_p(1)$, where $J_n=(J_{1n},J_{2n}')'\xrightarrow{d}J$, with $J=(J_1,J_2')'=N(0,\lim_{n\to\infty}\operatorname{E}(\mathbb{J}_n\mathbb{J}_n'))$. When $R_n<0$, the sign of $n^{1/4}(\hat{\phi}_n-\phi_{0n})$ must be chosen to minimize K_n in (2.7) (with ω_0 replaced by ω_{0n}). Note that $K_n=(\hat{\phi}_n-\phi_{0n})\mathbb{K}_n+o_p(n^{-1/4})$, where

$$\begin{aligned}\mathbb{K}_n&=2\big(\frac{1}{\sqrt{n}}(W_nX_n^{**},Z_n)\kappa+V_n\big)'W_n'H_nV_n\\&\quad+\begin{pmatrix}\sqrt{n}(\hat{\phi}_n-\phi_{0n})^2\\\sqrt{n}(\hat{\psi}_n-\psi_{0n})\end{pmatrix}'\Bigg[\begin{pmatrix}\frac{1}{3\sqrt{n}}(X_n\delta_0)'(W_n'^3H_n+3W_n'^2H_nW_n)V_n+\frac{1}{n}(W_n^2X_n\delta_0)'H_nW_n(W_nX_n^{**},Z_n)\kappa\\-2D_n'H_nW_n\big(\frac{1}{n}(W_nX_n^{**},Z_n)\kappa+\frac{1}{\sqrt{n}}V_n\big)\end{pmatrix}\\&\quad-\begin{pmatrix}\frac{1}{6n}(W_n^3X_n\delta_0)'H_n(-W_n^2X_n\delta_0,2D_n)\\0\end{pmatrix}\begin{pmatrix}\sqrt{n}(\hat{\phi}_n-\phi_{0n})^2\\\sqrt{n}(\hat{\psi}_n-\psi_{0n})\end{pmatrix}\Bigg]\\&=2\big(\frac{1}{\sqrt{n}}(W_nX_n^{**},Z_n)\kappa+V_n\big)'W_n'H_nV_n\\&\quad+J_n'\Bigg[\begin{pmatrix}\frac{1}{3\sqrt{n}}(X_n\delta_0)'(W_n'^3H_n+3W_n'^2H_nW_n)V_n+\frac{1}{n}(W_n^2X_n\delta_0)'H_nW_n(W_nX_n^{**},Z_n)\kappa\\-2D_n'H_nW_n\big(\frac{1}{n}(W_nX_n^{**},Z_n)\kappa+\frac{1}{\sqrt{n}}V_n\big)\end{pmatrix}\\&\quad-\begin{pmatrix}\frac{1}{6n}(W_n^3X_n\delta_0)'H_n(-W_n^2X_n\delta_0,2D_n)\\0\end{pmatrix}J_n\Bigg]+o_p(1).\end{aligned}$$

Thus, $\operatorname{P}(n^{1/4}(\hat{\phi}_n-\phi_{0n})\mathbb{K}_n<0|R_n<0)\to 1$ as $n\to\infty$. Equivalently, $\operatorname{P}(I(n^{1/4}(\hat{\phi}_n-\phi_{0n})<0)=I(\mathbb{K}_n>0)|R_n<0)\to 1$ as $n\to\infty$. Comparing the $\mathbb{K}_n$ above with that in the proof of Proposition 2.3, additional terms appear due to the drift in Assumption 8. Accounting for those additional terms, the asymptotic distribution of $\hat{\omega}_n$ in the proposition follows. □

Proof of Proposition 3.4. When $R_n<0$, taking a fourth order Taylor expansion of $\hat{Q}_n^*(\hat{\omega}_n)$ at ω_0, we have

$$\begin{aligned}\hat{Q}_n^*(\hat{\omega}_n)-\hat{Q}_n^*(\omega_0)&=\big(\frac{1}{2\sqrt{n}}\frac{\partial\hat{Q}_n^*(\omega_0)}{\partial\phi^2},\frac{1}{\sqrt{n}}\frac{\partial\hat{Q}_n^*(\omega_0)}{\partial\psi'}\big)\begin{pmatrix}\sqrt{n}(\hat{\phi}_n-\phi_0)^2\\\sqrt{n}(\hat{\psi}_n-\psi_0)\end{pmatrix}\\&\quad+\begin{pmatrix}\sqrt{n}(\hat{\phi}_n-\phi_0)^2\\\sqrt{n}(\hat{\psi}_n-\psi_0)\end{pmatrix}'\begin{pmatrix}\frac{1}{24n}\frac{\partial^4\hat{Q}_n^*(\omega_0)}{\partial\phi^4}&\frac{1}{4n}\frac{\partial^3\hat{Q}_n^*(\omega_0)}{\partial\phi^2\partial\psi'}\\\frac{1}{4n}\frac{\partial^3\hat{Q}_n^*(\omega_0)}{\partial\phi^2\partial\psi}&\frac{1}{2n}\frac{\partial^2\hat{Q}_n^*(\omega_0)}{\partial\psi\partial\psi'}\end{pmatrix}\begin{pmatrix}\sqrt{n}(\hat{\phi}_n-\phi_0)^2\\\sqrt{n}(\hat{\psi}_n-\psi_0)\end{pmatrix}+o_p(1)\\&=-\big(\frac{1}{\sqrt{n}}(W_nX_n^{**},Z_n)\kappa+V_n\big)'\mathbb{P}_{(-W^2X\delta_0,D)}\big(\frac{1}{\sqrt{n}}(W_nX_n^{**},Z_n)\kappa+V_n\big)+o_p(1);\end{aligned}$$

when $R_n > 0$, we have

$$\hat{Q}_n^*(\hat{\omega}_n) - \hat{Q}_n^*(\omega_0) = \frac{\partial \hat{Q}_n^*(\omega_0)}{\partial \psi'}(\hat{\psi}_n - \psi_0) + \frac{1}{2}(\hat{\psi}_n - \psi_0)'\frac{\partial^2 \hat{Q}_n^*(\omega_0)}{\partial \psi \partial \psi'}(\hat{\psi}_n - \psi_0) + o_p(1)$$
$$= -\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)'\mathbb{P}_D\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big) + o_p(1).$$

By the mean value theorem and (18), we have

$$\hat{Q}_n(\Psi_0, 0) - \hat{Q}_n(\tilde{\Psi}_n, 0) = \frac{1}{2}\sqrt{n}(\Psi_0 - \tilde{\Psi}_n)'\frac{1}{n}\frac{\partial^2 \hat{Q}_n(\check{\Psi}_n, 0)}{\partial \Psi \partial \Psi'}\sqrt{n}(\Psi_0 - \tilde{\Psi}_n)$$
$$= \big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)'\mathbb{P}_{(-WX\delta_0, X)}\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big) + o_p(1),$$

where $\check{\Psi}_n$ lies between Ψ_0 and $\tilde{\Psi}_n$. Thus,

$$\hat{Q}_n(\tilde{\theta}_n) - \hat{Q}_n(\hat{\theta}_n)$$
$$= I(R_n < 0)\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)'(\mathbb{P}_{(-W^2X\delta_0, D)} - \mathbb{P}_{(-WX\delta_0, X)})\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)$$
$$+ I(R_n > 0)\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)'(\mathbb{P}_D - \mathbb{P}_{(-WX\delta_0, X)})\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big) + o_p(1).$$

According to the proof of Proposition 3.1,

$$\hat{Q}_n(\tilde{\theta}_n) - \hat{Q}_n(\hat{\theta}_n)$$
$$= I(R_n < 0)\big[\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)'F_n\Pi_n^{-1/2}(P_B + P_A)\Pi_n^{-1/2}F_n'\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)\big]$$
$$+ I(R_n > 0)\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)'F_n\Pi_n^{-1/2}P_A\Pi_n^{-1/2}F_n'\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big) + o_p(1)$$
$$= I(R_n < 0)\big[V_n'F_n\Pi_n^{-1/2}P_B\Pi_n^{-1/2}F_n'V_n + \big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)'F_n\Pi_n^{-1/2}P_A\Pi_n^{-1/2}F_n'\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)\big]$$
$$+ I(R_n > 0)\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big)'F_n\Pi_n^{-1/2}P_A\Pi_n^{-1/2}F_n'\big(\frac{1}{\sqrt{n}}(W_n X_n^{**}, Z_n)\kappa + V_n\big) + o_p(1),$$

where the second equality holds because $B_n'\Pi_n^{-1/2}F_n'(W_n X_n^{**}, Z_n) = (W_n^2 X_n\delta_0)'F_n\Pi_n^{-1/2}M_{\Pi^{-1/2}F'D}\Pi_n^{-1/2}F_n'(W_n X_n^{**}, Z_n) = 0$. The expression for $\hat{Q}_n(\tilde{\theta}_n) - \hat{Q}_n(\hat{\theta}_n)$ is the same as that in the proof of Proposition 3.1 except for the additional terms due to the drift $\frac{1}{\sqrt{n}}\kappa$. Note that $(W_n X_n^{**}, Z_n)'F_n\Pi_n^{-1/2}P_A\Pi_n^{-1/2}F_n'(W_n X_n^{**}, Z_n) = (W_n X_n^{**}, Z_n)'F_n\Pi_n^{-1/2}(P_{\Pi^{-1/2}F'D} - P_{\Pi^{-1/2}F'(-WX\delta_0, X)})\Pi_n^{-1/2}F_n'(W_n X_n^{**}, Z_n) = (W_n X_n^{**}, Z_n)'F_n\Pi_n^{-1/2}(I_{k_f} - P_{\Pi^{-1/2}F'(-WX\delta_0, X)})\Pi_n^{-1/2}F_n'(W_n X_n^{**}, Z_n) = (W_n X_n^{**}, Z_n)'\mathbb{M}_{(-WX\delta_0, X)}(W_n X_n^{**}, Z_n)$, where the second equality uses the fact that $P_{\Pi^{-1/2}F'D}\Pi_n^{-1/2}F_n'(W_n X_n^{**}, Z_n) = \Pi_n^{-1/2}F_n'(W_n X_n^{**}, Z_n)$. Hence, the result in the proposition follows. □

Proof of Proposition 3.5. This proposition is proved in the main text in front of the proposition. □

Proof of Proposition 3.6. A first order Taylor expansion of $\frac{\partial \hat{Q}_n(\tilde{\Psi}_n, 0)}{\partial \Psi} = 0$ at $(\alpha_0 - n^{1/4}, \delta_0')'$ yields

$$0 = \frac{\partial \hat{Q}_n(\tilde{\Psi}_n, 0)}{\partial \Psi} = \frac{\partial \hat{Q}_n(\alpha_0 - n^{-1/4}, \delta_0, 0)}{\partial \Psi} + \frac{\partial^2 \hat{Q}_n(\alpha_0 - n^{-1/4}, \delta_0, 0)}{\partial \Psi \partial \Psi'}\begin{pmatrix}\tilde{\alpha}_n - \alpha_0 + n^{1/4} \\ \tilde{\delta}_n - \delta_0\end{pmatrix} + O_p(1).$$

By using the derivatives in (4)–(5) and (23)–(25), and the definition of a matrix exponential,

$$0 = 2(W_n X_n\delta_0, -X_n)'H_n\big(V_n - \frac{1}{2}n^{-1/2}W_n^2 X_n\delta_0\big) + 2(-W_n X_n\delta_0, X_n)'H_n(-W_n X_n\delta_0, X_n)\begin{pmatrix}\tilde{\alpha}_n - \alpha_0 + n^{-1/4} \\ \tilde{\delta}_n - \delta_0\end{pmatrix} + O_p(1).$$

Thus,

$$\sqrt{n}\begin{pmatrix}\tilde{\alpha}_n-\alpha_0+n^{-1/4}\\ \tilde{\delta}_n-\delta_0\end{pmatrix}=\big[\frac{1}{n}(-W_nX_n\delta_0,X_n)'H_n(-W_nX_n\delta_0,X_n)\big]^{-1}(-W_nX_n\delta_0,X_n)'H_n\big(\frac{1}{\sqrt{n}}V_n-\frac{1}{2n}W_n^2X_n\delta_0\big)+o_p(1). \tag{2.8}$$

A first order Taylor expansion of $\frac{\partial\hat{Q}_n(\tilde{\theta}_n)}{\partial\zeta}$ at $(\alpha_0-n^{-1/4},\delta_0',0)'$ yields

$$\begin{aligned}\frac{1}{\sqrt{n}}\frac{\partial\hat{Q}_n(\tilde{\theta}_n)}{\partial\zeta}&=\frac{1}{\sqrt{n}}\frac{\partial\hat{Q}_n(\alpha_0-n^{-1/4},\delta_0,0)}{\partial\zeta}+\frac{1}{n}\frac{\partial^2\hat{Q}_n(\alpha_0-n^{-1/4},\delta_0,0)}{\partial\zeta\partial\Psi'}\sqrt{n}\begin{pmatrix}\tilde{\alpha}_n-\alpha_0+n^{-1/4}\\ \tilde{\delta}_n-\delta_0\end{pmatrix}+o_p(1)\\&=-\frac{2}{\sqrt{n}}(W_nX_n^{**},Z_n)'H_n(e^{(\alpha_0-n^{-1/4})W_n}Y_n-X_n\delta_0)\\&\quad+\frac{2}{n}(W_nX_n^{**},Z_n)'H_n[-W_ne^{(\alpha_0-n^{-1/4})W_n}Y_n,X_n]\sqrt{n}\begin{pmatrix}\tilde{\alpha}_n-\alpha_0+n^{-1/4}\\ \tilde{\delta}_n-\delta_0\end{pmatrix}+o_p(1).\end{aligned}$$

Substituting (2.8) into the above equation yields

$$\begin{aligned}\frac{1}{\sqrt{n}}\frac{\partial\hat{Q}_n(\tilde{\theta}_n)}{\partial\zeta}&=-2(W_nX_n^{**},Z_n)'\mathbb{M}_{(-WX\delta_0,X)}\big(\frac{1}{\sqrt{n}}V_n-\frac{1}{2n}W_n^2X_n\delta_0\big)+o_p(1)\\&\xrightarrow{d}N\Big(\text{plim}_{n\to\infty}\frac{1}{n}(W_nX_n^{**},Z_n)'\mathbb{M}_{(-WX\delta_0,X)}W_n^2X_n\delta_0,\text{plim}_{n\to\infty}\frac{4}{n}(W_nX_n^{**},Z_n)'\mathbb{M}_{(-WX\delta_0,X)}(W_nX_n^{**},Z_n)\Big).\end{aligned}$$

Hence, the result in the proposition follows. □

Proof of Proposition 4.1. Let $\bar{g}_n(\theta)=\mathrm{E}[g_n(\theta)]$. By the definition of $\hat{\theta}_n$,

$$\frac{1}{n}\hat{Q}_n(\hat{\theta}_n)+\lambda_n\|\tilde{\zeta}_n\|^{-\mu}\|\hat{\zeta}_n\|\le\frac{1}{n}\hat{Q}_n(\theta_0)+\lambda_n\|\tilde{\zeta}_n\|^{-\mu}\|\zeta_0\|. \tag{2.9}$$

Note that $\frac{1}{n}\hat{Q}_n(\theta_0)=\frac{1}{n}V_n'F_n(\frac{1}{n}\hat{\Pi}_n)^{-1}\frac{1}{n}F_n'V_n=o_p(1)$. If $\zeta_0\neq0$, as $\tilde{\zeta}_n=\zeta_0+o_p(1)$, $\lambda_n\|\tilde{\zeta}_n\|^{-\mu}\|\zeta_0\|=O_p(\lambda_n)=o_p(1)$ under Assumption 11; if $\zeta_0=0$, $\lambda_n\|\tilde{\zeta}_n\|^{-\mu}\|\zeta_0\|=0$. As $\lambda_n>0$ and $\hat{Q}_n(\hat{\theta}_n)\ge0$, $\frac{1}{n}\hat{Q}_n(\hat{\theta}_n)=o_p(1)$ by (2.9). Since $\frac{1}{n}\hat{\Pi}_n-\frac{1}{n}\mathrm{E}(\Pi_n)=[\frac{1}{n}\hat{\Pi}_n-\frac{1}{n}\Pi_n]+[\frac{1}{n}\Pi_n-\frac{1}{n}\mathrm{E}(\Pi_n)]=o_p(1)$ by Proposition 2.4 and Lemma 1 in Qu and Lee (2012) and $\lim_{n\to\infty}\frac{1}{n}\mathrm{E}(\Pi_n)$ is nonsingular by Assumption 3, $\frac{1}{n}\hat{Q}_n(\hat{\theta}_n)\ge C\|\frac{1}{n}g_n(\hat{\theta}_n)\|^2$ w.p.a.1., where C is a finite positive constant. Then $\|\frac{1}{n}g_n(\hat{\theta}_n)\|=o_p(1)$. Since $\|\frac{1}{n}g_n(\hat{\theta}_n)\|\ge\|\frac{1}{n}\bar{g}_n(\hat{\theta}_n)\|-\|\frac{1}{n}g_n(\hat{\theta}_n)-\frac{1}{n}\bar{g}_n(\hat{\theta}_n)\|$ and $\sup_{\theta\in\Theta}\|\frac{1}{n}g_n(\theta)-\frac{1}{n}\bar{g}_n(\theta)\|=\sup_{\theta\in\Theta}\|\frac{1}{n}F_n'e^{(\alpha-\alpha_0)W_n}D_n\beta_0-\frac{1}{n}\mathrm{E}(F_n'e^{(\alpha-\alpha_0)W_n}D_n)\beta_0+\frac{1}{n}F_n'e^{(\alpha-\alpha_0)W_n}V_n-\frac{1}{n}[F_n'D_n-\mathrm{E}(F_n'D_n)]\beta\|=o_p(1)$ by Lemma 1, $\|\frac{1}{n}\bar{g}_n(\hat{\theta}_n)\|=o_p(1)$. As $\|\frac{1}{n}\bar{g}_n(\hat{\theta}_n)\|=\|\frac{1}{n}[\mathrm{E}(F_n'e^{(\hat{\alpha}_n-\alpha_0)W_n}D_n)\beta_0,\mathrm{E}(F_n'D_n)]\binom{1}{-\hat{\beta}_n}\|$, Assumption 5 implies that $\hat{\alpha}_n=\alpha_0+o_p(1)$. Then $\|\frac{1}{n}\bar{g}_n(\hat{\theta}_n)\|=\|\frac{1}{n}\mathrm{E}(F_n'D_n)(\beta_0-\hat{\beta}_n)+o_p(1)\|$. Assumption 5 further implies that $\hat{\beta}_n=\beta_0+o_p(1)$. Thus the result in the proposition holds. □

Proof of Proposition 4.2. With $\zeta_0=0$, by the definition of $\hat{\theta}_n$,

$$\frac{1}{n}\hat{Q}_n(\hat{\theta}_n)+\lambda_n\|\tilde{\zeta}_n\|^{-\mu}\|\hat{\zeta}_n\|\le\frac{1}{n}\hat{Q}_n(\theta_0)+\lambda_n\|\tilde{\zeta}_n\|^{-\mu}\|\zeta_0\|=\frac{1}{n}\hat{Q}_n(\theta_0). \tag{2.10}$$

Note that $\lambda_n>0$, $\frac{1}{n}\hat{Q}_n(\hat{\theta}_n)\ge C\|\frac{1}{n}g_n(\hat{\theta}_n)\|^2$ w.p.a.1. for some finite positive constant C, and

$$\frac{1}{n}\hat{Q}_n(\theta_0)=\frac{1}{n}V_n'F_n\big(\frac{1}{n}\hat{\Pi}\big)^{-1}\frac{1}{n}F_n'V_n=O_p(n^{-1}).$$

Then $\frac{1}{n}g_n(\hat{\theta}_n)=O_p(n^{-1/2})$. If $\hat{\zeta}_n\neq0$, then the AGLASSO criterion function is differentiable at $\hat{\theta}_n=(\hat{\alpha}_n,\hat{\delta}_n',\hat{\zeta}_n')'$ and we have the following first order condition with respect to ζ:

$$-\frac{2}{n}(W_nX_n^{**},Z_n)'F_n\big(\frac{1}{n}\hat{\Pi}_n\big)^{-1}\frac{1}{n}g_n(\hat{\theta}_n)+\lambda_n\|\tilde{\zeta}_n\|^{-\mu}\hat{\zeta}_n\|\hat{\zeta}_n\|^{-1}=0. \tag{2.11}$$

Note that $\frac{1}{n}(W_n X_n^{**}, Z_n)' F_n = O_p(1)$. As $\hat{\zeta}_n \neq 0$, there must be some component $\hat{\zeta}_{nj}$ of $\hat{\zeta}_n = (\hat{\zeta}_{n1}, \dots, \hat{\zeta}_{np})'$, where p is the length of ζ, such that $|\hat{\zeta}_{nj}| = \max\{|\hat{\zeta}_{nk}| : 1 \leq k \leq p\}$. Then $|\hat{\zeta}_{nj}|/\|\hat{\zeta}_n\| \geq 1/\sqrt{p} > 0$. Under Assumption 12, (2.11) cannot hold w.p.a.1., which is a contradiction to the first order condition. Hence the result in the proposition follows. □

Proof of Proposition 4.3. The first order derivative of the AGLASSO criterion function with respect to Ψ evaluated at $\hat{\theta}_n$ is

$$\frac{2}{n}(W_n e^{\hat{\alpha}_n W_n} y_n, -X_n)' F_n (\frac{1}{n}\hat{\Pi}_n)^{-1} \frac{1}{n} F_n' (e^{\hat{\alpha}_n W_n} y_n - D_n \hat{\beta}_n) = 0.$$

With $\zeta_0 = 0$, Proposition 4.2 shows that $\hat{\zeta}_n = 0$ w.p.a.1. Hence the following equation holds w.p.a.1:

$$\frac{2}{n}(W_n e^{\hat{\alpha}_n W_n} y_n, -X_n)' F_n (\frac{1}{n}\hat{\Pi}_n)^{-1} \frac{1}{n} F_n' (e^{\hat{\alpha}_n W_n} y_n - X_n \hat{\delta}_n) = 0.$$

This first order condition is exactly the same as the one derived from the corresponding N2SLS criterion function by imposing the constraint $\zeta = 0$. Thus the oracle property becomes apparent. By the mean value theorem,

$$0 = \frac{2}{n}(W_n e^{\hat{\alpha}_n W_n} y_n, -X_n)' F_n \big(\frac{1}{n}\hat{\Pi}_n\big)^{-1} \frac{1}{n} F_n' [e^{\alpha_0 W_n} y_n - X_n \delta_0 + (W_n e^{\ddot{\alpha}_n W_n} y_n, -X_n)(\hat{\Psi}_n - \Psi_0)],$$

where $\ddot{\Psi}_n$ lies between $\hat{\Psi}_n$ and Ψ_0. Thus,

$$\begin{aligned}
&\sqrt{n}(\hat{\Psi}_n - \Psi_0) \\
&= -\big[\frac{1}{n}(W_n e^{\hat{\alpha}_n W_n} y_n, -X_n)' F_n (\frac{1}{n}\hat{\Pi}_n)^{-1} \frac{1}{n} F_n' (W_n e^{\ddot{\alpha}_n W_n} y_n, -X_n)\big]^{-1} \frac{1}{n}(W_n e^{\hat{\alpha}_n W_n} y_n, -X_n)' F_n \big(\frac{1}{n}\hat{\Pi}_n\big)^{-1} \frac{1}{\sqrt{n}} F_n' V_n \\
&= \big[\frac{1}{n}(-W_n X_n \delta_0, X_n)' F_n \big(\frac{1}{n}\Pi_n\big)^{-1} \frac{1}{n} F_n' (-W_n X_n \delta_0, X_n)\big]^{-1} \frac{1}{n}(-W_n X_n \delta_0, X_n)' F_n (\frac{1}{n}\Pi_n)^{-1} \frac{1}{\sqrt{n}} F_n' V_n + o_p(1) \\
&\xrightarrow{d} N\big(0, \lim_{n\to\infty} \frac{1}{n}\big\{\mathrm{E}[(-W_n X_n \delta_0, X_n)' F_n] \bar{\Pi}_n^{-1} \mathrm{E}[F_n' (-W_n X_n \delta_0, X_n)]\big\}^{-1}\big),
\end{aligned}$$

where the asymptotic distribution follows by Lemma 2 in Qu and Lee (2012). □

Proof of Proposition 4.4. Note that

$$\begin{aligned}
&\frac{1}{n} g_n'(\hat{\theta}_n) \big(\frac{1}{n}\hat{\Pi}_n\big)^{-1} \frac{1}{n} g_n(\hat{\theta}_n) - \frac{1}{n} g_n'(\theta_0) \big(\frac{1}{n}\hat{\Pi}_n\big)^{-1} \frac{1}{n} g_n(\theta_0) \\
&= \frac{1}{n}[g_n(\hat{\theta}_n) - g_n(\theta_0)]' \big(\frac{1}{n}\hat{\Pi}_n\big)^{-1} \frac{1}{n}[g_n(\hat{\theta}_n) - g_n(\theta_0)] + \frac{2}{n}[g_n(\hat{\theta}_n) - g_n(\theta_0)]' \big(\frac{1}{n}\hat{\Pi}_n\big)^{-1} \frac{1}{n} g_n(\theta_0) \qquad (2.12) \\
&\geq C_1 \big\|\frac{1}{n} g_n(\hat{\theta}_n) - \frac{1}{n} g_n(\theta_0)\big\|^2 - C_2 \big\|\frac{1}{n} g_n(\hat{\theta}_n) - \frac{1}{n} g_n(\theta_0)\big\| \big\|\frac{1}{n} g_n(\theta_0)\big\|,
\end{aligned}$$

w.p.a.1., where C_1 and C_2 are finite positive constants, and the inequality follows by Assumption 3 and the Cauchy-Schwarz inequality. By the mean value theorem,

$$\frac{1}{n} g_n(\hat{\theta}_n) - \frac{1}{n} g_n(\theta_0) = \frac{1}{n}\mathrm{E}\big(\frac{\partial g_n(\theta_0)}{\partial \theta'}\big)(\hat{\theta}_n - \theta_0) + \frac{1}{n}\big[\frac{\partial g_n(\ddot{\theta}_n)}{\partial \theta'} - \mathrm{E}\big(\frac{\partial g_n(\theta_0)}{\partial \theta'}\big)\big](\hat{\theta}_n - \theta_0), \qquad (2.13)$$

where $\ddot{\theta}_n$ lies between $\hat{\theta}_n$ and θ_0. In addition, $\frac{1}{n}\frac{\partial g_n(\theta)}{\partial \theta'} - \frac{1}{n}\frac{\partial g_n(\theta_0)}{\partial \theta'} = [\frac{1}{n} F_n' W_n^2 e^{\alpha_1 W_n} Y_n (\alpha - \alpha_0), 0_{k_f \times k_d}]$ for some α_1 between α and α_0, where $\frac{1}{n} F_n' W_n^2 e^{\alpha W_n} Y_n = O_p(1)$ uniformly in a neighborhood of α_0 by Lemma 1. Furthermore, $\frac{1}{n}\frac{\partial g_n(\theta_0)}{\partial \theta'} - \frac{1}{n}\mathrm{E}\big(\frac{\partial g_n(\theta_0)}{\partial \theta'}\big) = [\frac{1}{n} F_n' W_n D_n \beta_0 - \frac{1}{n}\mathrm{E}(F_n' W_n D_n)\beta_0 + \frac{1}{n} F_n' W_n V_n, -\frac{1}{n} F_n' D_n + \frac{1}{n}\mathrm{E}(F_n' D_n)] = o_p(1)$ by

Lemma 1. Thus, $\frac{1}{n}[\frac{\partial g_n(\ddot{\theta}_n)}{\partial\theta'} - \mathrm{E}(\frac{\partial g_n(\theta_0)}{\partial\theta'})] = o_p(1)$. Since $\mathrm{E}(\frac{1}{n}\frac{\partial g_n(\theta_0)}{\partial\theta'}) = O(1)$ has full rank when $\zeta_0 \neq 0$ for large enough n, (2.13) implies that

$$C_3(1-a_n)\|\hat{\theta}_n - \theta_0\| \le \frac{1}{n}\|g_n(\hat{\theta}_n) - g_n(\theta_0)\| \le C_3(1+a_n)\|\hat{\theta}_n - \theta_0\|, \tag{2.14}$$

where C_3 is a finite positive constant, $a_n \ge 0$ and $a_n = o_p(1)$. By (2.9),

$$\frac{1}{n}g_n(\hat{\theta}_n)'\hat{\Pi}_n^{-1}g_n(\hat{\theta}_n) - \frac{1}{n}g_n(\theta_0)'\hat{\Pi}_n^{-1}g_n(\theta_0) \le \lambda_n\|\tilde{\zeta}_n\|^{-\mu}(\|\zeta_0\| - \|\hat{\zeta}_n\|) \le \lambda_n\|\tilde{\zeta}_n\|^{-\mu}\|\hat{\zeta}_n - \zeta_0\| \le C_4\lambda_n\|\hat{\theta}_n - \theta_0\| \tag{2.15}$$

w.p.a.1., where C_4 is a finite positive constant, and the last inequality follows since $\zeta_0 \neq 0$ and $\tilde{\zeta}_n = \zeta_0 + o_p(1)$. Combining (2.12), (2.14) and (2.15) yields

$$C_1C_3^2(1-a_n)^2\|\hat{\theta}_n - \theta_0\|^2 - C_2C_3(1+a_n)\|\hat{\theta}_n - \theta_0\| \cdot \frac{1}{n}\|g_n(\theta_0)\| \le C_4\lambda_n\|\hat{\theta}_n - \theta_0\|.$$

The above inequality can be written as

$$\|\hat{\theta}_n - \theta_0\| \cdot \big[C_1C_3^2(1-a_n)^2\|\hat{\theta}_n - \theta_0\| - C_2C_3(1+a_n)\big\|\frac{1}{n}g_n(\theta_0)\big\| - C_4\lambda_n\big] \le 0.$$

As $\frac{1}{n}g_n(\theta_0) = O_p(n^{-1/2})$, $\|\hat{\theta}_n - \theta_0\| = O_p(n^{-1/2} + \lambda_n)$. □

Proof of Proposition 4.5. If $\zeta_0 \neq 0$, the first order condition of the AGLASSO criterion function with respect to θ is:

$$\frac{2}{n}(W_ne^{\hat{\alpha}_nW_n}y_n, -D_n)'F_n\big(\frac{1}{n}\hat{\Pi}_n\big)^{-1}\frac{1}{n}g_n(\hat{\theta}_n) + \lambda_n\|\tilde{\zeta}_n\|^{-\mu}\|\hat{\zeta}_n\|^{-1}\begin{pmatrix}0\\ \hat{\zeta}_n\end{pmatrix} = 0. \tag{2.16}$$

By (2.14), Proposition 4.4 and Assumption 13, $\frac{1}{n}g_n(\hat{\theta}_n) = O_p(n^{-1/2})$. Then the first term on the l.h.s. of (2.16) has the order $O_p(n^{-1/2})$. As $\zeta_0 \neq 0$, $\tilde{\zeta}_n = \zeta_0 + o_p(1) = O_p(1)$ and $\hat{\zeta}_n = \zeta_0 + o_p(1) = O_p(1)$. By Assumption 13, the second term on the l.h.s. of (2.16) has the order $o_p(n^{-1/2})$. Hence, by the mean value theorem,

$$\frac{2}{n}(W_ne^{\hat{\alpha}_nW_n}y_n, -D_n)'F_n\big(\frac{1}{n}\hat{\Pi}_n\big)^{-1}\big[\frac{1}{\sqrt{n}}g_n(\theta_0) + \frac{1}{n}F_n'(W_ne^{\ddot{\alpha}_nW_n}y_n, -D_n)\sqrt{n}(\hat{\theta}_n - \theta_0)\big] + o_p(1) = 0,$$

where $\ddot{\alpha}_n$ lies between α_0 and $\hat{\alpha}_n$. If follows that

$$\begin{aligned}\sqrt{n}(\hat{\theta}_n - \theta_0) &= -\big[\frac{1}{n}(W_ne^{\hat{\alpha}_nW_n}y_n, -D_n)'H_n(W_ne^{\ddot{\alpha}_nW_n}y_n, -D_n)\big]^{-1}\frac{1}{\sqrt{n}}(W_ne^{\hat{\alpha}_nW_n}y_n, -D_n)'H_nV_n + o_p(1)\\ &= -\big[\frac{1}{n}(W_nD_n\beta_0, -D_n)'H_n(W_nD_n\beta_0, -D_n)\big]^{-1}\frac{1}{\sqrt{n}}(W_nD_n\beta_0, -D_n)'H_nV_n + o_p(1)\\ &\xrightarrow{d} N\big(0, \lim_{n\to\infty}\big\{\frac{1}{n}\mathrm{E}[(-W_nD_n\beta_0, D_n)'F_n]\bar{\Pi}_n^{-1}\mathrm{E}[F_n'(-W_nD_n\beta_0, D_n)]\big\}^{-1}\big).\end{aligned}$$

□

Proof of Proposition 4.6. We consider the following two cases separately: (1) $\zeta_0 \neq 0$, but $\hat{\zeta}_\lambda = 0$; (2) $\zeta_0 = 0$, but $\hat{\zeta}_\lambda \neq 0$.

Case 1: $\zeta_0 \neq 0$, but $\hat{\zeta}_\lambda = 0$. Let $\ddot{\theta}_n = (\ddot{\Psi}_n', 0)'$ be the restricted N2SLS estimator with the restriction $\zeta = 0$ imposed, where $\ddot{\Psi}_n = \arg\min_\Psi[(e^{\alpha W_n}y_n - X_n\delta)'\hat{H}_n(e^{\alpha W_n}y_n - X_n\delta)]$. As $\zeta_0 \neq 0$, $\bar{\theta} \equiv \mathrm{plim}_{n\to\infty}\ddot{\theta}_n \neq \theta_0$. Note that

$\frac{1}{n}g_n(\ddot{\theta}_n) = \frac{1}{n}\bar{g}_n(\ddot{\theta}_n)+o_p(1) = \frac{1}{n}\bar{g}_n(\bar{\theta})+o_p(1)$, where the first equality follows since $\sup_{\theta\in\Theta}\frac{1}{n}\|g_n(\theta)-\bar{g}_n(\theta)\| = o_p(1)$ as shown in the proof of Proposition 2.1, and the second equality follows by the mean value theorem. By Assumption 5, $\lim_{n\to\infty}\frac{1}{n}\bar{g}_n(\bar{\theta}) \neq 0$. Then $\frac{1}{n}\hat{Q}_n(\ddot{\theta}_n) \xrightarrow{p} C > 0$ for a constant C. By the definition of $\ddot{\theta}_n$ and the setting of Case 1, $h_n(\lambda) = \frac{1}{n}\hat{Q}_n(\hat{\theta}_\lambda) - \Gamma_n \geq \frac{1}{n}\hat{Q}_n(\ddot{\theta}_n) - \Gamma_n$. Thus, under Assumption 14, $h_n(\lambda) > C/2 > 0$ w.p.a.1. Furthermore, by Proposition 4.1, $\hat{\theta}_{\bar{\lambda}_n} = \theta_0 + o_p(1)$. Then $\frac{1}{n}g_n(\hat{\theta}_{\bar{\lambda}_n}) = \frac{1}{n}\bar{g}_n(\theta_0) + o_p(1) = o_p(1)$ and $\frac{1}{n}\hat{Q}_n(\hat{\theta}_{\bar{\lambda}_n}) = o_p(1)$. It follows that $h_n(\bar{\lambda}_n) = o_p(1)$. Therefore, $\mathrm{P}(\inf_{\lambda\in\{\lambda\in\Lambda:\zeta_0\neq 0,\text{ but }\hat{\zeta}_\lambda=0\}} h_n(\lambda) > h_n(\bar{\lambda}_n)) \to 1$ as $n\to\infty$.

Case 2: $\zeta_0 = 0$, but $\hat{\zeta}_\lambda \neq 0$. Under this setting, $h_n(\lambda) = \frac{1}{n}\hat{Q}_n(\hat{\theta}_\lambda)$. By Proposition 4.2, $\mathrm{P}(\hat{\zeta}_{\bar{\lambda}_n} = 0) \to 1$ as $n\to\infty$. Then w.p.a.1.,

$$
\begin{aligned}
n^{1/2}[h_n(\lambda) - h_n(\bar{\lambda}_n)] &= n^{-1/2}\hat{Q}_n(\hat{\theta}_\lambda) - n^{-1/2}\hat{Q}_n(\hat{\theta}_{\bar{\lambda}_n}) + n^{1/2}\Gamma_n \\
&\geq n^{-1/2}\hat{Q}_n(\tilde{\theta}_n) - n^{-1/2}\hat{Q}_n(\hat{\theta}_{\bar{\lambda}_n}) + n^{1/2}\Gamma_n,
\end{aligned}
\tag{2.17}
$$

where $\tilde{\theta}_n$ is the feasible N2SLS estimator (without penalty). By Proposition 2.4, $\tilde{\theta}_n = \theta_0 + O_p(n^{-1/4})$ when $\zeta_0 = 0$. Then by the mean value theorem, $n^{-3/4}g_n(\tilde{\theta}_n) = n^{-3/4}g_n(\theta_0) + \frac{1}{n}\frac{\partial g_n(\dot{\theta}_n)}{\partial\theta'}n^{1/4}(\tilde{\theta}_n - \theta_0) = n^{-3/4}F_n'V_n + \frac{1}{n}\frac{\partial g_n(\dot{\theta}_n)}{\partial\theta'}n^{1/4}(\tilde{\theta}_n - \theta_0)$, where $\dot{\theta}_n$ lies between θ_0 and $\tilde{\theta}_n$. As in the proof of Proposition 4.5, $\frac{1}{n}\frac{\partial g_n(\dot{\theta}_n)}{\partial\theta'} = \frac{1}{n}\mathrm{E}(\frac{\partial g_n(\theta_0)}{\partial\theta'}) + o_p(1) = O_p(1)$. Thus, $n^{-3/4}g_n(\tilde{\theta}_n) = O_p(1)$ and $n^{-1/2}\hat{Q}_n(\tilde{\theta}_n) = O_p(1)$. Since $\mathrm{P}(\hat{\zeta}_{\bar{\lambda}_n} = 0) \to 1$ as $n\to\infty$, $n^{-1/2}g_n(\hat{\theta}_{\bar{\lambda}_n}) = n^{-1/2}F_n'(e^{\hat{\alpha}_{\bar{\lambda}_n}W_n}y_n - X_n\hat{\delta}_{\bar{\lambda}_n})$ w.p.a.1. By Proposition 4.3, $\hat{\Psi}_{\bar{\lambda}_n} = \Psi_0 + O_p(n^{-1/2})$. Then by the mean value theorem, $n^{-1/2}g_n(\hat{\theta}_{\bar{\lambda}_n}) = n^{-1/2}g_n(\theta_0) + \frac{1}{n}\frac{\partial g_n(\ddot{\theta}_n)}{\partial\Psi'}n^{1/2}(\hat{\Psi}_{\bar{\lambda}_n} - \Psi_0) = n^{-1/2}F_n'V_n + \frac{1}{n}\frac{\partial g_n(\ddot{\theta}_n)}{\partial\Psi'}n^{1/2}(\hat{\Psi}_{\bar{\lambda}_n} - \Psi_0) = O_p(1)$, where $\ddot{\theta}_n$ lies between θ_0 and $\hat{\theta}_{\bar{\lambda}_n}$. Thus, $n^{-1/2}\hat{Q}_n(\hat{\theta}_{\bar{\lambda}_n}) = o_p(1)$. Since $n^{1/2}\Gamma_n \to \infty$ as $n\to\infty$ under Assumption 14, (2.17) implies that $\mathrm{P}(\inf_{\lambda\in\{\lambda\in\Lambda:\zeta_0=0,\text{ but }\hat{\zeta}_\lambda\neq 0\}} h_n(\lambda) > h_n(\bar{\lambda}_n)) \to 1$ as $n\to\infty$.

Combining the results in the above two cases, we then have the result in the proposition. □

References

Andrews, D.W.K., 1987. Asymptotic results for generalized Wald tests. Econometric Theory 3, 348–358.

Davidson, J., 1994. Stochastic limit theory: An introduction for econometricians. Oxford University Press, USA.

Kelejian, H.H., Prucha, I.R., 2007. HAC estimation in a spatial framework. Journal of Econometrics 140, 131–154.

Qu, X., Lee, L.F., 2012. LM tests for spatial correlation in spatial models with limited dependent variables. Regional Science and Urban Economics 42, 430–445.

Rotnitzky, A., Cox, D.R., Bottai, M., Robins, J., 2000. Likelihood-based inference with singular information matrix. Bernoulli 6, 243–284.

Ruud, P.A., 2000. An Introduction to Classical Econometric Theory. Oxford University Press.

White, H., 1994. Estimation, Inference and Specification Analysis. New York: Cambridge University Press.