Supplementary Internet Appendix: Structural Threshold Regression*

Andros Kourtellos[†] **Thanasis Stengos**[‡] **Chih Ming Tan**[§]

First Draft: May 2009
This Draft: February 26, 2015

Keywords: endogenous thresholds, sample splitting, nonlinear regression.

JEL Classification Codes: C13, C51

Abstract

This paper introduces the structural threshold regression (STR) model that allows for an endogenous threshold variable as well as for endogenous regressors. This model provides a parsimonious way of modeling nonlinearities and has many potential applications in economics and finance. Our framework can be viewed as a generalization of the simple threshold regression framework of Hansen (2000) and Caner and Hansen (2004) to allow for the endogeneity of the threshold variable and regime-specific heteroskedasticity. Our estimation of the threshold parameter is based on a two-stage concentrated least squares method that involves an inverse Mills ratio bias correction term in each regime. We derive its asymptotic distribution and propose a method to construct confidence intervals. We also provide inference for the slope parameters based on a generalized method of moments. Finally, we investigate the performance

*We thank the editor Peter C. B. Phillips, the co-editor Oliver Linton, and two anonymous referees whose comments greatly improved the paper. We also thank Bruce Hansen for helpful comments and seminar participants at the Athens University of Economics and Business, Hebrew University of Jerusalem, Ryerson University, Simon Fraser University, Universit libre de Bruxelles, University of Cambridge, University of Palermo, University of Waterloo, the University of Western Ontario, 10th World Congress of the Econometric Society in Shanghai, 27th Annual Meeting of the Canadian Econometrics Study Group in Vancouver, and 23rd $(EC)^2$ conference.

[†]Corresponding author: Department of Economics, University of Cyprus, P.O. Box 537, CY 1678 Nicosia, Cyprus, email: `andros@ucy.ac.cy`.

[‡]Department of Economics, University of Guelph, Guelph, Ontario N1G 2W1, Canada, email: `tstengos@uoguelph.ca`.

[§]Department of Economics, University of North Dakota, 293 Centennial Drive Stop 8369, Grand Forks, North Dakota, USA, email: `chihming.tan@business.und.edu`.

of the asymptotic approximations using a Monte Carlo simulation, which shows the applicability of the method in finite samples.

1 Monte Carlo Simulation Results

We explore two sets of simulation experiments that allow for the endogeneity of the threshold variable. The first set of simulations assumes a threshold regression model that allows for an endogenous threshold variable but retains the assumption of an exogenous slope variable (Model 1). The second set of simulations is based on a threshold regression model that allows for endogeneity in both the threshold and the slope variable (Model 2).

Model 1 is given by

$$y_i = \beta_1 + \beta_2 x_i + (\delta_1 + \delta_2 x_i) I\{q_i \leq \gamma\} + u_i, \tag{I.A.1}$$

where

$$q_i = 2 + z_{qi} + v_{qi}. \tag{I.A.2}$$

The threshold parameter is set at the center of the distribution of q_i, hence $\gamma = 2$. The instrumental variable z_{qi} is given by

$$z_{qi} = (w x_i + (1-w)\varsigma_{zi}) / \sqrt{w^2 + (1-w)^2} \tag{I.A.3}$$

and

$$u_i = 0.1\varsigma_{ui} + \kappa v_{qi}, \tag{I.A.4}$$

where x_i, v_{qi}, ς_{zi}, and ς_{ui} are independent *i.i.d.* $N(0,1)$ random variables. The degree of endogeneity of the threshold is controlled by κ. The degree of correlation between the instrumental variable z_{qi} and the included exogenous slope variable x_i is controlled by w. We fix $w = 0.5$, $\beta_1 = \beta_2 = 1$, and $\delta_1 = 0$ and vary δ_2 over the values of $1, 2, 3, 4, 5$, which correspond to a range of small to large threshold effects. We also vary κ over the values of 0.05, 0.50, 0.95 that correspond to low, medium, and large degrees of endogeneity of the threshold variable.

Model 2 is given by

$$y_i = \beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i} + (\delta_1 + \delta_2 x_{1i} + \delta_3 x_{2i}) I\{q_i \leq \gamma\} + u_i, \tag{I.A.5}$$

where q_i is given by equation (I.A.2) and

$$x_{1i} = z_{xi} + v_{xi},$$

where

$$z_{xi} = (w x_{2i} + (1-w)\varsigma_{zi}) \,/ \sqrt{w^2 + (1-w)^2}, \tag{I.A.6}$$

and

$$u_i = (c_{xu} v_{xi} + c_{qu} v_{qi} + (1 - c_{xu} - c_{qu})\varsigma_{ui}) \,/ \sqrt{c_{xu}^2 + c_{qu}^2 + (1 - c_{xu} - c_{qu})^2}, \tag{I.A.7}$$

where x_{2i}, ς_{zi} and ς_{ui} are independent *i.i.d.* $N(0,1)$ random variables. The degree of endogeneity of the threshold variable is controlled by the correlation coefficient between u_i and v_{qi} given by $c_{qu}/\sqrt{c_{xu}^2 + c_{qu}^2 + (1 - c_{xu} - c_{qu})^2}$. Similarly, the degree of endogeneity of x_{1i} is determined by the correlation between u_i and v_{xi} given by $c_{xu}/\sqrt{c_{xu}^2 + c_{qu}^2 + (1 - c_{xu} - c_{qu})^2}$. We fix c_{xu}, $w = 0.5$, $\beta_1 = \beta_2 = 1$, and $\delta_1 = \delta_2 = 0$. δ_3 varies over the values of $1, 2, 3, 4, 5$. c_{qu} varies over the values of 0.05, 0.25, 0.45 that correspond to correlations between q_i and u_i of about 0.07, 0.4, 0.7, respectively.

We consider sample sizes of 100, 250, 500, and 1000 using 1000 Monte Carlo replications simulations. In unreported exercises we also investigated alternative values of w and c_{xu} and found qualitatively similar results.

Tables I.A.1, I.A.2, and I.A.3 present the quantiles of the distribution of the STR (constrained) estimators for the threshold parameter, the slope coefficient of the upper regime, and the threshold effect, respectively. Table I.A.4 provides the 90% confidence interval coverage for the threshold parameter γ. Finally, Table I.A.5 presents the 95% confidence interval coverage for the slope coefficients β_2 and δ_2 in the case of Model 1 and β_3 and δ_3 in the case of Model 2.

Table I.A.1: Quantiles of the distribution of the STR threshold estimator $\widehat{\gamma}$

Model 1 - endogeneity only in the threshold variable

	$\delta_2 = 1$			$\delta_2 = 2$			$\delta_2 = 3$			$\delta_2 = 4$			$\delta_2 = 5$		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
Sample size							*low degree of endogeneity*								
100	1.890	1.976	2.022	1.897	1.976	2.000	1.898	1.976	1.999	1.898	1.976	1.999	1.899	1.976	1.999
250	1.955	1.991	2.009	1.958	1.990	2.000	1.958	1.990	2.000	1.958	1.990	2.000	1.958	1.990	2.000
500	1.977	1.995	2.005	1.977	1.995	2.000	1.978	1.995	2.000	1.978	1.995	2.000	1.978	1.995	2.000
1000	1.989	1.998	2.001	1.989	1.998	2.000	1.990	1.998	2.000	1.989	1.998	2.000	1.989	1.998	2.000
							medium degree of endogeneity								
100	1.802	1.982	2.134	1.872	1.979	2.058	1.877	1.977	2.043	1.883	1.978	2.026	1.888	1.977	2.019
250	1.922	1.992	2.059	1.950	1.991	2.026	1.955	1.991	2.014	1.956	1.991	2.010	1.956	1.990	2.002
500	1.962	1.997	2.029	1.972	1.996	2.010	1.975	1.995	2.006	1.976	1.995	2.006	1.976	1.995	2.004
1000	1.980	1.998	2.016	1.988	1.998	2.007	1.989	1.998	2.004	1.989	1.998	2.002	1.989	1.998	2.002
							high degree of endogeneity								
100	1.596	1.991	2.359	1.830	1.982	2.129	1.864	1.980	2.075	1.869	1.978	2.053	1.874	1.978	2.046
250	1.796	1.996	2.146	1.936	1.993	2.056	1.947	1.991	2.030	1.952	1.991	2.022	1.954	1.991	2.017
500	1.898	1.998	2.075	1.963	1.996	2.024	1.971	1.996	2.013	1.973	1.996	2.009	1.974	1.995	2.008
1000	1.942	1.999	2.038	1.981	1.998	2.015	1.985	1.998	2.008	1.988	1.998	2.007	1.989	1.998	2.004

Model 2 - endogeneity in both the threshold and slope variables

	$\delta_2 = 1$			$\delta_2 = 2$			$\delta_2 = 3$			$\delta_2 = 4$			$\delta_2 = 5$		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
Sample size							*low degree of endogeneity*								
100	1.097	1.964	2.842	1.516	1.971	2.483	1.744	1.976	2.203	1.802	1.975	2.127	1.834	1.976	2.098
250	1.352	1.988	2.608	1.824	1.992	2.186	1.900	1.991	2.088	1.924	1.991	2.056	1.941	1.991	2.044
500	1.635	1.997	2.324	1.898	1.996	2.063	1.948	1.996	2.036	1.960	1.996	2.029	1.969	1.996	2.019
1000	1.819	1.997	2.136	1.958	1.998	2.031	1.977	1.998	2.021	1.982	1.998	2.014	1.985	1.998	2.010
							medium degree of endogeneity								
100	1.079	1.937	2.856	1.392	1.964	2.485	1.709	1.975	2.223	1.808	1.976	2.138	1.840	1.976	2.112
250	1.223	1.968	2.601	1.776	1.989	2.186	1.894	1.991	2.094	1.918	1.991	2.056	1.938	1.991	2.046
500	1.361	1.988	2.436	1.874	1.995	2.067	1.940	1.995	2.036	1.958	1.996	2.024	1.967	1.996	2.021
1000	1.640	1.991	2.211	1.942	1.997	2.035	1.973	1.998	2.021	1.981	1.998	2.014	1.984	1.998	2.010
							high degree of endogeneity								
100	1.051	1.924	2.872	1.333	1.954	2.470	1.714	1.973	2.198	1.784	1.975	2.129	1.829	1.976	2.102
250	1.200	1.955	2.552	1.704	1.986	2.183	1.888	1.989	2.096	1.920	1.990	2.050	1.939	1.991	2.043
500	1.332	1.976	2.455	1.855	1.993	2.072	1.939	1.995	2.034	1.957	1.996	2.023	1.966	1.996	2.019
1000	1.549	1.977	2.235	1.926	1.997	2.037	1.974	1.998	2.022	1.980	1.998	2.014	1.983	1.998	2.010

Notes: Model 1 refers to equation (I.A.1) with $\gamma = 2$, $\beta_1 = \beta_2 = 1$, and $\delta_1 = 0$. Model 2 refers to equation (I.A.5) with $\gamma = 2$, $\beta_1 = \beta_2 = \beta_3 = 1$ and $\delta_1 = \delta_2 = 0$. For Model 1 'low', 'medium', and 'high' corresponds to κ=0.05, 0.50, 0.95 in equation (I.A.4) and for Model 2 to $c_{qu} = 0.05, 0.25, 0.40$ in equation (I.A.7).

Table I.A.2: Quantiles of the distribution of the STR estimator for the slope coefficient of the upper regime

Model 1: Quantiles of the distribution of $\widehat{\beta}_{2,LS}$

	$\delta_2 = 1$			$\delta_2 = 2$			$\delta_2 = 3$			$\delta_2 = 4$			$\delta_2 = 5$		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
Sample size							*low degree of endogeneity*								
100	0.969	1.000	1.035	0.969	1.000	1.035	0.969	1.000	1.035	0.969	1.000	1.036	0.969	1.000	1.036
250	0.980	1.000	1.021	0.980	1.000	1.021	0.980	1.000	1.021	0.980	1.000	1.021	0.980	1.000	1.021
500	0.986	1.000	1.015	0.986	1.000	1.015	0.986	1.000	1.015	0.986	1.000	1.015	0.986	1.000	1.015
1000	0.990	1.000	1.010	0.990	1.000	1.010	0.990	1.000	1.010	0.990	1.000	1.010	0.990	1.000	1.010
							medium degree of endogeneity								
100	0.861	1.002	1.176	0.863	0.998	1.177	0.863	0.997	1.177	0.863	0.997	1.177	0.863	0.997	1.177
250	0.910	1.003	1.102	0.910	1.002	1.102	0.908	1.002	1.102	0.907	1.003	1.101	0.907	1.003	1.101
500	0.936	0.998	1.065	0.936	0.998	1.065	0.935	0.998	1.065	0.935	0.998	1.065	0.935	0.998	1.065
1000	0.956	1.000	1.046	0.956	1.000	1.046	0.956	1.000	1.046	0.956	1.000	1.046	0.956	1.000	1.046
							high degree of endogeneity								
100	0.752	1.015	1.357	0.736	0.998	1.332	0.738	0.996	1.326	0.736	0.995	1.328	0.740	0.994	1.327
250	0.836	1.009	1.200	0.833	1.004	1.191	0.829	1.002	1.191	0.829	1.001	1.191	0.829	1.002	1.191
500	0.885	0.999	1.128	0.883	0.997	1.120	0.886	0.997	1.119	0.886	0.997	1.119	0.886	0.996	1.119
1000	0.919	1.000	1.088	0.918	0.999	1.086	0.918	0.998	1.085	0.918	0.999	1.085	0.918	0.999	1.085

Model 2: Quantiles of the distribution of $\widehat{\beta}_{3,GMM}$

Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
Sample size		$\delta_3 = 1$			$\delta_3 = 2$			$\delta_3 = 3$			$\delta_3 = 4$			$\delta_3 = 5$	
							low degree of endogeneity								
100	0.636	1.020	1.432	0.678	1.022	1.374	0.693	1.014	1.340	0.712	1.009	1.315	0.715	1.006	1.313
250	0.792	1.000	1.249	0.805	0.996	1.213	0.808	1.000	1.211	0.808	1.001	1.201	0.809	1.001	1.191
500	0.869	1.003	1.171	0.876	1.002	1.141	0.876	1.001	1.138	0.875	1.000	1.138	0.876	0.999	1.138
1000	0.903	1.004	1.104	0.906	1.002	1.097	0.909	1.002	1.095	0.909	1.001	1.095	0.910	1.001	1.096
							medium degree of endogeneity								
100	0.676	1.052	1.468	0.685	1.042	1.434	0.697	1.019	1.380	0.703	1.013	1.342	0.707	1.010	1.333
250	0.794	1.020	1.279	0.802	1.000	1.221	0.816	0.999	1.208	0.814	0.998	1.202	0.816	0.998	1.198
500	0.875	1.015	1.225	0.880	1.004	1.155	0.881	1.003	1.143	0.878	1.001	1.139	0.877	1.000	1.140
1000	0.911	1.015	1.158	0.909	1.003	1.102	0.911	1.002	1.094	0.910	1.002	1.095	0.910	1.001	1.094
							high degree of endogeneity								
100	0.680	1.076	1.483	0.703	1.048	1.491	0.708	1.024	1.403	0.706	1.017	1.371	0.706	1.010	1.369
250	0.813	1.045	1.308	0.814	1.017	1.238	0.818	1.002	1.209	0.819	1.000	1.207	0.818	1.000	1.206
500	0.882	1.032	1.250	0.880	1.011	1.169	0.877	1.005	1.143	0.877	1.004	1.140	0.877	1.003	1.140
1000	0.919	1.025	1.186	0.911	1.008	1.109	0.907	1.003	1.097	0.909	1.003	1.094	0.909	1.002	1.096

Notes: Model 1 refers to equation (I.A.1) with $\gamma = 2$, $\beta_1 = \beta_2 = 1$, and $\delta_1 = 0$. Model 2 refers to equation (I.A.5) with $\gamma = 2$, $\beta_1 = \beta_2 = \beta_3 = 1$ and $\delta_1 = \delta_2 = 0$. For Model 1 'low', 'medium', and 'high' corresponds to $\kappa = 0.05, 0.50, 0.95$ in equation (I.A.4) and for Model 2 to $c_{qu} = 0.05, 0.25, 0.40$ in equation (I.A.7).

Table I.A.3: Quantiles of the distribution of the STR estimator for the threshold effect

Model 1: Quantiles of the distribution of $\widehat{\delta}_{2,LS}$

	$\delta_2 = 1$			$\delta_2 = 2$			$\delta_2 = 3$			$\delta_2 = 4$			$\delta_2 = 5$		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
Sample size							*low degree of endogeneity*								
100	0.964	0.999	1.041	1.964	1.999	2.041	2.964	2.999	3.041	3.964	3.999	4.041	4.964	4.999	5.041
250	0.977	1.000	1.026	1.977	2.000	2.026	2.977	3.000	3.026	3.977	4.000	4.026	4.977	5.000	5.026
500	0.982	1.001	1.018	1.982	2.001	2.018	2.982	3.001	3.018	3.982	4.001	4.018	4.982	5.001	5.018
1000	0.988	1.000	1.012	1.988	2.000	2.012	2.988	3.000	3.012	3.988	4.000	4.012	4.988	5.000	5.012
							medium degree of endogeneity								
100	0.850	0.994	1.148	1.861	2.004	2.151	2.863	3.006	3.151	3.865	4.007	4.152	4.866	5.007	5.154
250	0.911	0.993	1.086	1.917	1.997	2.090	2.919	2.998	3.091	3.919	3.998	4.092	4.919	4.998	5.093
500	0.939	1.000	1.064	1.941	2.001	2.066	2.942	3.002	3.067	3.941	4.002	4.067	4.941	5.002	5.067
1000	0.960	1.001	1.040	1.961	2.001	2.041	2.961	3.001	3.041	3.961	4.001	4.041	4.961	5.001	5.041
							high degree of endogeneity								
100	0.651	0.957	1.250	1.731	1.991	2.283	2.737	3.002	3.288	3.749	4.009	4.287	4.749	5.011	5.286
250	0.804	0.970	1.149	1.842	1.989	2.161	2.846	2.994	3.161	3.849	3.994	4.163	4.850	4.994	5.166
500	0.870	0.989	1.115	1.888	2.001	2.120	2.891	3.002	3.124	3.891	4.003	4.123	4.893	5.003	5.122
1000	0.916	0.995	1.074	1.922	2.000	2.078	2.925	3.000	3.079	3.925	4.001	4.079	4.925	5.001	5.079

Model 2: Quantiles of the distribution of $\widehat{\delta}_{3,GMM}$

	$\delta_3 = 1$			$\delta_3 = 2$			$\delta_3 = 3$			$\delta_3 = 4$			$\delta_3 = 5$		
Quantile	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th	5th	50th	95th
Sample size							*low degree of endogeneity*								
100	0.3824	0.9801	1.4113	1.487	1.971	2.385	2.565	2.981	3.385	3.596	3.985	4.385	4.609	4.984	5.384
250	0.6874	0.996	1.2467	1.753	1.993	2.244	2.761	2.998	3.246	3.769	3.999	4.245	4.771	4.999	5.245
500	0.7937	0.9979	1.1659	1.825	1.996	2.163	2.829	3.000	3.164	3.827	4.000	4.166	4.826	5.001	5.164
1000	0.8757	0.9948	1.1169	1.881	1.997	2.115	2.881	3.000	3.111	3.882	3.999	4.115	4.881	5.000	5.116
							medium degree of endogeneity								
100	0.338	0.930	1.372	1.439	1.956	2.370	2.558	2.966	3.365	3.590	3.976	4.371	4.608	4.979	5.365
250	0.621	0.972	1.225	1.735	1.986	2.228	2.759	2.991	3.226	3.762	3.996	4.227	4.772	4.997	5.230
500	0.725	0.979	1.155	1.822	1.992	2.153	2.833	2.997	3.153	3.837	4.000	4.157	4.838	5.000	5.157
1000	0.823	0.979	1.108	1.881	1.991	2.112	2.886	2.994	3.116	3.884	3.994	4.116	4.887	4.994	5.116
							high degree of endogeneity								
100	0.396	0.898	1.309	1.423	1.930	2.329	2.572	2.973	3.341	3.620	3.984	4.343	4.630	4.991	5.353
250	0.619	0.938	1.181	1.719	1.970	2.202	2.769	2.985	3.205	3.789	3.992	4.204	4.790	4.993	5.211
500	0.707	0.952	1.123	1.819	1.988	2.140	2.852	2.996	3.138	3.856	4.000	4.142	4.856	5.001	5.145
1000	0.788	0.960	1.096	1.883	1.988	2.098	2.895	2.994	3.102	3.898	3.995	4.102	4.897	4.995	5.102

Notes: Model 1 refers to equation (I.A.1) with $\gamma = 2$, $\beta_1 = \beta_2 = 1$, and $\delta_1 = 0$. Model 2 refers to equation (I.A.5) with $\gamma = 2$, $\beta_1 = \beta_2 = \beta_3 = 1$ and $\delta_1 = \delta_2 = 0$. For Model 1 'low', 'medium', and 'high' corresponds to κ=0.05, 0.50, 0.95 in equation (I.A.4) and for Model 2 to $c_{qu} = 0.05, 0.25, 0.40$ in equation (I.A.7).

Table I.A.4: Nominal 90% confidence interval coverage for γ

	Model 1					Model 2				
	δ_2 1	2	3	4	5	δ_3 1	2	3	4	5
Sample size	*low degree of endogeneity*					*low degree of endogeneity*				
50	0.89	0.84	0.83	0.83	0.83	0.81	0.82	0.83	0.85	0.85
100	0.98	0.96	0.96	0.96	0.96	0.91	0.92	0.93	0.94	0.94
250	1.00	1.00	1.00	1.00	1.00	0.97	0.97	0.97	0.98	0.98
500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99	1.00
1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	medium degree of endogeneity					*medium degree of endogeneity*				
50	0.93	0.90	0.86	0.85	0.84	0.73	0.78	0.82	0.84	0.84
100	0.98	0.99	0.98	0.98	0.97	0.81	0.89	0.92	0.92	0.93
250	1.00	1.00	1.00	1.00	1.00	0.92	0.95	0.97	0.98	0.98
500	1.00	1.00	1.00	1.00	1.00	0.98	0.99	0.99	0.99	0.99
1000	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00
	high degree of endogeneity					*high degree of endogeneity*				
50	0.84	0.92	0.91	0.89	0.88	0.67	0.75	0.81	0.82	0.84
100	0.93	0.98	0.98	0.98	0.98	0.76	0.84	0.89	0.93	0.95
250	0.99	1.00	1.00	1.00	1.00	0.85	0.95	0.97	0.99	0.99
500	1.00	1.00	1.00	1.00	1.00	0.91	0.98	0.99	1.00	1.00
1000	1.00	1.00	1.00	1.00	1.00	0.94	1.00	1.00	1.00	1.00

Notes: Model 1 refers to equation (I.A.1) with $\gamma = 2$, $\beta_1 = \beta_2 = 1$, and $\delta_1 = 0$. Model 2 refers to equation (I.A.5) with $\gamma = 2$, $\beta_1 = \beta_2 = \beta_3 = 1$ and δ_1=$\delta_2 = 0$. For Model 1 'low', 'medium', and 'high' corresponds to κ=0.05, 0.50, 0.95 in equation (I.A.4) and for Model 2 to $c_{qu} = 0.05, 0.25, 0.40$ in equation (I.A.7).

Table I.A.5: Nominal 95% confidence interval coverage for the slope coefficients

		Model 1									
		Coverage for β_2					Coverage for δ_2				
	δ_2 / Sample size	1	2	3	4	5	1	2	3	4	5
		low degree of endogeneity									
	50	0.90	0.90	0.90	0.90	0.90	0.92	0.92	0.92	0.92	0.92
	100	0.92	0.92	0.92	0.92	0.92	0.94	0.94	0.94	0.94	0.94
	250	0.93	0.93	0.93	0.93	0.93	0.93	0.94	0.94	0.94	0.94
	500	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
	1000	0.95	0.95	0.95	0.95	0.95	0.96	0.96	0.96	0.96	0.96
		medium degree of endogeneity									
∞	50	0.77	0.79	0.79	0.79	0.79	0.88	0.90	0.90	0.90	0.90
	100	0.81	0.81	0.81	0.81	0.81	0.91	0.92	0.92	0.92	0.92
	250	0.85	0.85	0.85	0.85	0.86	0.94	0.95	0.95	0.95	0.95
	500	0.85	0.85	0.85	0.85	0.85	0.94	0.94	0.94	0.94	0.94
	1000	0.86	0.86	0.86	0.86	0.86	0.96	0.96	0.96	0.96	0.96
		high degree of endogeneity									
	50	0.74	0.76	0.78	0.78	0.78	0.84	0.88	0.89	0.90	0.90
	100	0.78	0.80	0.80	0.80	0.80	0.88	0.91	0.91	0.91	0.92
	250	0.83	0.84	0.84	0.85	0.84	0.93	0.95	0.95	0.95	0.95
	500	0.83	0.85	0.85	0.85	0.85	0.93	0.94	0.94	0.94	0.94
	1000	0.86	0.86	0.86	0.85	0.85	0.95	0.96	0.96	0.96	0.96

	Model 2									
	Coverage for β_3					Coverage for δ_3				
δ_3	1	2	3	4	5	1	2	3	4	5
	low degree of endogeneity									
50	0.80	0.84	0.87	0.88	0.89	0.80	0.84	0.87	0.88	0.89
100	0.83	0.88	0.91	0.92	0.92	0.83	0.88	0.91	0.92	0.92
250	0.91	0.93	0.93	0.94	0.94	0.91	0.93	0.93	0.94	0.94
500	0.91	0.93	0.94	0.93	0.93	0.91	0.93	0.94	0.93	0.93
1000	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
	medium degree of endogeneity									
50	0.78	0.80	0.83	0.85	0.86	0.79	0.83	0.86	0.88	0.89
100	0.81	0.84	0.89	0.89	0.90	0.81	0.86	0.90	0.91	0.92
250	0.82	0.89	0.91	0.91	0.91	0.86	0.92	0.94	0.94	0.94
500	0.83	0.92	0.93	0.93	0.93	0.85	0.92	0.93	0.93	0.93
1000	0.83	0.92	0.93	0.93	0.93	0.86	0.93	0.94	0.94	0.94
	high degree of endogeneity									
50	0.74	0.75	0.80	0.82	0.83	0.79	0.81	0.84	0.87	0.87
100	0.76	0.80	0.84	0.86	0.86	0.78	0.83	0.89	0.90	0.90
250	0.77	0.86	0.88	0.89	0.89	0.83	0.90	0.93	0.93	0.94
500	0.78	0.89	0.91	0.92	0.91	0.82	0.92	0.94	0.94	0.94
1000	0.76	0.89	0.90	0.91	0.90	0.79	0.93	0.94	0.94	0.94

Notes: Model 1 refers to equation (I.A.1) with $\gamma = 2$, $\beta_1 = \beta_2 = 1$, and $\delta_1 = 0$. Model 2 refers to equation (I.A.5) with $\gamma = 2$, $\beta_1 = \beta_2 = \beta_3 = 1$ and $\delta_1 = \delta_2 = 0$. For Model 1 'low', 'medium', and 'high' corresponds to κ=0.05, 0.50, 0.95 in equation (I.A.4) and for Model 2 to $c_{qu} = 0.05, 0.25, 0.40$ in equation (I.A.7).

2 Supplementary Proofs

Lemma I.A.1 *For some* $C < \infty$ *and* $\underline{\gamma} \leq \gamma' \leq \gamma \leq \bar{\gamma}$ *and* $r \leq 4$, *uniformly in* γ

$$Eh_i^r(\gamma,\gamma') \leq C|\gamma-\gamma'| \tag{I.A.8}$$

$$Ek_i^r(\gamma,\gamma') \leq C|\gamma-\gamma'| \tag{I.A.9}$$

Proof: Define $d_i(\gamma) = I_{\{q_i \leq \gamma\}}$ and $d_i^{\perp}(\gamma) = I_{\{q_i > \gamma\}}$. Define $h_i(\gamma,\gamma') = |\,(h_i(\gamma) - h_i(\gamma'))\, e_i|$ and $k_i(\gamma,\gamma') = |\,(h_i(\gamma) - h_i(\gamma'))\,|$. In the case of the STR model in equation (3.16) $h_i(\gamma) = (g_{xi}d_i(\gamma), \lambda_{1i}(\gamma)d_i(\gamma), \lambda_{2i}(\gamma)d_i(\gamma))'$ and thus $h_i(\gamma,\gamma')$ takes the form

$$h_i(\gamma,\gamma') = \begin{pmatrix} |g_{xi}e_i||d_i(\gamma)-d_i(\gamma')| \\ |\lambda_{1i}(\gamma)d_i(\gamma)e_i - \lambda_{1i}(\gamma')d_i(\gamma')e_i| \\ |\lambda_{2i}(\gamma)d_i(\gamma)e_i - \lambda_{2i}(\gamma')d_i(\gamma')e_i| \end{pmatrix}$$

From Lemma A.1 of Hansen (2000) we obtain that $E|g_{xi}e_i||d_i(\gamma)-d_i(\gamma')|^r \leq C_1|\gamma-\gamma'|$. Then it is sufficient to show that

$$E|\lambda_{1i}(\gamma)d_i(\gamma)e_i - \lambda_{1i}(\gamma')d_i(\gamma')e_i|^r \leq C_2|\gamma-\gamma'|$$

$$E|\lambda_{2i}(\gamma)d_i(\gamma)e_i - \lambda_{2i}(\gamma')d_i(\gamma')e_i|^r \leq C_3|\gamma-\gamma'|$$

Given that $\bar{\lambda}_{1i} = \sup_{\gamma\in\Gamma}|\lambda_{1i}(\gamma)|$, $\frac{d}{d\gamma}E(|\lambda_{1i}(\gamma)e_i|^r d_i(\gamma)) = E(|\lambda_{1i}(\gamma)e_i|^r|q = \gamma)f_q(\gamma) \leq [E(|\lambda_{1i}(\gamma)e_i|^4|q = \gamma)]^{r/4}f_q(\gamma) \leq [E(|\bar{\lambda}_{1i}e_i|^4|q = \gamma)]^{r/4}f_q(\gamma) \leq C^{r/4}\bar{f} \leq C_1$, where $C_1 = \max[1, C]$. Similarly, we obtain that $\frac{d}{d\gamma}E(|\lambda_{2i}(\gamma)e_i|^r d_i(\gamma)) \leq C_2$.

Then, by a first-order Taylor series expansion and Assumption 2.3, we show (I.A.8). Equation (I.A.9) follows analogously.

$$Eh_i^r(\gamma,\gamma') = \begin{pmatrix} E|g_{xi}e_i|^r d_i(\gamma) - E|g_{xi}e_i|^r d_i(\gamma')| \\ E|\lambda_{1i}(\gamma)e_i|^r d_i(\gamma) - E|\lambda_{1i}(\gamma')e_i|^r d_i(\gamma') \\ E|\lambda_{2i}(\gamma)e_i|^r d_i(\gamma) - E|\lambda_{2i}(\gamma')e_i|^r d_i(\gamma') \end{pmatrix} \leq C|\gamma - \gamma'|$$

where $C = (C_1', C_2', C_3')'$. ■

Lemma I.A.2 *Recall that $a_n = n^{1-2a}$. Let $H_i = (g_{xi}, e_i, \widehat{r}_i)$. Then,*

(i)

$$\sup_{\gamma\in\Gamma}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} H_i\widehat{r}_i' I(q_i \leq \gamma)\right| = O_p(1) \qquad \text{(I.A.10)}$$

(ii) For $0 < B < \infty$ such that for all $\varepsilon > 0$ and $\delta > 0$, there is a $\bar{v} < \infty$ and $\bar{n} < \infty$ such that for all $n \geq \bar{n}$:

$$P\left(\sup_{\bar{v}/a_n \leq |\gamma-\gamma_0| \leq B}\left|\frac{\sum_{i=1}^{n} H_i\widehat{r}_i'(I(q_i \leq \gamma) - I(q_i \leq \gamma_0))}{n^{1-\alpha}|\gamma - \gamma_0|}\right| \geq \delta\right) \leq \varepsilon \qquad \text{(I.A.11)}$$

(iii)

$$\sup_{|v|\leq\bar{v}} n^{-\alpha}\left|\sum_{i=1}^{n} H_i\widehat{r}_i'(I(q_i \leq \gamma_0 + v/a_n) - I(q_i \leq \gamma_0))\right| \xrightarrow{p} 0 \qquad \text{(I.A.12)}$$

Proof:

From Assumptions 1 and 2 and the linear first stage models (2.2) and (2.9) we get $\sqrt{n}(\Pi_x - \widehat{\Pi}_x) = O_p(1)$ and $\sqrt{n}(\pi_q - \widehat{\pi}_q) = O_p(1)$. Furthermore, given the continuity of the inverse Mills ratio terms we obtain $\sqrt{n}(\bar{\lambda}_{1i} - \widehat{\bar{\lambda}}_{1i}) = O_p(1)$, and $\sqrt{n}(\bar{\lambda}_{2i} - \widehat{\bar{\lambda}}_{2i}) = O_p(1)$. By letting $\breve{\Pi} = (\Pi_x, \bar{\lambda}_{1i}, \bar{\lambda}_{2i})'$ and $\widehat{\breve{\Pi}} = (\widehat{\Pi}_x, \widehat{\bar{\lambda}}_{1i}, \widehat{\bar{\lambda}}_{2i})'$ we obtain $\sqrt{n}(\breve{\Pi}_x - \widehat{\breve{\Pi}}_x) = O_p(1)$. Then using Lemma 1 of Caner and Hansen (2004) we establish (I.A.10), (I.A.11), and (I.A.12).

■

Lemma I.A.3 *Uniformly in $\gamma \in \Gamma$ as $n \longrightarrow \infty$*

$$\frac{1}{n}\widehat{X}^*(\gamma)'\widehat{X}^*(\gamma) = \frac{1}{n}\sum_{i=1}^{n}\widehat{x}_i^*(\gamma)\widehat{x}_i^*(\gamma)' \stackrel{p}{\longrightarrow} M(\gamma) \tag{I.A.13}$$

$$\frac{1}{n}\widehat{X}^*(\gamma_0)'G^*(\gamma_0) = \frac{1}{n}\sum_{i=1}^{n}\widehat{x}_i^*(\gamma_0)\widehat{x}_i^*(\gamma_0)' \stackrel{p}{\longrightarrow} M(\gamma_0) \tag{I.A.14}$$

$$\frac{1}{\sqrt{n}}\widehat{X}^*(\gamma)'\widetilde{e} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{x}_i^*(\gamma)\widetilde{e}_i = O_p(1) \tag{I.A.15}$$

Proof:

Let

$$M(\gamma) = \begin{pmatrix} M_\gamma(\gamma) & 0 \\ 0 & M_\perp(\gamma) \end{pmatrix}$$

where

$$M_\gamma(\gamma) = \begin{pmatrix} E\left(g_{xi}g_{xi}'I(q_i \le \gamma)\right) & E(\lambda_{1i}\left(\gamma\right)g_{xi}I(q_i \le \gamma)) & E(\lambda_{2i}\left(\gamma\right)g_{xi}I(q_i \le \gamma)) \\ E(\lambda_{1i}\left(\gamma\right)g_{xi}'I(q_i \le \gamma)) & E\left(\lambda_{1i}\left(\gamma\right)\right)^2 I(q_i \le \gamma) & E\lambda_{1i}\left(\gamma\right)\lambda_{2i}\left(\gamma\right)I(q_i \le \gamma) \\ E(\lambda_{2i}\left(\gamma\right)g_{xi}'I(q_i \le \gamma)) & E(\lambda_{2i}\left(\gamma\right)\lambda_{1i}\left(\gamma\right)I(q_i \le \gamma)) & E\left(\lambda_{2i}\left(\gamma\right)\right)^2 I(q_i \le \gamma) \end{pmatrix}$$

and

$$M_\perp(\gamma) = \begin{pmatrix} E\left(g_{xi}g_{xi}'I(q_i > \gamma)\right) & E(\lambda_{1i}\left(\gamma\right)g_{xi}I(q_i > \gamma)) & E(\lambda_{2i}\left(\gamma\right)g_{xi}I(q_i > \gamma)) \\ E(\lambda_{1i}\left(\gamma\right)g_{xi}'I(q_i > \gamma)) & E\left(\lambda_{1i}\left(\gamma\right)\right)^2 I(q_i > \gamma) & E\lambda_{1i}\left(\gamma\right)\lambda_{2i}\left(\gamma\right)I(q_i > \gamma) \\ E(\lambda_{2i}\left(\gamma\right)g_{xi}'I(q_i > \gamma)) & E(\lambda_{2i}\left(\gamma\right)\lambda_{1i}\left(\gamma\right)I(q_i > \gamma)) & E\left(\lambda_{2i}\left(\gamma\right)\right)^2 I(q_i > \gamma) \end{pmatrix}$$

To show (I.A.13) note that

$$\frac{1}{n}\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma)=\begin{pmatrix}\frac{1}{n}\widehat{X}_\gamma'\widehat{X}_\gamma & \frac{1}{n}\widehat{X}_\gamma'\widehat{\lambda}_{1\gamma}(\gamma) & \frac{1}{n}\widehat{X}_\gamma'\widehat{\lambda}_{2\gamma}(\gamma)\\ \frac{1}{n}\widehat{\lambda}_{1\gamma}(\gamma)'\widehat{X}_\gamma & \frac{1}{n}\widehat{\lambda}_{1\gamma}(\gamma)'\widehat{\lambda}_{1\gamma}(\gamma) & \frac{1}{n}\widehat{\lambda}_{1\gamma}(\gamma)'\widehat{\lambda}_{2\gamma}(\gamma)\\ \frac{1}{n}\widehat{\lambda}_{2\gamma}(\gamma)'\widehat{X}_\gamma & \frac{1}{n}\widehat{\lambda}_{2\gamma}(\gamma)'\widehat{\lambda}_{1\gamma}(\gamma) & \frac{1}{n}\widehat{\lambda}_{2\gamma}(\gamma)'\widehat{\lambda}_{2\gamma}(\gamma)\end{pmatrix}=$$

$$\begin{pmatrix}\frac{1}{n}\sum_i(\widehat{x}_i\widehat{x}_i'I(q_i\le\gamma)) & \frac{1}{n}\sum_i\widehat{\lambda}_{1i}(\gamma)\widehat{x}_iI(q_i\le\gamma) & \frac{1}{n}\sum_i\widehat{\lambda}_{2i}(\gamma)\widehat{x}_iI(q_i\le\gamma)\\ \frac{1}{n}\sum_i\widehat{\lambda}_{1i}(\gamma)\widehat{x}_i'I(q_i\le\gamma)) & \frac{1}{n}\sum_i(\widehat{\lambda}_{1i}(\gamma))^2I(q_i\le\gamma) & \frac{1}{n}\sum_i\widehat{\lambda}_{1i}(\gamma)\widehat{\lambda}_{2i}(\gamma)I(q_i\le\gamma)\\ \frac{1}{n}\sum_i\widehat{\lambda}_{2i}(\gamma)\widehat{x}_i'I(q_i\le\gamma)) & \frac{1}{n}\sum_i\widehat{\lambda}_{1i}(\gamma)\widehat{\lambda}_{2i}(\gamma)I(q_i\le\gamma) & \frac{1}{n}\sum_i(\widehat{\lambda}_{2i}(\gamma))^2I(q_i\le\gamma)\end{pmatrix}.$$

From (I.A.10) and Lemma 1 of Hansen (1996) we get

$$\begin{aligned}\frac{1}{n}\sum_i(\widehat{x}_i\widehat{x}_i'I(q_i\le\gamma)) = {} & \tfrac{1}{n}\textstyle\sum_i g_{xi}g_{xi}'I(q_i\le\gamma)-\tfrac{1}{n}\sum_i g_{xi}\widehat{r}_{xi}'I(q_i\le\gamma)\\ & -\tfrac{1}{n}\textstyle\sum_i\widehat{r}_{xi}g_{xi}'I(q_i\le\gamma)+\tfrac{1}{n}\sum_i\widehat{r}_{xi}\widehat{r}_{xi}'I(q_i\le\gamma)\\ & \xrightarrow{p} E\left(g_{xi}g_{xi}'I(q_i\le\gamma)\right).\end{aligned}$$

Additionally, for j=1,2 we have

$$\begin{aligned}\frac{1}{n}\sum_i\widehat{\lambda}_{ji}(\gamma)\widehat{x}_iI(q_i\le\gamma) = {} & \tfrac{1}{n}\textstyle\sum_i\lambda_{ji}(\gamma)g_{xi}'I(q_i\le\gamma)-\tfrac{1}{n}\sum_i\widehat{r}_{xi}\lambda_{ji}(\gamma)I(q_i\le\gamma)\\ & -\tfrac{1}{n}\textstyle\sum_i\widehat{r}_{\lambda_{ji}}\lambda_{ji}(\gamma)I(q_i\le\gamma)+\tfrac{1}{n}\sum_i\widehat{r}_{\lambda_{ji}}\widehat{r}_{xi}I(q_i\le\gamma)\\ & \xrightarrow{p} E\lambda_{ji}(\gamma)g_{xi}'I(q_i\le\gamma),\end{aligned}$$

$$\begin{aligned}\tfrac{1}{n}\textstyle\sum_i(\widehat{\lambda}_{ji}(\gamma))^2I(q_i\le\gamma) = {} & \tfrac{1}{n}\textstyle\sum_i(\lambda_{ji}(\gamma))^2I(q_i\le\gamma)-\tfrac{2}{n}\sum_i\widehat{r}_{\lambda_{ji}}\lambda_{ji}(\gamma)I(q_i\le\gamma)\\ & +\tfrac{1}{n}\textstyle\sum_i(\widehat{r}_{\lambda_{ji}})^2I(q_i\le\gamma)\\ & \xrightarrow{p} E(\lambda_{ji}(\gamma))^2I(q_i\le\gamma)),\end{aligned}$$

and

$$
\begin{aligned}
\tfrac{1}{n}\sum_i \widehat{\lambda}_{1i}(\gamma)\widehat{\lambda}_{2i}(\gamma)I(q_i \le \gamma) = \; & \tfrac{1}{n}\textstyle\sum_i \lambda_{1i}(\gamma)\lambda_{2i}(\gamma)I(q_i \le \gamma) - \tfrac{1}{n}\sum_i \lambda_{1i}(\gamma)\widehat{r}_{\lambda_{1i}}I(q_i \le \gamma) \\
& -\tfrac{1}{n}\textstyle\sum_i \lambda_{2i}(\gamma)\widehat{r}_{\lambda_{2i}}I(q_i \le \gamma) + \tfrac{1}{n}\sum_i \widehat{r}_{\lambda_{1i}}\widehat{r}_{\lambda_{2i}}I(q_i \le \gamma) \\
& \xrightarrow{p} E(\lambda_{1i}(\gamma)\lambda_{2i}(\gamma)I(q_i \le \gamma))
\end{aligned}
$$

Then, uniformly in $\gamma \in \Gamma$, we obtain $\frac{1}{n}\widehat{X}_{\gamma}(\gamma)'\widehat{X}_{\gamma}(\gamma) \xrightarrow{p} M_{\gamma}(\gamma)$. Similarly, we can show that $\frac{1}{n}\widehat{X}_{\perp}(\gamma)'\widehat{X}_{\perp}(\gamma) \xrightarrow{p} M_{\perp}(\gamma)$. Hence,

$$
\frac{1}{n}\widehat{X}^{*}(\gamma)'\widehat{X}^{*}(\gamma) \xrightarrow{p} M(\gamma) = \begin{pmatrix} M_{\gamma}(\gamma) & 0 \\ 0 & M_{\perp}(\gamma) \end{pmatrix}
$$

Equation (I.A.14) follows similarly.

To show (I.A.15) note that

$$
\frac{1}{\sqrt{n}}\widehat{X}_{\gamma}(\gamma)'\widetilde{e} = \begin{pmatrix} \frac{1}{\sqrt{n}}\sum\limits_i \widehat{x}_i\widetilde{e}_i I(q_i \le \gamma) \\ \frac{1}{\sqrt{n}}\sum\limits_i \widehat{\lambda}_{1i}(\gamma)\,\widetilde{e}_i I(q_i \le \gamma) \\ \frac{1}{\sqrt{n}}\sum\limits_i \widehat{\lambda}_{2i}(\gamma)\,\widetilde{e}_i I(q_i \le \gamma) \end{pmatrix}
$$

Using I.A.10, Lemma 2 of Caner and Hansen (2004), Lemma A.4 of Hansen (2000) we get

$$
\frac{1}{\sqrt{n}}\sum_i (\widehat{x}_i\widetilde{e}_i I(q_i \le \gamma)) = O_p(1)
$$

and for $j = 1, 2$

$$\frac{1}{\sqrt{n}}\sum_i \widehat{\lambda}_{ji}(\gamma)\widetilde{e}_i I(q_i \le \gamma) = \frac{1}{\sqrt{n}}\sum_i \lambda_{ji}(\gamma)e_i I(q_i \le \gamma) + \frac{1}{\sqrt{n}}\sum_i \widehat{r}_{\lambda_{ji}} e_i I(q_i \le \gamma) + \frac{1}{\sqrt{n}}\sum_i \lambda_{ji}(\gamma)\widehat{r}'_{xi}\beta_2 I(q_i \le \gamma) - \frac{1}{\sqrt{n}}\sum_i \widehat{r}_{\lambda_{ji}}\widehat{r}'_{xi}\beta_2 I(q_i \le \gamma) = O_p(1).$$

Therefore, $\frac{1}{\sqrt{n}}\widehat{X}_\gamma(\gamma)'\widetilde{e} = O_p(1)$. Similarly, we can show that $\frac{1}{\sqrt{n}}\widehat{X}_\perp(\gamma)'\widetilde{e} = O_p(1)$. Hence,

$$\frac{1}{\sqrt{n}}\widehat{X}^*(\gamma)'\widetilde{e} = O_p(1).$$

■

Lemma I.A.4 $a_n(\widehat{\gamma} - \gamma_0) = O_p(1)$.

Proof: First we establish that the unconstrained and the constrained problems share the same rate of convergence by exploiting the relationship between the constrained and unconstrained problems

$$S^R(\gamma) = S_n^U(\gamma) + (\vartheta - R'\beta)'(R'(\widehat{X}^*(\gamma)'\widehat{X}^*(\gamma))^{-1}R)^{-1}(\vartheta - R'\beta)$$

Then the proof proceeds in steps.

Let $\widehat{\beta}_\gamma$ denote the estimated coefficients of $\widehat{\beta}$ associated with the partitioned regressor matrix $\widehat{X}^*(\gamma)$, the unconstrained sum of squared residuals $S_n^U(\gamma)$, and threshold value γ. Let $\widehat{\beta}_{\gamma_0}$ denote the estimated coefficients of $\widehat{\beta}$ associated with the partitioned regressor matrix $\widehat{X}^*(\gamma_0)$, the unconstrained sum of squared residuals $S_n^U(\gamma_0)$, and threshold value γ_0. We also use the subscript 0 to denote the parameter at the true value.

Using Lemma A.2 of Perron and Qu (2006) we can deduce that

$$(\widehat{X}^*(\gamma)'\widehat{X}^*(\gamma))^{-1} = (\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1} + O_p(\frac{|\gamma - \gamma_0|}{n^2}) \qquad \text{(I.A.16)}$$

and

$$(R'(\widehat{X}^*(\gamma)'\widehat{X}^*(\gamma))^{-1}R)^{-1} = (R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1} + O_p(|\gamma - \gamma_0|). \qquad \text{(I.A.17)}$$

Consider

$$\begin{aligned}
\widehat{\beta}_\Delta &= \widehat{\beta}_\gamma - \widehat{\beta}_{\gamma_0} \\
&= (\widehat{X}^*(\gamma)'\widehat{X}^*(\gamma))^{-1}\widehat{X}^*(\gamma)'(G^*(\gamma_0)\beta_0 + e) - (\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}\widehat{X}^*(\gamma_0)'(G^*(\gamma_0)\beta_0 + e) \\
&= (\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}((\widehat{X}^*(\gamma) - \widehat{X}^*(\gamma_0))'G^*(\gamma_0)\beta_0 \\
&\quad + (\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}((\widehat{X}^*(\gamma) - \widehat{X}^*(\gamma_0))'e + |\gamma - \gamma_0|O_p(\tfrac{1}{n}) \\
&= (\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1/2}A_n
\end{aligned}$$

with

$$\begin{aligned}
A_n &= \widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1/2}(\widehat{X}^*(\gamma) - \widehat{X}^*(\gamma_0))'G^*(\gamma_0)\beta_0 \\
&\quad + (\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1/2}(\widehat{X}^*(\gamma)' - \widehat{X}^*(\gamma_0)')e + |\gamma - \gamma_0|O_p(\tfrac{1}{\sqrt{n}}) = |\gamma - \gamma_0|O_p(n^{-1/2}),
\end{aligned}$$

where the first equality uses (I.A.16). To get the second equality note that

$(\widehat{X}^*(\gamma) - \widehat{X}^*(\gamma_0))'G^*(\gamma_0) = |\gamma - \gamma_0|O_p(1)$,

$\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1/2}(\widehat{X}^*(\gamma) - \widehat{X}^*(\gamma_0))'G^*(\gamma_0)\beta_0 = |\gamma - \gamma_0|O_p(\frac{1}{\sqrt{n}})$, and

$(\widehat{X}^*(\gamma) - \widehat{X}^*(\gamma_0))'e = |\gamma - \gamma_0|O_p(1)$.

Therefore, $\widehat{\beta}_\Delta = |\gamma - \gamma_0|O_p(n^{-1})$.

Furthermore, note that $\widehat{\beta}_{\Delta}R{=}|\gamma-\gamma_0|O_p(n^{-1})$ and $(\vartheta{-}R'\beta)'=|\gamma-\gamma_0|O_p(n^{-1})$. Then,

$$S_n^R(\gamma)-S_n^R(\gamma_0)=[S_n^U(\gamma)-S_n^U(\gamma_0)]+[(\vartheta{-}R'\widehat{\beta}_{\gamma})'(R'(\widehat{X}^*(\gamma)'\widehat{X}^*(\gamma))^{-1}R)^{-1}(\vartheta{-}R'\widehat{\beta}_{\gamma})$$

$$-(\vartheta{-}R'\widehat{\beta}_{\gamma_0})'(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}(\vartheta{-}R'\beta_{\gamma_0})]$$

$$=[S_n^U(\gamma)-S_n^U(\gamma_0)]+[(\vartheta{-}R'\widehat{\beta}_{\gamma})'(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}(\vartheta{-}R'\widehat{\beta}_{\gamma})$$

$$-(\vartheta{-}R'\widehat{\beta}_{\gamma_0})'(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}(\vartheta{-}R'\widehat{\beta}_{\gamma_0})]+(\gamma-\gamma_0)^2O_p(n^{-1})$$

$$=[S_n^U(\gamma)-S_n^U(\gamma_0)]+(\widehat{\beta}_{\gamma_0}+\widehat{\beta}_{\Delta})'R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}R'(\widehat{\beta}_{\gamma_0}+\widehat{\beta}_{\Delta})$$

$$-\widehat{\beta}'_{\gamma_0}R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}R'\widehat{\beta}_{\gamma_0}-2\vartheta'R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}R'(\widehat{\beta}_{\gamma}-\widehat{\beta}_{\gamma_0})$$

$$+|\gamma-\gamma_0|^2O_p(n^{-1})$$

$$=[S_n^U(\gamma)-S_n^U(\gamma_0)]+2\widehat{\beta}'_{\Delta}R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}\widehat{X}^*(\gamma_0)'e$$

$$+\widehat{\beta}'_{\Delta}R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}R'\widehat{\beta}_{\Delta}$$

$$+2\widehat{\beta}'_{\Delta}R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}\widehat{X}^*(\gamma_0)'G^*(\gamma_0)\beta_0-\vartheta)$$

$$+|\gamma-\gamma_0|^2O_p(n^{-1})$$

$$=[S_n^U(\gamma)-S_n^U(\gamma_0)]+2\widehat{\beta}'_{\Delta}R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}\widehat{X}^*(\gamma_0)'e$$

$$+\widehat{\beta}'_{\Delta}R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}R'\widehat{\beta}_{\Delta}$$

$$+2\widehat{\beta}'_{\Delta}R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}\widehat{X}^*(\gamma_0)'(G^*(\gamma')-\widehat{X}^*(\gamma_0))\beta_0$$

$$+|\gamma-\gamma_0|^2O_p(n^{-1}).$$

Now consider the second term divided by $|\gamma-\gamma_0|$

$$||2\widehat{\beta}'_{\Delta}R(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R)^{-1}R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}\widehat{X}^*(\gamma_0)'e||/n^{2\alpha-1}(\gamma-\gamma_0)$$

$$=||A'_n((\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1/2}R\left(R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1}R\right)^{-1}R'(\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1/2})$$

$$\cdot((\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1/2}e)||/n^{2\alpha-1}(\gamma-\gamma_0)$$

$$\leq ||A_n'||||((\widehat{X}^*(\gamma_0)'\widehat{X}^*(\gamma_0))^{-1/2}e)||/n^{2\alpha-1}(\gamma-\gamma_0) = o_p(1)$$

Note that the third term is nonnegative and divided by $n^{2\alpha-1}(\gamma-\gamma_0)$ is also $o_p(1)$. The key object in the fourth term is $(G^*(\gamma') - \widehat{X}^*(\gamma_0))\beta_0$ which is also $o_p(1)$ when it is divided by $n^{2\alpha-1}(\gamma-\gamma_0)$.

Therefore,

$$\frac{S_n^R(\gamma) - S_n^R(\gamma_0)}{n^{2\alpha-1}(\gamma-\gamma_0)} \geq \frac{S_n^U(\gamma) - S_n^U(\gamma_0)}{n^{2\alpha-1}(\gamma-\gamma_0)} + o_p(1) \tag{I.A.18}$$

We can now focus on the unconstrained problem since the rates of convergence for the constrained and unconstrained problems are the same. Let the constants B, d, t be defined as $B > 0$, $0 < d < \infty$, $0 < t < \infty$. Let $\breve{M} = \sup_{|\gamma-\gamma_0|\leq B} |M_\gamma(\gamma)^{-1}|$ and $\breve{D} = \sup_{|\gamma-\gamma_0|\leq B} |D(\gamma)f(\gamma)|$. Define $\breve{M}^* = \breve{M} + \breve{M}^2\tau$. Fix $\epsilon > 0$, pick τ and reduce B so that

$$\tau + 3k\breve{M}^*(\bar{D}C + 2\tau)(1 + \breve{M}^*(M_0(\gamma_0) + \tau)) \leq d/12 \tag{I.A.19a}$$

$$\tau(M_0(\gamma_0) + \tau)\breve{M}^*(1 + 3t\breve{M}^*) \leq d/12 \tag{I.A.19b}$$

$$\tau^2\breve{M}^*(2 + 3t\breve{M}^*) \leq d/12 \tag{I.A.19c}$$

Without loss of generality assume $\tau \leq t$ and define $\Delta_i(\gamma) = I(q \leq \gamma) - I(q \leq \gamma_0)$.

By Lemma A.7 of Hansen (2000) and (I.A.11), there exist sufficiently large $\bar{\upsilon} = \upsilon(\epsilon) < \infty$ and $\bar{n} = \bar{n}(\epsilon) < \infty$ that for all $n \geq \bar{n}$, the following events given by equations (I.A.20a)-(I.A.20d) hold jointly with probability exceeding $1 - \epsilon/2$.

$$\sup_{\bar{\upsilon}/a_n \leq |\gamma-\gamma_0| \leq B} \frac{\sum |g_i(\gamma)|^2 \Delta_i(\gamma)}{n(\gamma-\gamma_0)} \leq 13d/12 \tag{I.A.20a}$$

$$\inf_{\bar{\upsilon}/a_n \leq |\gamma-\gamma_0| \leq B} \frac{\sum (c'g_i(\gamma))^2 \Delta_i(\gamma)}{n(\gamma-\gamma_0)} \geq 11d/12 \tag{I.A.20b}$$

$$\sup_{\bar{v}/a_n \le |\gamma-\gamma_0| \le B} \left| -\frac{\sum g_i(\gamma)\widehat{r}_i'\Delta_i(\gamma)}{n(\gamma-\gamma_0)} - \frac{\sum \widehat{r}_i g_i(\gamma)'\Delta_i(\gamma)}{n(\gamma-\gamma_0)} + \frac{\sum \widehat{r}_i\widehat{r}_i\Delta_i(\gamma)}{n(\gamma-\gamma_0)} \right| \le \tau \tag{I.A.20c}$$

$$\sup_{\bar{v}/a_n \le |\gamma-\gamma_0| \le B} \left| \frac{\sum (g_i(\gamma)-\widehat{r}_i)\widetilde{e}_i\Delta_i(\gamma)}{n^{1-\alpha}(\gamma-\gamma_0)} \right| \le \tau \tag{I.A.20d}$$

Additionally, by Proposition 1 the following events given by equations (I.A.21a)-(I.A.21e) hold jointly with probability exceeding $1-\epsilon$.

$$|\widehat{\gamma}-\gamma_0| \le B \tag{I.A.21a}$$

$$\sup_{\gamma\in\Gamma} \left| \frac{1}{n}\widehat{X}_\gamma(\gamma)\widehat{X}_\gamma(\gamma) - M_\gamma(\gamma) \right| \le \tau \tag{I.A.21b}$$

$$\sup_{\gamma\in\Gamma} \left| \left| \frac{1}{n}\widehat{X}_\gamma(\gamma)\widehat{X}_\gamma(\gamma) \right| - |M_\gamma(\gamma)| \right| \le \tau \tag{I.A.21c}$$

$$\left| \left| \frac{1}{n}\widehat{X}_0(\gamma_0)'G_0(\gamma_0) \right| - M_0(\gamma_0) \right| \le \tau \tag{I.A.21d}$$

$$\sup_{\gamma\in\Gamma} \left| \frac{1}{n^{1-\alpha}}\widehat{X}_\gamma(\gamma)\widetilde{e} \right| \le \tau \tag{I.A.21e}$$

Hence, the events (I.A.20a)-(I.A.21e) hold jointly with probability exceeding $1-\epsilon$.

We calculate

$$
\begin{aligned}
&G_0(\gamma_0)'(P^*(\gamma_0)-P^*(\gamma))G_0(\gamma_0)\\
&\qquad=(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma)-\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))\\
&\qquad-(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma)-\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))(I_l-(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma))^{-1}\widehat{X}_0(\gamma_0)'G_0(\gamma_0)\\
&\qquad-(I_l-G_0(\gamma_0)'\widehat{X}_0(\gamma_0)(\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))^{-1})\\
&\qquad\times(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma)-\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma))^{-1}\widehat{X}_0(\gamma_0)'G_0(\gamma_0)
\end{aligned}
\tag{I.A.22}
$$

$$
\begin{aligned}
&G_0(\gamma_0)'(P^*(\gamma_0)-P^*(\gamma))\widetilde{e}\\
&\quad=G_0(\gamma_0)'\widehat{X}_0(\gamma_0)(\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))^{-1}(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma)-\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))\\
&\quad\times(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma))^{-1}\widehat{X}_0(\gamma_0)'\widetilde{e}\\
&\quad-G_0(\gamma_0)'\widehat{X}_0(\gamma_0)(\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))^{-1}(\widehat{X}_\gamma(\gamma)'\widetilde{e}-\widehat{X}_0(\gamma_0)'\widetilde{e})
\end{aligned}
\tag{I.A.23}
$$

$$
\widetilde{e}'(P^*(\gamma_0)-P^*(\gamma))\widetilde{e}=\widetilde{e}'(P_0(\gamma_0)-P_\gamma(\gamma))\widetilde{e}+\widetilde{e}'(P_\perp(\gamma_0)-P_\perp(\gamma))\widetilde{e}
\tag{I.A.24}
$$

The first term of (I.A.24) is calculated as follows

$$
\begin{aligned}
&\widetilde{e}'(P_0(\gamma_0)-P_\gamma(\gamma))\widetilde{e}\\
&\quad=\widetilde{e}'\widehat{X}_0(\gamma_0)(\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))^{-1}(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma)-\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))\\
&\quad\times(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma))^{-1}\widehat{X}_0(\gamma_0)'\widetilde{e}\\
&\quad-2\widetilde{e}'\widehat{X}_0(\gamma_0)(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma))^{-1}(\widehat{X}_\gamma(\gamma)'\widetilde{e}-\widehat{X}_0(\gamma_0)'\widetilde{e})
\end{aligned}
\tag{I.A.25}
$$

The second term of (I.A.24) can be calculated similarly. Using definitions in Lemma I.A.3

we calculate the following decomposition

$$\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma) - \widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0) = \sum_{i=1}^n g_i(\gamma)g_i(\gamma)'\Delta_i(\gamma) - \sum_{i=1}^n g_i(\gamma)\widehat{r}_i'\Delta_i(\gamma) - \sum_{i=1}^n g_i(\gamma)\Delta_i(\gamma)'\widehat{r}_i + \sum_{i=1}^n \widehat{r}_i\widehat{r}_i'\Delta_i(\gamma) \tag{I.A.26}$$

Then, by applying Lemma 4 of Caner and Hansen (2004) and using equations (I.A.22), (I.A.23), (I.A.24), (I.A.25), and (I.A.26) we get that (I.A.20a)-(I.A.20d) and (I.A.21a)-(I.A.21e) imply the following:

$$\inf_{\bar{\varepsilon}/a_n \le |\gamma-\gamma_0| \le B} c'\left(\frac{G_0(\gamma_0)'(P^*(\gamma_0) - P^*(\gamma))G_0(\gamma_0)}{n(\gamma-\gamma_0)}\right)c \ge 5d/6 \tag{I.A.27}$$

$$\sup_{\bar{\varepsilon}/a_n \le |\gamma-\gamma_0| \le B} \left|\frac{c'G_0'(\gamma_0)(P^*(\gamma_0) - P^*(\gamma))\widetilde{e}}{n^{1-\alpha}(\gamma-\gamma_0)}\right| \le d/12 \tag{I.A.28}$$

$$\sup_{\bar{\varepsilon}/a_n \le |\gamma-\gamma_0| \le B} \left|\frac{\widetilde{e}'(P^*(\gamma_0) - P^*(\gamma))\widetilde{e}}{n^{1-2\alpha}(\gamma-\gamma_0)}\right| \le d/6 \tag{I.A.29}$$

where $d \in (0, \infty)$.

Using equation (A.3) of the Appendix in conjunction with the above inequalities in (I.A.27), (I.A.28), and (I.A.27) we can then write $S_n^U(\gamma) - S_n^U(\gamma_0)$ for $\bar{\upsilon}/a_n \le |\gamma-\gamma_0| \le C$ as

$$\frac{S_n^U(\gamma) - S_n^U(\gamma_0)}{n^{1-2\alpha}(\gamma-\gamma_0)} = \frac{\widetilde{e}'(P^*(\gamma_0) - P^*(\gamma))\widetilde{e}}{n^{1-2\alpha}(\gamma-\gamma_0)} + 2\frac{\widetilde{e}'(P^*(\gamma_0) - P^*(\gamma))G_0(\gamma_0)c}{n^{1-\alpha}(\gamma-\gamma_0)}$$

$$+\frac{c'G_0(\gamma_0)'(P^*(\gamma_0)-P^*(\gamma))G_0(\gamma_0)c}{n(\gamma-\gamma_0)}$$

$$\geq d/2 \tag{I.A.30a}$$

Since $S_n(\widehat{\gamma}) \leq S_n(\gamma_0)$, the joint events in equations (I.A.20a)-(I.A.20d) and (I.A.21a)-(I.A.21e) imply that $|\widehat{\gamma}-\gamma_0| \leq \bar{\varepsilon}/a_n$. Moreover, since (I.A.20a)-(I.A.20d) and (I.A.21a)-(I.A.21e) hold jointly with probability more than $1-\epsilon$ for all $n \geq \bar{n}$, we have that $P(n^{1-2\alpha}|\widehat{\gamma}-\gamma_0| > \bar{\varepsilon}) \leq \epsilon$ for $n \geq \bar{n}$. Hence, $a_n(\widehat{\gamma}-\gamma_0) = O_p(1)$.

■

Lemma I.A.5 *On* $[-\bar{\upsilon}, \bar{\upsilon}]$,

$$Q_n(\upsilon) = S_n^U(\gamma_0) - S_n^U(\gamma_0+\upsilon/a_n) \Rightarrow \mathcal{Q}(\upsilon)$$

where

$$\mathcal{Q}(\upsilon) = \begin{cases} -\mu|\upsilon| + 2\zeta_1^{1/2}\mathcal{W}_1(\upsilon), & \textit{uniformly on } \upsilon \in [-\bar{\upsilon}, 0] \\ -\mu|\upsilon| + 2\zeta_2^{1/2}\mathcal{W}_2(\upsilon), & \textit{uniformly on } \upsilon \in [0, \bar{\upsilon}] \end{cases}$$

with $\mu = c'Dcf$ *and* $\zeta_i = c'\Omega_i cf$, *for* $i = 1, 2$.

Proof: Our proof strategy follows Caner and Hansen (2004). Let us first reparameterize all functions of γ as functions of υ. For example, $X_\upsilon(\upsilon) = \widehat{X}_{\gamma_0+\upsilon/a_n}(\gamma_0+\upsilon/a_n)$, $P_\upsilon^*(\upsilon) = P^*_{\gamma_0+\upsilon/a_n}(\gamma_0+\upsilon/a_n)$ and for $\Delta_i(\gamma) = I(q_i \leq \gamma) - I(q_i \leq \gamma_0)$ we have $\Delta_i(\upsilon) = \Delta_i(\gamma_0+\upsilon/a_n)$. Then, using (A.3) of the Appendix we obtain

$$\begin{aligned} Q_n(\upsilon) &= S_n^U(\gamma_0) - S_n^U(\gamma_0+\upsilon/a_n) \\ &= (n^{-\alpha}c'G_0(\gamma_0)' + \widetilde{e}')P^*(\gamma_0)(G_0(\gamma_0)cn^{-\alpha} + \widetilde{e}) - (n^{-\alpha}c'G(\gamma_0)' + \widetilde{e}')P^*(\upsilon)(G_0(\gamma_0)cn^{-\alpha} + \widetilde{e}) \\ &= n^{-2a}c'G_0(\gamma_0)'(P^*(\gamma_0) - P^*(\upsilon))G_0(\gamma_0)c \qquad \text{(i)} \end{aligned}$$

$$+ n^{-a}c'G_0(\gamma_0)'(P^*(\gamma_0) - P^*(\upsilon))\widetilde{e} \qquad \text{(ii)}$$

$$+ \widetilde{e}'(P^*(\gamma_0) - P^*(\upsilon))\widetilde{e} \qquad \text{(iii)} \qquad \text{(I.A.31a)}$$

We proceed by studying the behavior of (i)-(iii).

(i) First, we establish that

$$n^{-2\alpha} \sup \left| \widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma) - \widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0) \right| = O_p(1) \qquad \text{(I.A.32)}$$

Using equations (I.A.26), (I.A.12), and Lemma A.10 of Hansen (2000) we get

$$\begin{aligned} n^{-2\alpha} \left| \widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon) - \widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0) \right| \quad &\leq \quad n^{-2\alpha} \sum_{i=1}^{n} |g_i(\upsilon)|^2 \Delta_i(\upsilon) + 2n^{-2\alpha} \left| \sum_{i=1}^{n} g_i(\upsilon)\widetilde{e}_i'\Delta_i(\upsilon) \right| \\ &\quad + n^{-2\alpha} \left| \sum_{i=1}^{n} \widetilde{e}_i\widetilde{e}_i'\Delta_i(\upsilon) \right| \\ &\Rightarrow \quad (|D_1 f|\,|\upsilon|)I(\upsilon < 0) + (|D_2 f|\,|\upsilon|)I(\upsilon > 0) \end{aligned}$$

This demonstrates equation (I.A.32).

Second, we obtain from equation (I.A.13) of Lemma I.A.3 that

$$\frac{1}{n}\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon) \Rightarrow M(\gamma_0) \qquad \text{(I.A.33)}$$

Then using equations (I.A.22), (I.A.32), (I.A.33), equation (I.A.12) of Lemma I.A.2, and Lemma I.A.3, we get that

$$n^{-2a}c'G_0(\gamma_0)'(P^*_{\gamma_0}(\gamma_0) - P^*_\upsilon(\upsilon))G_0(\gamma_0)c$$

$$= n^{-2\alpha}c'(\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon) - \widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))c$$

$$-n^{-2\alpha}c'(\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon)-\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))(I_l-(\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon))^{-1}(\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0)))c$$

$$-c'(I_m-G_0(\gamma_0)'\widehat{X}_0(\gamma_0)(\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))^{-1})$$

$$\times n^{-2\alpha}(\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon)-\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))(\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon))^{-1}\widehat{X}_0(\gamma_0))'G_0(\gamma_0)c$$

$$=n^{-2\alpha}\sum_{i=1}^{n}(c'g_i(\upsilon))^2\Delta_i(\upsilon)+o_p(1),\quad \text{uniformly in } \upsilon\in[-\bar{\upsilon},\bar{\upsilon}]. \tag{I.A.34}$$

Hence, using equation (I.A.34) and Lemma A.10 of Hansen (2000), uniformly in $\upsilon\in[-\bar{\upsilon},\bar{\upsilon}]$, we obtain that term (i) of $Q_n(\upsilon)$

$$n^{-2a}c'G_0(\gamma_0)'(P^*_{\gamma_0}(\gamma_0)-P^*_\upsilon(\upsilon))G_0(\gamma_0)c\Rightarrow\mu|\upsilon|. \tag{I.A.35}$$

(ii) First, note that using Lemma (I.A.3) and equation (I.A.33)

$$n^{\alpha}(\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon))^{-1}\widehat{X}_0(\gamma_0)\widetilde{e}=(n^{-1}(\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon))^{-1}n^{-1(1-\alpha)}\widehat{X}_0(\gamma_0)'\widetilde{e}=o_p(1) \tag{I.A.36}$$

Second, let $\mathcal{B}_1(\upsilon)$ and $\mathcal{B}_2(\upsilon)$ be independent one-sided vector Brownian motions with covariance matrices $\Omega_1 f$ and $\Omega_2 f$, respectively. Then, by equation (I.A.12) and Lemma A.11 of Hansen (2000) we have

$$\begin{aligned}
n^{-\alpha}(\widehat{X}_\upsilon(\upsilon)'\widetilde{e}-\widehat{X}_0(\gamma_0)'\widetilde{e})&\\
&=n^{-\alpha}\sum_{i=1}^{n}\widehat{g}_i(\upsilon)\widetilde{e}_i\Delta_i(\upsilon)\\
&=n^{-\alpha}\sum_{i=1}^{n}\widehat{g}_i(\upsilon)\widehat{r}_i'\beta^*\Delta_i(\upsilon)+n^{-\alpha}\sum_{i=1}^{n}g_i(\upsilon)e_i\Delta_i(\upsilon)-n^{-\alpha}\sum_{i=1}^{n}\widehat{r}_ie_i\Delta_i(\upsilon)\\
&=n^{-\alpha}\sum_{i=1}^{n}g_i(\upsilon)e_i\Delta_i(\upsilon)+o_p(1)
\end{aligned}$$

$$\Rightarrow \begin{cases} \mathcal{B}_1(\upsilon), & \text{uniformly on } \upsilon \in [-\bar{\upsilon}, 0] \\ \mathcal{B}_2(\upsilon), & \text{uniformly on } \upsilon \in [0, \bar{\upsilon}] \end{cases} \tag{I.A.37}$$

Therefore, using equations (I.A.23), (I.A.32), (I.A.33), (I.A.36) and (I.A.37) we obtain

$$\begin{aligned} & n^{-a} c' G_0(\gamma_0)'(P^*(\gamma_0) - P^*(\upsilon))\widetilde{e} \\ & \quad = G_0(\gamma_0)'\widehat{X}_0(\gamma_0)(\widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))^{-1} \\ & \quad \times n^{-2\alpha}(\widehat{X}_\gamma(\gamma)'\widehat{X}_\gamma(\gamma) - \widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0))n^{\alpha}(\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon))^{-1}\widehat{X}_0(\gamma_0)'\widetilde{e} \\ & \quad - G_0(\gamma_0)'\widehat{X}_0(\gamma)(\widehat{X}_\upsilon(\upsilon)'\widehat{X}_\upsilon(\upsilon))^{-1} n^{-\alpha}(\widehat{X}_\upsilon(\upsilon)'\widetilde{e} - \widehat{X}_0(\gamma_0)'\widetilde{e}) \\ & \quad = -n^{-\alpha}(\widehat{X}_\upsilon(\upsilon)'\widetilde{e} - \widehat{X}_0(\gamma_0)'\widetilde{e}) + o_p(1) \\ & \quad \Rightarrow \begin{cases} \mathcal{B}_1(\upsilon), & \text{uniformly on } \upsilon \in [-\bar{\upsilon}, 0] \\ \mathcal{B}_2(\upsilon), & \text{uniformly on } \upsilon \in [0, \bar{\upsilon}] \end{cases} \end{aligned}$$

Hence,

$$n^{-a} c' G_0(\gamma_0)'(P^*(\gamma_0) - P^*(\upsilon))\widetilde{e} \Rightarrow \begin{cases} 2\zeta_1^{1/2}\mathcal{W}_1(\upsilon), & \text{uniformly on } \upsilon \in [-\bar{\upsilon}, 0] \\ 2\zeta_2^{1/2}\mathcal{W}_2(\upsilon), & \text{uniformly on } \upsilon \in [0, \bar{\upsilon}] \end{cases} \tag{I.A.38}$$

where $\mathcal{W}_1(\upsilon)$ and $\mathcal{W}_2(\upsilon)$ are standard Brownian motions with variances $\zeta_1 = c'\Omega_1 c f$ and $\zeta_2 = c'\Omega_2 c f$, respectively.

(iii) Using equations (I.A.25), (I.A.32), (I.A.36), and (I.A.37)

$$
\begin{aligned}
&\widetilde{e}'(P_0(\gamma_0) - P_v(v))\widetilde{e} = \\
&\qquad n^{\alpha}\widetilde{e}'\widehat{X}_0(\gamma_0))(\widehat{X}_0(\gamma_0))'\widehat{X}_0(\gamma_0)))^{-1}n^{-2\alpha}(\widehat{X}_v(v)'\widehat{X}_v(v) - \widehat{X}_0(\gamma_0)'\widehat{X}_0(\gamma_0)) \\
&\qquad \times n^{\alpha}(\widehat{X}_v(v)'\widehat{X}_v(v))^{-1}\widehat{X}_0(\gamma_0)'\widetilde{e} \\
&\qquad - 2n^{-\alpha}(\widetilde{e}'\widehat{X}_v(v) - \widetilde{e}'\widehat{X}_0(\gamma_0))n^{\alpha}(\widehat{X}_v(v)'\widehat{X}_v(v))^{-1}\widehat{X}_0(\gamma_0)'\widetilde{e} \\
&\qquad = o_p(1), \quad \text{uniformly in } v \in [-\bar{v}, \bar{v}].
\end{aligned}
$$

Using a similar argument for $\widetilde{e}'(P_{\perp 0}(\gamma_0) - P_{\perp v}(v))\widetilde{e}$ together with equation (I.A.24) we get that the term (iii) of $Q_n(v)$

$$\widetilde{e}'(P^*(\gamma_0) - P^*(v))\widetilde{e} \Rightarrow 0. \tag{I.A.39}$$

Using equation of $Q_n(v)$ and (I.A.35)-(I.A.39) we get

$$Q_n(v) \Rightarrow \begin{cases} -\mu|v| + 2\zeta_1^{1/2}\mathcal{W}_1(v), & \text{uniformly on } v \in [-\bar{v}, 0] \\ -\mu|v| + 2\zeta_2^{1/2}\mathcal{W}_2(v), & \text{uniformly on } v \in [0, \bar{v}] \end{cases} \tag{I.A.40}$$

■

Lemma I.A.6 *If $\widetilde{W}_j \xrightarrow{p} W_j > 0$ for $j = 1, 2$ then the unconstrained estimators are asymptotically Normal*

$$\sqrt{n}(\widetilde{\beta}_1 - \beta_1) \xrightarrow{d} N(0, V_1)$$
$$\sqrt{n}(\widetilde{\beta}_2 - \beta_2) \xrightarrow{d} N(0, V_2)$$

where

$$V_1 = (S_1'W_1S_1)^{-1}S_1'W_1\Sigma_1W_1S_1(S_1'W_1S_1)^{-1} \tag{I.A.42a}$$

$$V_2 = (S_2' W_2 S_2)^{-1} S_2' W_2 \Sigma_2 W_2 S_2 (S_2' W_2 S_2)^{-1}. \tag{I.A.42b}$$

The constrained GMM class estimators are also asymptotically Normal

$$\sqrt{n}(\widehat{\beta}_C - \beta) \xrightarrow{d} N(0, V_C) \tag{I.A.43}$$

where

$$\begin{aligned} V_C \;&= V - W^{-1}R(R'W^{-1}R)^{-1}R'V - VR(R'W^{-1}R)^{-1}R'W^{-1} \\ &\quad -W^{-1}R(R'W^{-1}R)^{-1}R'VR(R'W^{-1}R)^{-1}R'W^{-1} \end{aligned} \tag{I.A.44}$$

and $V = diag(V_1, V_2)$.

Proof: First, we prove the asymptotic normality of the unconstrained estimators and in particular we start by providing details for the proof of $\widetilde{\beta}_1$. Let $\widehat{X}_\upsilon(\upsilon)$, $\widehat{X}_\perp(\upsilon)$, $\Delta\widehat{X}_\upsilon(\upsilon)$, $\widehat{Z}_\upsilon(\upsilon)$ denote the matrices obtained by stacking the following unconstrained vectors

$\widehat{x}_i(\gamma_0 + n^{-(1-2\alpha)}\upsilon)' I(q_i \leq \gamma_0 + n^{-(1-2\alpha)}\upsilon)$,

$\widehat{x}_i(\gamma_0 + n^{-(1-2\alpha)}\upsilon)' I(q_i > \gamma_0 + n^{-(1-2\alpha)}\upsilon)$,

$\widehat{x}_i(\gamma_0 + n^{-(1-2\alpha)}\upsilon)' I(q_i \leq \gamma_0 + n^{-(1-2\alpha)}\upsilon) - \widehat{x}_i(\gamma_0 + n^{-(1-2\alpha)}\upsilon) I(q_i \leq \gamma_0)$,

$\widehat{z}_i(\gamma_0 + n^{-(1-2\alpha)}\upsilon)' I(q_i \leq \gamma_0 + n^{-(1-2\alpha)}\upsilon)$.

Given that $\widehat{\pi}_q$ is consistent for π_q, we obtain $\widehat{\lambda}_{1i}(\gamma) \xrightarrow{p} \lambda_{1i}(\gamma)$ and $\widehat{\lambda}_{2i}(\gamma) \xrightarrow{p} \lambda_{2i}(\gamma)$ by applying the continuous mapping theorem. Furthermore, from Lemma 1 of Hansen (1996) and Lemmas A.4 and A.10 of Hansen (2000) we can deduce that uniformly on $\upsilon \in [-\bar{\upsilon}, \bar{\upsilon}]$ we obtain

$$n^{-1}\widehat{Z}_v'\widehat{X}_v(\upsilon) \xrightarrow{p} S_1 \tag{I.A.45}$$

$$n^{-1/2}\widehat{Z}_v(\upsilon)'e \Rightarrow N(0,\Sigma_1) \tag{I.A.46}$$

$$n^{-2\alpha}\widehat{Z}_v(\upsilon)'\Delta\widehat{X}_v(\upsilon) = O_p(1) \tag{I.A.47}$$

Let

$$\widetilde{\beta}_1(\upsilon) = (\widehat{X}_v(\upsilon)'\widehat{Z}_v(\upsilon)\widetilde{W}_1(\upsilon)\widehat{Z}_v(\upsilon)'\widehat{X}_v(\upsilon))^{-1}\widehat{X}_v(\upsilon)'\widehat{Z}_v(\upsilon)\widetilde{W}_1(\upsilon)\widehat{Z}_v(\upsilon)'Y,$$

$$\widetilde{\beta}_2(\upsilon) = (\widehat{X}_v(\upsilon)'\widehat{Z}_v(\upsilon)\widetilde{W}_2(\upsilon)\widehat{Z}_v(\upsilon)'\widehat{X}_v(\upsilon))^{-1}\widehat{X}_v(\upsilon)'\widehat{Z}_v(\upsilon)\widetilde{W}_2(\upsilon)\widehat{Z}_v(\upsilon)'Y$$

and write the unconstrained model as

$$Y = \widehat{X}_v(\upsilon)\beta_1 + \widehat{X}_\perp(\upsilon)\beta_2 - \Delta\widehat{X}_v(\upsilon)\delta_n + e \tag{I.A.48}$$

Using equation (I.A.48) we get

$$\sqrt{n}(\widetilde{\beta}_1(\upsilon) - \beta_1) =$$

$$((\tfrac{1}{n}\widehat{X}_v(\upsilon)'\widehat{Z}_v(\upsilon))\widetilde{W}_1(\upsilon)(\tfrac{1}{n}\widehat{Z}_v(\upsilon)'\widehat{X}_v(\upsilon)))^{-1}(\tfrac{1}{n}\widehat{X}_v(\upsilon)'\widehat{Z}_v(\upsilon)\widetilde{W}_1(\upsilon)(\tfrac{1}{\sqrt{n}}\widehat{Z}_v'u - \tfrac{1}{\sqrt{n}}\widehat{Z}_v(\upsilon)'\Delta\widehat{X}_v(\upsilon)\delta_n))$$

and by equations (I.A.45) - (I.A.47) we obtain uniformly on $\upsilon \in [-\bar{\upsilon},\bar{\upsilon}]$

$$\sqrt{n}(\widetilde{\beta}_1(\upsilon) - \beta_1) \Rightarrow (S_1'W_1S_1)^{-1}S_1W_1N(0,\Sigma_1).$$

Given Lemma 1, $\widehat{\upsilon} = n^{1-2\alpha}(\widehat{\gamma} - \gamma_0) = n^{1-2\alpha}(\widetilde{\gamma} - \gamma_0) = O_p(1)$, and using $\widetilde{\beta}_1 = \widetilde{\beta}_1(\widehat{\upsilon})$ we get

$$\sqrt{n}(\widetilde{\beta}_1 - \beta_1) = \sqrt{n}(\widetilde{\beta}_1(\upsilon) - \beta_1) \Rightarrow (S_1'W_1S_1)^{-1}S_1'W_1N(0,\Sigma_1) \sim N(0,V_1)$$

where $V_1 = (S_1' W_1 S_1)^{-1}(S_1' W_1 \Sigma_1 W_1 S_1)(S_1' W_1 S_1)^{-1}$. Similarly, we can get $\sqrt{n}(\widetilde{\beta}_2 - \beta_2) \Rightarrow N(0, V_2)$ with $V_2 = (S_2' W_2 S_2)^{-1}(S_2' W_2 \Sigma_2 W_2 S_2)(S_2' W_2 S_2)^{-1}$ as stated.

Next, we prove the asymptotic normality of the constrained estimator. First, note we can easily verify that

$$\sqrt{n}(\widetilde{\beta} - \beta) \xrightarrow{d} N(0, V) \tag{I.A.49}$$

where $V = diag(V_1, V_2)$.

Recall the relationship between the constrained and unconstrained estimators

$$\widehat{\beta}_C = \widetilde{\beta} - \widetilde{W} R (R' \widetilde{W} R)^{-1} (R' \widetilde{\beta} - \vartheta) \tag{I.A.50}$$

Therefore, given $rank(R) = r$ and $\widetilde{W} \xrightarrow{p} W > 0$ we obtain

$$\sqrt{n}(\widehat{\beta}_C - \beta) \xrightarrow{d} (I - WR(R'WR)^{-1}R')\sqrt{n}(\widetilde{\beta} - \beta) = N(0, V_C) \tag{I.A.51}$$

as stated.■