Supplementary Material for Monotonicity in Estimating Multiple Structural Breaks

In this supplement, we provide more simulation results.

A. The First Estimated Break

We first study the finite-sample performance of the *first* estimated break date when there are two breaks in DGP. Here we compare AGFL with SLS (Sequential LS), both of which are monotone procedures. We generate data from $y_t = \beta_t + \beta_t x_t + u_t$, where β_t has two breaks or three regimes. Since the location of the breaks may impact the performance of each method, we experiment with two specifications of $\mathcal{T}$, the set of break dates:

$$\begin{aligned} \text{Roughly even: } \mathcal{T} &= \left\{ \frac{T}{3} + i_1, \frac{2T}{3} + i_2 \right\}, \\ \text{Concentrated: } \mathcal{T} &= \left\{ \frac{T}{4} + i_1, \frac{T}{2} + i_2 \right\}, \end{aligned}$$

where i_1 and i_2 are random integers within $[-0.05T, 0.05T]$. In the roughly even case, break dates are fairly evenly spaced in the entire time span. In the concentrated case, break dates are concentrated in the first half of the time span. Note that in the simulations in the main text, break dates are exactly evenly spaced. The specifications here cover more realistic scenarios. The true values of β_t within each regime are randomly selected from the fixed set of $\{0, 0.5, 1\}$ without replacement. Hence there are two breaks with different jump sizes, such as ± 0.5, ± 1. We use the same specification of x_t and u_t as in the main text. Table 1 reports the mean absolute error (MAE) and standard deviations (STD) of the errors, which are calculated by subtracting the nearest true break fraction from the first estimated break fraction $(\min_k (T_1^{(1)} - T_k)/T)$.

We make the following observations from Table 1. First, when the noise level is low, SLS generally outperforms AGFL. But when the noise level is high, the conclusion reverses. This pattern is consistent across different specifications for x_t and u_t. Second, when the noise level is low, STD of AGFL is higher than that of SLS. But the noise level is high, the conclusion again reverses. Third, AGFL performs well when the break dates are concentrated in one half. Finally, in the roughly even case, the MAE of SLS declines faster than AGFL when the sample size increases. But in the concentrated case, the MAEs of both methods decline at about the same rate. This

Table 1: MAE and STD of the Error of the First Estimated Break Fraction.

		Roughly Even				Concentrated			
		SLS		AGFL		SLS		AGFL	
DGP	σ_u	MAE	STD	MAE	STD	MAE	STD	MAE	STD
T=60									
(1)	0.5	0.0208	0.0373	0.0259	0.0465	0.0219	0.039	0.0215	0.0368
(2)	0.5	0.0243	0.0448	0.0284	0.0501	0.0225	0.0394	0.0263	0.045
(3)	0.5	0.0316	0.0566	0.0337	0.0567	0.0306	0.0568	0.0319	0.0571
(4)	0.5	0.0283	0.0495	0.0311	0.0528	0.0278	0.0507	0.0294	0.0514
(5)	0.5	0.0233	0.0427	0.0284	0.0499	0.0254	0.0452	0.0254	0.0447
(1)	1	0.0514	0.0807	0.0435	0.0637	0.0618	0.1032	0.0455	0.0685
(2)	1	0.0593	0.0901	0.0485	0.0698	0.0665	0.1097	0.0489	0.0752
(3)	1	0.0699	0.1009	0.0563	0.0779	0.0837	0.1234	0.0676	0.0936
(4)	1	0.0673	0.0997	0.0506	0.0722	0.0799	0.1237	0.06	0.0891
(5)	1	0.058	0.0876	0.0474	0.068	0.0626	0.1038	0.0487	0.072
T=120									
(1)	0.5	0.0159	0.0335	0.0246	0.0483	0.0144	0.0272	0.0131	0.026
(2)	0.5	0.016	0.0336	0.0223	0.0445	0.0142	0.0279	0.0142	0.0272
(3)	0.5	0.0212	0.0421	0.0249	0.0475	0.02	0.0362	0.0201	0.0368
(4)	0.5	0.0194	0.0381	0.0232	0.0449	0.0163	0.0317	0.0174	0.0336
(5)	0.5	0.0149	0.0319	0.0213	0.0427	0.0152	0.0293	0.0146	0.0281
(1)	1	0.0321	0.0533	0.0312	0.0517	0.0314	0.0546	0.0276	0.0433
(2)	1	0.0344	0.0565	0.0339	0.0532	0.035	0.0586	0.0306	0.0477
(3)	1	0.0539	0.083	0.0433	0.0638	0.0573	0.0961	0.0471	0.0745
(4)	1	0.044	0.0697	0.0377	0.0575	0.0416	0.0706	0.0343	0.0535
(5)	1	0.0367	0.0612	0.0343	0.0558	0.0351	0.0635	0.0296	0.0492

indicates that the theoretical convergence rate of AGFL may be dependent on how breaks are located.

B. Estimation of Two Breaks

In this section, we study the performance of estimating two breaks when the true number of breaks is two. We use the same setup in the preceding section. So we may understand this exercise as a sequel to the previous section, which only studies the first estimated break. Furthermore, as in Table 2 in the main text, we compare AGFL not only to SLS (Sequential LS), but also non-monotone procedures such as Global LS, GFL, rDP, and ksPeak. To compare the accuracy of the *set* estimation, we calculate HD/T, the Hausdorff distance (HD) between the estimated and true sets divided by T. Note that HD/T generalizes the absolute error of estimated break fraction for the one-break case. We report not only the average HD/T, but also the median. Table 2 and 4 report the average HD/T for symmetric and asymmetric cases, respectively. And Table 3 and 5 report the median HD/T. We make the following observations:

First, the overall performance of AGFL compares favorably with SLS. Only when the noise level is low ($\sigma_u = 0.5$) and sample size is big ($T = 120$), SLS is stronger than AGFL. The mean and median results tell the same story. Second, the Global LS dominates SLS when the noise level is low. But when the noise level is high, SLS can sometimes outperform Global LS (e.g., when $\sigma_u = 1$ and $T = 60$). Third, AGFL dominates GFL and its variants. The post-processing procedures, especially ksPeak, improves on GFL substantially. When the noise level is high, ksPeak outperforms SLS and Global LS. Finally, similar to the case of the first estimated break, the average HD/T of SLS and Global LS declines faster than AGFL.

Table 2: Average HD/T of break-date estimation (2 breaks roughly evenly distributed)

		LS		PLS			
DGP	σ_u	Sequential	Global	AGFL	GFL	rDP	ksPeak
T=60							
(1)	0.5	0.0639	0.059	0.0543	0.1992	0.144	0.1147
(2)	0.5	0.0746	0.0735	0.0648	0.2029	0.1527	0.1281
(3)	0.5	0.1064	0.1033	0.0856	0.2198	0.1665	0.146
(4)	0.5	0.0954	0.0929	0.0769	0.2139	0.1667	0.145
(5)	0.5	0.0786	0.0773	0.0661	0.2018	0.1549	0.1318
(1)	1	0.2	0.2057	0.12	0.2165	0.1856	0.1705
(2)	1	0.2018	0.202	0.1305	0.218	0.1892	0.1775
(3)	1	0.207	0.2053	0.1481	0.2302	0.1998	0.1873
(4)	1	0.2051	0.2097	0.1389	0.2203	0.1911	0.1784
(5)	1	0.2014	0.2065	0.1271	0.2258	0.1971	0.1806
T=120							
(1)	0.5	0.029	0.0245	0.0345	0.2015	0.1398	0.1128
(2)	0.5	0.0374	0.0321	0.0388	0.2021	0.1461	0.1247
(3)	0.5	0.063	0.0583	0.0522	0.2181	0.1698	0.1464
(4)	0.5	0.0495	0.045	0.0467	0.2163	0.1639	0.1422
(5)	0.5	0.0327	0.0298	0.0382	0.2037	0.1513	0.125
(1)	1	0.1222	0.1205	0.0805	0.222	0.1943	0.1787
(2)	1	0.1283	0.1291	0.0827	0.2266	0.1949	0.1824
(3)	1	0.1737	0.1752	0.1169	0.231	0.2046	0.1925
(4)	1	0.1578	0.1577	0.1025	0.2286	0.2018	0.1907
(5)	1	0.1327	0.1356	0.0838	0.2229	0.1924	0.1768

Table 3: Median HD/T of break-date estimation (2 breaks roughly evenly distributed)

		LS		PLS			
DGP	σ_u	Sequential	Global	AGFL	GFL	rDP	ksPeak
T=60							
(1)	0.5	0.0333	0.0333	0.0333	0.2333	0.1014	0.0667
(2)	0.5	0.05	0.05	0.05	0.2333	0.1179	0.085
(3)	0.5	0.0667	0.0667	0.0527	0.25	0.1509	0.1179
(4)	0.5	0.0601	0.0527	0.0527	0.2478	0.1509	0.1167
(5)	0.5	0.0373	0.0373	0.0373	0.2333	0.1179	0.085
(1)	1	0.1863	0.1951	0.0972	0.2291	0.1818	0.1537
(2)	1	0.1799	0.1833	0.1014	0.2333	0.1833	0.1537
(3)	1	0.1863	0.1833	0.1269	0.2333	0.19	0.1675
(4)	1	0.1886	0.1951	0.1179	0.2173	0.1833	0.1675
(5)	1	0.1756	0.1875	0.1	0.2478	0.1947	0.1667
T=120							
(1)	0.5	0.0167	0.0167	0.0177	0.2428	0.092	0.0437
(2)	0.5	0.0186	0.0186	0.0243	0.2501	0.1	0.0621
(3)	0.5	0.0333	0.03	0.0333	0.2588	0.1569	0.109
(4)	0.5	0.0264	0.0264	0.03	0.2585	0.1397	0.0886
(5)	0.5	0.0186	0.0167	0.0236	0.2506	0.1167	0.0667
(1)	1	0.075	0.0717	0.0583	0.253	0.2015	0.1718
(2)	1	0.0833	0.0837	0.0583	0.2585	0.204	0.1797
(3)	1	0.1436	0.143	0.0917	0.2501	0.215	0.1918
(4)	1	0.117	0.1174	0.0755	0.2539	0.2084	0.1903
(5)	1	0.0886	0.0892	0.0583	0.2534	0.1981	0.1671

Table 4: Average HD/T of break-date estimation (2 breaks concentrated in the first half)

		LS		PLS			
DGP	σ_u	Sequential	Global	AGFL	GFL	rDP	ksPeak
T=60							
(1)	0.5	0.0741	0.0712	0.0629	0.1656	0.1331	0.1206
(2)	0.5	0.0917	0.0866	0.0874	0.1777	0.1517	0.142
(3)	0.5	0.1226	0.1173	0.1232	0.1868	0.166	0.1593
(4)	0.5	0.1021	0.0985	0.0981	0.1863	0.1551	0.1455
(5)	0.5	0.096	0.0914	0.0892	0.171	0.1423	0.1308
(1)	1	0.2062	0.2012	0.1558	0.199	0.1846	0.1836
(2)	1	0.2138	0.2069	0.1592	0.201	0.1869	0.1794
(3)	1	0.2376	0.2354	0.1942	0.2228	0.2053	0.2012
(4)	1	0.2302	0.2191	0.1814	0.2193	0.2045	0.1991
(5)	1	0.2257	0.2211	0.174	0.2109	0.1951	0.1913
T=120							
(1)	0.5	0.0324	0.0272	0.0299	0.1746	0.1324	0.1156
(2)	0.5	0.0363	0.0316	0.0387	0.1697	0.138	0.1211
(3)	0.5	0.0719	0.0672	0.0684	0.1836	0.1534	0.1435
(4)	0.5	0.0572	0.0516	0.0554	0.1809	0.1496	0.1392
(5)	0.5	0.0369	0.0327	0.0386	0.1773	0.1392	0.1243
(1)	1	0.127	0.1227	0.0987	0.1883	0.1725	0.1663
(2)	1	0.1362	0.1332	0.1062	0.1948	0.18	0.1746
(3)	1	0.206	0.1996	0.1663	0.2098	0.1969	0.1921
(4)	1	0.1791	0.1765	0.1366	0.2006	0.1862	0.1815
(5)	1	0.1478	0.1441	0.1148	0.195	0.1791	0.175

Table 5: Median HD/T of break-date estimation (2 breaks concentrated in the first half)

		LS		PLS			
DGP	σ_u	Sequential	Global	AGFL	GFL	rDP	ksPeak
T=60							
(1)	0.5	0.0373	0.0373	0.0333	0.1833	0.1167	0.0833
(2)	0.5	0.05	0.05	0.05	0.1893	0.1462	0.1118
(3)	0.5	0.0697	0.0667	0.0833	0.2	0.1509	0.1344
(4)	0.5	0.0634	0.0601	0.0527	0.2	0.1344	0.1167
(5)	0.5	0.05	0.05	0.05	0.1841	0.1213	0.1
(1)	1	0.1667	0.1667	0.1054	0.2	0.1675	0.1641
(2)	1	0.1818	0.1675	0.1167	0.2007	0.1716	0.1537
(3)	1	0.2167	0.2007	0.1667	0.2028	0.1863	0.1833
(4)	1	0.1951	0.1795	0.1491	0.2108	0.1841	0.174
(5)	1	0.2	0.1875	0.1333	0.2108	0.1841	0.1675
T=120							
(1)	0.5	0.0186	0.0167	0.0177	0.2005	0.125	0.0779
(2)	0.5	0.0236	0.0186	0.0236	0.2	0.1336	0.0907
(3)	0.5	0.0333	0.03	0.0333	0.2083	0.1526	0.1253
(4)	0.5	0.0264	0.025	0.0264	0.2004	0.1583	0.1251
(5)	0.5	0.0186	0.0167	0.025	0.206	0.1417	0.0986
(1)	1	0.0752	0.075	0.0534	0.2016	0.1788	0.1667
(2)	1	0.0833	0.0833	0.0601	0.2085	0.1917	0.1758
(3)	1	0.1669	0.1584	0.1251	0.2098	0.188	0.183
(4)	1	0.1346	0.1336	0.0917	0.2085	0.185	0.1752
(5)	1	0.095	0.0932	0.0712	0.2084	0.1838	0.175