

realdata_estimation_results_log

N	T	ρ	$\sigma_a^2_estimate$	$\sigma_a^2_start$	$\sigma_u^2_estimate$	$\sigma_u^2_start$	LL
4	60	0.8785	0.138	0.1	16.727	0.1	-2.60E+11
4	60	0.8785	0.103	0.1	16.747	1	-2.60E+11
4	60	0.8785	0.101	0.1	16.749	2	-2.60E+11
4	60	0.8785	0.1	0.1	16.749	5	-2.60E+11
4	60	0.8785	0.1	0.1	16.749	10	-2.60E+11
4	60	0.8785	1.069	1	16.206	0.1	-2.68E+11
4	60	0.8785	1.035	1	16.224	1	-2.68E+11
4	60	0.8785	1.019	1	16.232	2	-2.68E+11
4	60	0.8785	1.005	1	16.24	5	-2.68E+11
4	60	0.8785	1.001	1	16.242	10	-2.68E+11
4	60	0.8785	2.07	2	15.698	0.1	-2.77E+11
4	60	0.8785	2.046	2	15.71	1	-2.77E+11
4	60	0.8785	2.031	2	15.717	2	-2.76E+11
4	60	0.8785	2.012	2	15.726	5	-2.76E+11
4	60	0.8785	2.003	2	15.731	10	-2.76E+11
4	60	0.8785	5.067	5	14.439	0.1	-3.00E+11
4	60	0.8785	5.053	5	14.444	1	-3.00E+11
4	60	0.8785	5.042	5	14.448	2	-3.00E+11
4	60	0.8785	5.022	5	14.455	5	-3.00E+11
4	60	0.8785	5.006	5	14.461	10	-3.00E+11
4	60	0.8785	10.062	10	12.931	0.1	-3.35E+11
4	60	0.8785	10.051	10	12.934	1	-3.34E+11
4	60	0.8785	10.043	10	12.936	2	-3.34E+11
4	60	0.8785	10.025	10	12.94	5	-3.34E+11
4	60	0.8785	10.007	10	12.945	10	-3.34E+11
4	60	0.8785	20.056	20	11.068	0.1	-3.90E+11
4	60	0.8785	20.045	20	11.069	1	-3.90E+11
4	60	0.8785	20.038	20	11.07	2	-3.90E+11
4	60	0.8785	20.023	20	11.072	5	-3.90E+11
4	60	0.8785	20.003	20	11.075	10	-3.90E+11
4	60	0.8785	50.049	50	8.547	0.1	-5.05E+11
4	60	0.8785	50.034	50	8.547	1	-5.05E+11
4	60	0.8785	50.029	50	8.548	2	-5.05E+11
4	60	0.8785	50.014	50	8.548	5	-5.05E+11
4	60	0.8785	100.048	100	6.877	0.1	-6.27E+11
4	60	0.8785	100.027	100	6.878	1	-6.27E+11
4	60	0.8785	100.021	100	6.878	2	-6.27E+11
4	60	0.8785	100.007	100	6.878	5	-6.27E+11

realdata_estimation_results_log

4	60	0.8785	200.052	200	5.49	0.1	-7.85E+11
4	60	0.8785	200.02	200	5.49	1	-7.85E+11
4	60	0.8785	200.014	200	5.49	2	-7.85E+11
4	60	0.8785	200.001	200	5.49	5	-7.85E+11

Notes:

1. All estimation runs had the same lower and upper parameter bounds: [0,0,0] and [1, inf, inf]
2. Iterations included all combinations of starting values for sigma's of [0.1, 1, 2, 5, 10, 20, 50, 100, 200]
3. The rows that are missing (e.g., sigma_a^2 starting at 200, sigma_u^2 starting at 10) did not converge and reported -inf values for the log likelihood. This could indicate the need for more decimals in the precision of the numbers being used – we're currently limited to double precision, but there are ways this could be increased if we think the true values of the parameter may be in the areas hidden by this precision issue (see note 4).
4. A Contour plot of the log likelihood resulted in "2022-01-30 LL plot of starting values.png", where the blank areas in the top right of the plot are where the LL returned -inf
5. Note that the LL is getting more negative as the starting value of sigma_a^2 increases. The total search range is unchanged (see note 1), which to me is an indication that the minimization function is not finding the global minimum – it is minimizing the negative LL, and -2.6E11 larger than -7.85E11, so the first row should be a better minimum than the last row. But the algorithm does not find it's way back to (0.1, 0.1) when starting from (200, 5) for example. This is an indication to me that there's some needed tweaking of the default gradient descent method and that our function might be lumpy.
6. I double checked that the ending estimation of the region fixed effects and time trends are identical (down to some 8th decimal) between all the valid results listed here. Which means, the residuals used in all the estimations above are pretty much identical. So the only difference between all the above estimations is the starting value for the search