Notes for Data Replication[1]

I. Monte Carlo Experiments\Table 1 - Parameter Estimates[2]

- Each of the reported experiments depends upon two files (…\Code): 1) the specification of the simulated data (a script) and 2) the likelihood (a function). There are 4 versions of the former (labeled: \mc_sprobit_tscs…) included which generate the data used in the 4 reported experiments (though the code varies only in its specification of the ρ and φ parameters). There is a single version of the likelihood (labeled: \sprobit…) included that is called by each of the 4 data generating codes and which does not require any adjustment.

- Little should need to be changed prior to running the scripts (which automatically call the function). It is currently set to make use of distributed processing in a multi-core environment, so these specifications would need be changed if one were replicating the results locally.[3] In particular, 'matlabpool open 12' and 'matlabpool close' should be removed and the 'parfor' loop should be changed to 'loop'

- Finally, the `output' object is assigned values for the mean estimates, RMSE, etc…which should be similar to those depicted in Table 1 in the text (with minor differences arising from the simulation of new data). The results from our experiments are saved in the \Data folder.

II. Monte Carlo Experiments\Table 2 & Figure 1 – Effects Estimates

- What is noted as '*True Effect'* in Table 2 is generated using the \Code\CF_effects_truth… script for the relevant values of ρ and φ. This should not require any additional modifications (save modifying the directory from which to load the estimation data on your machine) and will produce results similar to those saved in the \Results\sp_tscs_phi_rho_fx_true results files and indicated in Table 2.

- The STP-MSL results in Table 2 are produced by the \Code\CF_effects_estimates… script which again requires little additional modifications and saves the output in \Results\sp_tscs_phi_rho_fx_est.

- Given the limited availability of Matlab to perform standard regime-switching and discrete-time hazard analysis, we estimated these models (and effects) in Stata. This is done using the Code\effects_estimatse_naive do files, which: 1) import the same data generated in Matlab and used for STP-MSL, 2) estimate the parameter estimates for the naïve models, and 3) estimate the effects estimates for the naïve models (calling the same units for the counter factual as in the STP-MSL models). The `sum' command on the saved results of these trials provides the information included in Table 2.

---

[1] Unless explicitly noted (as for the some of the effects estimation) all programming is done is Matlab.

[2] All subsequent references to file locations are located in this directory.

[3] If desired we can also include the bash scripts we used to run this on a unix cluster (though there is nothing novel there, it simply specifies the time to allocate for processing and calls the matlab scripts).

- Figure 1, summarizing the STP-MSL results, is generated in R using the \Code\figure1_acc_fx script. As specified this will produce exactly what is included in the text. Unfortunately, if attempting to produce a new figure corresponding to new inputs (from new simulated results) the replicator will have to input these results manually as we did not automate this portion.

III. Application – Civil War

- Table 3 indicates the results from the purely spatial models which is generated using the \Code\hb_ksg_rep_sprobit_SAM_rho script (and association likelihood function). For the application we do not use the `matlabpool' environment (given there is only a single trial) and both the naïve and STP-MSL results are produced. You need to import the outcome-attribute data (\hb_ksg_rep_Africa.csv), weights matrix (\hb_ksg_contig_W.csv), and time lag matrix (\hb_ksg_contig_TL.csv), so be sure to correctly specify the directory you're importing the data from.

- Table 4 is produced identically to Table 3 using the \Code\hb_ksg_rep_sprobit_SAM_rho_phi script.

- The estimated-response paths for Figure 2 are generated using the \Code\CF_effects_app_LRRP script and plotted in R using the figure2_senegal script. The only modification a user should have to make in the code is specifying the location to load the estimation data from.

- The Area Under the Curve (AUC) estimates and the ROC Curves in Figure 3 are produced using the \Code\ROC_app script in Matlab. Here one needs to import the STP-MSL results from the prior estimation stage, so again care needs to be taken in importing these values.