Development Paper Analysis Read-Me
September 5, 2011

In the original submission, I relied on the data set created in development_data_00.do. This data set contains only 22 villages, where migrants were followed-up, and uses the remittance variable (rmg) reported by migrants themselves.

In the revisions, I am creating an alternative data set in development_data_00_51v.do, which contains all 51 villages. (I use the combine_hh_rosters_dec10.do code first.) Here, I will use the remittance variable (rmg_hh) reported by household members. This approach has the advantages of (i) containing information from more villages, and (ii) using hh-reported remittance information, which is less likely to be missing. (Recall only 50% of migrants were followed up, so we have missing remittance information on migrants who were not found in destination.)

In some of the supplementary analysis necessary to justify our instrument (% migrating or remitting in the sibling network), we need to show that migrants do not send remittances to their siblings' households. If we go with the latter data set, we will still need to rely on the information from migrants (i.e., only from 22 villages). But we can show this to be the case in those villages, and extrapolate from that.

Another downside of using the latter, larger data set, is that we will not be able to compute cumulative trip measures at the hh and village levels. But we can still compute basic indicators for the proportion of hh members who have migrated before or village prevalence.

The justification for using the latter data set in the revision memo: We have cleaned up a larger portion of the data set for a different set of analyses after our submission. Here we are taking advantage of that larger data set.

Development_data_00.do: Creates cross-sectional data from 22 villages. Remittance information is based on migrants' accounts. Uses id_lh (not nrpid) to identify individuals. Generates a data set called CS9400_Devt.dta.

Development_data_00_51v.do: Creates cross-sectional data from 51 villages. The prelude code is stored in combine_hh_rosters_dec10.do. Remittance information is based on households' accounts. Uses nrpid to identify individuals. Generates a data set called CS944_Devt_51v.dta.

Devt_analysis_12Oct2011_part1.do: Runs PCA analysis and saves the results

Devt_analysis_12Oct2011_part2.do: Runs all analyses and saves the final tables.