1 Uganda 20% PMT (n = 2671)

In my opinion the results are somewhat promising, looking at similar papers which have applied ML/Data Science methods to PMT, our dataset is larger and is unique in that we have cross sectional data and panel data for numerous countries. If SVM is the avenue we decide to go down, maybe there is some potential for altering the objective function or defining custom loss functions. I think there is some room to survey other models such as more tree/decision based or neural networks, since SVM seems to be outperformed by XGBoost, out of the box at least. Maybe with some variable selection/analysis of the variables, we could achieve similar, but this sort of reminds me of p-hacking. My main question lies is will we it be more beneficial to use the full 40k dataset and benefit from large sample or whether to run the models for each subset of countries. I also want to add that possibly like the data science example of classifying titanic survivors in which a simple rule of women and children often outperforms even the most complicated models. Similarly here since "On average, targeting to households with elderly or disabled members, widows or children does as well as Basic PMT", there is potential for similar here.

2 SVM

2.0.1 Basic

	Kernel	Accuracy	Precision	sensitivity	F1_Score	AUC
1	linear	0.74	0.62	0.17	0.27	0.57
2	radial	0.75	0.67	0.16	0.26	0.57
3	polynomial	0.74	0.62	0.16	0.26	0.56
4	sigmoid	0.67	0.37	0.27	0.31	0.55

(sigmoid confusion matrix)

		predicted	
Actual	0		1
0	320		66
1	108		39

This is the results of SVM using basic regression variables for Uganda (n = 2671). I had to drop Christian and Muslim as they were constant here. I think there is some room for improvement using double selection lasso on variables in the whole dataset/extended regression.

2.1 radial w/ some tuning

5 fold cross validation and light cost and sigma grid search.

		predicted	
Actual	0		1
Class0	374		12
Class1	123		24

	Kernel	Cost	Gamma	Accuracy	Precision	Sensitivity	F1_Score	AUC
1	Radial	1.00	0.08	0.75	0.67	0.16	0.26	0.66

could be potential to look at original literature and further coding/programming to optimise or bayesian optimisation. But the improvement over basic is promising.

3 XGBoost

3.0.1 Basic

	Accuracy	Precision	sensitivity	F1_Score	AUC
1	0.74	0.60	0.20	0.30	0.74

		predicted	
Actual	0		1
0	365		20
1	119		30

This is the results of XGBoost on the same dataset with basic setup. It predicts the non poor better but actual poor worse (39 → 30). Again I think there is room for improvement if we consider different variables.

3.1 tweaking threshold

threshold	Accuracy	Precision	sensitivity	F1_Score	AUC
0.19	0.59	0.39	0.88	0.54	0.74

		predicted	
Actual	0		1
0	183		202
1	18		131

This is just an exploration of looking at thresholds, clearly if you lower the threshold you predict more are poor and therefore cover more of the poor. Nothing we didnt know before.

3.2 grid search for highest auc

	Accuracy	Precision	sensitivity	F1_Score	AUC
1	0.74	0.60	0.20	0.30	0.73

	0	1
0	365	20
1	119	30

3.3 grid search for highest sensitivity

	Accuracy	Precision	sensitivity	F1_Score	AUC
1	0.59	0.39	0.84	0.53	0.73

		predicted	
Actual	0		1
0	188		197
1	24		125

4 Other Methods/thoughts

I think the panel section of our data should definitely be looked at. There is potential for causal inference here or synthetic analysis of our targeting? Or potentially some room for economic models? Especially with the shocks variables that are used in some of the more detailed econometric analysis further in the paper. I think it could also be interesting to look at those who fall in and out of poverty under the various lines, their relationships with shocks etc. With regards to other classification methods, XGBoost seems to do relatively well out of the box, there is potential for ensemble methods/neural networks but I am not sure if this in an avenue for us, but I would be interested in looking at applying some of Susan Athey's work on causal ML here or Bayesian Structural Time Series, if possible.