Inward Growth Site Selection

ONE ACRE FUND

29-Nov-2018

Introduction

For the past two years we have recommended inward growth sites to Field Leaders (FLs) based on One Acre Fund's existing market penetration in a ward. The underlying theory was that each area has a market penetration 'ceiling' - a maximum number of clients that we could enroll. As we approach the ceiling the number of new clients enrolled by a new site will decrease (diminishing marginal returns). This approach implied that we should be adding sites in areas where current market penetration is relatively low.

However, when we analyzed the performance of inward growth sites early this year, we found limited evidence of the negative relationship between enrollment success and market penetration:

- We used two different regression equations to test this hypothesis. One produced statistically significant results the other did not.
- In both cases there was only a mildly negative "effect" on the number of new clients enrolled in an inward growth site as market penetration at the ward level increased.

Additional thoughts on what could have affected the accuracy of the analysis and what we were able to do to verify the results of the analysis can be found in the second appendix.

Given the apparent ineffectiveness of the 'market penetration' approach for inward growth site placement, we propose an alternative method based on prior enrollment performance of an area.

Why didn't market penetration appear to work?

Our results suggest that market penetration should not be the most important consideration when making site recommendations to FLs. There are a few possible reasons for this:

- It is possible that we were wrong to assume that marginal net client gain from inward growth sites would diminish linearly as we approach the 'ceiling'. Perhaps we observe very mild diminishing marginal returns until we are very close to the 'ceiling'.
- We have evidence from a new, likely more reliable, dataset that market penetration is lower in almost all areas of operation than it appears from the census data. When we recalculated market penetration using the new dataset, we found that OAF market penetration was on average 5% lower across all inward growth areas. Based on this data, it is possible we are still far enough from the market penetration 'ceiling' that it is not yet an important factor for determining site success.
- It is possible that available datasets on market penetration are still not accurate or large enough to detect the "effect" of market penetration on net client gain from inward growth. In this case it is still not wise to use these datasets to make our site recommendations.

Inward Growth Site Selection

29-Nov-2018

Guiding Questions

- Do you agree with the principle behind the proposed site selection method?
- It is likely that the proposed method will result in an even greater proportion of inward growth sites being located in Western Province than the current method. Are we comfortable with the possibility of restricting growth in Nyanza and Rift?

Proposed inward growth site selection method

Principle: Adding the minimum number of inward growth sites to reach our LR 2020 scale targets.

Application of Principle: We propose to prioritize adding inward growth sites by performance of existing sites in the same geographic area and district. The rationale is that most factors driving enrollment performance of inward growth site success are linked by district or geographic area.

- Factors linked to district: mentorship, district culture, leadership, positive reputation.
- Factors linked to geographic area: agro-ecological zone, socioeconomic demographics, popular crop-mixes.

Therefore, adding sites based on past performance in the same district or geographic area can help us reach our scale goals with the minimum number of inward growth sites.

Pros	Cons
Likelihood of a greater number of opportunities for promotion available in Western for high performing staff	Likely pushback from FLs in Western to adding more inward growth sites than previous years
Likelihood of reduced exposure to areas with weak value proposition and poor district cultures	Slows the path to building an experienced staff body and skilled leadership in areas outside Western.
	Risk of strain on the promotions and hiring process in Western Province

Steps for Site Selection: The proposed inward growth site selection method has been created with these pros and cons in mind. Additional detail on the plan are available in the first appendix but broadly it will involve three steps:

1. Ranking districts/areas with high total transaction size per FO.
2. Using maps and geospatial datasets to recommend site placements wherever there is a sufficient population base to support a site and no existing sites. These site suggestions are ranked according to the ranking from part 1. The visualization of placements may also help reduce pushback from FLs in Western.
3. Capping the inward growth site recommendations in successful areas and replacing them with sites in less successful areas based on the field staff pipeline to reduce the strain on promotions

and hiring processes in Western and increase the number of opportunities for promotion in areas outside of Western.

Appendix: Explanation of proposed method

Part 1: Identifying successful areas

- **Deadline: December 13, 2018**
- **Rationale:** We have no clear evidence that market penetration is a significant determinant of site success according to a new geospatial dataset. That dataset has shown that we are probably on average about 5% less dense than the estimates from the census dataset. This suggests a couple things:
    - That we are not close enough to the real market penetration ceilings in the vast majority of our areas to see significantly diminishing returns.
    - For the vast majority of the areas that we operate in other factors besides market penetration will be the most important determinants of a site enrollment success. Some examples include FL experience, value proposition to prospective clients, and popularity/reputation of the program.

  Based on this reasoning it makes sense that we can achieve scale most efficiently by focusing on adding inward growth sites to areas we are already more successful in without trying to understand exactly what drives site success. We could use clients per FO by district and geographic region to identify the areas we should bet on.
- **Key data points:**
    - Average clients/FO and variance of clients/FO at the district level.
    - Heat map of clients per FO.
    - Red flags for districts that did especially poorly in repayment.
- **Anticipated questions:**
    - *Why use by the district and the geographic approach?* We can group factors affecting site performance into two categories. There are those related to district management and those related to the suitability of our program to a geographic area. So we should look at it in both ways. Examples of the former: mentorship, district culture, leadership, positive reputation. Examples of the latter: agro-ecological zone, socioeconomic demographics, popular crop-mixes.
    - *How will you choose between the two methods to arrive at a final list of sites?* We will have to look at the data to see which is more useful for producing site recommendations. If we find that geographic area overrules districts (ex. heat map shows bands of high performance that cut across district areas and/or there is high variance within districts) then we will prioritize geographic area. If we find that performance is much more consistent within districts (ex. the heat map aligns with district boundaries and/or the variance within districts is low) then we will focus on adding sites to a particular district and may avoid adding many new districts. We may also choose to go for a combined approach where we prioritize geography in some areas and districts in others.

- *How will you arrive at the total number of sites to add?* Regardless of whether we prioritize geographic area or districts we will be able to come up with a model that uses the performance of nearby sites or districts to calculate the number of new clients we can expect to get from any portfolio of inward growth sites. In part 3 we will factor in the field staff pipeline to cap the number of sites we add. The final site list will be compared against the scale goals for LR 2020 to help us adjust the total number of inward growth sites or reset targets if the amount of growth that needs to come from inward growth is unrealistic.
- *Does this mean that we will place very few sites in Rift and Nyanza?* Western will probably receive a higher proportion of inward growth sites than it has in the past because our most successful areas are in Western. However, we will also be considering whether the field staff pipeline in each of these areas can support the growth in those areas. The number of strong field staff we can promote will limit the number of sites we will be able to place in Western. This means that we will still need to add sites in Nyanza and Rift to reach our ambitious scale goals. In addition outward growth sites will mostly be in Rift and Nyanza.
- *Why not just look at inward growth site performance?* Inward growth site performance gives us a sense of performance in the first year not the long-term prospects of these areas. Overall site performance captures this best. Therefore it is unlikely that looking at inward growth site performance in isolation will be more useful than performance across all sites. One situation where it might be useful is if we think that we have saturated an area and will experience rapidly diminishing clients per FO with each additional site. Even though the data we have does not show evidence that this is an important concern, FL feedback should be able to catch this.

- **Steps for Execution:**
  1. Johnna produces the heat map of clients per FO.
  2. Sarv collates the district-level summary statistics.
  3. Johnna and Sarv compare the heatmap and the district-level summary statistics to decide whether and how to prioritize geographic area or district.

Part 2: First draft of site recommendations to share with Field Staff Management

- **Deadline: December 21, 2018**
- **Rationale:** Now that we have the districts and areas we want to focus on we need to create a list of recommended sites that we can share with Field Staff Management as they analyze the field staff pipeline. In this step we will be using two more factors to create the algorithm for site selection:
  - Increasing Coverage: We will be prioritizing adding sites to open areas of the map. We have chosen to focus on this approach for a few reasons:
    - We do not have to worry as much about how close we are to the market penetration ceiling. This would be more of a concern if we decided to pack sites together or overlap them
    - We will be less likely to get push back from FLs that there are no more clients left in that area or that having too many sites too close together will cause conflicts between FOs.

        - Scouting becomes less complex for FLs. FLs do not have to worry about how many clients remain in a particular area. They only need to focus on whether there are enough people who grow maize.
    - Checking Population Base: Here we check the underlying population base in the open areas using the new geospatial population map that has recently become available for Kenya. Basically, this will help us avoid areas with very sparse population or market centers, towns and peri-urban areas.
- **Key data points:**
    - Geospatial population map: This is the same data source that we used to compare against the census data to verify the results of the last analysis.
    - Average site size: We will have to run a short survey to get a rough sense of how big sites are. We will use this data to find the open areas on the map.
    - Ward Boundaries: We need this layered on top of our population map and map of site geotags to describe the location of the new site to FLs.
    - Terrain/Road Map: Useful for describing scouting locations to FLs.
- **Anticipated questions:**
    - *How will we identify the open areas?* We will get an estimate of site size from the survey of FOs. Using this information and the existing GPS coordinates of site centers we will draw small circles to represent sites on a map. Then we will identify GPS locations that are far enough away from all existing sites to accommodate a site of average size.
    - *How will we know that there are enough people in an area to support a site?* We will essentially look at the population bases that are supporting nearby site and use that as a benchmark. That can be done by looking at the population that lies within the site circles.
- **Steps for Execution:**
    - Survey of FOs to come up with an estimate of average site size by region (Western, Nyanza, Rift).
    - Johnna and Sarv align on an algorithm that will produce site recommendations using district vs. geographic and the additional factors from above.
    - Johnna implements the site selection algorithm producing a list of site recommendations to be shared with Field Staff Management that has the following pieces of information:
        - Current district association
        - GPS coordinates
        - List of nearby sites
        - Government ward

Part 3: Incorporating feedback from Field Staff Management

- **Deadline: January 31, 2019**
- **Rationale:** Adding many new sites will force us to create new districts with new leaders, which could affect the performance of sites in their areas. On top of this, our new method will put disproportional strain on the field staff pipeline in successful areas. Therefore, we want to optimize our recommendations by considering a few additional factors related to field staff pipelines and the number of new districts that we will need to create. These factors include:

  - Optimal proportion of old and new sites: We should limit the number of new sites in a new district without taking too many of the best performing sites from surrounding districts.
  - Total number of sites in district: We will want to make sure that all districts are within an acceptable range of total sites under management.
  - Geography: We will want to consider proximity to the center of the district and to other nearby sites to minimize travel to Monday meetings and because we try to promote FMs from the area they will serve.
- **Key data points:**
  - Checking new/old site performance in new/old districts.
  - Total number of sites vs. site performance.
  - Field staff pipeline.
- **Steps for Execution:**
  - Sarv will compare new/old site performance relative to age of district and total number of sites within the district to come up a range of acceptable inward growth sites to old and new districts.
  - Dilanthi and Sarv will look at the site list and the field staff pipeline to cap the number of inward growth sites we add in specific areas assuming some recommendations will be dropped after FL feedback.
  - Johnna and Sarv produce the final list of inward growth site recommendations to be shared with FLs for feedback pending approval from the Steering Committee.

Appendix: Verification of inward growth analysis

Verification of these results: There could have been a number of possible reasons we were not seeing the results that we were expecting from the inward growth analysis. We were able to check that data quality was not the issue but we were not able to verify that a larger sample size or a better proxy for market penetration would have changed the results.

Possible Issue	Checks
Data Quality Issue: Poor population estimates from the census, and/or poor ward location information from roster.	- Checked the population estimates from the census against another dataset that was thought to have better quality data. We found that population estimates from the census were lower than the estimates from the other data set by around 5-7%. - The number of errors in the ward location data was not large enough to affect the outcome of this analysis.
Insufficient Sample Size: It is possible that our sample size was too small to accurately measure the relationship.	- There is no way to check this and no solution right now. We can repeat the same analysis as last year only after the qualification deadline.
Weak proxies for market penetration: We used an estimate of rural population as the maximum number of clients available in an area. Ideally, we would need	- We have no better proxy for market penetration. We do not have the data to come up with an estimate for the number of people in an area that would be willing to enroll.

an estimate of the number of clients actually interested in our program.	