Entrepreneurship Exploration and Analysis

Nathaniel Coffin, David Swart, Sanjay Tamrakar, Jie Wang,
STA 660
for Christopher Sutter

1. Introduction

The main purpose of this project is to help entrepreneurial ventures in developing countries to survive, grow and thrive with the help of organizational sponsorship. TechnoServe specifically chose small to middle-sized businesses in these countries as a means of promoting poverty-alleviation because their goal focused around promoting growth. More specifically, this study was to explore which programs influence venture growth in the context of poverty. We are trying to explore differences in entrepreneurial growth by constructing reasonable models of log monthly sales among several different venture characteristics. Monthly sales data was collected from 3 years of pre-, during-, and post-training of 139 firms in three countries. Training for countries 2 and 3 started in July 2014 and training for country 1 started in Aug 2014. The months are coded such that Jan. 2013 was month 1, Feb. 2014 was 2 and so forth, because base sales represents sales during months 1-12. The structure of the program is such that the firms were monitored for 19 months (countries 1 and 3) or 20 months (country 2) before the training started, for 30 months during training, and then for 12 months after training ended, though many dropped out before the 12 post-training months had transpired. Advisors were assigned to provide aftercare during the period following the training. In all, more than 10 predictors were explored and they are summarized in Table 1.

Different hypothesis will be considered: (1) Firms to which female advisors are assigned as trainers will experience reduced entrepreneurial growth gained from organizational sponsorship compared to those assigned male advisors; (2) Female-led ventures in contexts of poverty will experience less entrepreneurial growth from organizational sponsorship than male-led ventures; (3) Business training in contexts of poverty will experience more entrepreneurial growth from organizational sponsorship than business model. (4) The older a firm, the less entrepreneurial growth will be evident as a result of organizational sponsorship. We explored three modeling strategies related to these hypotheses.

There are several challenges in the data analysis:

1. There are missing values in several predictors (Table 2);
2. Seasonality or periodicity should be explored considering the data was collected over time;
3. Other possible problems during data collection: Some original firm age numbers were 0, so those firms didn't exist during the entire base year. We ended up

excluding all of these observations, mostly because we excluded all firms with 0 base sales (see Section 2 and 3 for additional discussion);

4. Selection bias is evident, which limits conclusions that can be made from our data analysis. Additional discussion is provided in the next paragraph.

We distinguish two categories of variables for our concern of selection bias. Category 1 includes variables that represent some trait of the firm or entrepreneur that would exist without intervention from Technoserve. For example, Firm Age. Category 2 includes variables that were imposed by Technoserve. An example of this would be advisor gender. We would need some separate information about the population of firms for appropriate evaluation of selection bias for category 1. We would compare firms selected for the Technoserve program to other firms and search for distinctions. We suspect selection bias for category 1 because Technoserve uses several selection criteria for which firms are included. Selection bias for category 2 happens because traits of the firm or entrepreneur are strongly related to treatments imposed by Technoserve. We can evaluate this as the correlation of variables of category 1 and variables of category 2, though this method misses selection bias from variables not included in the data (example: entrepreneur intelligence). Many variables related to category 2 represent something that started after data collection did. For example, advisors did not begin advising until training had ended. We can evaluate selection bias as anything that seems like the effect of a category 2 variable before it existed. For example, in country 1 before training started the firms that would receive business model training earned more sales than firms that would receive business plan training. Technoserve could eliminate selection bias for Category 2 by using a randomized experiment.

Table 1: predictors considered

Predictors	Description
Sex of Entrepreneur (ent_female)	1 indicates firm’s entrepreneur is female, 0 indicates it is male
Sex of Advisor (adv_female)	1 indicates the advisor assigned to a firm was female, 0 if they were male.
Firm Age (firm_age)	Firm age in years
Method (Mod.v.Plan)	1 indicates the firm received business planning training, 0 means they received the traditional business model training. The Business Model is the traditional approach whereas the business planning model is the new approach.

In Training (Training.State1)	1 indicates the firm is currently undergoing Technoserv's training program, 0 indicates the firm has yet to begin training with Technoserv or already finished
Post Training (Training.State2)	1 indicates the firm has completed Technoserv's training program, 0 in both Training.State1 and Training.State2 indicates the firm has yet to begin training with Technoserv
Sectors	Levels: Construction;Food; Furniture/Craft; Industry; Light Manufacturing; Other; Services; Technology; Trade; Agriculture/Biz
Month	Time in months
Country	1, 2, and 3
log base sales	Logged version of the monthly sales, a modified version is used as the 'response' in Model 3 (Section 3)
log average base sales over training status	Logged average version of the monthly sales over training status, used as the response in Model 1,2 (Section 3)
Log Monthly Employees (lemployees)	log(number of workers) which is log(employees + 1), used in Model 3 (Section 3)
Log Average Period Employees (log_avg_period_employees)	The log of the average number of employees during the months of a particular training period (pre/during/post). Used in Models 1 and 2 in Sections 2 and 3

In Table 2, we summarize the missing observations in the dataset. In this report we present three model approaches, and they handle these missing observations in various ways. Models 1 (Section 3.1) and 2 (Section 3.2) use the same approach to analyze the data: Regression model with AR(1) error structure, using the log of the average of the disguised sales over each training state as the response. There are two main differences between Model 1 and Model 2. In Model 2, all missing values were removed before modeling. In Model 1, we only eliminated the observations when base sales = 0. Firm_age was included in Model 2 but not in Model 1 because Model 1 used many observations for which Firm_age was 0. In Model 3

(Section 3.3), log of monthly disguised sales was used as a response, and a multivariate normal and conditional mean formula was used to deal with the missing values and to analyze the data. This is similar to regression, but allows an analysis in the presence of missing observations.

Table 2: Summary of missing values in the data

Predictor	Missing Proportion (%)
ent_age	2.28
firm_age	34.67
basecredit	39.04
base_employees	39.04

2. Exploratory Data Analysis

In this section, we will provide some visualization of the data and discuss what we can (and can't) learn from them.

Figure 1 shows that the male entrepreneurs made the most log(monthly sales) for Country 1 in all the training states (before, during, after). For Country 2, there is a gradual increase in the log monthly sales for the female entrepreneurs that eventually surpasses the males, suggesting that the training may have an effect on the log(monthly sales). For Country 3, male entrepreneurs did better than the female for the first two training state, but by the end of the training the female entrepreneurs are narrowing the gap. We do not have data for the Country 3 after training, so we aren't sure if it would have followed the same trend as Country 2 and Country 3. We can also notice some selection bias on this plot for both Country 2 and Country 3, as the log(monthly sales) didn't start from the same point for both genders. Also, the hypotheses have to do with growth due to the Technoserve program, so the key thing to look at is "during" and "after" to see if there's evidence of a difference in growth between the males and females for these periods.

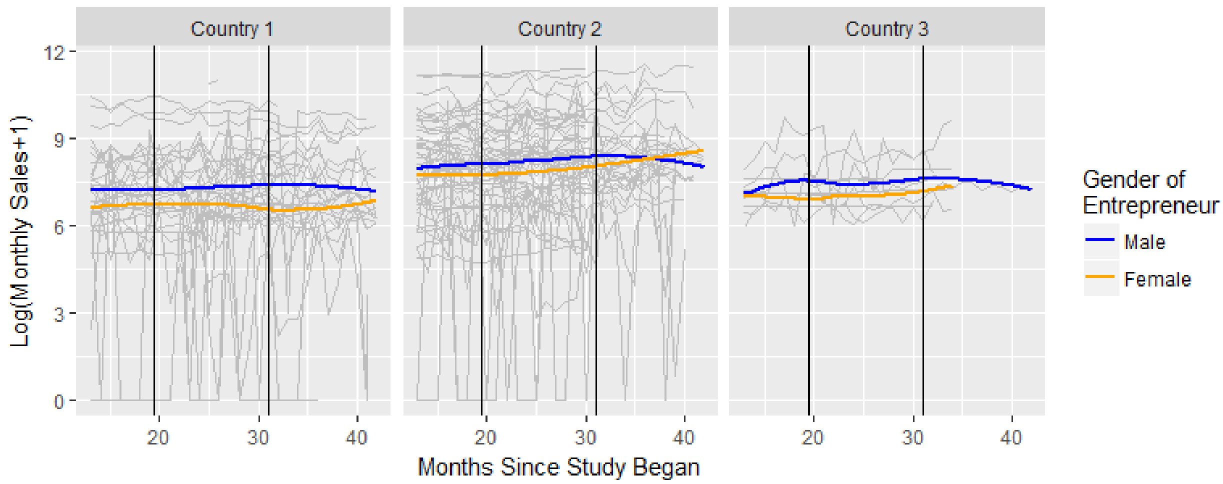

Figure 1: Log(Monthly Sales+1) by Month and Country, with smoothed lines by Gender of Entrepreneur. The vertical lines represent the three different training states for each of the countries: the first line represents the month when the training started and the second line represents the month when the training ended.

Figure 2 shows that training method apparently has a bigger impact for Country 1. After the training begins the log(monthly sales) starts to go up for Business Planning and eventually moves higher than that of the Business Model training, whereas on the other hand as soon as the training begins the Business Model approach produces a downward trend. We can see a similar trend for Country 2. Country 3 shows a somewhat different trend, but since there is little data after training part, it is hard to draw conclusions.

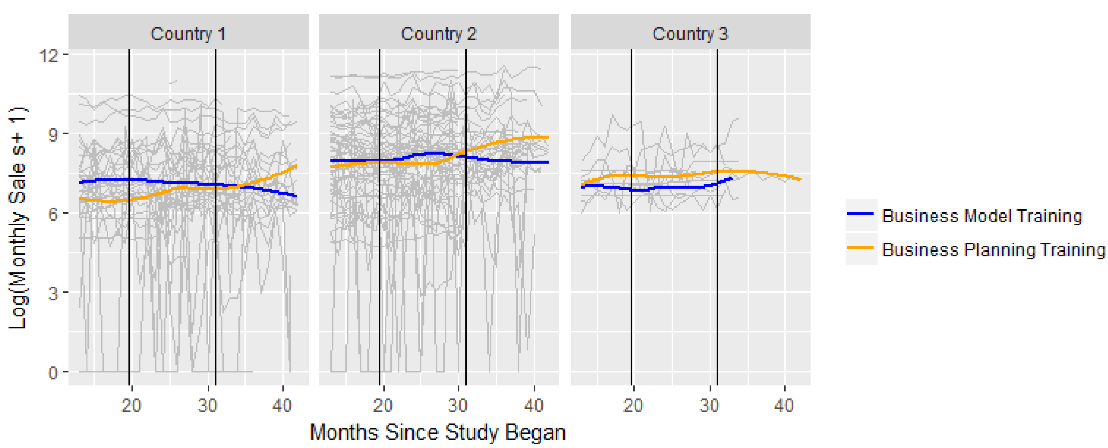

Figure 2: Log(Monthly Sales) by Month and Country, with smoothed lines by Training Method.

Figure 3 shows that the gender of advisor appears to be associated with log(monthly sales). After the training ended, there was an advisor appointed to the entrepreneur, so we can see that for the Country 1, male advisors seem to have a more positive association than the female advisors. For Country 2, male advisors seem to have a small positive association with the log(monthly sales), though a significant amount of data are missing after training. For

Country 3, it is hard to conclude much given the small amount of data. Since the advisor didn't begin interacting with the firm until after training, any differences observed before the training ends is evidence of selection bias.

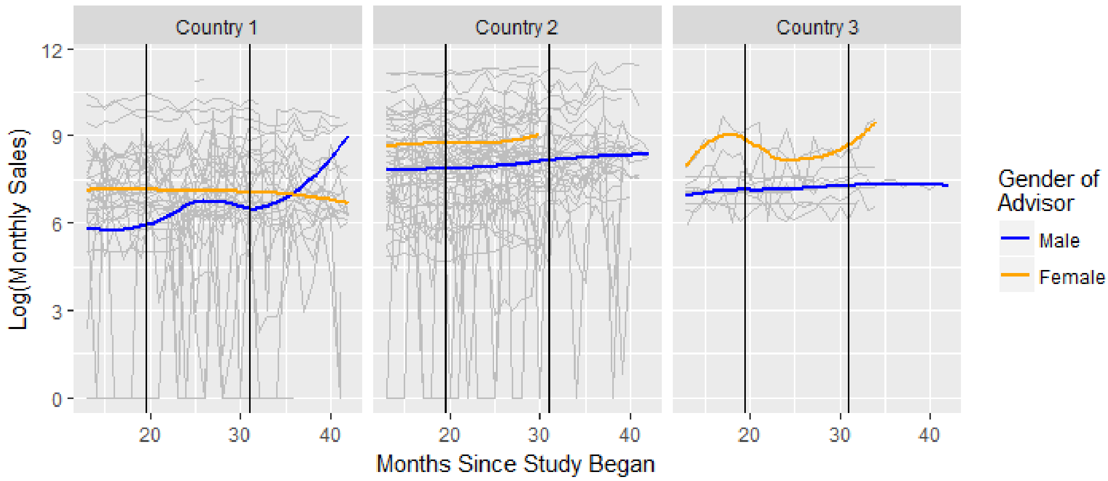

Figure 3: Log(Monthly Sales) by Month and Country, with smoothed lines by gender of advisor.

Figure 4 provides a justification for taking the log of the sales and employees. Most firms have relatively small values of these variables, but a few have large values. Thus, the variables are normalized by taking logs. Hence, we prefer to use the log(disguised sales + 1) and log(employees + 1) for our model, since after taking log of both the variables, they seem to be spread out instead of clustered. We've added 1 to disguised sales because some monthly sales were 0, and added 1 to employees because it then includes the entrepreneur.

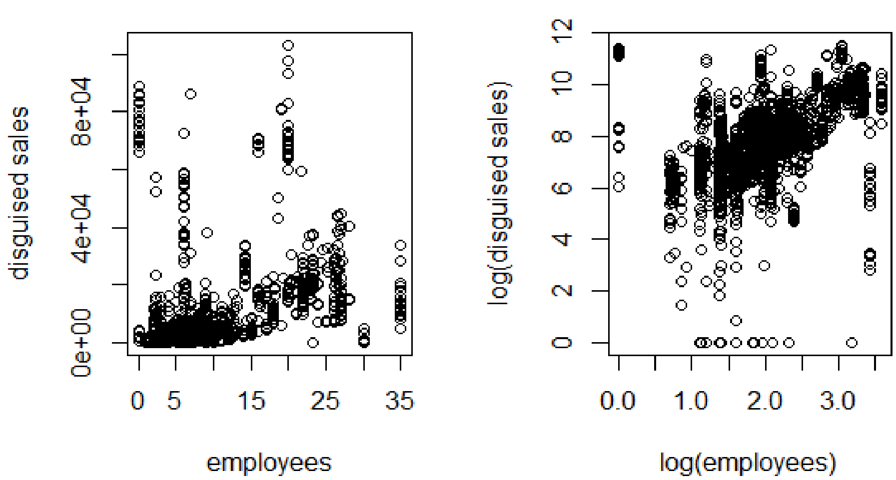

Figure 4: Scatterplot of Employees vs. Disguised Sales and Scatter Plot of log(employees +1) and log(disguised sales +1)

3. Models

This data was quite complicated, so we used several strategies to model the log of disguised sales. The first two models used the log of the average of the disguised sales over each training state, while the third model used log of monthly disguised sales.

3.1 Model without Firm Age as a Predictor and Using Average Within Training States

3.1.1 Data Preprocessing

In this model, we didn't use firm age as a predictor because it had nearly 35% of its observations with "0" (see Table 2), which were new firms that we wouldn't want to consider in our analysis. But of course this means we will not be able to test for the hypothesis related to firm age since firm age isn't included as a predictor in this model. However, because the criterion for the selection of the company was that it should have a revenue of at least $25,000 per year, we did remove the observations with base sales of 0. We simplified the modeling by computing log of the average monthly sales for each training state (before, during, after), instead of modeling the monthly sales directly. This is the same transformation procedure used for Log Average Period Employees. We did this because monthly sales included many 0's which are difficult to model. The downside of this approach is that we lose the information in each month individually. Three hypotheses were considered in this model:

1. Female-led ventures in contexts of poverty will experience less entrepreneurial growth from organizational sponsorship than male-led ventures.
2. Female-advised ventures in contexts of poverty will experience less entrepreneurial growth from organizational sponsorship than male-advised ventures.
3. Business training in contexts of poverty will experience more entrepreneurial growth from organizational sponsorship than business model.

Note that in order for any interpretations of the sort given in the hypotheses, we must assume that there is no selection bias except that which is captured by the log(base sales). Since the experiment was not randomized, we suspect selection bias.

3.1.2 Accounting for correlation within the observations from the same firm

Since we are making regression models using repeated measurements on a firm of log(average sales) in the three training states, it's possible that these measurements within firm are correlated. There are different ways to model this correlation structure. An autoregressive (AR) model predicts future behavior based on past behavior. It's used for forecasting when there is some correlation between values in a time series and the values that precede and follow them. It is basically a linear regression of the data in the current series against one past value in the same series, hence AR(1). In this case, we have three measurements on each firm, so we assume that they are correlated in a way that is captured by the AR(1) structure, and indeed the models improved when we added it.

3.1.3 Modeling

Our model ended up with 13 relevant predictors and an intercept, three of which were two-factor interaction effects. The response variable was the log of the average disguised sales within each training period. Since this is exploratory modeling, we used an informal forward selection process to find a reasonable model. We started with the predictors related to the hypotheses (sex of entrepreneur, sex of advisor, business model vs. business plan), training state, along with interaction terms as sex of entrepreneur*training state, sex of advisor*training state, and business model vs business plan*training state. These interactions are included in order to measure the change in sales across training states. The AIC (Akaike Information Criterion) was used as a criterion to pick the best model. This information-theoretic criterion is commonly used to compare the quality of various models (Bozdogan, 1987). In addition to the 4 predictors and 3 interactions, we added other possible predictors one by one to see if the AIC decreased. We used the rule of thumb that says a 2-point decrease in AIC is enough to warrant including the predictor in the model, whereas if the AIC is not increased by at least 2, we did not keep this predictor in the model (Burnham et al 2002). We repeated the process for all the remaining predictors, and found out that that the AIC decreased when adding log(average period employees), so we decided to add this predictor to our starting model. After adding that term, we again added the predictors one by one to see if the AIC decreased and it decreased for log(base sales disguised); so we decided to add these predictors to our starting model. After making these changes in the model, we further looked at other two-factor interactions in the model to see if the AIC changed by more than 2 points, but none of them did. Results are in Table 3. Note that we did not include p-values to emphasize that the modeling is exploratory and the p-values in such a case are not reliable. However, this model could be a starting point for future analysis of future data. We also do discuss a few p-values relevant to the hypotheses in the next subsection.

Table 3: Predictors Chosen for Model 1

Predictor	**Coefficient**	**Predictor**	**Coefficient**
intercept	0.9283	Mod.v.Plan* Training.State1	-0.0239
adv_female	-0.0995	Mod.v.Plan* Training.State2	0.1568
ent_female	-0.0237	ent_female* Training.State1	-0.1160

log(base sales disguised)	0.6259	ent_female* Training.State2	-0.1226
Mod.v.Plan	0.2194	adv_female* Training.State1	-0.0929
Training.State1	0.1988	adv_female* Training.State2	-0.2042
Training.State2	0.2707	Log_avg_period_ employees	0.2409
Parameter estimate(phi) for AR(1) and residual standard error			
Parameter estimate(phi)	0.349	Residual Standard Error	0.7135

*These coefficients are on a logarithmic scale, so their effects on the response are exponential rather than simply additive.

3.1.5 Conclusion

H1: Female-led ventures in contexts of poverty will experience less entrepreneurial growth from organizational sponsorship than male-led ventures.

For Female-led ventures:

before training : Y*= $\beta_1$ent_female + other terms + error

after training, Y*= $\beta_1$ent_female +β_7Training.State2 + β_{12}ent_female*Training.State2 + other terms + error,

so the difference between after and before is $\beta_7 + \beta_{12}$

For male-led ventures:

before training: ent_female and the training state variables are all 0,

after training : ent_female=0, Training.State1=0, and Training.State2=1,

So the difference between after and before is β_{12}

So if we want to test the difference between male- and female-led firms, in terms of the change between average after and average before training, we simply would test whether β_{12}=0 vs. $\beta_{12} \neq 0$.

In this case, the p-value associated with this parameter was 0.3423, so we do not have sufficient evidence to conclude that the mean change in firm growth differs between male and female-led ventures. Since the hypothesis actually suggests the alternative that female-led firms would grow more slowly, we are interested in a one-sided alternative hypothesis: $\beta_{12}<0$. Since the test to obtain the p-value is based on the t-distribution, which is symmetric, the p-value for this alternative is simply half of the p-value for the two-sided alternative, or approximately 0.1712. Note again that because of the exploratory nature of the study, this p-value may not be valid. It would be more reliable if our suggested model was used from the beginning and tested.

H2: Female-advised ventures in contexts of poverty will experience less entrepreneurial growth from organizational sponsorship than male-advised ventures.

By following the same method as for H1, we could just test whether $\beta_{13} = 0$ vs. $\beta_{13}<0$.

Conclusion: p-value of 0.1461/2=0.0731, so here is some evidence of a difference between male and female-led ventures, though again we provide the caveat that the modeling is exploratory.

H3: Business training in contexts of poverty will experience more entrepreneurial growth from organizational sponsorship than business model.

By following the same method as for H1, we could just test $\beta_8 = 0$ vs. the alternative that $\beta_8 > 0$.

Conclusion: p-value of 0.2512/2=0.1256, we do not have much evidence to conclude that Business Training method leads to more entrepreneurial growth than Business Model.

Based on the model described above, we have little evidence of differential growth due to these three factors, when accounting for possible covariates in the model. This conclusion is based upon the p-values associated with the three variables of interest in the hypotheses.

3.1.6: Response vs Predicted response plot and Residual Plot

The Y vs Y-predicted plot is linear without too much noise, which demonstrates our model fit is reasonably good. The residual plot looks good, except for the bottom 3 points, which are clustered together; when we identified those points, we noticed that these points came from Country 3, which makes sense as we only had a few data points for Country 3.

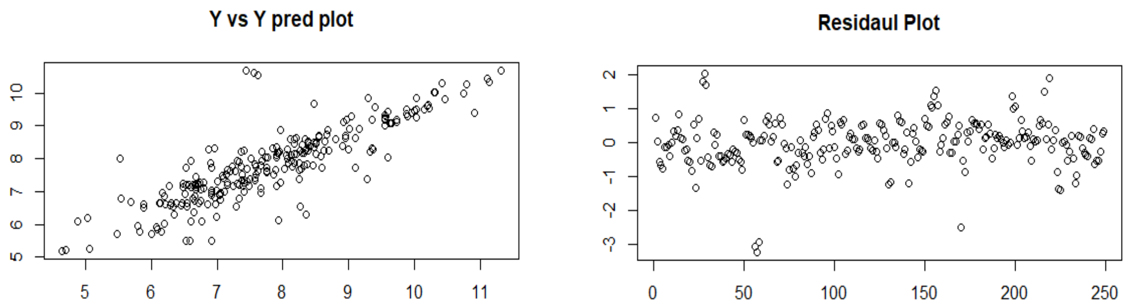

Figure 5: Y vs Y pred plot and Residual plot for the model without the firm as a predictor

3.2 Modeling with Firm Age as a Predictor

3.2.1 Data Preprocessing

Since firm age is included as a variable of interest in the set of hypotheses, we wanted to explore it as a predictor in the model for log(sales). However, there were 739 observations with either missing values or a 0 for firm age. In order to accommodate firm age in the model, we dropped these 739 observations Needless to say, changing the data set changed our exploratory data analysis results. Our goal is to test the same hypotheses as the previous model, but we would also like to test whether firm age has an effect on log average sales for a given period.

3.2.2 Exploratory Data Analysis in the Reduced Data Set

Removing the 739 observations had a large effect on some of our graphical analysis. The general trends across the months for each of the three countries, and the plots with trends for sex of entrepreneur looked similar (see Appendix 1 for the plot of the reduced data set), but when we plotted trend lines for sex of advisor and training method, things looked very different. These plots can be found in Figure 5. When we cut out the observations for which firm age is missing or 0, we lose all of the firms with female advisors for Country 2. The trends stay mostly the same, but the fact that we lost an entire category from one country suggests more selection bias. For the training method, the trend line for Business Planning Training is still increasing for Country 1 and Country 2, but an oscillation is introduced in Country 3, likely due to the small number of firms remaining in the data. The trend lines for Business Model Training are a little steeper, yet roughly the same for Country 1 and Country 2, but the line is completely absent for Country 3. There is no clear explanation for these phenomena, since we do not have a clear explanation for why those firm age values were 0 or missing.

3.2.3: Variable selection

As with the previous model, we used an informal version of forward selection to find worthwhile predictors for the model, and we used AIC as a basis for whether a newly introduced predictor was worth keeping. We ended up with a somewhat messy final model. Like the previous model, we introduced an autocorrelation structure to the model, since there were three repeated measures on each firm. This correlation structure greatly improved AIC.

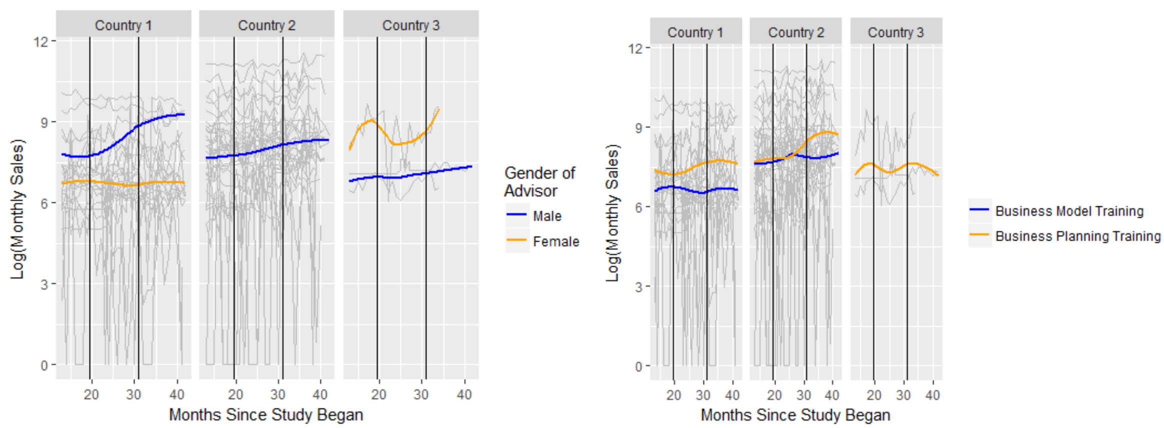

Figure 6: Log(Monthly Sales+1) by Month and Country, with smoothed lines for Advisor Sex and Training Method.

3.2.3 Modeling

Our model ended up with 15 relevant predictors and an intercept, several of which were interaction effects. The response variable was the log of the average disguised sales within each training period. The predictors and their associated coefficients can be found in Table.3. The model would look like a normal multiple linear regression model, where the predicted log of average period sales during a specific training period is equal to the sum of the predictors times their estimated coefficients.

Table 4: Predictors Chosen for Model 2

Predictor	Coefficient	Predictor	Coefficient
intercept	-0.4497	Mod.v.Plan*Training.State1	-0.102
adv_female	-0.1025	Mod.v.Plan*Training.State2	0.1199
ent_female	0.4216	ent_female*Training.State1	-0.1393

firm_age	-0.007	ent_female*Training.State2	-0.0544
Mod.v.Plan	0.3558	adv_female*Training.State1	-0.1797
Training.State1	0.3243	adv_female*Training.State2	-0.4149
Training.State2	0.4842	firm_age*Training.State1	0.0002
log(basesales...disguised)	0.7788	firm_age*Training.State2	-0.0065

*These coefficients are on a logarithmic scale, so their effects on the response are exponential rather than simply additive.

3.2.4: Conclusion

The process by which we tested the hypotheses is analogous to the procedure outlined in Section 3.1.5.

H1: Older ventures in contexts of poverty will experience more entrepreneurial growth from organizational sponsorship than younger ventures.

To test whether the age of a firm has an effect on log average period sales, we need to test whether β_{15}=0, the parameter associated with the firm_age*Training.State2 predictor. The R output used to compute the estimated model parameters also gives a two-sided t significance test, and since the t-distribution is symmetric we can just divide that p-value by two. With a p-value of 0.2037, we have insufficient evidence to conclude that the mean change in firm growth is lower for older firms.

H2: Female-led ventures in contexts of poverty will experience less entrepreneurial growth from organizational sponsorship than male-led ventures.

With a p-value of 0.3775, we have insufficient evidence to conclude that the mean change in firm growth is lower for female-led ventures.

H3: Female-advised ventures in contexts of poverty will experience less entrepreneurial growth from organizational sponsorship than male-advised ventures.

With a p-value of 0.0109, we conclude that the mean change in firm growth is lower for female-advised ventures, reflected by the negative coefficient on adv_female*Training.State2. Any actual difference in these parameters may be explained by selection bias. If this model is used subsequently on new data and firms randomly chosen from the population, the resulting p-value would be more meaningful.

H4: Business training in contexts of poverty will experience more entrepreneurial growth from organizational sponsorship than business model.

With a p-value of 0.2525, we fail to conclude that Business Model training lead to a greater mean change in firm growth than the Business Planning training.

3.3 Linear Regression Model to Predict Monthly Sales Compensating for Seasonality and Missing Data

3.3.1 Data Preprocessing

The purpose of this model is to evaluate firm growth as measured by increasing sales. This model uses interactions of month and other variables as a way of modeling how those variables are associated with firm growth. This model also uses the three time periods (before, during, and after training) as a way of identifying selection bias, since 'effects' before the variable existed are actually selection bias. The predictors for this model were created as functions of the original data instead of being chosen using AIC as in Models 1 and 2. Some of these predictors were formed to directly test one of the client's particular hypotheses. The other predictors were formed to explain nuisance variability and therefore increase accuracy. The analysis produces a coefficient for each predictor variable. The magnitude of the coefficients indicates the size of the effect relative to the hypotheses. The resulting model could be used to test the hypothesis using data from the next several iterations of the Technoserve program.

3.3.2 Modeling

First, some of the data was removed from the complete dataset for this model. The program required $25000 base sales, but there were firms with 0 reported for base sales. So any firm with 0 base sales was excluded from the data for the model. Data pertaining to any months of 0 sales for a specific firm were also excluded. For the remaining monthly sales, the log(sales) was used as the response variable. Monthly sales are right-skewed, so a log transformation is reasonable, except for the fact that there are months with 0 sales. If we created a model where 0 sales were included, a good model would assign a discrete positive probability to 0 sales and a continuous probability distribution for positive sales. This could be done well with two models used together. One model predicts the probability of 0 sales and the other model predicts the sales if positive. It is important to note that modeling the serial correlation between monthly observations would be very important for the model that determines the probability of 0 sales (because we expect 0's to be clustered together), but it turned out to be much less important for the model that determines the sales if positive. Because we do not require predictions, but are only using the model for evaluating the hypotheses, we only need the positive sales model. Thus, we fit a model that omitted the 0 monthly sales.

In order to get a robust seasonality model, we made eight separate regression models for seasonality. We created a categorical variable, monthname, for the name of the month (examples: January, November, June), and also included the following variables: Firmid, month (since January 2013), country, and sector in these models. Interactions were included. One difference among the models was whether they included country and sector. Another difference was whether the seasonality was sinusoidal or categorical. The reason for using eight models is because we do not know what it really is and using all eight together gives a conservative estimate of the seasonal effects. Because the data covers only about three years, we cannot get a good estimate of the seasonality for each firm separately, but we can get a joint estimate of seasonality. The predictors Firmid and month were used in the process of estimating parameters, to account for possible differences among firms, though the seasonality predictions are made without assuming differences among firms. The average predictions of the eight models was used as an estimate of the seasonal effect. We defined the new response variable as log(disguised sales) minus the seasonal effect. In the end, the numerical importance of the seasonality was quite small. Appendix 2 gives a little more information about the seasonality models.

A linear regression-type model was created using the predictors listed in Table 3. The method we used produces results with data sets that include missing values, whereas, ordinary regression would not. To create this model, we used a multivariate normal model where the conditional mean formula was used to calculate the coefficients. This procedure would give the same results as the corresponding regression model if the data did not include missing values. This method works for datasets with missing data, because it uses correlation or covariance matrices to characterize the relationship between the response and the predictors, and these matrices exist even if data is missing. The resulting coefficients are given in Table 3. The standardized coefficients represent the coefficients if we divide each variable by its standard deviation including the response variable. For comparison, a value of 1 or -1 represents a perfect fit in a model with only 1 predictor. The standardized coefficients are somewhat like correlations.

Instead of choosing a model using forward selection as in the previous models, this model was chosen at the outset. It includes log(basesales...disguised) as a way of compensating for the different sizes of each firm. This model also includes lemployees because it is not constant for several firms and it explains some quick increases or decreases of sales. The other terms are required for the hypotheses. There were so many predictors that we didn’t want any more for this model because they could cause overfitting because we have data from only 93 firms for this model. Most variables in the data set, including many used for the model and many others, are highly correlated, so they could ruin the interpretability of our model. The prospect of keeping them constant while changing a predictor of interest seems unlikely.

Table 5: Predictors for this model

Predictor	Coefficient	Standardized Coefficients
Intercept	-6.166322	
log(basesales...disguised)	0.6796672210	0.729665163
lemployees	0.2885190831	0.137723461
Mod.v.Plan1	-0.1502838889	-0.050634452
month*Mod.v.Plan1	0.0018790341	0.017995426
(Month-duringstart)*Mod.v.Plan1*(Training.State1 + Training.State2)	-0.0170837187	-0.074215948
Firm age	0.0002747610	0.002072429
Month* Firm age	-0.0004130205	-0.093610186
adv_female1	-0.3958431141	-0.128210719
month* adv_female1	-0.0036487686	-0.032936869
(Month- afterstart)* adv_female1 * Training.State2	-0.0348212361	-0.047288477
ent_female1	-0.4418373540	-0.149746076
month* ent_female1	0.0093327479	0.087299417
adv_female1* ent_female1	0.8404238528	0.202644865
Month* ent_female1* adv_female1	-0.0168287968	-0.113368139
(Month- afterstart)* ent_female1* adv_female1 * Training.State2	0.0599412258	0.057746028

Month	0069856548	0.037013177
(Month- duringstart)*(Training.State1 + Training.State2)	0.0134825589	0.072425970
(Month- afterstart)* * Training.State2	0.0041953738	0.007919832

This was an exploratory study and the estimated coefficients are small, so we aren't sure that the true coefficients are not zero. We didn't find any strong evidence that the training does anything. Selection bias seems like a substantial part of the effect. If you trust the coefficients' sign or get bigger coefficients with data from future iterations of the program, the interpretation of the coefficients should be as follows.

The variable Month*ent_female1*adv_female1 is important because month indicates that this variable is a measure of firm growth. The other two terms indicate female entrepreneur and advisor. Because the coefficient for this term is negative it suggests that firms with a female entrepreneur that would be assigned a female advisor grew slower than other firms for the time period before being assigned an advisor (because both Training.State1 and Training.State2 are 0). This would indicate selection bias because before the advisor is assigned there is no reason to think that the firms should develop any differently overall.

The coefficient for (Month- afterstart)*ent_female1*adv_female1*Training.State2 is positive which means that female advisors increased firm growth more for firms with female entrepreneurs, because Training.State2 indicates that only the state after training has ended is included in this variable.

The variable, Month*adv_female1, is important because month indicates that this variable is a measure of firm growth. The other term indicates advisor. Because the coefficient for this term is negative it means that firms that would be assigned a female advisor grew more slowly than other firms for the time period before being assigned an advisor. Therefore, this indicates selection bias. Again, we emphasize that this (and other) effects are relatively small so this interpretation is far from definitive.

The coefficient for (Month- afterstart)*adv_female1*Training.State2 is negative which means that female advisors decreased firm growth more. The coefficient for month*ent_female1 is positive which means that firms owned by female entrepreneurs grew faster.

The coefficient for month*Mod.v.Plan1 is positive which means that firms that would receive the business model training grew faster before training started.

The coefficient for (Month- duringstart)*Mod.v.Plan1*(during + after) is negative which means that business model training decreased firm growth more than biz plan.

The coefficient for month* ent_female1 is positive which means that firms owned by female entrepreneurs grew faster.

3.3.3 Simpler analysis for similar interpretations

The method described above for Model 3 is somewhat complicated and can't be implemented easily in a standard statistical software package. Here we present steps that could be taken to implement a method quite similar.

- Ignore seasonality and just use log(sales) as the response variable.
- Fix the data collection procedure so missing data isn't a problem
- Use ordinary regression with all the predictors used in the results above.

The results would be interpreted in the same way as the interpretations above.

4. Conclusion and Discussion

4.1 Model comparison

Table 5 shows that two predictors including female/male-led ventures and Business Model and Business Planning training do not lead to obvious changes in firm growth in both Model 1 and Model 2. However, Model 2 provides some evidence that the mean change in firm growth, as measured by log(sales), differs between male and female-advised ventures; Model 1 has less evidence to that effect.

Table 5: Comparison among models

	Model 1		Model 2	
	P value	strength against null	P value	strength against null
Female/Male-led ventures	0.1712	weak	0.3775	very weak

Female/Male-advised ventures	0.0731	moderate	0.0109	fairly strong
Business Model VS Business Planning	0.1256	weak	0.2525	weak
Venture age	Not included		0.2037	weak

*Model 3 did not generate relevant p-values, so its results could not be compared with the other models in that regard. Instead, one should consult the standardized coefficients in Table 5. One can compare the size of the standardized coefficients of the variables associated with the hypotheses with the standardized coefficients of the log(base sales), for instance, which has a relatively large value. Since the relevant coefficients are so small, none of the predictors appeared to offer very strong evidence for any of the hypotheses. We do note that Model 3 suggested that female-advised ventures grow more slowly, as do the first two models. However, for the entrepreneur gender and business model vs. business planning, Model 3 has opposite effect directions. This could be due to many things, including the small, noisy estimates of the parameters in the models.

4.2 Discussion

In the model using the full data set (Model 1), there is little evidence of differences between Female/Male-led ventures, Female/Male-advised ventures and Business Model, Business Planning for venture growth respectively. Firms owned and advised by both sexes did not change the venture growth significantly, and both training methods yielded similar quality of sales. The log of base sales and the log average number of employees during a given training state seemed to be the only variables that had a large effect on log average sales during a training period, implying the size of a firm was what really mattered for determining sales.

The model using the reduced data set (Model 2) only had one of these measures of company size, log base sales, and the sector from which a firm comes also had a significant effect on log average sales during a training period. More interesting though is that the sex of the entrepreneur and firm age, as well as training method also proved nonsignificant with the reduced data set. However, there was some evidence that female-advised firms grew more slowly than male-advised firms, though this may be due to selection bias. Firm age ended up not being significant in the model, but we still had to reduce the data set to even consider it in the first place, since there were a lot of missing or nonsensical 0 values that could throw off the analysis.

For Model 3, which modeled monthly sales instead of average monthly sales over the three training states, there doesn’t appear to be strong conclusions possible regarding the hypotheses, though the effect direction for advisor gender is the same as in Models 1 and 2.

Throughout the study we had to alter the data set numerous times due to missing data and observations with 0 for a variable that should not have been 0, such as base sales. The data was messy, and results could have differed significantly if we had cleaned it differently or if we had excluded more or fewer observations. Briefly, the conclusion we got from either Models 1 or 2 may be explained by selection bias considering these two models aren't able to separate selection bias from factor effects. Model 3 does attempt to separate these components, and indicates that none of the predictors appeared to offer very strong evidence for any of the hypotheses, considering the small, relevant coefficients (Table 5).

If Technoserv wishes to conduct more studies like this in the future, we have two major suggestions.

- The first, as one might guess after reading the previous paragraph, is an overhaul of their data collection and recording process. We realize the ideal goal of having a data set without any missing values is nearly impossible, but the current dataset had so much missing data that analysis was very difficult. The original data set had 3,683 observations. Filtering out observations where base sales were 0 brought us down to the 2,245 observations of the "full" data set. That's 1,438 valuable points of data lost because of one variable. We also lost 739 more observations because of firm age, and we may have lost more if we carried out Dr. Sutter's original goal of testing whether entrepreneur age was significant as well. We know there might be some cultural taboos and response bias that prevents complete accuracy in these responses, but improvement in the data collection process will allow more meaningful results.
- Second, Technoserv needs to do a completely randomized experiment. We understand total randomization could be a financial risk, but as it stands these results are only applicable to the firms Technoserv gathered data from. We do not encourage using either of the models provided to predict how a new firm might perform during training, since they would likely not fit into the exact framework built around the firms considered in the current study. In an ideal world, Technoserv would randomly assign training and aftercare combinations based on sex of the entrepreneur, sector, country, and maybe even age range, but if that would be too costly, they still might be able to get some meaningful results just by randomizing among all of their firms.

If Technoserv follows both of these points of advice, we believe they could get some meaningful results that could be used to great benefits for the countries in which these firms operate. Helping these third-world economies is a noble goal, and we would like them to find great success in that regard.

Appendix 1

Model 2:
H1: Venture age will experience less entrepreneurial growth gained from organizational Sponsorship.

- Firms before based on age: $\beta_3$firm_age
- Firms after based on firm age: $\beta_3$firm_age + β_6Training.State2 + β_{15}firm_age*Training.State2
- Y* = β_{15}firm_age*Training.State2

- Test H_0: $\beta_{15} = 0$
- P-value: 0.2037

H2.1 Female-led ventures in contexts of poverty will experience less entrepreneurial growth from organizational sponsorship than male-led ventures.

- Female entrepreneur before: $\beta_2$ent_female
- Female entrepreneur after: $\beta_2$ent_female + β_6Training.State2 + β_{11}ent_female*Training.State2
- Y* = β_6Training.State2 + β_{11}ent_female*Training.State2

- Male entrepreneur before: 0
- Male entrepreneur after: β_6Training.State2
- Y* = β_6Training.State2

- Test H_0: $\beta_{11} = 0$
- P-value: 0.3775

H2.2 Female-advised ventures in contexts of poverty will experience less entrepreneurial growth from organizational sponsorship than male-advised ventures.

- Female advisor before: $\beta_1$adv_female
- Female advisor after: $\beta_1$adv_female + β_6Training.State2 + β_{13}adv_female*Training.State2
- Y* = β_6Training.State2 + β_{13}adv_female*Training.State2

- Male advisor before: 0
- Male advisor after: β_6Training.State2
- Y* = β_6Training.State2

- Test H_0: $\beta_{13} = 0$
- P-value: 0.0109

H3:Business training in contexts of poverty will experience more entrepreneurial growth from organizational sponsorship than business model.

- Business planning before: β_4Mod.v.Plan
- Business planning after: β_4Mod.v.Plan + β_6Training.State2 + β_9Mod.v.Plan*Training.State2
- Y* = β_6Training.State2 + β_9Mod.v.Plan*Training.State2

- Business model before: 0
- Business model after: β_6Training.State2
- Y* = β_6Training.State2

- Test H_0: β_9
- P-value: 0.2525

Appendix 2: Model 3 seasonality models

The seasonality models all include Firmid and month and the interaction of Firmid and month and month name. Some treat monthname as categorical; others use a sine/cosine curve. Some include sector and the interaction of sector and monthname. Some include country and the interaction of country and month name.

References:

Bozdogan, Hamparsum. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, 52 (1987), pp. 345-370.

Burnham, K. P., and Anderson, D.R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed. Springer-Verlag. ISBN 0-387-95364-7.