

Do Cooperatives Benefit Small Guatemalan Coffee Farmers?

The Competitive Role of Guatemalan Coffee Cooperatives on Farm-Gate Coffee Prices

Seth B. Morgan
University of San Francisco
2008

***Abstract-** Economic theory avers commodity markets are typically comprised of a few intermediary purchasers which, in oligopsony, offer a below competitive price to the often numerous and poor producers of the good. Cooperative economic theory, however, establishes that marketing cooperatives, which are able to offer higher prices through collective bargaining, can act as “competitive yardsticks” in commodity markets forcing colluding private intermediaries to raise their price. In effect, the cooperative can exert a pro-competitive role on the market increasing the economic welfare of the small producers. Although this theory is widely accepted as accurate, it has never been tested empirically. The research presented below does not attempt to tackle such a large task, at least not directly. What it does endeavor to do is, through a series of regressions, to model cooperative and private intermediary coffee pricing behavior as well as the determinants of coffee cooperative membership in San Pedro Necta, Huehuetenango, Guatemala. The results, together with pertinent qualitative data, are then used to formulate educated conjectures to suggest the degree to which one would expect coffee cooperatives to play a pro-competitive role on farm-gate coffee prices in western Guatemala.*

Introduction & Research Question

Nearly 70% of all coffee is grown on farms of less than 10 hectares. The numerous farmers are typically poor and often incur market-limiting obstacles such as weak bargaining power, inaccessible credit, asymmetric information, and poor infrastructure. Cooperative economic theory suggests that local purchasers to whom small producers sell their commodities (coffee in this case) are typically a few intermediaries with substantial market power. The barriers to entry into the intermediary coffee market would allow a few colluding buyers to either enjoy monopolistic or oligopsonistic control over the purchasing of coffee produced by the small farmers. The price offered to the producer is therefore assumed to be below what the market price would be under perfect competition.

It is further theorized that the existence of coffee cooperatives can, to some degree, restore competitiveness to the market. Cooperatives are able to offer stronger lobbying power, access to market information as well as direct purchasing agreements with foreign importers translating into higher prices for its member producers. As a result, it is speculated that private intermediaries would be forced to offer a higher price to non-cooperative producers in areas where cooperatives operate so as not to lose market share, thus to some degree, restoring competitive prices to the market increasing the economic welfare of the its producers.

The principal goal of the presented research is to provide empirical evidence with which to suggest if Guatemalan coffee cooperatives are expected to have played or to continue to actively play a pro-competitive role on farm-gate coffee pricing (i.e. is the existence of the cooperatives forcing intermediary purchasers to raise their price towards one that would be seen under perfect competition). Certainly, to statistically prove that the coffee cooperatives in Guatemala have caused or are currently causing an upward trend in private intermediary pricing, price data from cooperative and non-cooperative members alike would need to be collected over at least a ten year period. Unfortunately, no such data exists given the nature of the commodity’s production (principally by small,

impoverished farmers) and the lack of record keeping by private buyers. Also, an extremely sophisticated spatial econometric model would need to be employed to determine the reach and intensity of cooperative pricing influence per region under observation. This is beyond the means of the present research. However, the research will endeavor to provide reasoned judgments as to what one would expect to be observed from research that would directly model the pro-competitive role of Guatemalan coffee cooperatives.

The presented research attempts to accomplish this task by offering a series of regressions modeling prices and cooperative membership. The first regression utilizes panel price data from 2005 to 2007 to ascertain if cooperatives are indeed offering higher prices to its members versus the prices private buyers offer outside the cooperative. The results, combined with collected qualitative field data, will then be used to make suggestions as to whether, one, oligoposny is present at the farm-gate level (i.e. is a below competitive price actually being offered by private intermediaries) and, two, if the cooperative is offering a higher price, should one anticipate a rise in the private intermediary price over time (i.e. is the cooperative playing a pro-competitive role on farm-gate coffee prices). The second regression will look to model the determinants of cooperative membership. During the data collection period of this research, it was passively observed that the smallest and poorest farmers typically were not cooperative members like their small to medium size coffee producing cohorts. If farm size or other variables are seen to be significant in determining cooperative membership, such a results can give further insight as to whether a higher cooperative price is likely to drive up the private intermediary price over time.

Methodology

It would have been ideal to conduct the research throughout all coffee growing regions of Guatemala. However, due to time and funding constraints, the study has been confined to the municipality of San Pedro Necta in the department of Huehuetenango, Guatemala¹. This region was selected for the large amount of coffee grown in the area, the heterogeneity of its high quality coffee² and the observed population of both cooperative-member and non-member small coffee producers.

The data collected was gathered firsthand by interviewing both cooperative members and non-members directly through field surveys. In regards to establishing a random sample, three communities were randomly selected from the municipality of San Pedro Necta: Jolimex/ Tuiscuz, Los Coles and the town of San Pedro Necta itself. The goal was to establish a complete list of all coffee producers within the communities from which to randomly select those farmers to be interviewed. This was a difficult task given the nature of coffee production in Guatemala. No such list of coffee producers exists on

¹ The department of Huehuetenango (7,403 square kilometers) is located in Region VII of Guatemala with the departmental head of the city of Huehuetenago. It shares its northern and western borders with Mexico. To east lies the department of El Quiche while the departments of San Marcos, Quetzaltenango and Totonicapan border it to the south.

² This allows us to be free of controlling for differences in quality as all coffee from this region is deemed Strictly Hard Bean (SHB), the highest classification of quality based on altitude.

file within the municipality or within the communities themselves. As such, local leaders such as auxiliary mayors and other community leaders with intimate knowledge of their populations were convoked in order to enumerate such lists.

This methodology proved successful in a few communities but impractical in others. Thus, in an effort to expand and strengthen the sample population, other methodologies were implemented including geographic as well as the blanket sampling of neighboring communities. As such, although the sample frame cannot be deemed a true random sample, it can be considered a quasi-random sample without imposing obstinate limitations on the conclusions of the research.

The Pricing Model: Regression I

In order to determine if the existence of cooperatives in western Guatemala does indeed force private intermediary buyers to raise their price, farm-gate price data would need to be collected over a substantial time period. At least ten years of data would be required in order to have enough variation to accurately observe price trends. Unfortunately, no such data exists as transactions between farmers and private buyers are not recorded and are often not accurately archived on the farmer level within cooperatives themselves. Even if a surveyed producer offered such data for prices she received ten years ago, such data (without accurate records) would not be reliable. Naturally, such data could presently start to be collected over a ten-year period but is much beyond the means of this project. Also, contrary to the economic theory presented above, the existence of only a few private buyers on the first level of the value chain was not observed to be true in San Pedro Necta. In fact, there are numerous private buyers at the farm level. Given that there is no licensing scheme to track who is buying and selling coffee at this level, it is impossible to determine how many buyers there indeed are as well as impossible to track entry and exit of these buyers in and out the market. Such information would be needed along with an intricate spatial econometric model in order to be able to directly model the pro-competitive role western Guatemala's coffee cooperatives play on farm-gate coffee prices.

As such, the first half of this research will employ a simpler pricing model from which inferences can be drawn (supported by qualitative data) about the pro-competitive role of cooperatives in San Pedro Necta, Huehuetenango, Guatemala. The main focus will be to see if selling to a cooperative or association³ does fetch a higher price or not. If the coefficients of these dummy variables come out to be very close to zero or insignificant (i.e. a farmer receives the same price for her coffee inside and outside of the cooperative), it can perhaps be concluded that a competitive price has already been established in the region. If the coefficients come out to be large, positive and significant, then with supporting quantitative and qualitative data, educated conclusions can be formulated as to whether the private price should increase in the future (i.e. the cooperative is actively playing a competitive role on farm-gate coffee prices) or whether there are other particular reasons as to why a private intermediary could continue to offer a lower price.

The data collected spans three years (2005- 2007) though it must be noted that prices were recorded transaction by transaction within each year⁴. As such, the model employs panel data divided crosssectionally by farmer and in time by year (technically, by

³ Note that sales to a cooperative and to an association will be controlled for separately. Although the structure and function of the two economic entities are very similar, an association typically engages in direct exporting allowing it to capture more value added and, therefore, gives them the luxury of offering a higher price to its members.

⁴ This is a key characteristic of the first model. Whether the farmer sold to a cooperative/ association or not is controlled for by transaction. As such, the first model does not control for cooperative/ association membership directly although it is obvious that if a farmer sold coffee to a cooperative/ association, she is a member of that institution by requirement. The model must be formulated in this way as a cooperative member can sell coffee within and without of the cooperative during the same harvest year.

each transaction within each year). The initial model to be used is therefore the following⁵:

$$\text{price}_{it} = \beta_0 + \beta_1 \text{quantsold}_{it} + \beta_2 \text{coop}_{it} + \beta_3 \text{asoc}_{it} + \beta_4 \text{y2006}_{it} + \beta_5 \text{y2007}_{it} + a_i + u_{it}$$

Dependent Variable:

price_{it}- Farm-gate price received by the producer from the purchaser (private intermediary, cooperative or association) measured in Quetzales (Q), the currency of Guatemala⁶.

Independent Variables:

quantsold_{it}- the quantity of coffee sold by the farmer measured in quintals, the local measurement representing 100kg. The quantity sold by the producer is hypothesized to positively affect the farm-gate price received. The data suggests that medium size producers are more likely to sell to a cooperative from which it is, at least at the outset, assumed they receive a higher price.

coop_{it}- a dummy variable indicating if the coffee of the respective transaction was sold to a cooperative or not (a value of 0 if no, 1 if yes). Congruent with economic theory, selling to a cooperative is at the start hypothesized to positively affect price received (i.e. a farmer should expect a higher price from a cooperative).

asoc_{it}- a dummy variable indicating if the coffee of the respective transaction was sold to an association or not (a value of 0 if no, 1 if yes). Congruent with economic theory, selling to an association is initially hypothesized to positively affect price received (i.e. a farmer should expect a higher price from an association).⁷ It is also expected that the coefficient will be larger than the coefficient on *coop* given direct exporting.

y2006_{it}- a dummy variable indicating if the coffee of the respective transaction was sold in the year of 2006 or not (a value of 0 if no, 1 if yes). Given that commodity coffee prices in the New York Stock Exchange⁸ have been consistently rising over the past few years, selling coffee in 2006 should fetch a higher price versus selling the same quantity in 2005, our base year. The coefficient is thus expected to be positive.

y2007_{it}- a dummy variable indicating if the coffee of the respective transaction was sold in the year of 2007 or not (a value of 0 if no, 1 if yes). Again, by the same logic as above

⁵ Note that more key variables (introduced below) will be included in the model specification section.

⁶ 1 USD = Q7.6

⁷ Note that if *coop* and *asoc* both take on values of 0, then the base outcome is that it was sold to a private buyer.

⁸ The New York Stock Exchange handles the trade of Arabica coffee, the type grown in Guatemala, while the London Stock Exchange handles Robusta.

a sale in 2007 should fetch a higher price for the same quantity sold in 2005, our base year. The coefficient is again expected to be positive and larger than that for y_{2006} .

size_i- the size of the farmer's plot of land dedicated to coffee production measured in cuerdas. This variable is assumed to be time invariant for this model. As mentioned before, medium sized growers have been observed to more likely be cooperative members (where a higher price is offered). A positive coefficient is therefore expected.⁹

α_i - time invariant component associated with panel data.

*'Service' Dependent Variables*¹⁰:

contract_{it}- a dummy variable controlling for if the transaction was carried out by a previously signed contract (a value of 0 if no, 1 if yes). It is undetermined as to what the sign of the related coefficient should be.

advance_{it}- a dummy variable controlling for if the transaction was accompanied by an earlier cash advance (a value of 0 if no, 1 if yes). The coefficient on this variable is hypothesized to be negative given that this a service provided at a cost to the buyer.

payment_{it}- a dummy variable controlling for if full payment was immediately received by the farmer upon handing the produce over to the buyer (a value of 0 if no, 1 if yes). The coefficient on this variable is hypothesized to be negative given that buyers offering this service should be able to offer a lower price than buyers not offering the service.

pickup_{it}- a dummy variable controlling for if the coffee of the transaction was picked up by the buyer directly from the producer's farm or not (a value of 0 if no, 1 if yes). Like above, the coefficient on this variable is hypothesized to be negative given that buyers offering this service should be able to offer a lower price than buyers not offering the service.

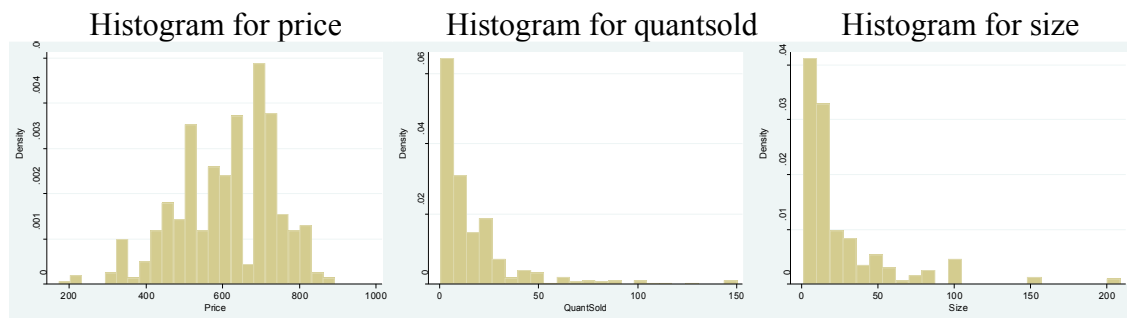
⁹ This variable is not included in our initial model given that it is suspected to introduce multicollinearity if indeed farm size (and therefore quantity sold) is highly correlated with cooperative/ association membership indirectly captured by our *coop* and *asoc* dummy variables. Its inclusion will therefore be under the model specification section.

¹⁰ The dummy variables controlling for the services offered by the buyer are not included in the initial model for similar reasons as *size*. According to the data collected, only cooperatives and associations offer contracts for sales (and only by contract) while private intermediaries never do. Similarly, only cooperatives/ associations offer cash advances and never offer full payment upon turning in the coffee to the institution. Private buyers, conversely, never offer cash advances and always pay in full immediately upon purchasing the coffee. Finally, cooperatives never offer to pick up the produce from the farm directly while private buyers may or may not do so. Therefore, whether the producer received these services per transaction is hypothesized to be indirectly controlled for by the *coop* and *asoc* dummies. Introducing them is likely to introduce strong multicollinearity and is best left to model specification.

Summary of hypotheses¹¹:

<u>Variable</u>	<u>Coefficient</u>	<u>Alternative Hypothesis</u>	<u>Expected Sign</u>
quantsold _{it}	β_1	>0	+
coop _{it}	β_2	>0	+
asoc _{it}	β_3	>0	+
y2006 _{it}	β_4	>0	+
y2007 _{it}	β_5	>0	+
size _i	β_6	>0	+
contract _{it}	β_7	?	?
advance _{it}	β_8	<0	-
payment _{it}	β_9	<0	-
pickup _{it}	β_{10}	<0	-

Data Transformations



The dependent variable, price, is observed to be somewhat negatively skewed while both quantsold and size are positively skewed. Therefore, price is squared while both quantsold and size are logged to correct the skewness.

```

-----
stats |      price  price_2
-----+-----
skewness | -.4911415  -.013675
-----

stats | quants~d  ln_qua~d
-----+-----
skewness |  3.111915  -.057937
-----

stats |      size  ln_size
-----+-----
skewness |  2.676181  .0164043
-----

```

¹¹ Note that the null hypotheses are that each respective coefficient is equal to zero.

The initial OLS pricing model regressed with the data transformations:

```
. regress price_2 ln_quantsold coop asoc y2006 y2007
```

Source	SS	df	MS	Number of obs	=	540
Model	3.0170e+12	5	6.0339e+11	F(5, 534)	=	36.58
Residual	8.8087e+12	534	1.6496e+10	Prob > F	=	0.0000
Total	1.1826e+13	539	2.1940e+10	R-squared	=	0.2551
				Adj R-squared	=	0.2481
				Root MSE	=	1.3e+05

price_2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ln_quantsold	34807.24	5167.753	6.74	0.000	24655.62 44958.86
coop	75412.82	12935.73	5.83	0.000	50001.65 100824
asoc	152265.9	33638.77	4.53	0.000	86185.38 218346.5
y2006	30179.72	14507.45	2.08	0.038	1681.048 58678.39
y2007	72803.77	14065.68	5.18	0.000	45172.91 100434.6
_cons	249639.7	14645.33	17.05	0.000	220870.1 278409.2

Given that the coefficient of a logged variable is difficult to interpret when regressed by a squared dependent variable, the regression of simply price on the same dependent variables will be run as well.

```
. regress price ln_quantsold coop asoc y2006 y2007
```

Source	SS	df	MS	Number of obs	=	540
Model	2269577.7	5	453915.54	F(5, 534)	=	38.55
Residual	6288079.52	534	11775.4298	Prob > F	=	0.0000
Total	8557657.22	539	15876.9151	R-squared	=	0.2652
				Adj R-squared	=	0.2583
				Root MSE	=	108.51

price	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ln_quantsold	30.50881	4.366203	6.99	0.000	21.93177 39.08585
coop	61.03173	10.92932	5.58	0.000	39.56199 82.50146
asoc	114.3905	28.42119	4.02	0.000	58.5594 170.2215
y2006	35.70379	12.25725	2.91	0.004	11.62544 59.78214
y2007	74.57854	11.88401	6.28	0.000	51.2334 97.92369
_cons	483.1662	12.37375	39.05	0.000	458.859 507.4734

Since the t-stats, the significance level of the coefficients and the goodness of fit are nearly unaffected and also given that *price* is not drastically negatively skewed, the second model (without squaring price) will be used in order to directly interpret the marginal effects of the independent variables on price.

Endogeneity

Theory would suggest that *price*, the dependent variable, would have explanatory power over *quantsold*, the first independent variable. Obviously, holding all else constant, a

farmer would choose to sell more coffee to a buyer offering a higher price. This implies that *quantsold* may be endogenous introducing bias into the model.

The Regression Test:

The variable *size* has been chosen as the best instrumental variable from the data available. Farm size should be highly correlated with the quantity produced and, therefore, quantity sold. Farm size should also not be correlated with the offered price¹².

```
. reg ln_quantsold ln_size coop asoc y2006 y2007
. predict fsresid, residuals
. reg price ln_quantsold fsresid coop asoc y2006 y2007
```

Source	SS	df	MS	Number of obs =	485
Model	2278956.82	6	379826.137	F(6, 478) =	32.36
Residual	5610610.73	478	11737.6794	Prob > F =	0.0000
				R-squared =	0.2889
				Adj R-squared =	0.2799
Total	7889567.55	484	16300.7594	Root MSE =	108.34

	price	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ln_quantsold		135.4449	31.89704	4.25	0.000	72.76919 198.1207
fsresid		-108.2369	32.29816	-3.35	0.001	-171.7008 -44.77298
coop		-37.69656	31.95166	-1.18	0.239	-100.4796 25.0865
asoc		-23.32098	54.97959	-0.42	0.672	-131.3525 84.71057
y2006		30.17785	13.35328	2.26	0.024	3.939475 56.41623
y2007		55.46091	14.99475	3.70	0.000	25.99713 84.92469
_cons		310.4784	51.96977	5.97	0.000	208.361 412.5959

Notice that when we run price on the predicted residuals from the regression of *ln_quantsold* (along with all the other exogenous variables), its coefficient comes out to be highly significant (at a 1% level) indicating that the variable of quantity sold is endogenous. Further notice, however, that when the Two Stage Least Squares/ Instrumental Variable regression (presented below) is run with *size* as the IV, our coefficients on *coop* and *asoc* become extremely insignificant and the coefficient values drop drastically (the coefficients become negative in fact) negating any significant explanatory power of these two key dummy variables.

```
. reg ln_quantsold ln_size coop asoc y2006 y2007
. predict ln_quantsoldhat
. reg price ln_quantsoldhat coop asoc y2006 y2007
. ivreg price (ln_quantsold=ln_size) coop asoc y2006 y2007
```

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs =	485
Model	-3042480.57	5	-608496.115	F(5, 479) =	16.64
Residual	10932048.1	479	22822.6474	Prob > F =	0.0000
				R-squared =	.

¹² One may argue that a trend of price increases over a few years would induce an increase in farm size. However, the sample data shows that from 2005 to 2007 farm size dedicated to coffee production was invariant.

```
-----+-----
Total | 7889567.55 484 16300.7594
Adj R-squared = .
Root MSE = 151.07
```

```
-----+-----
price | Coef. Std. Err. t P>|t| [95% Conf. Interval]
-----+-----
ln_quantsold | 129.0416 41.84452 3.08 0.002 46.82011 211.2631
coop | -33.27244 42.84582 -0.78 0.438 -117.4614 50.91654
asoc | -20.3502 75.65472 -0.27 0.788 -169.0063 128.3059
y2006 | 31.10301 18.52828 1.68 0.094 -5.303742 67.50977
y2007 | 58.00726 20.34885 2.85 0.005 18.02323 97.9913
_cons | 321.3167 68.10676 4.72 0.000 187.4918 455.1417
-----+-----
```

```
Instrumented: ln_quantsold
Instruments: coop asoc y2006 y2007 ln_size
-----+-----
```

Looking at the data set, it can be seen that medium size farmers are more likely to join cooperatives/ associations than their smaller producing peers¹³. If so, including *size* in the model may be introducing multicollinearity.

The Hausman Test for Endogeneity:

Given that there is no better IV available than *ln_size*, the Hausman test for endogeneity will be run to verify the results of the regression test for endogeneity.

```
. regress price ln_quantsold coop asoc y2006 y2007
. estimates store ols
. ivreg price (ln_quantsold=ln_size) coop asoc y2006 y2007
. estimates store iv
. hausman iv ols
```

```
-----+-----
---- Coefficients ----
| (b) (B) (b-B) sqrt(diag(V_b-V_B))
| iv ols Difference S.E.
-----+-----
ln_quantsold | 129.0416 30.50881 98.53281 41.61611
coop | -33.27244 61.03173 -94.30417 41.42842
asoc | -20.3502 114.3905 -134.7407 70.11328
y2006 | 31.10301 35.70379 -4.600776 13.89449
y2007 | 58.00726 74.57854 -16.57128 16.51805
-----+-----
```

```
b = consistent under Ho and Ha; obtained from ivreg
B = inconsistent under Ha, efficient under Ho; obtained from regress
```

```
Test: Ho: difference in coefficients not systematic
```

```
chi2(5) = (b-B)' [(V_b-V_B)^(-1)] (b-B)
= 6.17
Prob>chi2 = 0.2904
```

The null is not rejected indicating (according to this test), the log of quantity sold is not endogenous. Given the aforementioned problems with the preferred IV, *ln_size*, the

¹³ This will be formally tested in the second regression.

results of the regression test for endogeneity will be ignored in favor of the Hausman results.

Panel Data Issues

As this data is a panel data set, the time invariant error component, represented by a_i , of the model introduces bias into the OLS model if the correlation between it and the other exogenous variables is not equal to zero. Even if it is equal to zero, the time invariant error component causes the model to suffer from serial correlation. Random effects would be the preferred technique in correcting for serial correlation as random effects does not remove one time period from the data set (a serious concern here given that the data set is across only three years). This also means that time invariant variables could still be used if necessary. Random effects, however, does not correct for any correlation between a_i and the exogenous variables. Therefore, a Hausman test must be run in order to determine if random effects (versus a fixed effects technique) will leave bias in the model due to the time invariant error component.

```
. xtreg price ln_quantsold coop asoc y2006 y2007, fe
. estimates store fe
. xtreg price ln_quantsold coop asoc y2006 y2007, re
. estimates store re
. hausman fe re
```

---- Coefficients ----				
	(b) fe	(B) re	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
ln_quantsold	32.82212	30.29641	2.52571	8.859473
coop	31.72964	43.85068	-12.12104	8.165101
asoc	89.97115	102.4243	-12.45317	21.56072
y2006	45.12452	43.11829	2.006228	1.845434
y2007	79.61159	78.6829	.928693	2.252264

b = consistent under H_0 and H_a ; obtained from xtreg
 B = inconsistent under H_a , efficient under H_0 ; obtained from xtreg

Test: H_0 : difference in coefficients not systematic

chi2(5) = (b-B)' [(V_b-V_B)^(-1)] (b-B)
 = 4.38
 Prob>chi2 = 0.4967

The null of the Hausman test is not reject indicating that random effects will not be biased and, in fact, can be shown to be more efficient in its estimates than fixed effects. Therefore, random effects will be the technique employed for this model.

```
. xtreg price ln_quantsold coop asoc y2006 y2007, re
```

```
Random-effects GLS regression           Number of obs   =       540
Group variable (i): id                 Number of groups =       143

R-sq:  within = 0.1997                 Obs per group:  min =        1
        between = 0.2612                                     avg  =       3.8
        overall = 0.2616                                     max  =        9

Random effects u_i ~ Gaussian           Wald chi2(5)     =    150.64
corr(u_i, X) = 0 (assumed)             Prob > chi2      =     0.0000
```

	price	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ln_quantsold		30.29641	5.660601	5.35	0.000	19.20184	41.39099
coop		43.85068	11.39251	3.85	0.000	21.52177	66.1796
asoc		102.4243	28.66639	3.57	0.000	46.23923	158.6094
y2006		43.11829	9.624325	4.48	0.000	24.25496	61.98162
y2007		78.6829	9.484422	8.30	0.000	60.09377	97.27202
_cons		482.3793	14.80602	32.58	0.000	453.36	511.3985
sigma_u		72.288					
sigma_e		82.813125					
rho		.4324513	(fraction of variance due to u_i)				

Model Specification

Although the goodness-of-fit presented in the above random effects model is not terribly low (around .26), it will be of use to include the service variables in a model specification manner (i.e. one at a time) to see if the model with those appropriate service variables can predict more of the variation in price received. Due to suspect multicollinearity, these variables were not included in the original model (see footnote 9 above).

```
. xtreg price ln_quantsold coop asoc y2006 y2007 contract, re
```

```
Random-effects GLS regression           Number of obs   =       540
Group variable (i): id                 Number of groups =       143

R-sq:  within = 0.1999                 Obs per group:  min =        1
        between = 0.2632                                     avg  =       3.8
        overall = 0.2645                                     max  =        9

Random effects u_i ~ Gaussian           Wald chi2(6)     =    151.11
corr(u_i, X) = 0 (assumed)             Prob > chi2      =     0.0000
```

	price	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ln_quantsold		29.62687	5.731409	5.17	0.000	18.39351	40.86022
coop		28.82658	23.5768	1.22	0.221	-17.3831	75.03627
asoc		89.61339	33.64819	2.66	0.008	23.66416	155.5626
y2006		43.21119	9.634192	4.49	0.000	24.32853	62.09386
y2007		78.67394	9.49289	8.29	0.000	60.06822	97.27966
contract		16.8966	23.15537	0.73	0.466	-28.48708	62.28029
_cons		482.7079	14.80684	32.60	0.000	453.687	511.7287
sigma_u		72.15646					

```

sigma_e | 82.897784
rho | .43105624 (fraction of variance due to u_i)
-----
rho | .43105624 (fraction of variance due to u_i)
-----
sigma_u | 68.117748
sigma_e | 75.629666
rho | .44788456 (fraction of variance due to u_i)
-----

```

```

. corr coop asoc contract
(obs=552)

```

```

      |      coop      asoc contract
-----+-----
coop |      1.0000
asoc |     -0.1359      1.0000
contract |     0.8207     0.1865      1.0000

```

The dummy variable *contract* is shown to be highly correlated with *coop* and somewhat correlated with *asoc*. The correlation is positive as expected (private buyers do not offer contracts). Including it in the model does introduce multicollinearity making the key explanatory variable, *coop*, now insignificant. The new variable is insignificant in any case so it will not be included.

```

. xtreg price ln_quantsold coop asoc y2006 y2007 advance, re

```

```

Random-effects GLS regression           Number of obs   =       525
Group variable (i): id                 Number of groups =       142

R-sq:  within = 0.2079                 Obs per group:  min =         1
      between = 0.2698                   avg =         3.7
      overall  = 0.2694                   max =         9

Random effects u_i ~ Gaussian           Wald chi2(6)     =       153.58
corr(u_i, X) = 0 (assumed)              Prob > chi2      =       0.0000

```

```

-----+-----
price |      Coef.   Std. Err.   z   P>|z|   [95% Conf. Interval]
-----+-----
ln_quantsold | 29.66742   5.745057   5.16  0.000   18.40731   40.92752
coop | 30.05615  18.63904   1.61  0.107   -6.475697  66.58799
asoc | 95.34181  29.70546   3.21  0.001   37.12018  153.5634
y2006 | 43.417    9.83773   4.41  0.000   24.13541   62.6986
y2007 | 80.6084   9.706527  8.30  0.000   61.58396   99.63284
advance | 19.01038  17.80811   1.07  0.286   -15.89288  53.91364
_cons | 481.263   14.82571  32.46  0.000   452.2051  510.3208
-----+-----
sigma_u | 70.718191
sigma_e | 83.43291
rho | .41807518 (fraction of variance due to u_i)
-----

```

```

. corr coop asoc advance
(obs=537)

```

```

      |      coop      asoc  advance
-----+-----
coop |      1.0000
asoc |     -0.1389      1.0000
advance |     0.7975     0.0712      1.0000

```

The dummy variable *advance* is shown to be highly correlated with *coop* and somewhat correlated with *asoc*. The correlation is, again, positive as expected as private buyers do not offer cash advances. Including it in the model does introduce multicollinearity making the key explanatory variable, *coop*, less significant. The new variable is also insignificant so it will not be included.

```
. xtreg price ln_quantsold coop asoc y2006 y2007 payment, re

Random-effects GLS regression                Number of obs   =       539
Group variable (i): id                      Number of groups =       143

R-sq:  within = 0.2007                      Obs per group:  min =         1
        between = 0.2645                      avg =         3.8
        overall = 0.2603                      max =         9

Random effects u_i ~ Gaussian                Wald chi2(6)    =      150.41
corr(u_i, X) = 0 (assumed)                  Prob > chi2     =       0.0000
```

	price	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
ln_quantsold		29.772	5.706984	5.22	0.000	18.58652 40.95749
coop		29.72113	18.19044	1.63	0.102	-5.931485 65.37374
asoc		84.84022	33.23945	2.55	0.011	19.69209 149.9883
y2006		43.06969	9.615093	4.48	0.000	24.22446 61.91493
y2007		78.73236	9.485188	8.30	0.000	60.14173 97.32299
payment		-18.5052	18.8759	-0.98	0.327	-55.50128 18.49087
_cons		502.1814	24.82343	20.23	0.000	453.5284 550.8344
sigma_u		72.78847				
sigma_e		82.489178				
rho		.43776957	(fraction of variance due to u_i)			

```
. corr coop asoc payment
(obs=551)
```

	coop	asoc	payment
coop	1.0000		
asoc	-0.1362	1.0000	
payment	-0.6650	-0.2461	1.0000

The dummy variable *payment* is shown to be highly correlated with *coop* and somewhat correlated with *asoc*. The correlation is, again, positive as predicted as private buyers do not offer cash advances. Including it in the model does introduce multicollinearity making the key explanatory variable, *coop*, less significant. The new variable is once again insignificant so it will not be included.

```
. xtreg price ln_quantsold coop asoc y2006 y2007 pickup, re

Random-effects GLS regression                Number of obs   =       540
Group variable (i): id                      Number of groups =       143

R-sq:  within = 0.1963                      Obs per group:  min =         1
        between = 0.3137                      avg =         3.8
        overall = 0.2845                      max =         9

Random effects u_i ~ Gaussian                Wald chi2(6)    =      162.33
corr(u_i, X) = 0 (assumed)                  Prob > chi2     =       0.0000
```

```

-----
      price |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
ln_quantsold |    31.4352   5.538255     5.68   0.000    20.58042   42.28998
      coop |    25.96418  13.02311     1.99   0.046     .4393526   51.48902
      asoc |    87.28291  28.92801     3.02   0.003    30.58506   143.9808
      y2006 |    42.83868   9.660844     4.43   0.000    23.90377   61.77359
      y2007 |    79.16241   9.516724     8.32   0.000    60.50998   97.81485
      pickup |   -36.7403   12.79257    -2.87   0.004   -61.81328  -11.66732
      _cons |    498.995  15.55584    32.08   0.000   468.5061   529.4838
-----+-----
sigma_u |    68.054046
sigma_e |    82.887382
rho |    .40266785   (fraction of variance due to u_i)
-----

```

```

. corr coop asoc pickup
(obs=552)

```

```

-----+-----
      |      coop      asoc      pickup
-----+-----
coop |    1.0000
asoc |   -0.1359    1.0000
pickup |  -0.4051   -0.0851    1.0000
-----+-----

```

The dummy variable *pickup* is shown to be highly correlated with *coop* and somewhat correlated with *asoc*. The correlation is negative as expected as cooperatives and associations do not offer the service of picking up the coffee from the production site (the farm). Including it in the model does introduce some multicollinearity into the model making the key explanatory variable, *coop*, less significant (though only slightly; it is still significant at a 5% level). The new variable itself is significant and does add explanatory power (the goodness-of-fit is higher compared to the original re model with no service dummies). Intuitively, since the service of picking up the coffee directly from the farm is often variable with private purchasers, it makes sense that this variable would indeed have explanatory power (and negatively as predicted) over the offer price. The benefits of its addition to the model outweigh the slight multicollinearity it carries and will thus be included in the regression.

Farm size is another variable that was originally excluded due to assumed multicollinearity (see footnote 8 above). Also, note that it is time invariant and would have been dropped if a fixed effects model had been more appropriate.

```

. xtreg price ln_quantsold coop asoc y2006 y2007 pickup ln_size, re

Random-effects GLS regression              Number of obs   =      485
Group variable (i): id                    Number of groups =      125

R-sq:  within = 0.2200                     Obs per group:  min =      1
        between = 0.3448                      avg   =      3.9
        overall = 0.3108                      max   =      9

Random effects u_i ~ Gaussian              Wald chi2(7)    =    165.24
corr(u_i, X) = 0 (assumed)                 Prob > chi2     =    0.0000

```

```

-----
      price |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----

```

```

-----+-----
ln_quantsold |   30.58369   5.988516    5.11  0.000    18.84641   42.32096
      coop |    23.8408   13.75567    1.73  0.083   -3.119816   50.80141
      asoc |   101.8114   31.58627    3.22  0.001   39.90346   163.7194
      y2006 |    48.40689   10.2013    4.75  0.000   28.41272   68.40107
      y2007 |    86.68935   10.06515    8.61  0.000   66.96202  106.4167
      pickup |   -39.75007   13.35172   -2.98  0.003  -65.91897  -13.58118
      ln_size |    12.79592    6.790542    1.88  0.060   -5.132982   26.10514
      _cons |    465.1797   22.6672   20.52  0.000   420.7528   509.6066
-----+-----
      sigma_u |    68.88575
      sigma_e |    82.721546
      rho |    .40949303   (fraction of variance due to u_i)
-----+-----

```

```

. corr coop asoc ln_size
(obs=540)

```

```

-----+-----
      |      coop      asoc  ln_size
-----+-----
      coop |    1.0000
      asoc |   -0.1108    1.0000
      ln_size |    0.1558   -0.0007    1.0000
-----+-----

```

The log of farm size is not as correlated with *coop* as initially predicted. The new variable is significant a 10% level. Also, notice that its inclusion does increase the goodness-of-fit and does not drastically affect the coefficients of our key variables. Thus, it will be included in our final regression of the pricing model.

Heteroskedasticity

Before presenting the final random effects regression to be used in determining the influence of selling to cooperatives on price, the model must be tested for and, if necessary, corrected for heteroskedasticity. Given that any heteroskedasticity in the model would most likely be from the time invariant error component (a_i), if the basic OLS model suffers from heteroskedasticity, then the random effects model will fall victim as well (a_i is not eliminated under random effects). Thus, a Breusch-Pagan test for heteroskedasticity will be run on the OLS model.

```

. reg price ln_quantsold coop asoc y2006 y2007 pickup ln_size
. hettest, rhs

```

```

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: ln_quantsold coop asoc y2006 y2007 pickup ln_size

      chi2(7)      =    56.89
      Prob > chi2  =    0.0000

```

As the null is rejected, there is heteroskedasticity in the OLS model and is therefore assumed to be present in the random effects model. This will be corrected for by running the random effects model with robust standard errors.

```
. xtreg price ln_quantsold coop asoc y2006 y2007 pickup ln_size, re robust
```

```
Random-effects GLS regression           Number of obs   =       485
Group variable (i): id                 Number of groups =       125

R-sq:  within = 0.2200                  Obs per group:  min =         1
      between = 0.3448                    avg =         3.9
      overall = 0.3108                    max =         9

Random effects u_i ~ Gaussian          Wald chi2(7)     =       177.73
corr(u_i, X) = 0 (assumed)            Prob > chi2     =       0.0000
```

```
-----+-----
```

	price	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
ln_quantsold		30.58369	6.122174	5.00	0.000	18.58445	42.58293
coop		23.8408	12.6491	1.88	0.059	-.9509837	48.63257
asoc		101.8114	24.49858	4.16	0.000	53.79507	149.8277
y2006		48.40689	12.15252	3.98	0.000	24.58839	72.22539
y2007		86.68935	11.5924	7.48	0.000	63.96867	109.41
pickup		-39.75007	12.13995	-3.27	0.001	-63.54395	-15.9562
ln_size		12.79592	7.385578	1.73	0.083	-1.679547	27.27139
_cons		465.1797	28.07825	16.57	0.000	410.1473	520.212
-----+-----							
sigma_u		68.88575					
sigma_e		82.721546					
rho		.40949303	(fraction of variance due to u_i)				
-----+-----							

Interpreting the Results of the Pricing Model

The following table summarizes the statistics of the final random effects model (adjusted for heteroskedasticity) just presented:

Independent Variable	Coefficient (S.E.)	Z-score	P-value
Log of quantity sold	30.58369 (6.122174)	5.00	0.000
Cooperative	23.8408 (12.6491)	1.88	0.000
Association	101.8114 (24.49858)	4.16	0.000
2006	48.40689 (12.15252)	3.98	0.000
2007	86.68935 (11.5924)	7.48	0.000
Pick up service	-39.75007 (12.13995)	-3.27	0.001
Log of farm size	12.79592 (7.385578)	1.73	0.083
Constant	465.1797 (28.07825)	16.57	0.000

R-squared	With-in: .2200		
	Between: .3448		
	Overall: .3108		

All coefficients are highly significant (the coefficient on log of farm size being the only one not significant at a 1% level though still significant at 10%). The coefficient on log of quantity sold is shown to be positive as expected by the theory presented above. Its coefficient indicates that each additional percent increase in quantity sold (again,

measured in quintals) will induce a one-unit change (i.e. an additional Quetzal) in price. This should be reinterpreted to mean that those farmers selling larger quantities of coffee receive a higher price per quintal. As mentioned previously, this is most likely due to the observation that mid-size producers were most likely to be cooperative/ association members where a higher price is obtained.

The coefficients of the cooperative and association dummy variables also come out to be positive statistically supporting the theory that a higher price is attained by selling to the cooperative/ association. The coefficients can be interpreted as each quintal sold to a cooperative or association receives about 24 and 102 Quetzales more in revenue respectively. Notice that selling to an association earns roughly Q78 more per quintal than selling to a cooperative. This is most likely due to the fact that associations engage in direct exporting (unlike cooperatives) thereby capturing more value added which is passed down to the coffee producers.

The coefficients of the year dummies, as predicted, are positive increasing in value by year. This was to be expected as the commodity price of Arabica coffee has been increasing over the past few years. Roughly speaking, each quintal sold in 2006 or 2007 would fetch Q48 and Q87 more (respectively) than in 2005¹⁴.

Receiving the service of coffee pick up, however, negatively affects price. This was anticipated as picking up the coffee from the farm is a cost to the buyer. Specifically, the model predicts that any transaction in which this service is offered allows the buyer to purchase each quintal for Q38 less than if the buyer required the farmer to deliver the coffee.

Finally, note that the coefficient on the log of farm size is positive as envisaged. To reiterate, relatively larger producers were observed to be more likely to sell to cooperatives/ associations (i.e. be members of such) where a higher price is received indirectly shown through this coefficient (a literal interpretation of each additional cuerda per farmer translates into Q13 more per quintal would theoretically make little sense).

Conclusions of the Pricing Model

The results presented here suggest that selling to cooperatives and/ or associations does translate into significantly higher coffee prices for farmers in San Pedro Necta, Huehuetenango, Guatemala. Given these results, should the higher cooperative/ association price induce private intermediary purchasers to raise their price in subsequent years? In other words, will these higher prices allow the cooperative institution to play a pro-competitive role on farm-gate coffee pricing? Combining the regression results with qualitative observations obtained in country, the suggested answer is no. First, there are many intermediary purchasers in San Pedro Necta offering to buy coffee from its many small coffee producers. These buyers are not in collusion (they act and make decisions independently) and therefore can be said to be in competition (i.e. oligoposony is not present). Thus, by definition, these private buyers should not be holding the offer price below a competitive one. Second, even though cooperatives and associations engage in direct exporting and/ or can bargain for higher prices with domestic exporters, not all producers can take advantage of the higher price per quintal offered by cooperatives due

¹⁴ Note that the price received was corrected for inflation within the raw data.

to certain costs (and, as we will see in the second regression, certain barriers) involved with membership. A few minor costs include membership fees, participation requirements (with fines if not fulfilled) and perhaps loss of a price premium for higher quality coffee given that cooperatives mix all member coffee. More importantly, some producers can see the difference in the services offered or not offered by cooperatives/ associations vs. private purchasers as a cost. For example, for isolated communities, the cost of carrying the coffee to the cooperative (which does not offer pick up service) cannot necessarily be compensated by the higher price received (as shown in the regression results with a negative and significant coefficient on *pickup*). Also, given that most small producers are very poor, they cannot afford to wait for all of the coffee to be liquidated by the cooperative before receiving full payment. The delay in payment from the collective institution is therefore a cost to the producer. Conversely, the fact that cooperatives provide cash advances to members while private purchasers do not could be seen as a service attracting sales to the institution. However, the interest charged on the cash advances is typically quite high and accumulates until the coffee is liquidated often times three months after turning over the produce. In summation, the private buyers should be able to continue to offer a lower price not because of the lack of competition or a below competitive offer price but because of other costs associated with cooperative membership; the most important of these are the delivery cost born by the farmer and the delay in payment by the cooperative/ association forcing members to sacrifice present consumption for future consumption without any compensation by way of interest. The suggested conclusion is therefore that the private purchasers are already offering a competitive price amongst themselves and that the price gap between them and the cooperatives/ associations will continue stably due to the service costs. In other words, the research suggests that the coffee cooperatives in western Guatemala are not actively playing a pro-competitive role on farm-gate pricing.

MODELING COOPERATIVE MEMBERSHIP: REGRESSION II

Beyond modeling the determinants of offer price by institution, it will also be of interest to design a regression modeling the determinants of cooperative membership. The results of such a regression, with cooperative membership as a binary dependent variable, would serve to assist in bolstering provisional suggestions as to the degree coffee cooperatives in western Guatemala are expected to play a pro-competitive role on farm-gate prices.

The data from the year 2007¹⁵ will be employed exclusively for this second regression. Upon looking at the raw data, one notices that cooperative membership is time invariant across the three years for all surveyed farmers. The independent variables are seen to be constant over time as well. The regression will then be a cross-sectional logit/ probit model with cooperative membership as the binary dependent variable as follows:

$$\text{memb07}_i = \beta_0 + \beta_1 \text{expr}_i + \beta_2 \text{size}_i + \beta_3 \text{educhh}_i + \beta_4 \text{flang}_i + \beta_5 \text{spkspan}_i + \beta_6 \text{litspan} + u_i$$

Dependent Binary Variable:

memb07_i- a binary (dummy) dependent variable indicating if the farmer was a registered member of a cooperative or association in 2007 (a value of 0 if a nonmember and 1 if member).

Independent Variables:

expr_i- the number of coffee growing years of experience per farmer. The expected sign of the coefficient is ambiguous as there is no theory hypothesizing whether more experienced farmers tend to be cooperative members or not.

size_i- the size of the farmer's plot of land dedicated to coffee production measured in *cuerdas*. As aforementioned, medium sized growers had qualitatively been observed to more likely be cooperative members (where a higher price is offered). A positive coefficient is therefore expected.

educhh_i- the number of years of education obtained by the farmer (head of household). Again, the expected sign of the coefficient is undetermined (although a more educated farmer may have confront less barriers to cooperative membership given the conceivable need to read and/ or speak Spanish).

flang_i- a dummy variable controlling for the first language of the farmer (a value of 0 for those speaking an indigenous language as their mother tongue and a value of 1 for

¹⁵ All variables in this second regression are time invariant except years of coffee growing experience. Therefore, any year from the panel data would give the same values for each respective variable (excluding experience). This consequently establishes a cross-sectional dataset for this regression. The year 2007 was chosen so as to have the most cumulative values for years experience.

Spanish). This variable is expected to have positive explanatory power over membership as those who speak Spanish as their first language do not have to learn another language to conduct business within a cooperative¹⁶.

spkspan_i- a dummy variable specifying if the farmer speaks Spanish or no (a value of 0 if she does not and a value of 1 if she does). By the same reasoning expressed in footnote 18, the inability to speak Spanish may be a limiting factor in a farmer's decision to join a cooperative. Therefore, the sign of the coefficient should be positive.

spanlit_i- a dummy variable controlling for if the farmer can read and write Spanish (a value of 0 if she is unable and a value of 1 if she is able). As cooperatives and associations operate by contracts written in Spanish, the inability to (especially) read Spanish may deter a farmer from joining a cooperative. Therefore, a positive coefficient is once again predicted for this variable.

Summary of hypotheses:

Variable	Coefficient	Alternative Hypothesis	Expected Sign
expr _i	β_1	?	?
size _i	β_2	>0	+
educhh _i	β_3	>0	+
flang _i	β_4	>0	+
spkspan _i	β_5	>0	+
spanlit _i	β_6	>0	+

The initial logit cooperative membership regression¹⁷:

```
. logit memb07 expr ln_size educhh flang spkspan litspan
```

```
Logistic regression                Number of obs   =        105
                                LR chi2(6)      =        45.79
                                Prob > chi2       =        0.0000
Log likelihood = -49.84181         Pseudo R2      =        0.3148
```

memb07	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
expr	.1240535	.0263495	4.71	0.000	.0724096 .1756975
ln_size	.7605388	.2748168	2.77	0.006	.2219078 1.29917
educhh	.0069058	.1171441	0.06	0.953	-.2226924 .236504
flang	.5602917	.5902177	0.95	0.342	-.5965137 1.717097
spkspan	.2177752	.7753345	0.28	0.779	-1.301852 1.737403
litspan	.6501837	.7600442	0.86	0.392	-.8394755 2.139843

¹⁶ One should note that most of the cooperative directors speak the regional indigenous language, Mam. Therefore, those speaking Spanish as a first language actually may not have an advantage when seeking cooperative membership. However, given that contracts, seminars, etc. are all in Spanish, the fact that a farmer would not have to learn Spanish as a second language should have some advantages when considering membership.

¹⁷ The natural log of *size* will be utilized just like in the first regression.

```

      _cons | -5.289983   1.211188   -4.37   0.000   -7.663867   -2.916099

```

Model Specification

The first variable to be removed will be education as it is the most insignificant followed by first language:

```
. logit memb07 expr ln_size flang spkspan litspan
```

```

Logistic regression                               Number of obs   =          107
                                                    LR chi2(5)      =          39.49
                                                    Prob > chi2     =          0.0000
Log likelihood = -54.303323                       Pseudo R2      =          0.2667

```

memb07	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
expr	.1044393	.0228578	4.57	0.000	.0596387 .1492398
ln_size	.6463686	.2512795	2.57	0.010	.1538699 1.138867
flang	.3122792	.5486592	0.57	0.569	-.7630731 1.387632
spkspan	.5689056	.7478931	0.76	0.447	-.8969379 2.034749
litspan	.5086572	.5555891	0.92	0.360	-.5802775 1.597592
_cons	-4.814387	1.106307	-4.35	0.000	-6.982708 -2.646065

```
. logit memb07 expr ln_size spkspan litspan
```

```

Logistic regression                               Number of obs   =          107
                                                    LR chi2(4)      =          39.17
                                                    Prob > chi2     =          0.0000
Log likelihood = -54.466122                       Pseudo R2      =          0.2645

```

memb07	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
expr	.103775	.0226402	4.58	0.000	.059401 .1481491
ln_size	.643103	.250839	2.56	0.010	.1514676 1.134738
spkspan	.6615906	.7291221	0.91	0.364	-.7674625 2.090644
litspan	.5255219	.5543043	0.95	0.343	-.5608945 1.611938
_cons	-4.789401	1.100974	-4.35	0.000	-6.947269 -2.631532

```
. corr spkspan litspan
(obs=118)
```

	spkspan	litspan
spkspan	1.0000	
litspan	0.5301	1.0000

In determining which variable to remove next (Spanish spoken or Spanish literacy), there are a few points to consider. First, it may initially appear appropriate to remove Spanish literacy even though it is slightly more significant. One could argue that with the aptitude to speak Spanish, the farmer could at least “get by” in cooperative transactions. However, note the high correlation between the two variables. This correlation suggests that Spanish literacy encompasses Spanish fluency in many cases thus surmounting the theory just given. Moreover, Spanish literacy is technically more significant, thus, is

statistically more appropriate to be left in the model. Adhering to such logic, spoken Spanish is removed resulting in the final model to be used.

```
. logit memb07 expr ln_size litspan
```

```
Logistic regression                Number of obs   =       107
                                   LR chi2(3)         =       38.33
                                   Prob > chi2         =       0.0000
Log likelihood = -54.88601          Pseudo R2       =       0.2588
```

memb07	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
expr	.1029836	.022684	4.54	0.000	.0585239 .1474433
ln_size	.662854	.2505629	2.65	0.008	.1717598 1.153948
litspan	.7838135	.4796189	1.63	0.102	-.1562222 1.723849
_cons	-4.423691	1.010179	-4.38	0.000	-6.403605 -2.443776

```
. lstat
```

```
Logistic model for memb07
```

Classified	True		Total
	D	~D	
+	37	13	50
-	14	43	57
Total	51	56	107

```
Classified + if predicted Pr(D) >= .5
True D defined as memb07 != 0
```

Sensitivity	Pr(+ D)	72.55%
Specificity	Pr(- ~D)	76.79%
Positive predictive value	Pr(D +)	74.00%
Negative predictive value	Pr(~D -)	75.44%
False + rate for true ~D	Pr(+ ~D)	23.21%
False - rate for true D	Pr(- D)	27.45%
False + rate for classified +	Pr(~D +)	26.00%
False - rate for classified -	Pr(D -)	24.56%
Correctly classified		74.77%

```
. mfx
```

```
Marginal effects after logit
y = Pr(memb07) (predict)
= .46781914
```

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]	X
expr	.0256392	.00565	4.54	0.000	.014561 .036717	20.6449
ln_size	.1650271	.0622	2.65	0.008	.043125 .28693	2.61975
litspan*	.1921564	.11436	1.68	0.093	-.031989 .416302	.551402

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Interpreting the Results of the Cooperative Membership Model

The following table summarizes the (mfx) statistics of the final cooperative membership model just presented:

<u>IndependentVariable</u>	<u>dy/dx (S.E.)</u>	<u>Z-score</u>	<u>P-value</u>
Experience	.0256392 (.00565)	4.54	.000
Log of farm size	.1650271 (.0622)	2.65	.008
Spanish literacy	.1921564 (.11436)	1.68	.093
Sensitivity: 72.55%			
Specificity: 76.79			
Correctly classified: 74.77%			

The marginal effect of experience is significant at a 1% level. Its interpretation is that one more year experience (from the mean) in growing coffee will increase the probability of joining a cooperative by 2.5%. This is an interesting finding given that no hypothesis as to the direction of the marginal effect was predicted beforehand. One may conjecture that experienced coffee farmers prefer cooperative membership because perhaps they are better at exploiting cooperative cash advances in efficiently buying and applying fertilizer (the most common item bought with the advance) amongst other experiential wisdom.

The marginal effect of the log of farm size is also significant at a 1% level. Its marginal effect indicates that each additional *cuerda* of land dedicated to coffee production (from the mean) increases the probability of membership by 16.5%. This finding is equivalent to what was deduced from collecting the raw data in country. From this, one can suggest that relatively larger producing farmers are more likely to join cooperatives because, one, if it is assumed that a medium size producer is wealthier than a small producer, the former is more financially capable to withstand the delay in payment than the later. Given this capability, the relatively larger producer rationally chooses cooperative membership in order to receive a higher price per quintal of coffee. The inability of a smaller, poorer producer to endure the delayed payment may in fact be a key obstacle to cooperative membership and, therefore, an impediment to fetching a higher price. Second, a relatively larger producer (again, assuming this implies greater income) may be able to bare the interest costs of receiving cash advances from the cooperative. Given that a cash advance can be used to by fertilizer and other production inputs (such as herbicides, etc.) to enhance output, the rational farmer will join a cooperative to take advantage of the credit market provided by the institution if she can financially bare the burden of interest on any loan. Likewise, the inability to bear interest fees may be another disincentive for smaller, poorer farmers in joining cooperatives.

Finally, the marginal effect of Spanish literacy is significant at a 10% level and can be interpreted as those farmers literate in Spanish are 19.2% more likely to be cooperative members. Congruent with the hypothesis previously presented, the incapacity to read and write Spanish, not to mention the inability to speak the language (see correlation results above), may be a key barrier to cooperative participation.

Conclusions of the Cooperative Membership Model

The key finding that relatively larger producers are more likely to be cooperative members further supports the suggestion¹⁸ that the demand for services (or perhaps for poorer farmers, the inability to demand certain services) is crucial in understanding why private purchasers may be able to continue to offer a lower price despite the existence of a higher cooperative price in the market. For example, if a smaller, poorer producer is not financially capable of waiting for the liquidation of all cooperative member coffee (typically months after turning in the coffee) despite the lure of a higher offer price from the institution, private buyers can offer a lower price for the “service” of payment on delivery. This is not to imply that the private price is noncompetitive. Indeed, it most likely is given the numerous private purchasers operating independently. What it does imply is that the difference in services (or “disservices”) between the cooperative and the private market allows for the unequal, yet, stable price disparity between the two purchasing agents. Again, these findings suggest that the cooperatives of western Guatemala are not actively playing a pro-competitive role on farm-gate coffee prices.

It is also interesting to discover that the inability to read and write Spanish may be a key obstacle to cooperative membership. The results of this second model suggest that if, for example, an indigenous farmer who only speaks her native tongue or a Spanish-speaking farmer of little or no education may not have access to cooperative membership and, consequently, the higher price per quintal the institution offers. Economically, one can propose that this proficiency can be considered a necessity in being able to access a market (in this case the cooperative market) where a higher price is to be found. Those without the aptitude are thus only able to access a market where a lower (though competitive within that specific market) is to be obtained.

Finally, the implications that more experienced farmers are more likely to be cooperative members conjures the suggestion that there are some unobservables as to why more intimate familiarity with the production process attracts producers to collective institutions. Perhaps these farmers, from experience, know that there are certain technological “spillovers” associated with membership. Or, as asserted above, more experienced farmers may be able to more efficiently take advantage of the services (or able to more efficiently weather the “disservices”) offered by the cooperative increasing their net benefit of their respective likelihood functions in electing cooperative versus non-cooperative membership.

¹⁸ See the conclusions from the first model.

FINAL CONCLUSIONS AND SUGGESTED POLICY

In combining the implications formulated from the results of the two models, the research collectively suggests that the cooperatives in western Guatemala do not play a competitive role in farm-gate coffee pricing despite a price gap between the higher cooperative price and the lower private price. This suggestion should not be interpreted as a failure of the cooperative institution.

First and foremost, oligopsonic collusion does not appear to exist amongst the private intermediary purchasers of the market. There are many private buyers operating in the market operating quite independently. Thus, at least between those buyers outside the cooperative structure, a competitive price is most likely being offered.

Second, it appears that the differences in services offered (or the lack there of) might support a stable dual-priced market. This is most readily seen by the negative coefficient associated with the pick up service offered by private buyers and not cooperatives. In effect, a lower price combined with the pick up service offered by a private buyer may, in terms of total utility, outweigh the higher price offered by the cooperative. This may especially be true for isolated, mountainous communities lacking adequate infrastructure connecting to the local where the cooperative is operating.

Third certain barriers may prevent some farmers from becoming cooperative members where a higher price is to be received. The research suggests illiteracy (as well as perhaps simply the inability to speak Spanish) is most likely a barrier to cooperative membership. Thus, these farmers are left to the residual private purchasing market. As private intermediaries are an additional agent in the value chain (which marketing cooperatives as an institution attempt to cut out), they capture some of the value added thus resulting in a lower price offered to their market share.

Finally, the poorest and smallest producers may not be able bear some of the additional costs associated with cooperative membership. Cooperatives do not pay its members immediately once their coffee is delivered to the institution. The cooperative typically waits a few months to liquidate some or all of the coffees thus making its members wait for full payment. The poorest producers cannot afford to wait for the payment as a matter of subsistence despite the higher price to eventually be received. These poorer farmers will thus opt to accept a lower price from a private buyer offering the “service” of full payment upon handing over the coffee. Also, many farmers choose to become cooperative members for the sole fact of having access to cooperative credit (cash advances) with which to buy key production inputs and engage in consumption smoothing. However, even though the cooperative attempts to correct for the credit market failure in rural Guatemala, the interest rates are high (relative to a bank or micro credit loan) and are accrued until the coffee is liquidated not when the coffee is handed in. This most likely decreases the net benefit of one of the principal attractions of cooperative membership.

Collectively, these arguments offer the conjecture that the price disparity will continue stably (moving with world prices) between cooperative/ associations and private buyers. Even though this suggests that cooperatives do not play a pro-competitive role in farm-gate pricing, this in no manner implies cooperatives are an insignificant economic institution. The research suggests that they do capture more value added for their members, especially those members of associations. The collective institution also

provides seminars, capacity building workshops as well as knowledge sharing and comradery all of benefit to members.

Therefore, in terms of policy, the public and private sectors should support cooperatives, especially in terms of improving membership access. For example, low interest loans could be made available to cooperatives allowing the institution to offer lower interest loans to farmers or allowing the cooperative pay in full before all the coffee is liquidated. Also, investment in human capital (including public schools and literacy campaigns) of the rural communities where small producers operate could help to decrease the educational/ language barriers to membership. Finally, the success of direct exporting by associations could be a useful template in establishing direct labeling for the coffee farmers of the region. The quality of coffee grown in that part of Guatemala is one of the highest in the world. Farmers could receive more revenue per quintal if it was directly exported and marketed as such.