Statistical Issues Related to the Low-Carbon Fuel Standard
Submitted by
David M. Rocke, PhD
October 31, 2014
Under contract 13-405 (2014)

Analysis of Simulations for ILUC

Two separate simulation methodologies were employed by CARB to help determine factors to which Indirect Land Use Change (iLUC) is sensitive. The iLUC impact of biofuels relates to the unintended increase of carbon emissions due to land-use changes around the world induced by the expansion of croplands for production of biofuels such as ethanol in response to the increased global demand for these fuels. If more biofuels are needed, in general the price of the feedstock would rise compared to other uses of the land. This in turn may result in forests or other uncropped land being converted to agricultural use. Because natural lands, such as rainforests and grasslands, store carbon in their soil and in biomass as plants grow each year, clearance of wilderness for new farms translates to a net increase in greenhouse gas emissions. Due to this change in the carbon stock of the soil and the biomass, indirect land use change has consequences in the greenhouse-gas emissions balance of a biofuel.

Both sets of simulations are based on the Global Trade Analysis Project (GTAP) database and the Agro-ecological Zone Emission Factor (AEZ-EF) Model. One method was to use varying specific values of some parameters as sensitivity analysis. For example, this could consist of YDEL, the price elasticity of yield, ETL1, the elasticity of transformation between forest, cropland, and pasture, ETL2, the elasticity of transformation among crops, PAEL_US, the yield elasticity for cropland/pasture in the US, and PAEL_Brazil, yield elasticity for cropland/pasture in Brazil. The other simulation method used the Monte Carlo methodology in which values for a large number of parameters were chosen at random repeatedly.

In order to determine the most influential factors, we conducted a statistical analysis of the iLUC factor for corn ethanol in terms of the input variables in a simulation with 600 variables and 3,000 trials. This was done using stepwise regression, but since all the parameters were chosen independently in the Monte Carlo (except CDGC and CDGS, which were highly correlated), the coefficient estimates were almost orthogonal, so the results of a single analysis of the 600 variable model would have been very similar, except for CDGC and CDGS. Table 1 gives the results of this analysis. The most influential factors in terms of contribution to the sum of squares were YDEL, the price elasticity of yield, the ESBV parameters, the elasticity of substitution between primary input factors in production, ETA, the elasticity of effective hectares with respect to harvested area, and ETL1, the elasticity of transformation among crops.

Table 1. Statistical Analysis of Corn Ethanol ILUC Factor in a Monte Carlo Simulation

Response: ilucFactor

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
ESBV.11.0.	1	68324	68324	4989.7281	< 2.2e-16	***
YDEL	1	65612	65612	4791.7008	< 2.2e-16	***
ETA	1	37960	37960	2772.2342	< 2.2e-16	***
ESBV.13.0.	1	17097	17097	1248.6237	< 2.2e-16	***
ETL1	1	13970	13970	1020.2320	< 2.2e-16	***
CDGC	1	13886	13886	1014.0667	< 2.2e-16	***
croplandPastureEmissionRatio	1	7214	7214	526.8437	< 2.2e-16	***
ESBV.12.0.	1	4978	4978	363.5544	< 2.2e-16	***
N2O_N_EF	1	2975	2975	217.2690	< 2.2e-16	***
PAEL.3.0.	1	2268	2268	165.6035	< 2.2e-16	***
pastureSoil_C.0.1.	1	2089	2089	152.5737	< 2.2e-16	***
croplandSoil_C	1	2034	2034	148.5450	< 2.2e-16	***
youngStandAglb	1	1471	1471	107.4001	< 2.2e-16	***
SUBP.0.18.	1	1356	1356	98.9945	< 2.2e-16	***
EFED	1	946	946	69.0674	< 2.2e-16	***
SUBP.0.1.	1	874	874	63.8461	1.934e-15	***
totalTree_C.0.4.	1	890	890	64.9935	1.094e-15	***
croplandLandUseFactor.5.0.	1	752	752	54.9003	1.661e-13	***
PAEL.1.0.	1	694	694	50.7027	1.354e-12	***
SUBP.0.2.	1	644	644	47.0584	8.416e-12	***
totalTree_C.0.1.	1	627	627	45.8145	1.572e-11	***
carbonNitrogenRatio	1	639	639	46.6822	1.016e-11	***
SUBP.0.3.	1	562	562	41.0261	1.751e-10	***
deadwoodByLatitude_C.3.1.	1	525	525	38.3264	6.844e-10	***
croplandLandUseFactor.10.0.	1	488	488	35.6556	2.646e-09	***
deadwoodByRegion_C.4.1.	1	515	515	37.5940	9.912e-10	***
deadwoodByRegion_C.1.1.	1	473	473	34.5168	4.715e-09	***
totalTree_C.0.2.	1	385	385	28.1390	1.215e-07	***
forestSoil_C.0.18.	1	383	383	27.9501	1.339e-07	***
forestSoil_C.0.4.	1	367	367	26.8051	2.407e-07	***
oldStandAglb	1	313	313	22.8335	1.856e-06	***
pastureSubsoilLossFraction	1	323	323	23.5576	1.277e-06	***
totalTree_C.0.18.	1	253	253	18.4775	1.777e-05	***
croplandLandUseFactor.6.0.	1	246	246	17.9905	2.291e-05	***
forestLitter_C.10.1.	1	218	218	15.9474	6.677e-05	***
pastureAgb.6.0.	1	211	211	15.4370	8.732e-05	***
understory_C	1	202	202	14.7871	0.0001230	***
GWP_N2O	1	177	177	12.9423	0.0003267	***
pastureSoil_C.0.19.	1	175	175	12.8020	0.0003520	***
ETL2	1	171	171	12.4815	0.0004175	***
EPSR	1	170	170	12.3870	0.0004391	***
foregoneGrowthRate	1	152	152	11.1033	0.0008727	***
croplandLandUseFactor.4.0.	1	149	149	10.8470	0.0010016	**
ESBM.4.0.	1	143	143	10.4288	0.0012547	**
ESBM.2.0.	1	124	124	9.0317	0.0026764	**
ESBV.25.0.	1	119	119	8.7089	0.0031924	**
pastureSoil_C.0.12.	1	115	115	8.4070	0.0037663	**
pastureSoil_C.0.3.	1	117	117	8.5596	0.0034642	**
ESBV.30.0.	1	105	105	7.6970	0.0055672	**
forestLitter_C.15.1.	1	108	108	7.8711	0.0050571	**
ELEN.9.0.	1	102	102	7.4502	0.0063818	**

ELEN.26.0.	1	103	103	7.5010	0.0062047	**
cropCarbonAnnualizationFactor	1	87	87	6.3746	0.0116303	*
ELEG.19.0.	1	88	88	6.4184	0.0113473	*
pastureSubsoil_C.0.1.	1	86	86	6.2890	0.0122040	*
forestLitter_C.13.1.	1	86	86	6.2485	0.0124856	*
ELNC.16.0.	1	83	83	6.0512	0.0139554	*
ESBM.46.0.	1	76	76	5.5190	0.0188785	*
forestLitter_C.9.1.	1	72	72	5.2607	0.0218848	*
SUBP.0.13.	1	76	76	5.5662	0.0183778	*
pastureSoil_C.0.8.	1	72	72	5.2931	0.0214824	*
ELEN.2.0.	1	71	71	5.1593	0.0231958	*
totalTree_C.0.6.	1	65	65	4.7814	0.0288496	*
ESBV.2.0.	1	68	68	4.9825	0.0256817	*
ELEG.3.0.	1	65	65	4.7447	0.0294704	*
ELKE.10.0.	1	68	68	4.9421	0.0262881	*
deforestedFraction.11.0.	1	64	64	4.6579	0.0309946	*
ELNE.7.0.	1	63	63	4.6191	0.0317009	*
croplandLandUseFactor.15.0.	1	64	64	4.6402	0.0313146	*
forestRootShootRatio	1	63	63	4.5786	0.0324578	*
deadwoodByRegion_C.18.1.	1	59	59	4.2837	0.0385692	*
deforestedFraction.8.0.	1	59	59	4.2987	0.0382306	*
ELKE.37.0.	1	57	57	4.1496	0.0417355	*
pastureSubsoil_C.0.3.	1	57	57	4.1742	0.0411345	*
ELEN.29.0.	1	57	57	4.1843	0.0408909	*
pastureSoil_C.0.18.	1	58	58	4.2081	0.0403236	*
deforestedFraction.13.0.	1	55	55	4.0201	0.0450553	*
hwpFraction.9.0.	1	52	52	3.7859	0.0517839	.
forestLandUseFactor.11.0.	1	52	52	3.7882	0.0517122	.
forestSoil_C.0.13.	1	52	52	3.7649	0.0524376	.
ELNE.22.0.	1	48	48	3.4933	0.0617215	.
totalTree_C.0.12.	1	51	51	3.7565	0.0527010	.
ESBM.41.0.	1	49	49	3.5807	0.0585568	.
ELHL	1	48	48	3.5264	0.0605018	.
croplandLandUseFactor.3.0.	1	47	47	3.4426	0.0636396	.
forestLitter_C.17.1.	1	46	46	3.3286	0.0681885	.
ELNC.13.0.	1	45	45	3.2580	0.0711825	.
ELNE.4.0.	1	43	43	3.1227	0.0773172	.
ESBV.1.0.	1	44	44	3.1827	0.0745296	.
ELNC.19.0.	1	43	43	3.1486	0.0760975	.
forestSoil_C.0.11.	1	42	42	3.0762	0.0795527	.
SUBP.0.4.	1	44	44	3.1855	0.0743993	.
ELEG.2.0.	1	42	42	3.0802	0.0793588	.
PAEL.11.0.	1	41	41	3.0253	0.0820827	.
ELNC.5.0.	1	41	41	2.9984	0.0834557	.
forestBurningEF	1	41	41	2.9782	0.0844994	.
ELKE.15.0.	1	42	42	3.0370	0.0814919	.
pastureSubsoil_C.0.8.	1	39	39	2.8725	0.0902161	.
ESBM.16.0.	1	39	39	2.8535	0.0912852	.
croplandLandUseFactor.1.0.	1	42	42	3.0817	0.0792853	.
ELKE.1.0.	1	39	39	2.8257	0.0928772	.
deforestedFraction.7.0.	1	37	37	2.7211	0.0991387	.
ELVL	1	37	37	2.7172	0.0993831	.
forestSubsoil_C.0.8.	1	39	39	2.8846	0.0895377	.
forestSubsoil_C.0.18.	1	37	37	2.7202	0.0991942	.
ELNE.24.0.	1	39	39	2.8418	0.0919521	.
ELEN.4.0.	1	40	40	2.9344	0.0868207	.
ELNE.6.0.	1	37	37	2.7386	0.0980619	.

forestSoilLossFraction	1	35	35	2.5360	0.1113837
forestLandUseFactor.3.0.	1	36	36	2.6196	0.1056590
ELEG.7.0.	1	33	33	2.3757	0.1233479
ELKE.36.0.	1	32	32	2.3144	0.1282875
ESBM.33.0.	1	36	36	2.6437	0.1040686
ELNC.26.0.	1	35	35	2.5444	0.1107993
ELEN.6.0.	1	36	36	2.5966	0.1072009
ELNE.34.0.	1	32	32	2.3068	0.1289195
PAEL.6.0.	1	32	32	2.3672	0.1240167
ESBV.28.0.	1	32	32	2.3410	0.1261183
pastureAgb.10.0.	1	37	37	2.6804	0.1017002
ELNE.16.0.	1	33	33	2.3810	0.1229333
forestSubsoil_C.0.14.	1	31	31	2.2673	0.1322385
pastureSoil_C.0.16.	1	33	33	2.3782	0.1231485
ELHB	1	33	33	2.3743	0.1234546
ELNC.1.0.	1	33	33	2.3922	0.1220537
ELKE.18.0.	1	35	35	2.5512	0.1103183
ELNC.17.0.	1	30	30	2.1732	0.1405476
ESBV.19.0.	1	31	31	2.2578	0.1330512
ELEN.31.0.	1	33	33	2.4252	0.1195113
pastureAgb.12.0.	1	30	30	2.1670	0.1411076
ELKE.34.0.	1	33	33	2.4155	0.1202515
ELNE.33.0.	1	32	32	2.3370	0.1264439
ELNE.32.0.	1	32	32	2.3271	0.1272524
ESBM.22.0.	1	32	32	2.3090	0.1287354
ELKE.41.0.	1	30	30	2.2042	0.1377488
SUBP.0.5.	1	34	34	2.4534	0.1173836
ELNC.2.0.	1	31	31	2.2766	0.1314507
ELNE.14.0.	1	28	28	2.0659	0.1507380
ELEN.7.0.	1	28	28	2.0718	0.1501589
forestSubsoil_C.0.11.	1	31	31	2.2497	0.1337495
ELNE.18.0.	1	31	31	2.2353	0.1350028
ELNE.17.0.	1	27	27	1.9797	0.1595262
ELNC.14.0.	1	29	29	2.1052	0.1469068
deforestedFraction.1.0.	1	29	29	2.0978	0.1476215
ELEG.11.0.	1	28	28	2.0785	0.1494954
ESBM.21.0.	1	28	28	2.0808	0.1492744
Residuals	2854	39080	14		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Price Elasticity of Yield (YDEL)

In view of the importance of YDEL in the analysis, and in view of the conflicting results in the literature on its likely size, the next part of the project undertaken was to analyze one of the data sets upon which these estimates have been based. The data were used in a 2012 dissertation of Juan Francisco Rosas Pérez (also given as Juan Francisco Rosas in a 2014 paper by Rosas, Hayes, and Lence, apparently taken from the dissertation). In these works, the price elasticity of yield was estimated from data on corn (maize) in Iowa for 1960–2004, and was said to be in the range of 0.29. The data set was publicly available so it was used for a re-analysis. The analysis used by Rosas Pérez, was complex, and can be criticized for insufficiently handling autocorrelation in the series. Therefore, a simpler analysis was conducted that should have similar results to the more complex analysis if the latter is not flawed.

The data set used was the one supplied with the Rosas Pérez dissertation, though there is no good data dictionary and the meaning of some of the statistics was less than clear. The most clearly relevant variables were a corn price index series (here called corn.price) and a corn supply index series (corn.supply) and their natural logarithms (lcorn.price and lcorn.supply). There do not seem to be good data on land devoted to corn, or perhaps land at all, since the variable Z4 = Q Land is equal to 1 for all years, so this analysis was aimed at the price elasticity of supply not the price elasticity of yield; this would tend to overestimate the effect of price on supply given that land substitution is often an easier response to greater potential profit from a crop than is attempting to increase yield.

The quantity of interest then would be the ratio of the percentage change in supply to the percentage change in price. Roughly, the percentage change is equal to the actual change on the natural log scale. For example (110 – 100)/100 = 0.10 while log(110) – log(100) = 0.0953, so we will proceed to relate the change on the log scale of supply to the change on the log scale of price.

Without participating in debates about the proper functional form of multi-equation models of the agricultural economy, we can go back to statistical basics using the following principles:

1. All other things being equal, the price elasticity of supply can be estimated by regressing log(supply) on log(price).
2. In regressions with autocorrelated time series, it is important to account for the self-effects of the series being predicted before asking if another series has an effect. This is sometimes called Granger causality analysis.

In fact, both series are autocorrelated in a plausibly autoregressive way, with the ACF function declining slowly and the PACF function dropping of more quickly (see Figures 1 and 2 for the supply series later in the document). As can be seen from the output in Table 2, there is no significant relationship of supply to current or past prices after accounting for last year's supply. In fact, the estimated coefficients are not even positive.

While there may exist alternative explanations of these results with respect to omitted factors, it is hard to find such modeling aspects that provide effects in the direction of reducing the apparent response of supply to price and that themselves could explain a large elasticity that is so hidden. The best interpretation of these results is that

1. The price elasticity of yield implied by the Iowa corn data is likely close to 0 and very unlikely to be as large as 0.10 or 0.20.
2. The results obtained by Rosas Pérez showing an apparently higher elasticity is likely caused by mishandling the autocorrelation in the time series.

As documented in Berry (2011), Berry and Schlenker (2011), and Roberts and Schlenker (2013), much of the literature providing purported estimates of the price elasticity of yield is deeply methodologically flawed. In addition to the problems of endogeneity and autocorrelation that are badly handled, there are other important issues. In Goodwin, Michele Marra, Piggott, and Mueller (2012), for example, 15 years of data are multiplied into 405 data points by considering 27 different districts. But there are still only 15 price values and it is hard to believe that the strong relationships of weather, price, and technology within a given year can be handled by econometric tricks. The analyses, such as those in Roberts and Schlenker (2013), that are methodologically sound all show small to zero price elasticities of yield.

Table 2. Regression Analysis for Price Elasticity of Supply for Iowa Corn

```
> anova(lm(lcorn.supply~lcorn.supply1+lcorn.price+lcorn.price1))

Analysis of Variance Table

Response: lcorn.supply
              Df  Sum Sq Mean Sq F value    Pr(>F)    
lcorn.supply1  1 1.58085 1.58085 30.5328 2.191e-06 ***
lcorn.price    1 0.00558 0.00558  0.1078    0.7444    
lcorn.price1   1 0.01618 0.01618  0.3125    0.5793    
Residuals     40 2.07103 0.05178                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(lm(lcorn.supply~lcorn.supply1+lcorn.price+lcorn.price1
           +lcorn.price2))

Analysis of Variance Table

Response: lcorn.supply
              Df  Sum Sq Mean Sq F value    Pr(>F)    
lcorn.supply1  1 1.39173 1.39173 26.6904 7.889e-06 ***
lcorn.price    1 0.00466 0.00466  0.0894    0.7666    
```

```
lcorn.price1   1 0.01436 0.01436  0.2755    0.6027
lcorn.price2   1 0.07523 0.07523  1.4428    0.2371
Residuals     38 1.98145 0.05214
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> summary(lm(lcorn.supply~lcorn.supply1+lcorn.price+lcorn.price1))

Call:
lm(formula = lcorn.supply ~ lcorn.supply1 + lcorn.price +
lcorn.price1)

Residuals:
     Min       1Q   Median       3Q      Max
-0.64342 -0.11119  0.01966  0.14210  0.52123

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.71117    0.24967   2.848  0.00691 **
lcorn.supply1  0.62929    0.13427   4.687 3.19e-05 ***
lcorn.price   -0.02265    0.23289  -0.097  0.92301
lcorn.price1  -0.12364    0.22116  -0.559  0.57925
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2275 on 40 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.4362,    Adjusted R-squared:  0.394
F-statistic: 10.32 on 3 and 40 DF,  p-value: 3.676e-05
```

Figure 1. Autocorrelation of Corn Supply in Iowa

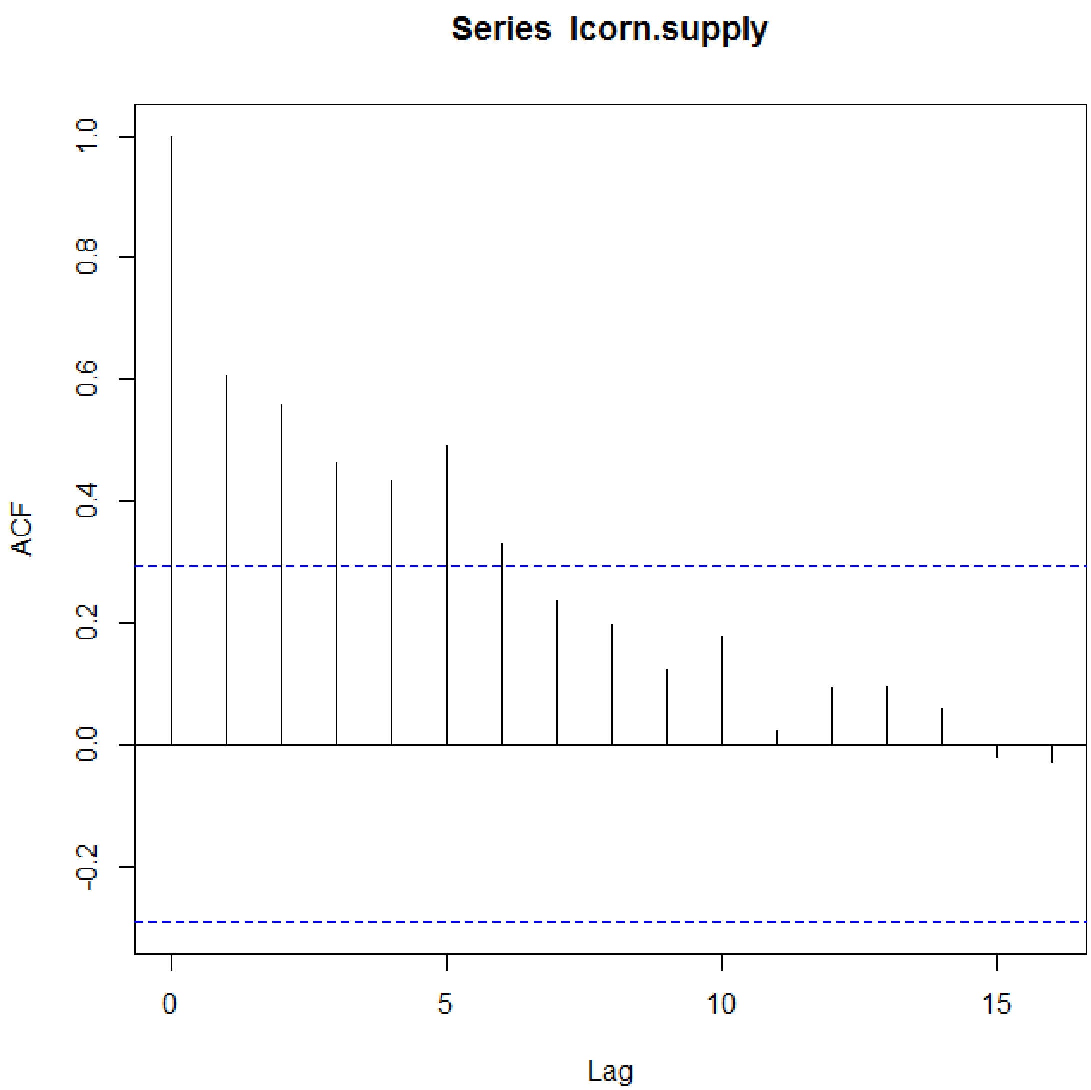
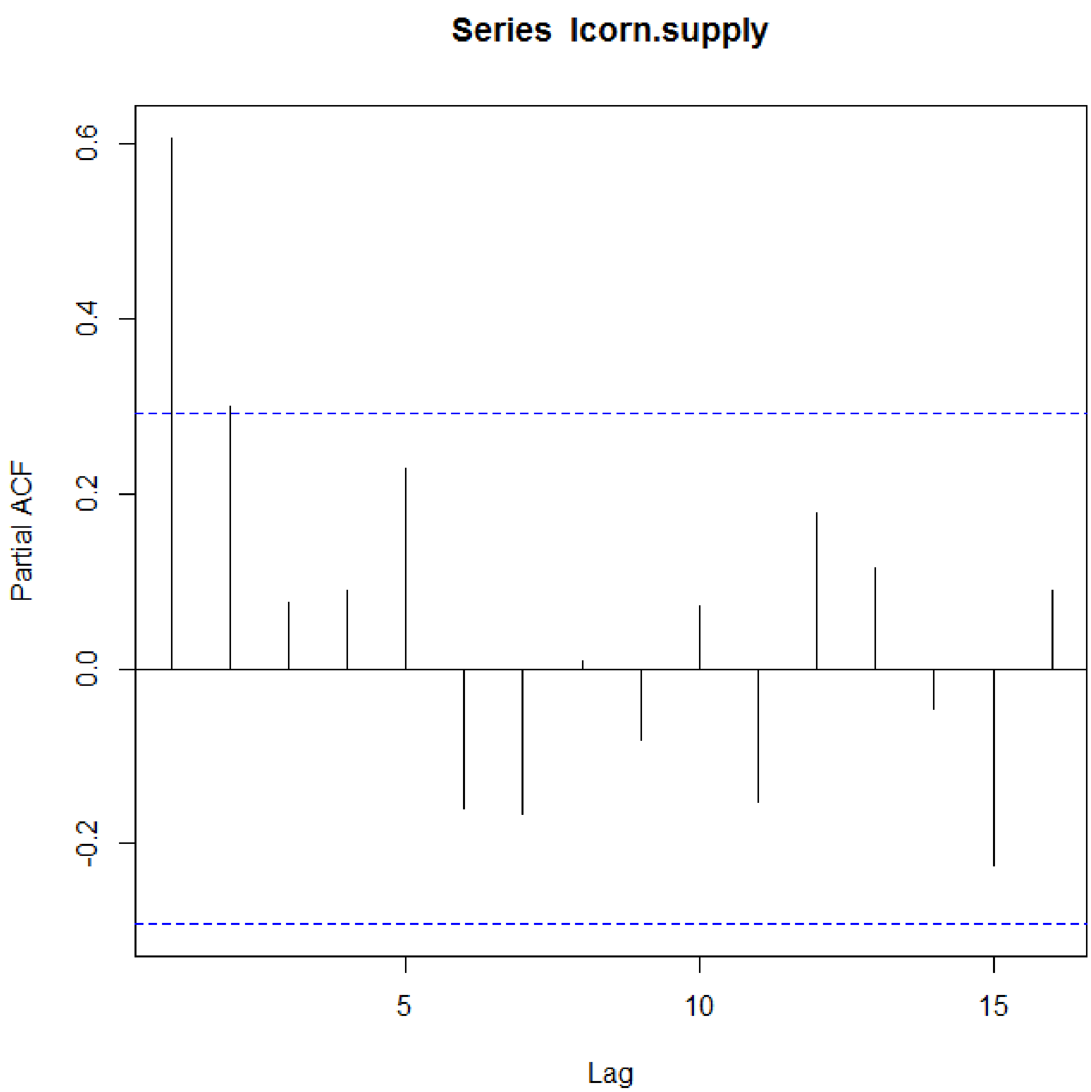


Figure 2. Partial Autocorrelation of Corn Supply in Iowa



Bibliography

Berry, Steven T. (2011) “Biofuels Policy and the Empirical Inputs to GTAP Models,” manuscript.

Berry, Steven and Wolfram Schlenker (2011) “Technical Report for the ICCT: Empirical Evidence on Crop Yield Elasticities,” manuscript.

Goodwin, Barry K., Michele Marra, Nicholas Piggott and Steen Mueller (2012) “Is Yield Endogenous to Price? An Empirical Evaluation of Inter- and Intra-Seasonal Corn Yield Response,” manuscript.

Pérez, Juan Francisco Rosas, (2012) *Essays on the Environmental Effects of Agricultural Production*, Diss. Iowa State University.

Roberts, Michael J. and Wolfram Schlenker (2013) “Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate,” *American Economic Review*, 103, 2265–2295.

Rosas, Juan Francisco, Dermot J. Hayes, and Sergio H. Lence (2014) “Response of Crop Yields to Output Prices: A Bayesian Approach to Complement the Duality Theory Econometrics,” manuscript.