The impact of weather and soil on crop yield in the Midwestern US in the past and in the future

Erik Nelson, Bowdoin College

I. Data

For each county in the six Midwestern states of Indiana, Iowa, Illinois, Michigan, Minnesota, and Ohio I calculate a soil capability score.

$$S_c = 4S_{c1} + 3S_{c2} + 2S_{c3} + S_{c4} \quad (1)$$

where S_{c1} is the fraction of county *c*'s area in land capability classes (LCCs) 1 and 2, S_{c2} is the fraction of county *c*'s area in LCCs 3 and 4, S_{c3} is the fraction of county *c*'s area in LCCs 5 and 6, and S_{c4} is the fraction of county *c*'s area in LCCs 7 and 8. The lower the capability class, the more appropriate (or less limiting) the soil is for crop growth. The composite S_c score can range from 1 to 4 where higher numbers indicate that the profile of soil in the county is more concentrated in more capable soils. The soil map is from Radeloff et al. (2012). The distribution of S_c across the six states' counties is given in Figure 1.

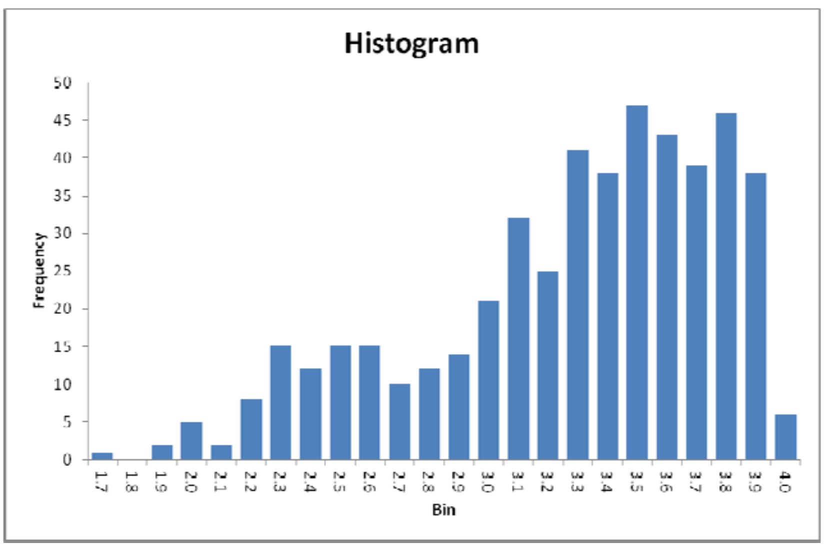

Figure 1

I collected and calculated county-level crop production data and growing season weather for the years 1950 through 2008. Let A_{jct} indicate the percentage of county *c*'s land used to harvest crop *j* in year *t*. The index *j* includes maize, soybeans, and wheat. Let y_{jct} indicate the average per acre yield of crop *j* in county *c* in year *t* (USDA-NASS 2011). I also calculated several county-level growing season weather variables. First, I collected monthly temperature averages and precipitation levels for the years 1950 through 2008 for each 0.5 degree grid cell in the 6 state area (CRU 2010). Then, using a gridded map that gives growing season dates for crop *j* (Sacks et al. 2010), I calculated *j*'s growing degree days (GDD) and growing season precipitation in each cell for the years 1950 through 2008. Temperature readings only added to the GDD measure if they were 5 degrees Celsius or greater and they occurred during the crop's growing season. The code to convert monthly daytime temperature averages and monthly precipitation amounts to GDDs and growing season precipitation comes from Jamie Gerber, Institute of the Environment, University of Minnesota. Finally, a county's time series of growing season weather was set equal to that of the grid cell closest to the county's centroid. Let G_{jct} indicate growing degree days for crop *j* in county *c* in year *t* and R_{jct} indicate growing season precipitation (mm) for crop *j* in county *c* in year *t*. To capture the non-linear impact growing season weather can have on crop growth I also use squared terms of G_{jct} and R_{jct} in the yield model described below. I assume that planting and harvesting dates remained static from 1950 to 2008.

I divide the dataset of counties into soil quintiles or class. The 20% of counties with the most capable soils as measured by the soil statistic *S* are grouped together; the 20% of counties with the next best soils are grouped together, etc. Let the set of counties *c* in each class be defined by $c \in C_q$ where *q* = 1, 2, 3, 4, 5 indexes the soil class. In Table 1 I give the range in the S statistic that defines each class.

Table 1

Soil class	S
1	[0.000, 2.785]
2	(2.785, 3.218]
3	(3.218, 3.437]
4	(3.437, 3.680]
5	(3.681, 4.000]

A map of soil classes is in Figure 2.

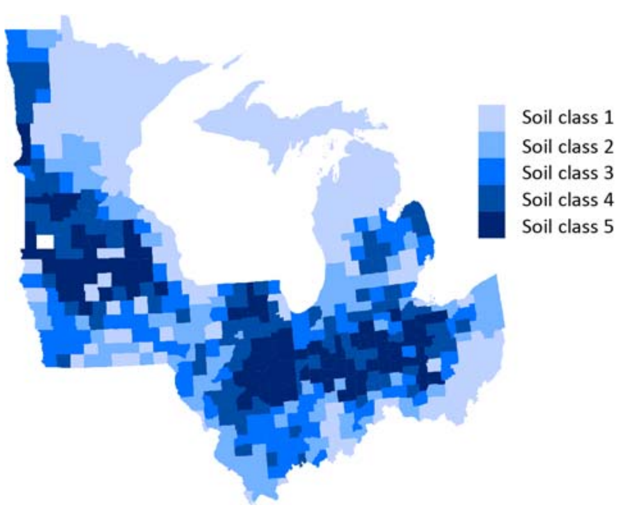

Figure 2

II. Yield models and estimation

For each group of counties in soil class *q* I regress crop *j*'s yield in county *c* in year *t* on growing season weather in county *c* in year *t* and the distribution of land use in county *c* in year *t*.

$$y_{jct} = \alpha_c + \beta_{0j} + \beta_{1j}t + \beta_{2j}G_{jct} + \beta_{3j}G_{jct}^2 + \beta_{4j}R_{jct} + \beta_{5j}R_{jct}^2 + \beta_{6j}A_{jct} + \beta_{7j}A_{kct} + \beta_{8j}A_{mct} \quad \text{for } c \in C_q \qquad (2)$$

where *k* and *m* index the other two crops and *t* ranges from 1 (1950) to 59 (2008). The coefficient α_c "fixes" the unique effect of each county on yield (I fix counties in this model to control for all unobserved variables that are specific to the biophysical, economic, political, and cultural aspects of each county). The omitted land use category in model (2) is the percentage of land in all other uses in county *c* in year *t*. This statistic is given by $100 - A_{jct} - A_{kct} - A_{mct}$.

The *A* variables are included in the model to help explain how crop *j* is allocated over *c*'s soil profile. For example, as $A_{maize,ct}$ increases and approaches 1 it is more probable that the soil profile used to generate maize in *c* at time *t* is well described by the county's soil summary statistic S_c. As $A_{maize,ct}$ falls it is more likely that maize production occupies a niche space of soil that is less well described by S_c. In other words, a lower A_{jct} increases the possibility that production of *j* in *c* is on soils not representative of the soil class that *c* is in. This niche space effect on yield is unknown. A crop that occupies less of a county's space could have a higher density of production on the upper end of the county's soil profile, a higher density of production on the lower end, or, serendipitously, a density that matches the county's overall soil profile. The area variables, therefore, are meant to control for any yield variation that is explained by production on soils in *c* that are in a different soil class than the one that *c* has been assigned to.

I suspect that A_{jct} varies over time due to market forces. I would conjecture that as the price of *j* increases relative to the prices of *k* and *m* and/or the cost of producing *j* falls relative to the cost of producing *k* and *m*, A_{jct} increases relative to A_{kct} and A_{mct}, all else equal. Therefore, in an indirect manner, the *A* variables are controlling for the variation in yield in a county due to market forces.

I estimate model (2) for each soil class q, crop combination. Estimated model results are given in the Tables 2 – 4.

Table 2: Estimate of model (2) for maize

Soil Class		1		2		3		4		5	
Variable	**Coeff.**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**
Time	β_{1j}	1.34	0.00	1.57	0.00	1.55	0.00	1.66	0.00	1.77	0.00
G	β_{2j}	0.14	0.00	0.13	0.00	0.17	0.00	0.23	0.00	0.29	0.00
G^2	β_{3j}	-3.4×10^{-5}	0.00	-3.2×10^{-5}	0.00	-4.2×10^{-5}	0.00	-5.5×10^{-5}	0.00	-6.9×10^{-5}	0.00

R	β_{4j}	0.13	0.00	0.16	0.00	0.16	0.00	0.20	0.00	0.27	0.00
R^2	β_{5j}	$-1.1\text{x}10^{-4}$	0.00	$-1.4\text{x}10^{-4}$	0.00	$-1.4\text{x}10^{-4}$	0.00	$-1.8\ \text{x}10^{-4}$	0.00	$-2.5\text{x}10^{-4}$	0.00
A_{maize}	β_{6j}	1.87	0.00	0.97	0.00	0.86	0.00	0.93	0.00	0.94	0.00
$A_{soybeans}$	β_{7j}	1.86	0.00	0.33	0.00	0.33	0.00	0.08	0.12	-0.03	0.57
A_{wheat}	β_{8j}	-1.80	0.00	-0.09	0.52	-0.20	0.02	-0.29	0.00	0.24	0.01
Con.	α_c	-2756	0.00	-3197	0.00	-3202	0.00	-3507	0.00	-3804	0.00
N		4927		5622		5708		5827		5782	
R^2											
within		0.77		0.79		0.81		0.83		0.84	
between		0.71		0.55		0.61		0.56		0.19	
overall		0.72		0.75		0.78		0.79		0.80	

Across all soil classes maize yield exhibits an inverted U shape response to growing season weather. Warmer and wetter seasons have a positive impact on yield, up to a point. Eventually too much warmth and wetness begins to drag yield down. Over time average yields have increased more across counties with the best soils than those with lesser soils. This would suggest that technological development in maize production has utilized the best soils and the management practices used on the better soils versus poorer soils and its related management practices. The positive effect of A_{maize} on yield across all soil classes suggests that the upper end of the county's soil profile is utilized during periods of additional maize cropping in a county. The use of the upper ends of the soil profiles in the lower soil class counties appears to be a competition between maize and wheat production. For example, in soil class 1 as the area of wheat in a county increases average maize yield falls. The sum of $A_{maize,ct}$, $A_{soybeans,ct}$, and $A_{wheat,ct}$ is much higher in the higher soil class categories (see the appendix). Therefore, it is not surprising that the niche space effects are smaller and in some cases not statistically significant in the higher soil classes. In these counties the soil class statistic *S* is much more representative of the soils actually used for crop production.

Table 3: Estimate of model (2) for soybeans

	Soil class	**1**		**2**		**3**		**4**		**5**	
Variable	**Coeff.**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**
Time	β_{1j}	0.40	0.00	0.39	0.00	0.39	0.00	0.39	0.00	0.41	0.00
G	β_{2j}	0.06	0.00	0.08	0.00	0.09	0.00	0.09	0.00	0.09	0.00
G^2	β_{3j}	$-1.5\text{x}10^{-5}$	0.00	$-1.8\text{x}10^{-5}$	0.00	$-2.0\text{x}10^{-5}$	0.00	$-2.1\text{x}10^{-5}$	0.00	$-2.1\text{x}10^{-5}$	0.00
R	β_{4j}	0.05	0.00	0.05	0.00	0.05	0.00	0.07	0.00	0.09	0.00
R^2	β_{5j}	$-4.2\text{x}10^{-5}$	0.00	$-4.9\text{x}10^{-5}$	0.00	$-4.9\text{x}10^{-5}$	0.00	$-6.7\text{x}10^{-5}$	0.00	$-8.7\text{x}10^{-5}$	0.00
A_{maize}	β_{6j}	0.55	0.00	0.24	0.00	0.24	0.00	0.20	0.00	0.23	0.00
$A_{soybeans}$	β_{7j}	0.15	0.00	-0.003	0.87	0.01	0.37	0.02	0.11	$-2.1\text{x}10^{-3}$	0.87
A_{wheat}	β_{8j}	-0.56	0.00	-0.12	0.00	$3.9\text{x}10^{-3}$	0.86	0.01	0.77	0.08	0.00
Con.	α_c	-845	0.00	-843	0.00	-844	0.00	-856	0.00	-898	0.00
N		3368		5581		5706		5810		5782	
R^2											
within		0.71		0.75		0.77		0.78		0.78	
between		0.65		0.67		0.82		0.78		0.67	

overall		0.68		0.73		0.78		0.77		0.77	

The trends in soybean yield as explained by weather , soil, and niche effects are remarkably similar to maize's with one exception. Technological improvement in soybean yield over time has been consistent across soil categories and has not biased towards farms with better soils.

Table 4: Estimate of model (2) for wheat

		1		2		3		4		5	
		Est. Coeff.	**p-value**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**	**Est. Coeff.**	**p-value**
Time	β_{1j}	0.54	0.00	0.58	0.00	0.67	0.00	0.70	0.00	0.79	0.00
G	β_{2j}	$4.6\text{x}10^{-3}$	0.49	$-3.1\text{x}10^{-3}$	0.77	0.01	0.27	$3.3\text{x}10^{-3}$	0.81	0.01	0.49
G^2	β_{3j}	$-3.0\text{x}10^{-6}$	0.15	$-1.7\text{x}10^{-6}$	0.57	$-6.3\text{x}10^{-6}$	0.07	$-3.9\text{x}10^{-6}$	0.34	$-5.8\text{x}10^{-6}$	0.23
R	β_{4j}	0.02	0.00	0.01	0.10	0.03	0.00	0.02	0.01	0.02	0.02
R^2	β_{5j}	$-2.9\text{x}10^{-5}$	0.00	$-2.9\text{x}10^{-5}$	0.00	$-6.0\text{x}10^{-5}$	0.00	$-4.5\text{x}10^{-5}$	0.00	$-3.8\text{x}10^{-5}$	0.00
A_{maize}	β_{6j}	0.17	0.02	0.08	0.01	0.02	0.52	0.01	0.82	-0.07	0.03
$A_{soybeans}$	β_{7j}	0.34	0.00	0.11	0.00	0.01	0.76	0.01	0.59	0.03	0.38
A_{wheat}	β_{8j}	0.06	0.71	0.10	0.10	-0.04	0.28	-0.06	0.06	-0.12	0.01
Con.	α_c	-1043	0.00	-1096	0.00	-1288	0.00	-1347	0.00	-1518	0.00
N		3977		4869		4758		5107		4378	
R^2											
within		0.66		0.67		0.70		0.71		0.74	
between		0.45		0.07		0.01		0.00		0.02	
overall		0.60		0.56		0.58		0.58		0.61	

Estimates of the wheat yield model differ from the previous two estimates in several ways. First, variation in GDD has had no statistically significant effect on wheat yield across all soil categories. Second, the niche effects are more muted in wheat production across all soil classes. Wheat production is similar to maize and soybean production in its response to growing season precipitation and similar to maize production in that the impact of technological change has been biased towards better soils.

III. Establishing a baseline – recent estimated yields

To estimate the effects that changes in climate and cropland soil quality could have on crop yield now and in the future I first establish a baseline of predicted average yields from 2000 to 2008. To do this I use observed explanatory variable levels from 2000 to 2008 and the estimated model results given in the tables above. For example, let $\bar{y}_{maize,q}$ indicate the average 2000 to 2008 maize yield across counties that make up soil class q,

$$\bar{y}_{maize,q} = \frac{\sum_{t=51}^{59}\sum_{c=1}^{C_q}\hat{y}_{maize,ct}}{9\times\sum_{c=1}^{C_q}1} \tag{3}$$

where

$$\hat{y}_{maize,ct} = \hat{\sigma}_{maize,c} + \hat{\beta}_{1,maize}t + \hat{\beta}_{2,maize,}G_{maize,ct} + \hat{\beta}_{3,maize}G^2_{maize,ct} + \hat{\beta}_{4,maize}R_{maize,ct} + \hat{\beta}_{5,maize}R^2_{maize,ct} + \hat{\beta}_{6,maize}A_{maize,ct} + \beta_{7,maize}A_{soybeans,ct} + \hat{\beta}_{8,maize}A_{wheat,ct}$$

$$\text{for } c \in C_q \qquad (4)$$

$\hat{\sigma}_{maize,c}$ indicates the "average constant coefficient (STATA uses an estimated constant coefficient that averages the constant and fixed effect coefficients; see http://www.stata.com/support/faqs/stat/xtreg2.html) and the "^" indicates the estimated coefficients for the given soil class q (see Table 2). I calculate $\bar{y}_{soybeans,q}$ and $\bar{y}_{wheat,q}$ in the same manner. Values for $\bar{y}_{maize,q}$, $\bar{y}_{soybeans,q}$ and $\bar{y}_{wheat,q}$ and their standard deviations are given in Table 5.

Table 5: Predicted average annual yields from 2000 through 2008 (bushels / acre)

Soil class	**1**		**2**		**3**		**4**		**5**	
	Estimate	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**
$\bar{y}_{maize,q}$	121.3	19.93	133.9	15.85	141.4	14.00	147.1	14.68	156.4	11.40
$\bar{y}_{soybeans,q}$	39.0	5.13	40.6	3.93	42.5	3.90	43.8	3.54	46.7	2.77
$\bar{y}_{wheat,q}$	50.2	2.15	53.4	2.59	57.6	2.48	60.3	2.39	66.1	2.21

Table 5 establishes that an average acre of maize, soybeans, and wheat grown on an average acre of land in areas with better soils have better yields, all else equal. In other words, better soils lead to better yields, all else equal. To what extent better yields on better soils are explained by better or more intensive management practices is unknown; I cannot determine that with my model. If greater yields on better soils are partly explained by better or more intensive management practices then better soils generate a positive production externality as well: farmers are more willing to risk time and expense on farming if they know that they are growing crops on better soils. The data in Table 5 is graphed in Figures 2 – 4.

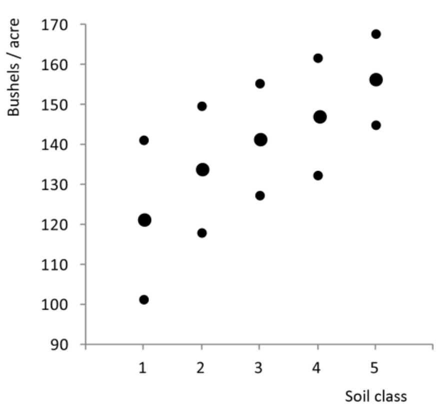

Figure 3: Maize

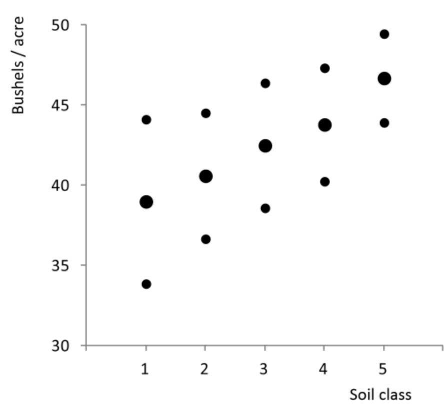

Figure 4: Soybeans

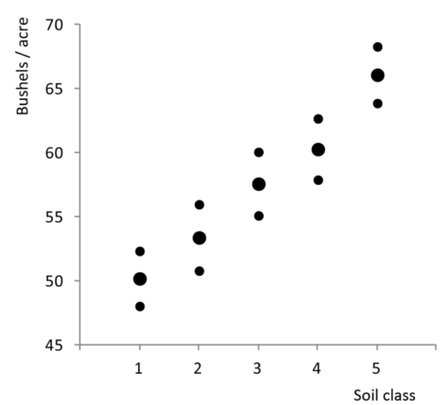

Figure 5: Wheat

IV. The capacity of soil to increase current production

How can American farmers utilize soil resources right now to increase the efficiency of crop production? One method would be to grow more crops on the soil types, LCCs 1 and 2, that are primarily used for cropland in soil class 5. Presumably if non-cropped areas of LCCs 1 and 2 are brought into production they would be able to mimic crop yields found in soil class 5 areas (assuming management practices associated with soil class 5 are also used). The acres in LCCs 1 and 2 that are not currently cropped and not in urban use across each soil class as of 2001 are given in Table 6.[1] As of 2001 there were more than 19 million undeveloped and available acres in the most capable soils in the counties that form soil class 1.

Table 6: Acres available for cropping in LCCs 1 or 2 as of 2001

Soil Class	Acres
5	2,088,003
4	3,362,076
3	4,240,432
2	5,414,875
1	19,183,846

Notes: Data comes from Radeloff et al. (2012)

In Table 7 I present the average number of acres used annually for maize, soybean, and wheat harvest from 2000 to 2008 in each soil class. The data in Tables 6 and 7 suggest that there is more than enough good soils in soil class 1 areas to place all soil class 1 production on the best soils.

Table 7: Average number of acres used annually for harvest from 2000 to 2008

Soil Class	Maize	Soybeans	Wheat	Total
5	12,843,674	11,809,598	723,452	25,376,724
4	10,089,931	9,804,525	1,411,379	21,305,835
3	8,295,655	8,165,158	1,114,306	17,575,119
2	5,575,536	5,201,073	583,584	11,360,193
1	1,314,666	885,613	120,113	2,320,392

For example, suppose all maize, soybeans, and wheat in soil class 1 areas were grown on soils that mimicked those found in soil class 5 (and management practices were also changed accordingly). To estimate the yield impact of this change we can use the weather from soil class 1 counties with soil class 5's estimated yield function where A_{jct} values are set equal to the average values observed across soil class 5 category counties. We use the average A_{jct} values from soil class 5 because they are the values that minimize the niche effects. Niche effects will be minimal when we are targeting specific soils.

[1] Land available for cropland includes protected land formally cropped, protected and unprotected pasture, protected and unprotected forest, and protected and unprotected range.

Let us look at maize specifically. The expected average maize yield in county *c* in soil class 1 for any year from 2000 to 2008 (*t* = 51 through *t* = 59) assuming it uses the best undeveloped soils in its area is given by,

$$\ddot{y}_{maize,ct} = \hat{\sigma}_{maize,5} + \hat{\beta}_{1,maize,5}t + \hat{\beta}_{2,maize,5}G_{maize,ct} + \hat{\beta}_{3,maize,5}G^2_{maize,ct} + \hat{\beta}_{4,maize,5}R_{maize,ct} + \hat{\beta}_{5,maize,5}R^2_{maize,ct} + \hat{\beta}_{6,maize,5}\bar{A}_{maize,5} + \hat{\beta}_{7,maize,5}\bar{A}_{soybeans,5} + \hat{\beta}_{8,maize,5}\bar{A}_{wheat,5}$$

$$\text{for } c \in C_1 \qquad (5)$$

where $\hat{\beta}_{k,maize,5}$ indicates that the estimated coefficient is from the soil class 5 maize model and $\bar{A}_{j,5}$ is the average percentage of county area in *j*'s production in soil class 5 counties from 2000 to 2008. Finally, let $\ddot{y}_{maize,1}$ indicate the average 2000 to 2008 maize yield across counties that make up soil class 1 given the targeted use of soils,

$$\ddot{y}_{maize,1} = \frac{\sum_{t=51}^{59}\sum_{c=1}^{C_1}\ddot{y}_{maize,ct}}{9\times\sum_{c=1}^{C_1}1} \qquad (6)$$

I calculate $\ddot{y}_{soybeans,1}$ and $\ddot{y}_{wheat,1}$ in the same manner. Values for $\ddot{y}_{maize,1}$, $\ddot{y}_{soybeans,1}$ and $\ddot{y}_{wheat,1}$ and their standard deviations are given in Table 8.

Table 8: Targeted use of best soils across soil class 1 areas

	Average annual yield on best soils from 2000-08 (bu / acre)		Observed average annual yield from 2000-08 (bu / acre)		Average annual harvested area from 2000-08 (acres)	Gain in annual harvest (bushels)
	Estimate	SD	Estimate	SD		
$\ddot{y}_{maize,1}$	147.9	12.16	121.3	19.93	1,314,666	34,970,116
$\ddot{y}_{soybeans,1}$	43.4	3.83	39.0	5.13	885,613	3,896,697
$\ddot{y}_{wheat,1}$	66.1	2.21	50.2	2.15	120,113	1,909,797

The first four columns of Table 8 are graphed in Figures 6-8.

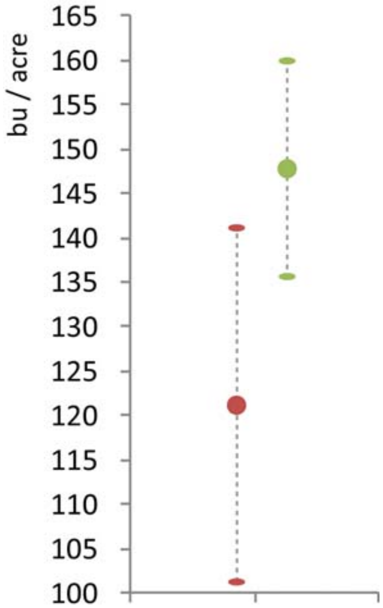

Figure 6: Maize

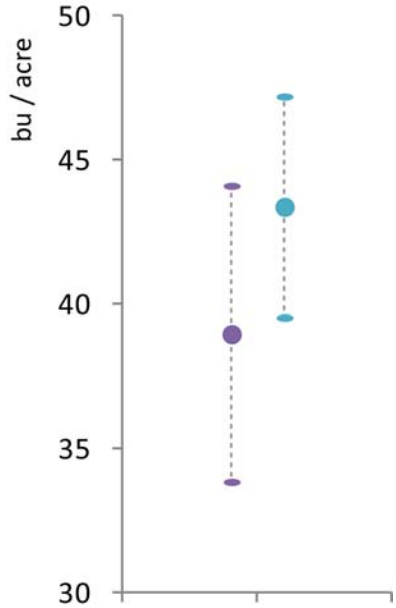

Figure 7: Soybeans

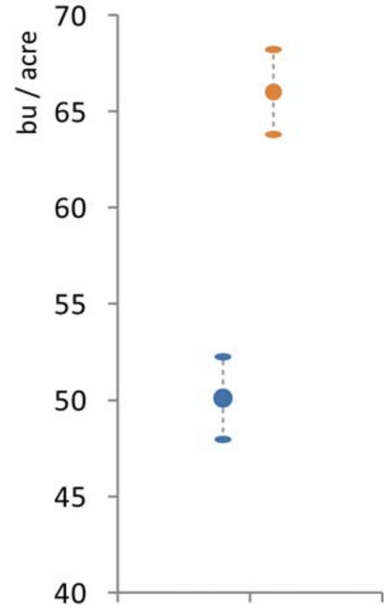

Figure 8: Wheat

How practical would such a reallocation of land in soil class 1 areas be? Some soil class 1 cropland is surely already on the best soils, so not all 2,320,392 cropped acres in soil class 1 would have to move to mimic soil use in soil class 5. However, the gain in crop production might not be worth the loss in ecosystems services from converting the undeveloped land to highly managed cropland.

Another way to leverage better soils would be to marginally improve the soils already used for cropland (I assume a non-marginal improvement may be physically impossible and/or prohibitively expensive). For example, suppose soils in counties in soil class 1 are marginally improved such that these counties now have *S* scores that would move them into soil class 2. What would be the impact of such an improvement? Further, suppose soils in class 2 are marginally improved such that they became equivalent to class 3, etc.

Let us look at maize specifically. The expected average annual maize yield in county *c* in soil class $q - 1$ for any year from 2000 to 2008 (t = 51 through t = 59) assuming its soil improves such that it becomes a member of soil class q is given by,

$$\acute{y}_{maize,ct} = \hat{\sigma}_{maize,q} + \hat{\beta}_{1,maize,q}t + \hat{\beta}_{2,maize,q}G_{maize,ct} + \hat{\beta}_{3,maize,q}G^2_{maize,ct} + \hat{\beta}_{4,maize,q}R_{maize,ct} + \hat{\beta}_{5,maize,q}R^2_{maize,ct} + \hat{\beta}_{6,maize,q}\bar{A}_{maize,q} + \hat{\beta}_{7,maize,q}\bar{A}_{soybeans,q} + \hat{\beta}_{8,maize,q}\bar{A}_{wheat,q}$$

$$\text{for } c \in C_{q-1} \qquad (7)$$

where $\hat{\beta}_{k,maize,q}$ indicates that the estimated coefficient is from the soil class q maize model and $\bar{A}_{j,q}$ is the average proportion of county area in j's production in soil class q counties across the years 2000 through 2008. We use the q^{th} class' annual average A_{jct} values with their estimated coefficients because they control for the observed niche effects in that soil class. I do not know how farmers will spatially react to marginally better soils. The data on $\bar{A}_{j,q}$ is the best information I have.

Finally Let $\acute{y}_{maize,q-1}$ indicate the average 2000 to 2008 maize yield across counties that make up soil class q – 1 given that their soil profile has reached class q status,

$$\acute{y}_{maize,q-1} = \frac{\sum_{t=51}^{59}\sum_{c=1}^{C_{q-1}} \acute{y}_{maize,ct}}{9\times\sum_{c=1}^{C_{q-1}} 1} \qquad (8)$$

I calculate $\acute{y}_{soybeans,q-1}$ and $\acute{y}_{wheat,q-1}$ in the same manner. Values for $\acute{y}_{maize,q-1}$, $\acute{y}_{soybeans,q-1}$ and $\acute{y}_{wheat,q-1}$ and their standard deviations are given in Table 9.

Table 9: Predicted average annual yields (bu / acre) from 2000 through 2008 with marginal soil improvement

	***q* = 2**		***q* = 3**		***q* = 4**		***q* = 5**	
	Estimate	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**
$\bar{y}_{maize,q}$	136.4	6.87	135.9	12.08	140.3	11.26	151.3	10.59
$\bar{y}_{soybeans,q}$	39.2	3.21	41.4	2.61	42.9	2.30	45.4	2.28
$\bar{y}_{wheat,q}$	54.6	3.23	57.5	2.66	60.2	2.78	67.0	2.37

Data from Table 9 (and from Table 5 for comparison) are presented in Figures 9 – 11. Figure 9 gives maize results, Table 10 gives soybean results, and Table 11 gives wheat results.

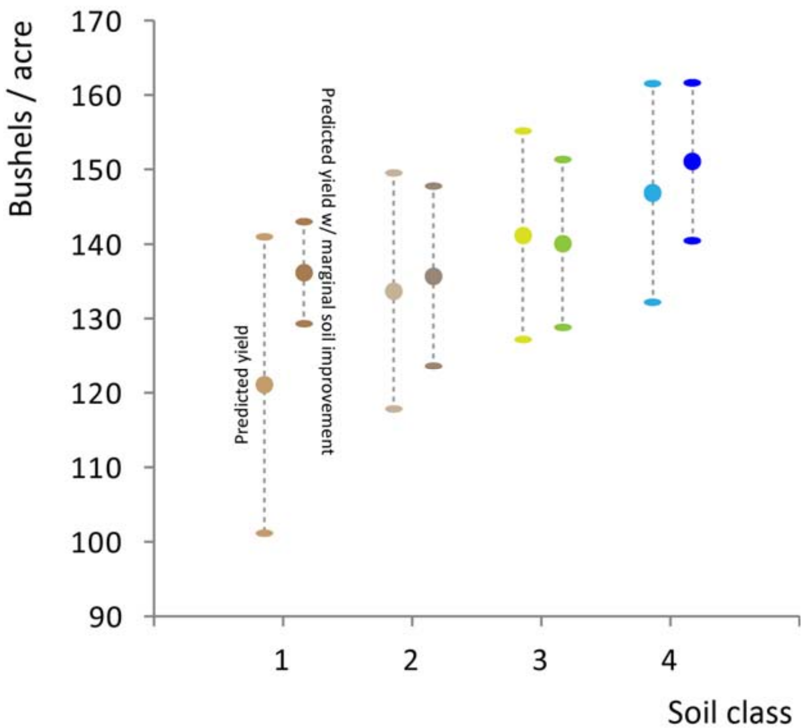

Figure 9: Maize

TBA

Figure 10

TBA

Figure 11

V. Climate change

In Table 10 I give the 1950 through 1958 and 2000 through 2008 GDD and growing season precipitation averages across soil categories and crop types. For example, let $\bar{G}_{maize,c,5058}$ indicate the average GDD from 1950 to 1958 across counties in soil class q during the maize growing season.

$$\bar{G}_{maize,q,5058} = \frac{\sum_{t=1}^{9}\sum_{c=1}^{C_q} G_{maize,ct}}{9\times\sum_{c=1}^{C_q} 1} \tag{9}$$

The variable $\bar{G}_{maize,q,0008}$ indicate the average GDD from 2000 to 2008 across counties in soil class q during the maize growing season.

$$\bar{G}_{maize,q,0008} = \frac{\sum_{t=51}^{59}\sum_{c=1}^{C_q} G_{maize,ct}}{9\times\sum_{c=1}^{C_q} 1} \tag{10}$$

I calculate $\bar{R}_{maize,q,5058}$, $\bar{R}_{maize,q,0008}$, $\bar{G}_{soybeans,q,5058}$, $\bar{G}_{soybeans,q,0008}$ $\bar{R}_{soybeans,q,5058}$, $\bar{R}_{soybeans,q,0008}$, $\bar{G}_{wheat,q,5058}$, $\bar{G}_{wheat,q,0008}$, $\bar{R}_{wheat,q,5058}$, and $\bar{R}_{wheat,q,0008}$

In the last rows of Table 10 I indicate the change in growing season weather across those two time periods for each soil class, crop combination. Recall that I assume planting dates have not changed across this time period.

Table 10

		$\bar{G}_{maize,q}$	$\bar{R}_{maize,q}$	$\bar{G}_{soybeans,q}$	$\bar{R}_{soybean,q}$	$\bar{G}_{wheat,q}$	$\bar{R}_{wheat,q}$
1950-1958	1	2084	450	2108	410	1624	345
	SD	389	99	269	102	285	92
	2	2404	464	2175	403	1767	360
	SD	342	114	290	104	202	102
	3	2387	450	2165	391	1764	346
	SD	308	113	267	103	191	100
	4	2272	444	2066	387	1719	351
	SD	270	107	231	97	157	88
	5	2281	460	2071	403	1788	367
	SD	213	110	188	101	152	92
2000-2008	1	2178	499	2052	436	1616	335
	SD	315	107	259	99	215	94
	2	2427	517	2196	440	1764	374
	SD	329	111	276	98	179	89
	3	2391	512	2167	437	1754	374
	SD	282	108	241	96	163	90
	4	2292	503	2082	431	1726	372
	SD	245	103	210	93	139	85
	5	2301	521	2092	448	1796	398
	SD	189	101	164	90	125	91
		Percentage change between 1950-1958 and 2000-2008					
	1	4.5%	10.7%	-2.7%	6.4%	-0.5%	-2.8%
	2	1.0%	11.5%	0.9%	9.1%	-0.2%	3.9%
	3	0.2%	14.0%	0.1%	11.7%	-0.6%	8.3%
	4	0.9%	13.2%	0.8%	11.2%	0.4%	6.0%
	5	0.9%	13.2%	1.0%	11.4%	0.5%	8.3%

Note: Standard deviation is calculated over all n x T observations

Between the time period of 1950-1958 and 2000-2008, GDD across all 2 crops and all 5 soil classes barely changed. The greatest change was a 4.5% increase in average GDD for maize across soil class 1 areas. Changes in average annual growing season precipitation were much more dramatic, increasing by up to 14% for maize across soil class 3 areas. Figure 12 graphs the changes in GDD and growing season precipitation average from 1950-1958 and 2000-2008.

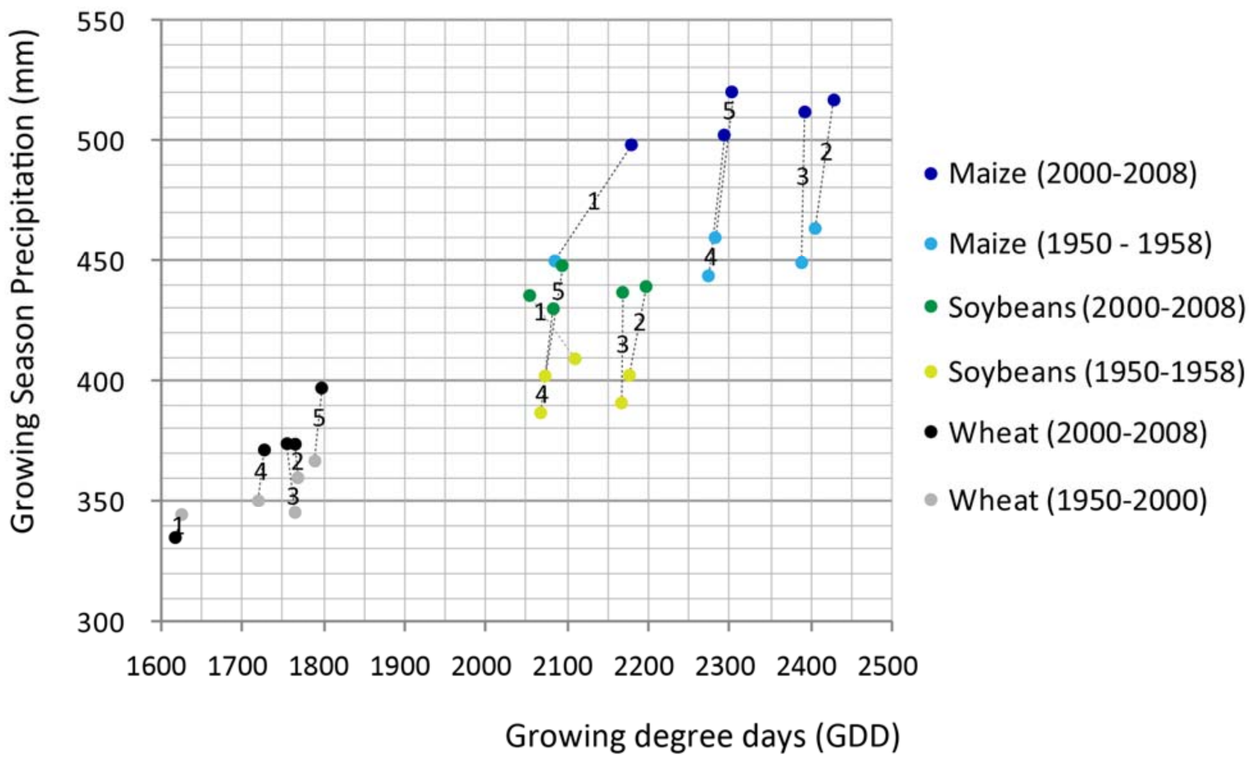

Figure 12

Most climate models predict much more rapid climate change over these 5 states in the next 50 years. In addition, can we expect rates of yield improvement as explained by time (a proxy for technological progress) to continue at the rate seen from 1950 to 2008?

In Tables 11 and 12 I present expected yield for each crop over the period 2050 to 2058 under various assumptions of climate change and technological progress.

To generate these values for maize in soil class q, for example, I calculate the following for each c and t from 51 through 59,

$$\begin{aligned}\tilde{y}_{maize,ct} = \hat{\sigma}_{maize,c} + \hat{\beta}_{0,maize,} + \hat{\beta}_{1,maize}(t+x) + \hat{\beta}_{2,maize,}\left(\gamma G_{maize,ct}\right) \\ +\hat{\beta}_{3,maize}\left(\gamma G_{maize,ct}\right)^2 + \hat{\beta}_{4,maize}\left(\theta R_{maize,ct}\right) \\ +\hat{\beta}_{5,maize}\left(\theta R_{maize,ct}\right)^2 + \hat{\beta}_{6,maize} A_{maize,ct} \\ +\beta_{7,maize} A_{soybeans,ct} + \hat{\beta}_{8,maize} A_{wheat,ct} \\ \text{for } c \in C_q \qquad (11)\end{aligned}$$

where x = 50 if I assume technological progress will continue into the future at the rate that it did from 1950 to 2008 and x = 25 if I assume technological progress will continue into the future at half the rate it did from 1950 to 2008. The parameters γ and θ indicate the expected change in annual GDD and growing season precipitation, respectively. For example, $\gamma = 1.20$ would increase each observation of maize GDD by 20%.

For example, assume *t* = 51 (the year 2000). Assume we expect annual maize GDD and growing season precipitation to increase 20% between 2000-2008 and 2050-2058 and technology to progress at the historical rate (*x* = 50). Thus $\tilde{y}_{maize,c,51}$is predicted maize yield in 2050 in county *c* assuming $\gamma = \theta = 1.20$ and *x* = 50. Now assume *t* = 51, $\gamma = \theta = 1.20$, and x = 25. In this case, $\tilde{y}_{maize,c,51}$is predicted maize yield in 2050 in county *c* assuming $\gamma = \theta = 1.20$ and technology in 2050 is such that is grew at half the historical rate. In both these examples I assume that for $A_{maize,ct}$, $A_{soybeans,ct}$, and $A_{wheat,ct}$ in 2050 is the same as it was in 2000.

When I calculate for a given γ, θ, and x,

$$\tilde{y}_{maize,q} = \frac{\sum_{t=51}^{59}\sum_{c=1}^{C_{maize,q}} \tilde{y}_{maize,ct}}{9\times\sum_{c=1}^{C_{maize,q}} 1} \quad (12)$$

I am finding expected average annual yield for maize in soil class *q* from 2050 to 2058 assuming the 2000 to 2008 observations for $A_{maize,ct}$, $A_{soybeans,ct}$, and $A_{wheat,ct}$ and 50-year change in climate and technology given by γ, θ, and *x*. I find $\tilde{y}_{soybeans,q}$ and $\tilde{y}_{wheat,q}$ in a similar way. $\tilde{y}_{maize,q}$, $\tilde{y}_{soybeans,q}$, and $\tilde{y}_{wheat,q}$ for various combinations of γ, θ, and *x* are given in Tables 11 and 12.

Table 11: Historic technological change (*x* = 50)

		1		2		3		4		5	
$\gamma = \theta =$ 1 (No CC)	Maize	188.5	19.93	212.3	15.85	218.6	14.00	230.3	14.68	244.9	11.40
	Soybeans	59.1	5.13	60.2	3.93	61.9	3.90	63.4	3.54	67.14	2.77
	Wheat	77.3	2.15	82.2	2.59	91.0	2.48	95.5	2.39	105.6	2.21
$\gamma = \theta = 1.1$	Maize	184.4	20.23	203.6	19.70	209.4	17.20	222.1	17.38	234.8	15.53
	Soybeans	58.7	5.28	58.8	4.99	60.9	4.12	62.8	3.60	66.7	2.91
	Wheat	76.3	2.29	80.0	2.86	88.6	3.14	93.0	2.79	103.2	2.48
$\gamma = \theta = 1.2$	Maize	176.6	22.95	190.3	25.72	194.6	23.34	207.2	23.61	215.9	22.73
	Soybeans	56.8	6.20	55.4	7.20	57.7	5.75	60.1	4.97	66.0	4.10
	Wheat	75.0	2.57	77.6	3.21	85.5	3.97	90.1	3.28	100.3	2.85
$\gamma = 1.1; \theta = 1.2$	Maize	184.4	20.18	202.8	19.92	208.7	17.32	221.2	17.51	232.8	16.21
	Soybeans	58.8	5.23	58.8	4.95	61.0	4.05	62.7	3.56	66.4	3.07
	Wheat	76.1	2.44	79.4	3.08	87.7	3.62	92.2	3.14	102.5	2.74
$\gamma = 1.2; \theta = 1.1$	Maize	176.6	22.71	191.2	25.55	195.3	23.33	208.1	23.52	217.9	23.37
	Soybeans	56.7	6.23	55.4	7.25	57.7	5.80	60.1	5.01	64.3	4.02
	Wheat	75.2	2.42	78.2	3.00	86.4	3.53	90.9	2.95	101.0	2.62

Table 12: Slowed technology growth (*x* = 25)

		1		2		3		4		5	
$\gamma = \theta =$ 1 (No CC)	Maize	154.9	19.93	173.1	15.85	180.0	14.00	188.7	14.68	200.6	11.40
	Soybeans	49.1	5.13	50.4	3.93	52.2	3.90	53.6	3.54	56.9	2.77
	Wheat	63.8	2.15	67.8	2.59	74.3	2.48	77.9	2.39	85.8	2.21
$\gamma = \theta = 1.1$	Maize	150.8	20.23	164.4	19.70	170.8	17.20	180.5	17.38	190.5	15.53
	Soybeans	48.6	5.28	49.0	4.99	51.2	4.12	53.0	3.60	56.5	2.91
	Wheat	62.7	2.29	65.6	2.86	71.9	3.14	75.4	2.79	83.5	2.48
$\gamma = \theta = 1.2$	Maize	143.0	22.75	151.1	25.72	156.0	23.34	165.6	23.61	171.6	22.73
	Soybeans	46.8	6.20	45.6	7.20	48.1	5.75	50.3	4.97	53.8	4.10
	Wheat	61.4	2.57	63.2	3.21	68.8	3.97	72.5	3.28	80.6	2.85
$\gamma = 1.1; \theta = 1.2$	Maize	150.8	20.18	163.6	19.92	170.1	17.32	179.6	17.51	188.5	16.21
	Soybeans	48.7	5.23	49.0	4.95	51.3	4.05	53.0	3.56	56.2	3.07
	Wheat	62.5	2.44	65.0	3.08	71.0	3.62	74.6	3.14	82.8	2.74
$\gamma = 1.2; \theta = 1.1$	Maize	143.0	22.71	152.0	25.55	156.7	23.2	166.5	23.52	173.6	22.37
	Soybeans	46.7	6.23	45.5	7.25	47.0	5.80	50.3	5.01	54.1	4.02
	Wheat	61.6	2.42	63.8	3.00	69.7	3.53	73.3	2.95	81.3	2.62

Given climate change what would the impact be of marginally improving the soils used for cropland? This is the same analysis as we did above given current weather / climate conditions (equations (7)-(8) and Table 9) except now we are considering a future with climate change. For example, suppose soils in counties in soil class 1 are marginally improved such that these counties now have *S* scores that would move them into soil class 2. What would be the impact of such an improvement be in the future given assumptions for climate change and technological improvement? In this case we assume that technological improvement in crop production is moderate (x = 25) and GDD and growing season precipitation increase 20% across the entire study area.

Let us look at maize specifically. The expected average annual maize yield in county *c* in soil class *q* – 1 for any year from 2050 to 2058 assuming its soil improves such that it becomes a member of soil class *q* is given by,

$$\dddot{y}_{maize,ct} = \hat{\sigma}_{maize,q} + \hat{\beta}_{1,maize,q}(t + 25) + \hat{\beta}_{2,maize,q}\left(1.2G_{maize,ct}\right) + \hat{\beta}_{3,maize,q}\left(1.2G_{maize,ct}\right)^2 + \hat{\beta}_{4,maize,q}1.2R_{maize,ct} + \hat{\beta}_{5,maize,q}\left(1.2R_{maize,ct}\right)^2 + \hat{\beta}_{6,maize,q}\bar{A}_{maize,q} + \hat{\beta}_{7,maize,q}\bar{A}_{soybeans,q} + \hat{\beta}_{8,maize,q}\bar{A}_{wheat,q}$$

$$\text{for } c \in C_{q-1} \qquad (13)$$

where $\hat{\beta}_{k,maize,q}$ indicates that the estimated coefficient is from the soil class *q* maize model and $\bar{A}_{j,q}$ is the average proportion of county area in *j*'s production in soil class *q* counties across the years

2000 through 2008. As before we use the q^{th} class' annual average A_{jct} values with their estimated coefficients because they control for the observed niche effects in that soil class. I do not know how farmers will spatially react to marginally better soils. The data on $\bar{A}_{j,q}$ is the best information I have. Therefore, $\ddot{y}_{maize,c,51}$ can be interpreted as the expected yield of maize in county *c* in the year 2050 assuming technological improvement in crop production is moderate and climate change is described by $\gamma = \theta = 1.20$

Finally Let $\dddot{y}_{maize,q-1}$ indicate the average 2050 to 2058 maize yield across counties that make up soil class *q* – 1 given that their soil profile has reached class *q* status,

$$\dddot{y}_{maize,q-1} = \frac{\sum_{t=51}^{59} \sum_{c=1}^{C_{q-1}} \ddot{y}_{maize,ct}}{9 \times \sum_{c=1}^{C_{q-1}} 1} \tag{14}$$

I calculate $\dddot{y}_{soybeans,q-1}$ and $\dddot{y}_{wheat,q-1}$ in the same manner. Values for $\dddot{y}_{maize,q-1}$, $\dddot{y}_{soybeans,q-1}$ and $\dddot{y}_{wheat,q-1}$ and their standard deviations are given in Table 13.

Table 13: Predicted average annual yields (bu / acre) from 2050 through 2058 with marginal soil improvement ($\gamma = \theta = 1.2$ and *x* = 25)

	***q* = 2**		***q* = 3**		***q* = 4**		***q* = 5**	
	Estimate	**SD**	**Estimate**	**SD**	**Estimate**	**SD**	**Estimate**	**SD**
$\bar{y}_{maize,q}$	166.3	17.74	148.1	29.6	151.5	30.14	167.7	29.18
$\bar{y}_{soybeans,q}$	48.7	4.36	46.2	5.80	47.4	6.36	52.8	4.69
$\bar{y}_{wheat,q}$	65.0	4.43	68.6	4.24	72.3	3.96	82.3	3.17