Impacts of Non-Grade Quality Factors on North Dakota Origin Soybean Basis Values: A Panel Regression Analysis

by

David W. Bullock, William W. Wilson, and
Ryan Thompson

Suggested citation format:
Bullock, D. W., W. W. Wilson, and R. Thompson. 2021. "Impacts of Non-Grade Quality Factors on North Dakota Origin Soybean Basis Values: A Panel Regression Analysis." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. [http://www.farmdoc.illinois.edu/nccc134].

Impacts of Non-Grade Quality Factors on North Dakota Origin Soybean Basis Values: A Panel Regression Analysis

David W. Bullock, William W. Wilson, and Ryan Thompson[*]

Paper presented at the NCCC-134 Conference on Applied Commodity Analysis, Forecasting, and Market Risk Management, 2021.

---

[*] Research Associate Professor, and University Distinguished Professor and CHS Endowed Chair in Risk and Trading in the Department of Agribusiness and Applied Economics, North Dakota State University; and Market Risk Specialist at The Gavilon Group, LLC, respectively. Address all correspondence to David W. Bullock, NDSU Dept 7610, P.O. Box 6050, Fargo, ND 58108-6050 or by e-mail at david.w.bullock@ndsu.edu. Funding for this project was provided by a grant from the North Dakota Soybean Council.

Impacts of Non-Grade Quality Factors on North Dakota Origin Soybean Basis Values: A Panel Regression Analysis

The impact of protein and essential amino acid (cysteine, lysine, methionine, threonine, and tryptophan) content upon North Dakota origin basis values was examined using a panel dataset covering eight Crop Reporting Districts (CRDs) over ten marketing years (2009/10 through 2018/19). Fixed-effect panel regression models were estimated based upon marketing year and quarterly averages to capture seasonal effects related to the timing of quality report releases. Principal component analysis (PCA) was applied to the amino acid quality measurements as a data reduction technique and to correct for multicollinearity in the variables. The regression results indicated a statistically significant positive relationship between three essential amino acids (methionine, threonine, and tryptophan) and local basis values in the quarters corresponding with and just following the release of the annual quality reports. Protein content was found to have little to no effect upon local basis values when considered with the amino acid measurements.

Key Words: soybean quality, origin basis, principal components analysis, panel regression, hedonic modeling.

Introduction

There has been substantial growth and changes in the production of soybeans within North America over the past 20 years. Soybean production area has grown northwesterly of traditional producing regions. For illustration, area planted in North Dakota has grown from 640,000 acres in 1995 to 2.9 million acres in 2005 and 5.8 million acres in 2020 (USDA-NASS, 2021). North Dakota ranked 9th in U.S. soybean production in 2019 (Jantzi, Hagemeister, and Krupich, 2020), and ranked 9th in value of soybean exports in 2017 (USDA-ERS, 2021). Similar changes are occurring with increased production in Canada. This growth has been driven by a number of factors including improved technology relative to cereals (notably but not limited to genetic modification), changes in the length of growing season (i.e., days between last spring frost and first fall frost), strong growth in demand driven primarily by China, and the adoption of more efficient logistical systems favoring the Pacific Northwest (PNW) export market, among others. North Dakota is one of the primary origin markets for PNW soybeans.

Coinciding with this growth has been an increase in quality heterogeneity. Soybeans grown in the northern Midwest region (including North Dakota) have historically tended to have lower and more variable protein levels when compared to soybeans from other regions of the United States (Breene et al., 1988; Hurburgh et al., 1990); however, in recent years the gap has narrowed (Naeve and Miller-Garvin, 2019).

Additionally, the current domestic marketing system for soybeans has failed to widely adopt a more differentiated (e.g., specifications, testing, premiums and discounts for quality deviations, segregation, blending, etc.) procurement strategy such as is common in other crops such as wheat (Hertsgaard, Wilson, and Dahl, 2019). Most soybeans are bought and sold on grades, and some buyers may place a limit on protein levels. This is notwithstanding that specific levels of essential amino acids (EAA) are required by many soybean end-users, including the very large Chinese export market (Wilson, Dahl, and Hertsgaard, 2020).

The combination of greater quality variability and demand specifications results in a high degree of uncertainty for buyers regarding end-use quality. As a result, it is becoming more common for soybean buyers to apply discounts to broad purchase areas. As an example, it is common for Pacific Northwest (PNW) soybeans to be discounted up to 40 cents per bushel relative to the U.S. Gulf on a China delivery basis (Wilson, 2016). Other buyers preclude specific origins (e.g., PNW) due to their perception of historically lower protein levels.

For years, end-users have used crude protein content as a non-grade measure of soybean quality, particularly as a proxy for the livestock feed use value in soybean meal. However, recent studies have found crude protein to be a poor predictor of overall feed quality of soybean meal (Ravindran, Abdollahi, and Bootwalla, 2014). In particular, lower crude protein soybeans tend to have a higher proportion of the five most critical essential amino acids (EAA) including lysine, cysteine, methionine, threonine, and tryptophan that are critical for livestock digestibility when compared to higher crude protein soybeans (Thakur and Hurburgh, 2007; Medic, Atkinson, and Hurburgh, 2014).

Despite the lack of differentiated marketing practices regarding value in soybeans, buyers are aware of quality differentials and may procure accordingly. For the past 15 years, the U.S. Soybean Export Council (USSEC) and the North Dakota Soybean Council (NDSC) have collected soybean quality data (including protein and amino acid content) by USDA Crop Reporting Districts (CRD's). This data is released in a summary report on an annual basis (Northern Crops Institute, 2019). The report is generally released to the public during January / February following harvest. In addition to these reports, some crushing firms and/or handlers develop their own internal geographic profiles of soybean end-use quality. Since this data is available, it may impact the local soybean basis values. Explicit premium/discount schedules for protein and essential amino acid content are not reported at terminal markets. Consequently, it is likely that price differentials would be reflected in the origin basis levels. The hypothesis is that protein and EAA conscious buyers would be able to view this public data on a regional basis and adjust their local price (basis) bids in order to acquire the desired quality of soybean.

The purpose of this paper is to develop an econometric model to estimate the impacts of quality variability on regional basis values in soybeans. Specifically, we develop a hedonic model and estimated using cross-sectional, time series panel dataset covering 8 North Dakota CRD's over a period of 10 years (2009/10 to 2018/19) for a total of 80 observations. The data includes quality observations on protein content and five essential amino acids (cysteine, lysine, methionine, threonine, tryptophan). In addition, this data was supplemented with soybean basis data (weekly) within each CRD. To control for external effects, additional data on PNW export basis (to control for international market effects), secondary car values (to control for transportation cost effects), and nearby futures spreads (to control for domestic supply/demand factors) were also be included in the dataset.

Using the dataset, five panel regression models were identified and estimated: one using marketing year averages and four using quarterly averages (to account for seasonality). Principal components analysis (PCA) was applied to the five EAA variables to reduce the number of variables in the panel regression and to control for the strong presence of multicollinearity in the EAA variables. The dependent variable in each regression was the average local basis value with quality (EAA and protein) and control variables (to account for additional market and logistics factors) as the independent variables. Coefficient signs and t-tests were used to test for the significance of the quality (protein and EAA) variables with regards to their impact upon the local basis levels.

This paper has two major contributions. One is that despite the importance of quality variability and buyer demands, in addition to a multitude of technical studies, there have been few economic studies on how the market values these non-grade quality attributes. Second, the model specified in this paper builds upon the hedonic modeling literature, which historically has mostly been applied to cereals and other crops, by estimating hedonic market values in soybeans.

Background and Previous Studies

Soybean quality has been analyzed in recent years from both a livestock feeding (Mourtzinis et al., 2018) and human food (Sudarić, 2020) basis. In the United States, quality variability from a protein perspective (Breene et al., 1988; Hurburgh et al., 1990) has been well documented as soybean production expanded into the northernmost states of the Midwest region. The protein variability issue was heightened in 2017 when the U.S. soybean crop had lower protein content across all regions and drew concern for meeting the par delivery requirements for CBOT soybean meal futures along with shifting Chinese demand away from the United States and towards Brazil (Plume, 2017; Plume, 2018).

It is important to note that, for the 2018/19 to 2020/21 marketing years, exports are projected to account for approximately 45% and 27% of total use by volume for U.S. raw soybeans and soymeal respectively (Wilson, Golden, and Hubbs, 2021). According to the Soy Transportation Coalition (2017), the Pacific Northwest (PNW) comprised approximately 24.2% of total U.S. soybean exports (by volume) in 2017 – second to the U.S. Gulf at a 60.2% share. Much of the growth in the PNW soybean export market can be attributed to the growth in prominence of China as an importer of U.S. soybeans. In 2014, China accounted for 60.7% of U.S. soybean exports (by volume), making it the largest and most dominant export destination for U.S. soybeans (Hart and Schulz, 2015).

The existing marketing practices for soybeans have not fully adapted to the impacts of non-grade quality variability. Most purchases of soybeans at origin country elevators are normally based solely on USDA grade factors with testing and segregation by non-grade quality factors more the exception rather than the rule (Wilson, 2016). However, due to concerns of the lower protein level, some end-use buyers have started adding clauses regarding none-grade quality deviations to their purchase contracts. The primary origins tributary to the PNW export market are primarily located in the regions known for producing lower protein soybeans. Therefore, many quality conscious end-use buyers have avoided purchasing soybeans from the PNW. As a result, it is not uncommon in trading for PNW values to be a discount to the US Gulf, which is normally at a discount to Brazil, for quality reasons (Wilson, Dahl, and Hertsgaard, 2020).

Protein and EAA measures are quality attributes and knowing and/or limiting their values is important to buyers. Marketing practices for soybeans differ from other agricultural commodities (e.g., wheat, durum, barley and other cereals) in which important attributes are measured and are contract terms throughout the marketing system (Wilson and Dahl, 2009). As examples, in the case of hard wheats it is common to specify grade, and inter-grade limits on individual grade factors, protein, color, falling number, vomitoxin, etc. These are measured throughout the marketing system and provide important signals to all market participants. These signals have impacts on grower agronomic decisions (variety choice, plating time, fertilizer etc.), handlers on segregation and shipping, to end-users, in addition to breeders. These are important signals to the entire marketing system. Similar practices have not yet emerged in the soybean sector.

Quality Sourcing Strategies for Soybeans

The more recent studies on the economics of soybean quality include Hertsgaard, Wilson, and Dahl (2019) and Wilson, Dahl, and Hertsgaard (2020) which addressed issues related to alternative specifications for soybean in order to achieve end-use quality. This was premised that most buyers specified protein or may not even do that. They developed a model to determine the optimal testing strategy, including costs and risks.

These studies were prompted in part due to the common industry practice of addressing protein deficiencies by purchasing shipments from different origins and blending them to meet end-use requirements. In this case, exporters would buy shuttle trains of soybeans from North Dakota, typically at a lower delivered cost; and blend them with higher-cost, higher-protein soybeans from Nebraska. Alternatively, foreign importers can (and do) buy simultaneously from the PNW and US Gulf and blend at the import destination.

This strategy was evaluated in an extended model (Wilson, Dahl, and Hertsgaard, 2020) that is similar to models previously developed to examine wheat quality strategies (Wilson, 1989; Wilson and Dahl, 2006; Wilson, Dahl, and Jabs, 2007; Wilson and Dahl, 2009). Results indicated that buyer risk can be reduced from 4.42% to 0.25% (interpreted as the probability of not meeting the buyer requirements) by testing and blending PNW shipments with a share from the US Gulf. This results in a monumental strategy of managing logistics and timing, but it is a practice.

Finally, given the growth of blockchain and its potential applications to agriculture it may be possible to mitigate some buyer risks through use of that technology. Lakkakula, Bullock, and Wilson (2021) analyzed the benefit of using blockchain smart contracts to mitigate asymmetric information between a PNW buyer and a North Dakota origin seller regarding the protein content of a soybean shipment for export. A decision tree model that captured the effects of asymmetry of information between the buyer and seller was developed. Results indicated that blockchain could effectively reduce and mitigate the asymmetry of information which was previously a barrier to completing the transaction, resulting in accrued price premiums to the seller.

Hedonic Modeling of Quality Characteristics

The specification used in this study builds on the literature of hedonic pricing models (sometimes referred as the input characteristics model). These have a long history in economics (Lancaster, 1966) and in agricultural economics (Ladd and Martin, 1976; Ladd, 1978). Ultimately the goal of this specification is to determine the marginal value of a change in a product attribute using regression analysis. In general, the model is of the form $P = f(X_1, X_2, \ldots, X_n, \mathbf{Y})$ where P is price, X_i are quality attributes, and $\mathbf{Y}$ are all other factors. The regression coefficients of the X_i (β_i) represents the marginal value of quality attribute i.

This type of market valuation is very important in agriculture. The hedonic value of attributes impacts growers with regards to agronomic decisions, buyers and handlers with regards to segregation and blending, and breeders with regards to trait selection. For many commodities it was a historically held view that the market doesn't pay for quality attributes. Therefore, the historical focus of breeding efforts was on increasing yields, and this has been the case historically for soybeans. If a hedonic value can be attributed to a quality characteristic, it can be included in breeding targets.

Hedonic models have a long history in the agricultural economics literature. These include applications in malting barley (Wilson, 1984), wheat (Wilson, 1989; Espinosa and Goodwin, 1991), international wheat (Ahmadi-Esfahani and Stanmore, 1995), and rice (Brorsen, Grant, and Rister, 1984) among others. In some cases, these commodities have explicit premiums (e.g., protein in wheat) in which case, the use of hedonic econometric analysis is not so appropriate. However, hedonic analysis is particularly important for attributes that are not measured and for which there are implicit prices (e.g., location of production, variety in malting barley, etc.).

Our approach builds on these models. Specifically, we seek to estimate the hedonic value of some characteristics not normally measured in the marketing system. These include protein, as well as the essential amino acid (EAA) attributes, which are important and when information becomes available, there is an implicit value reflected in its price.

Data and Methodology

For crop quality and local basis, the dataset is comprised of observations across 8 North Dakota Crop Reporting Districts (CRDs) and covering 10 marketing years from 2009/10 through 2018/19 for a total of 80 observations. The soybean marketing year covers the months from September through the following August. A map of the North Dakota CRDs is in Figure 1 (Appendix A).

The soybean crop quality data includes measurements of the percentage (dry matter basis in decimal format) of five essential amino acids (cysteine, lysine, methionine, threonine, and tryptophan) that were identified as important in feed rations by Hertsgaard, Wilson, and Dahl (2019) along with the percentage (13% moisture basis in % format) of protein. These values were obtained from annual *North Dakota Soybean Quality Survey* reports produced for the North Dakota Soybean Council by the Northern Crops Institute and the University of Minnesota. This data is summarized by North Dakota CRDs with the exception of the Southwest district (due to insufficient soybean production).

A timeseries boxplot (distribution across CRDs) summary of the soybean quality data is illustrated in Figure 2. The 2009 crop year was characterized by an exceptionally low level of cysteine and lysine with high levels of methionine and threonine. The growing season was characterized by delayed spring planting due to wet weather followed by frost damage due to a rain delayed harvest (Jeradechachai and Tulbek, 2010). The 2014 crop year had exceptionally low levels of all amino acids. This was another year characterized by a late planted crop with cool, wet conditions prevailing throughout the growing season (Miller-Garvin, Orf, and Naeve, 2014). The 2015 crop year shows a structural shift to a higher average protein level for all ND soybeans. This was likely due to the release of a new higher protein variety of soybean (ND1406HP) by North Dakota State University (2014).

For local basis, a single country elevator location was chosen within each CRD based upon availability of data. Table 1 (Appendix B) shows the chosen locations for each of the 8 CRDs. The basis data were from DTN ProphetX (in $ per bushel format).

To control for other macro factors that might influence local soybean basis, additional data was gathered on export basis (PNW in $/bushel), railroad tariffs and fuel surcharges (from Jamestown, ND to PNW in $/railcar), secondary railcar market values (in $/railcar), and the nearby soybean futures spread (1^{st} deferred minus nearby price in $/bushel). The export basis were daily observations of 'track' values from TradeWest Brokerage and were used for their influence (representing international market factors) upon local basis values. The railroad tariff rate (to PNW), were obtained from the Burlington Northern Santa Fe railroad website. The fuel surcharge, and secondary railcar market values were weekly observations and were from TradeWest Brokerage. These were included due to their direct impact on the local basis. The soybean futures price data was daily observations derived from DTN ProphetX and the nearby spread was used as a proxy for the influence of current domestic demand upon the local basis.

All of the basis, price, and transportation data were aggregated into marketing year and quarterly averages were derived. The quarterly averages were used to analyze the potential seasonal effects upon the origin basis. For example, the first quarter (Q1) represents the average of the daily or weekly observations across the months of September through November (i.e., first quarter of the marketing year), while Q2 represents the average across December through February, and so forth. The soybean quality data is released once each year and typically the report is published in late January / early February following the harvesting of the crop.

Principal Component Analysis of Quality Data

Table 2 shows the Spearman rank-order correlation among each of the five essential amino acids and protein content by CRD over the 10-year study period. Only one value was found to be significant (with tryptophan in the Southeast CRD) out of the 40 comparisons. Therefore, protein was treated as a separate quality factor in each of the regression models. This result is also important as it illustrates that protein would be a poor proxy for end-use performance in purchasing strategies.

To test for multicollinearity in the amino acid data, variance inflation factors (VIF) were calculated on the amino acid percent data in column-wise format sorted first by CRD and then by marketing year. Figure 3 shows the estimated VIF factors for each of the five amino acids.

Values greater than 10 generally indicate the present of multicollinearity in the series data. Four of the five amino acids had VIF values greater than 10 which is indicative of multicollinearity in the data.

To account for the multicollinearity in the amino acid variables and also as a data reduction technique, principal components analysis (PCA) was applied to the sorted data for the five amino acids using the XLStats (Addinsoft, 2020) statistical analysis software. From the analysis, two component factor variables (F1 and F2) were retained which accounted for 97.56 percent of the total variability in the five series. Table 3 provides a summary of the eigenvector weightings, correlations, percent contribution, and squared cosines for factor variables.

The eigenvectors represent the linear weights applied to each amino acid variable to construct the principal component factor variable. The correlation listed is the Pearson correlation between the amino acid variable and the factor variable. The contribution percentage shows what percentage of the factor variable is represented by each amino acid variable. For F1, it is evenly represented between all of the amino acids except for cysteine which is highly represented in the F2. The squared cosine essentially shows what percentage of the variability of each amino acid is represented by each of the factor variables. Note that they don't necessarily sum to one since a small amount of variability may be in the discarded factor variables (F3 through F5).

Figure 4 shows a correlation graph of each amino acid with the two retained factor variables. The first factor variable (F1) represents 80.32 percent of the total variability and is positively correlated with all five amino acid variables with lysine having the highest correlation. The second factor variable (F2) is most positively correlated with cysteine and also has a positive correlation with lysine. It has a negative correlation with methionine and threonine and near zero correlation with tryptophan.

Conceptual Model

To determine the impact of non-grade soybean quality factors upon North Dakota origin basis, the following panel regression model was estimated:

$$b_{it} = \alpha + \mathbf{x}_{it}\beta + \mathbf{z}_t\gamma + c_i + u_{it}, \qquad (1)$$

where b_{it} is the average basis observed at origin i and marketing year t, $\mathbf{x}_{it}$ is the observed quality panel data (amino acid factor variables and protein), $\mathbf{z}_t$ is the observed macro factor control variables (PNW basis, rail shipping costs, secondary railcar values, and nearby futures spread) that only vary by marketing year, c_i is the unobserved panel (regional) effect, and u_{it} is the equation residual term. The β in equation 1 represent hedonic values and indicate the change in basis value for a unit change in the associated quality variable. In this study, the form (pooled OLS, random effects, or fixed effects) of the regression model is identified and first estimated using the marketing year average for t. To account for seasonal effects, the identified model is then estimated using quarterly averages for t which results in a total of five regression models (i.e., marketing year, Q1, Q2, Q3, and Q4) estimated.

To identify the form of the panel regression model, a series of tests were conducted. First, a modified Wald test for groupwise heteroscedasticity (H0: $\sigma_i^2 = \sigma^2$ for all 8 regions) was applied.

The test statistic is distributed as chi-squared with 8 degrees of freedom. Rejection of the null hypothesis indicates the presence of heteroscedasticity. The test was applied to the fixed effects regression of equation 1 using the *xttest3* procedure in *Stata* (Statacorp, 2021). If groupwise heteroscedasticity was detected, the remaining tests and the final regressions were conducted using the *vce(robust)* option in Stata.

The second test applied was the Breusch and Pagan (1980) lagrange multiplier test to determine if pooled OLS is the correct model specification. Rejection of the null hypothesis indicates that either the random or fixed effects models would be preferred over pooled OLS. To test this hypothesis the Stata *xttest0* procedure was applied to the estimated random effects model.

Normally, the Hausman (1978) specification test can be utilized to determine the appropriate form (i.e., random or fixed effects) of the panel regression model. Wooldridge (2010) indicates that models of the form of equation 1 can be estimated by either random or fixed effects but notes that the Hausman test cannot be employed for models with regressors that only change with time. This problem is due to singularity in the asymptotic variance matrix of the difference between the random and fixed effects estimators.

Instead, this study utilizes an alternative test proposed by Mundlak (1978) which proposes estimating the following augmented model specification to avoid the singularity issue:

$$b_{it} = \alpha + \mathbf{x}_{it}\beta + \mathbf{z}_t\gamma + \bar{\mathbf{x}}_i\delta + c_i + u_{it}, \tag{2}$$

where the $\bar{\mathbf{x}}_i$ represent the average of the panel quality variables across all time periods. Equation 2 was estimated using random effects GLS estimation and a Wald chi-square test of the null hypothesis that the $\delta_k = 0$ for all *k* quality variables was conducted. Rejection of the null hypothesis would indicate that the fixed effects model was preferred over the random effects.

A summary of the panel regression identification tests is given in Table 4. The modified Wald test supports the presence of groupwise heteroscedasticity in the panel dataset; therefore, models were estimated using the robust error correction option in Stata. The Breusch and Pagan lagrange multiplier test rejects the use of the pooled OLS estimator. Finally, the Mundlak augmented regression test supports using the fixed effects estimation over random effects. Therefore, the estimated models presented in the next section were estimated using the *xtreg* (i.e., panel regression) procedure in Stata with the *fe* (i.e., fixed effects) and *vce(robust)* (i.e., robust error correction) options activated.

Estimation Results

The fixed effects panel regression results are presented in Table 5. The models are presented using the marketing year (MY) and quarterly (Q1-Q4) averages for the dependent (origin basis) and macro variables (transportation costs, PNW export basis, and nearby futures spreads). The quality variables are reported once per year, so they remain the same across all regressions.

For the explanatory variables, *F1* and *F2* represent the amino acid factor variables (extracted via principal components), *Protein* is the average protein content, *Tariff_FSC* is the sum of the railroad tariff and fuel surcharge from Jamestown ND to PNW, *PNW_Basis* is the export basis for the Pacific Northwest (PNW) market, *Sec_Car_Value* is the secondary railcar market value, and *NB_Fut_Sprd* is the nearby CBOT soybean futures price spread. Coefficients followed by three stars are significant at the 99% confidence level, two stars at the 95% confidence level, and one star for the 90% confidence level.

The coefficient estimates on the amino acid factor variables are both highly significant for the two quarters corresponding and immediately subsequent to (Q2 and Q3) the release of the *North Dakota Soybean Quality Survey* reports. Both coefficients are significant at the 99% confidence level with the dominant factor (F1 with 80.3% of total amino acid variability) having a positive sign with the secondary factor (F2 with 17.2% of variability) having a negative sign. Taken together, these would lend support to the hypothesis that the release of essential amino acid information has a statistically significant effect upon the regional North Dakota basis values. The positive sign on the major factor coefficient indicate a general positive correlation between the essential amino acid content and the local basis values. These coefficients are also significant and of the same sign for the overall marketing year average basis value.

The regression results indicate that protein has only a marginal, but positive, impact upon local soybean basis, particularly in the first two quarters of the marketing year. For the most part, the quality variables have a negligible impact upon regional soybean basis in the 1st and 4th quarters of the marketing year. The low overall R^2 value for the Q4 equation indicates that the model has very little explanatory power and is probably due to the fact that most soybean shipments have occurred in the earlier quarters of the marketing year.

In addition to essential amino acid factors, the PNW basis and secondary railcar market values have highly significant coefficient values and are of the correct sign. A large percentage of North Dakota soybeans flow into the Pacific Northwest export market which has a large influence upon local basis values and prices. Previous research (Bullock and Wilson, 2020; Wilson, Bullock, and Lakkakula, 2020) has shown that secondary railcar market values vary more frequently when compared to rail tariffs and therefore, have a greater influence upon basis values for grains and oilseeds.

In terms of the nearby futures price spread, the results appear to be mostly insignificant with the exception of Q3. Bullock and Wilson (2020) found that the nearby futures spread had a strong and statistically significant influence upon the PNW export basis; therefore, it is likely that the futures spread effect is likely picked up through the export basis coefficient.

To determine the impact of each of the individual essential amino acids in the regression models, the eigenvectors from Table 3 were used to calculate the coefficient values and coefficient standard errors. The formula for the individual amino acid coefficients was:

$$\hat{\varphi}_j = \varepsilon_{j1} \cdot \hat{\beta}_{F1} + \varepsilon_{j2} \cdot \hat{\beta}_{F2}, \tag{3}$$

where $\hat{\varphi}_j$ is the estimated regression coefficient (hedonic value) for amino acid j, ε_{jk} is the eigenvector factor loading of amino acid j upon factor k, and $\hat{\beta}_{Fk}$ is the estimated regression coefficient for factor k in Table 5. The formula for the individual amino acid regression coefficient standard errors was:

$$se(\hat{\varphi}_j) = \sqrt{\varepsilon_{j1}^2 \cdot se(\hat{\beta}_{F1})^2 + \varepsilon_{j2}^2 \cdot se(\hat{\beta}_{F2})^2}, \tag{4}$$

where *se(.)* is the regression coefficient standard error.

Table 6 shows the converted regression coefficients and t-statistics (along with significant level) for the individual amino acids. The results show that the primary positive impact upon origin basis is through methionine, threonine, and tryptophan which are all significant at the 99% confidence level for the MY, Q2, and Q3 regressions. Cysteine has a lesser (from a statistical significance perspective) but negative effect upon origin basis. Lysine appears to have an insignificant effect in the three regressions. Cysteine and lysine both have a positive and significant impact in the Q1 regression and a lesser but negative effect in the Q4 regression. In the Q1 regression, methionine and threonine have a marginally significant (90% confidence level) negative impact upon the origin basis.

The negative coefficient value for cysteine may be due to the fact that it is already produced naturally by animals and humans (Karau and Grayson, 2014). Methionine, like cysteine, is one of the sulfur amino acids and is not produced naturally by animals and humans but must come from external sources. Adding cysteine to a ration can reduce the requirement for methionine; however, both amino acids are generally supplied at a fixed ratio in broiler (Pacheco et al., 2018) and other animal rations. Part of the reason for this is that methionine is converted into cysteine in the digestive process and too much of the sulfur amino acids can have a negative impact upon animal growth (Bin, Huang, and Zhou, 2017).

Besides cysteine, the other four amino acids (methionine, lysine, threonine, and tryptophan) are considered the main fermentative amino acids in animal feed and human nutrition (Karau and Grayson, 2014). With regards to the insignificance of the regression coefficients for lysine in the MY, Q2, and Q3 regressions; perhaps the best explanation is that lysine competes with threonine and methionine for synthesis through the aspartate biosynthesis pathway in cereal and legume production (Hacham et al., 2007). Also, there is evidence of optimal ratios of methionine to lysine in lactation rations (Wei et al., 2019).

Summary and Conclusions

Over the past 20 years, North Dakota soybean production has grown significantly due to improved seed technology with the advent of GM varieties (which has displaced non-GM cereal production), climate change, better logistics, and the advent of China as a major importer of soybeans out of the Pacific Northwest (PNW) export market. Coinciding with this growth has been an increase in quality heterogeneity, particularly with respect to non-grade factors such as protein and essential amino acid (EAA) content. As these quality issues have become more prominent, it has induced efforts by the U.S. Soybean Export Council (USSEC) and state-level soybean organizations to begin annual reporting of non-grade quality factors. In North Dakota, this effort began with the 2009 crop year and includes quality information summarized by Crop Reporting District (CRD).

The main purpose of this study is to examine and measure the impact that this information may have on North Dakota origin basis values. The hypothesis is that buyers monitor the release of this summary information and adjust their local basis bids accordingly in order to source the desired quality of soybean in terms of protein and EAA. This "origin targeting" strategy has been used historically in wheat procurement (Wilson and Dahl, 2009) to acquire the desired quality.

In this study, panel quality data on protein content and five essential amino acids (cysteine, lysine, methionine, threonine, and tryptophan) was collected from 8 North Dakota CRD's over 10 marketing years (2009/10 through 2018/19). Within each CRD, a local elevator was chosen and weekly basis data was collected over the same time period. To control for other non-market factors, weekly data was also collected on the PNW export basis, transportation costs from Central North Dakota to PNW (tariff, fuel surcharge, and secondary railcar market value), and the nearby CBOT soybean futures spread.

Analysis of Spearman rank-order correlations between protein content and each of the 5 EAA's indicated no significant relationship; therefore, protein was considered as a separate quality factor in the panel regression models. To preserve degrees of freedom and due to the presence of strong multicollinearity, the five EAA measurement variables were condensed down into two factor variables using principle components analysis (PCA). The two factor variables explained over 97% of the total variability in the five EAA variables. The first factor variable (80.3% of variability) has a strong correlation with all five EAA variables with the second factor variable (17.2% of variability) had positive correlation with cysteine and lysine, and negative correlation with methionine and threonine, indicating a tradeoff in production between the four factors (tryptophan had a near zero correlation with the second factor).

A total of five panel regression models were estimated – one for the marketing year (MY) average of the weekly origin basis and control factor variables, and four quarterly (Q1, Q2, Q3, and Q4) average models to measure seasonal effects related to the timing of the crop harvest and release of the quality report information. To identify the correct panel regression model to estimate, a series of three tests were performed. First, groupwise (between CRD's) heteroscedasticity was tested using a modified Wald test. The null hypothesis was rejected at the 99% level for all five models supporting the presence of groupwise heteroscedasticity. Therefore, subsequent modeling used robust estimators.

The Breusch and Pagan (1980) lagrange multiplier test was then applied to test the null hypothesis that the pooled OLS estimator was the correct model (over fixed and random effects). This was rejected at the 99% confidence level for all five models. Due to the presence of panel invariant time series variables (control variables) in the independent data set, the Hausman (1978) specification test cannot be correctly applied due to singularity issues in the asymptotic variance matrix of the difference between fixed and random effects estimators. Therefore, an alternative augmented regression test first proposed by Mundlak (1978) was applied to test between fixed and random effects estimators. The test results strongly supported (95% confidence level or higher) the use of the fixed effects over the random effects estimator.

The panel regression results indicated that the two EAA factor variables were highly significant (95% or better) in the MY, Q2, and Q3 models with a positive value for the first EAA factor and a negative value for the second factor. The Q2 and Q3 models are notable in that they correspond with the annual time period corresponding (Q2) and just following (Q3) the release of the annual soybean quality report. Protein content had the correct anticipated sign (positive) in all five models; however, the coefficient was only statistically significant (90% confidence level) in just the Q1 and Q2 models. The PNW basis (representing influence of export market) was highly significant (99% confidence level) in all five models and had the correct positive coefficient sign. Additionally, the secondary railcar market value was also highly significant (99% confidence level) in all five models and also had the correct negative coefficient sign. The fraction of the variance explained by panel (CRD location) effects was 60% or higher in the MY, Q2, and Q3 regression mode.

The PCA eigenvectors were used to convert the coefficient and standard error values from the two EAA factors into equivalent values for the five EAA values. The results indicated that three (methionine, threonine, and tryptophan) of the five EAA variables had a significant (99% confidence level) positive impact upon the average origin basis for the MY, Q2, and Q3 models while cysteine had a significant (95% confidence level or better) negative value in all three models. Lysine was insignificant in the three models and had mixed results in the Q1 (positive at 95% level) and Q4 (negative at 95% confidence) models. Cysteine also had mixed results in the Q1 (positive at 99%) and Q4 (negative at 90%) models while methionine and threonine had insignificant results in Q4 and negative (90% significance) results in Q1. Tryptophan was insignificant in the Q1 and Q4 models.

The results of this study strongly support the hypothesis that the release of protein and EAA quality data has had a significant impact upon North Dakota soybean origin basis values. In particular, increased values of three of the EAA's (methionine, threonine, and tryptophan) correspond to increases in the hedonic value of local soybean prices as measured by stronger origin basis bids. Cysteine appears to have a negative hedonic value and this is likely due to two factors: (1) the substitution effect between cysteine and methionine / threonine in production as evidenced by the second EAA factor variable, and (2) the fact that cysteine is already produced naturally by animals and humans with methionine converted into cysteine by animals and humans through the digestive process. This increase in the sulfur amino acids (cysteine and methionine) has been shown to eventually reach a level that can have negative implications for animal growth. Lysine also competes with methionine and threonine in soybean production (as evidenced by the second factor) through the aspartate biosynthesis pathway in cereal and legume production.

The fact that hedonic values are significant is important to the evolution of the marketing system. One would expect that overtime, these could become more important features of the marketing system and pricing. More transparent hedonic values would have a positive impact throughout the marketing system. Growers would be impacted in variety and agronomic choices; handlers would be impacted by segregation and blending, as well as in pricing; end-users would be better able to reflect their values in purchasing strategies; and breeders could incorporate values into their breeding targets. None of these would be pertinent so long as the hedonic values are non-transparent.

Of moderate surprise was the relatively lower significance of protein content in the regression results. However, as has been noted in recent published studies (Hertsgaard, Wilson, and Dahl, 2019; Wilson, Dahl, and Hertsgaard, 2020) soybean buyers have generally looked at blending soybeans from other higher protein producing regions with lower protein soybeans from the northern states (such as North Dakota) to achieve the desired quality characteristics. Also, there is evidence that protein conscious export buyers would generally prefer the U.S. Gulf over the PNW export market as evidenced by market discounts.

References

Addinsoft. 2020. *XLSTAT Statistical and Data Analysis Solution.* (version 2020.1.3). New York, USA: Addinsoft. Available online at https://www.xlstat.com.

Ahmadi-Esfahani, F.Z., and R.G. Stanmore. 1995. “Quality Premiums for Australian Wheat in the Growing Asian Markets.” *Australian Journal of Agricultural Economics* 38(3): 237–250. doi: 10.1111/j.1467-8489.1995.tb00588.x.

Bin, P., R. Huang, and X. Zhou. 2017. “Oxidation Resistance of the Sulfur Amino Acids: Methionine and Cysteine.” *BioMed Research International* 2017: 1–6. doi: 10.1155/2017/9584932.

Breene, W.M., S. Lin, L. Hardman, and J. Orf. 1988. “Protein and Oil Content of Soybeans from Different Geographic Locations.” *Journal of the American Oil Chemists’ Society* 65(12): 1927–1931. doi: 10.1007/BF02546009.

Breusch, T.S., and A.R. Pagan. 1980. “The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics.” *The Review of Economic Studies* 47(1): 239–253. doi: 10.2307/2297111.

Brorsen, B.W., W.R. Grant, and M.E. Rister. 1984. “A Hedonic Price Model for Rough Rice Bid/Acceptance Markets.” *American Journal of Agricultural Economics* 66(2): 156–163. doi: 10.2307/1241032.

Bullock, D.W., and W.W. Wilson. 2020. “Factors Influencing the Gulf and Pacific Northwest Soybean Export Basis: An Exploratory Statistical Analysis.” *Journal of Agricultural and Resource Economics* 45(2): 317–334. doi: 10.22004/AG.ECON.302457.

Espinosa, J.A., and B.K. Goodwin. 1991. “Hedonic Price Estimation for Kansas Wheat Characteristics.” *Western Journal of Agricultural Economics* 16(1): 72–85.

Hacham, Y., L. Song, G. Schuster, and R. Amir. 2007. “Lysine Enhances Methionine Content by Modulating the Expression of S-Adenosylmethionine Synthase: Lysine Affects the Expression Level of SAMS and Methionine Content.” *The Plant Journal* 51(5): 850–861. doi: 10.1111/j.1365-313X.2007.03184.x.

Hart, C., and L. Schulz. 2015. *China’s Importance in U.S. Ag Markets*. Ames, IA: Iowa State University, Center for Agricultural and Economic Development. Available online at https://www.card.iastate.edu/ag_policy_review/pdf/fall-2015.pdf. [Accessed Mar. 25, 2021].

Hausman, J.A. 1978. “Specification Tests in Econometrics.” *Econometrica* 46(6): 1251–1271. doi: 10.2307/1913827.

Hertsgaard, D.J., W.W. Wilson, and B. Dahl. 2019. “Costs and Risks of Testing and Blending for Essential Amino Acids in Soybeans.” *Agribusiness* 35(2): 265–280. doi: 10.1002/agr.21576.

Hurburgh, C.R., T.J. Brumm, J.M. Guinn, and R.A. Hartwig. 1990. “Protein and Oil Patterns in U.S. and World Soybean Markets.” *Journal of the American Oil Chemists’ Society* 67(12): 966–973. doi: 10.1007/BF02541859.

Jantzi, D., K. Hagemeister, and B. Krupich. 2020. *North Dakota Agricultural Statistics 2020*. Ag Statistics 89. Fargo, ND: USDA-National Agricultural Statistics Service, Northern Plains Region, North Dakota Field Office.

Jeradechachai, T., and M. Tulbek. 2010. *North Dakota Soybean Quality Survey: 2009*. Northern Crops Institute. Available online at https://static1.squarespace.com/static/5335991ce4b0a8020cab3cbb/t/5a90999f419202959300652d/1519425952585/2009+North+Dakota+Soybean+Quality+Survey.pdf. [Accessed Mar. 25, 2021].

Karau, A., and I. Grayson. 2014. “Amino Acids in Human and Animal Nutrition.” In H. Zorn and P. Czermak, eds. *Biotechnology of Food and Feed Additives*. Advances in Biochemical Engineering/Biotechnology. Berlin, Heidelberg: Springer, pp. 189–228. Available online at https://doi.org/10.1007/10_2014_26.

Ladd, G.W. 1978. *Research on Product Characteristics: Models, Applications, and Measures*. Research Bulletin 584. Ames, IA: Iowa Agriculture and Home Economics Experiment Station. Available online at http://lib.dr.iastate.edu/researchbulletin/vol36/iss584/1. [Accessed Mar. 24, 2021].

Ladd, G.W., and M.B. Martin. 1976. “Prices and Demands for Input Characteristics.” *American Journal of Agricultural Economics* 58(1): 21–30. doi: 10.2307/1238573.

Lakkakula, P., D.W. Bullock, and W.W. Wilson. 2021. “Asymmetric Information and Blockchains in Soybean Commodity Markets.” *Applied Economic Perspectives and Policy* Forthcoming. doi: 10.1002/aepp.13159.

Lancaster, K. 1966. “A New Approach to Consumer Theory.” *Journal of Political Economy* 74(2): 132–157.

Medic, J., C. Atkinson, and C.R. Hurburgh. 2014. “Current Knowledge in Soybean Composition.” *Journal of the American Oil Chemists’ Society* 91(3): 363–384. doi: 10.1007/s11746-013-2407-9.

Miller-Garvin, J., J. Orf, and S.L. Naeve. 2014. *United States Soybean Quality: Annual Report 2014*. U.S. Soybean Export Council and University of Minnesota. Available online at https://28vp741fflb42av02837961y-wpengine.netdna-ssl.com/wp-content/uploads/2015/10/2014-Final-US-Commodity-Soybean-Quality-Report.pdf#:~:text=QUALITY%20REPORT:%202014%201%20SUMMARY%20The%20American%20Soybean,decisions.%202014%20A%20REAGE,%20YIELDS,%20AND%20TOTAL%20PRODUTION. [Accessed Mar. 25, 2021].

Mourtzinis, S., B.S. Borg, S.L. Naeve, J. Osthus, and S.P. Conley. 2018. “Characterizing Soybean Meal Value Variation across the United States: A Swine Case Study.” *Agronomy Journal* 110(6): 2343–2349. doi: 10.2134/agronj2017.11.0624.

Mundlak, Y. 1978. “On the Pooling of Time Series and Cross Section Data.” *Econometrica* 46(1): 69–85. doi: 10.2307/1913646.

Naeve, S.L., and J. Miller-Garvin. 2019. *United States Soybean Quality: Annual Report 2019*. St. Paul, MN: University of Minnesota Department of Agronomy and Plant Genetics.

North Dakota State University. 2014. “NDSU Releases New Soybean Variety.” *NDSU News Release*, October 22. Available online at https://www.ndsu.edu/news/view/detail/14827/. [Accessed Mar. 25, 2021].

Northern Crops Institute. 2019. *2019 North Dakota Soybean Quality Survey*. North Dakota Soybean Council and Northern Crops Institute. Available online at https://ndsoybean.org/wp-content/uploads/2020/02/2019NDsoybean012920-1.pdf. [Accessed Nov. 4, 2020].

Pacheco, L.G., N.K. Sakomura, R.M. Suzuki, J.C.P. Dorigam, G.S. Viana, J. Van Milgen, and J.C. Denadai. 2018. “Methionine to Cystine Ratio in the Total Sulfur Amino Acid Requirements and Sulfur Amino Acid Metabolism Using Labelled Amino Acid Approach for Broilers.” *BMC Veterinary Research* 14(1): 364. doi: 10.1186/s12917-018-1677-8.

Plume, K. 2017. “Low-Protein Soy Crop Dents Meal Quality, May Lift Feed Costs.” *Reuters*, October 30. Available online at https://www.reuters.com/article/us-usa-soybeans-protein/low-protein-u-s-soy-crop-dents-meal-quality-may-lift-feed-costs-idUSKBN1CZ0D6. [Accessed Mar. 23, 2021].

Plume, K. 2018. “Protein Plight: Brazil Steals U.S. Soybean Share in China.” *Reuters,* January 28. Available online at https://www.reuters.com/article/us-usa-soybeans-protein-insight/protein-plight-brazil-steals-u-s-soybean-share-in-china-idUSKBN1FE0FM. [Accessed Mar. 23, 2021].

Ravindran, V., M. Abdollahi, and S. Bootwalla. 2014. “Nutrient Analysis, Apparent Metabolisable Energy and Ileal Amino Acid Digestibility of Full Fat Soybean for Broilers.” *Animal Feed Science and Technology* 197: 233–240. doi: 10.1016/j.anifeedsci.2014.08.014.

Soy Transportation Coalition. 2017. *Soybean, Corn, and Wheat Exports by Port Region (2017)*. Available online at https://www.soytransportation.org/Stats/Ports_ExportsByPortRegion.pdf. [Accessed Mar. 25, 2021].

Statacorp. 2021. *Stata: Statistics / Data Analysis, Special Edition* (version 16.1). College Station, TX: Statacorp, LLC. Available online at https://www.stata.com.

Sudarić, A. 2020. “Introductory Chapter: Soybean - Quality and Utilization.” In A. Sudarić, ed. *Soybean for Human Consumption and Animal Feed*. Rijeka, Croatia: IntechOpen, pp. Ch. 1. doi: 10.5772/intechopen.93942.

Thakur, M., and C.R. Hurburgh. 2007. "Quality of US Soybean Meal Compared to the Quality of Soybean Meal from Other Origins." *Journal of the American Oil Chemists' Society* 84(9): 835–843. doi: 10.1007/s11746-007-1107-8.

USDA-ERS. 2021. *Annual State Agricultural Exports Interactive Chart*. USDA - Economic Research Service. Available online at https://www.ers.usda.gov/data-products/state-agricultural-trade-data/annual-state-agricultural-exports/. [Accessed Mar. 23, 2021].

USDA-NASS. 2021. *Quick Stats Online Database*. USDA - National Agricultural Statistics Service. Available online at https://quickstats.nass.usda.gov/. [Accessed Mar. 23, 2021].

Wei, H., X. Zhao, M. Xia, C. Tan, J. Gao, J.K. Htoo, C. Xu, and J. Peng. 2019. "Different Dietary Methionine to Lysine Ratios in the Lactation Diet: Effects on the Performance of Sows and Their Offspring and Methionine Metabolism in Lactating Sows." *Journal of Animal Science and Biotechnology* 10(1): 76. doi: 10.1186/s40104-019-0373-2.

Wilson, C., D. Golden, and T. Hubbs. 2021. *Oil Crops Outlook: March 2021*. Situation and Outlook Report OCS-21c. Washington, DC: USDA-Economic Research Service. Available online at https://www.ers.usda.gov/webdocs/outlooks/100686/ocs-21c.pdf?v=3681.2. [Accessed Mar. 25, 2021].

Wilson, W.W. 1984. "Hedonic Prices in the Malting Barley Market." *Western Journal of Agricultural Economics* 9(1): 29–40.

Wilson, W.W. 1989. "Differentiation and Implicit Prices in Export Wheat Markets." *Western Journal of Agricultural Economics* 14(1): 67–77.

Wilson, W.W. 2016. "China: The World's Largest Soybean Market: How Vital Is It to North Dakota?" Invited Presentation to the North Dakota Soybean Council Expo, February 2, Fargo, ND.

Wilson, W.W., and B. Dahl. 2006. "Costs and Risks of Segregating GM Wheat in Canada." *Canadian Journal of Agricultural Economics* 54(3): 341–359. doi: 10.1111/j.1744-7976.2006.00054.x.

Wilson, W.W., and B. Dahl. 2009. "Grain Contracting Strategies to Induce Delivery and Performance in Volatile Markets." *Journal of Agricultural and Applied Economics* 41(2): 363–376. doi: 10.1017/S1074070800002844.

Wilson, W.W., B.L. Dahl, and E. Jabs. 2007. "Optimal Supplier Testing and Tolerance Strategies for Genetically Modified (GM) Wheat." *Agricultural Economics* 36(1): 39–48. doi: 10.1111/j.1574-0862.2007.00175.x.

Wilson, W.W., D. Bullock, and P. Lakkakula. 2020. *Dynamic Changes in Rail Shipping Mechanisms for Grain*. Agribusiness & Applied Economics Report AER No. 794. Fargo, ND: North Dakota State University.

Wilson, W.W., B. Dahl, and D. Hertsgaard. 2020. “Soybean Quality Differentials, Blending, Testing and Spatial Arbitrage.” *Journal of Commodity Markets* 18: 100095. doi: 10.1016/j.jcomm.2019.100095.

Wooldridge, J.M. 2010. *Econometric Analysis of Cross Section and Panel Data*, 2nd ed. Cambridge, MA: The MIT Press.

Appendix A: Figures

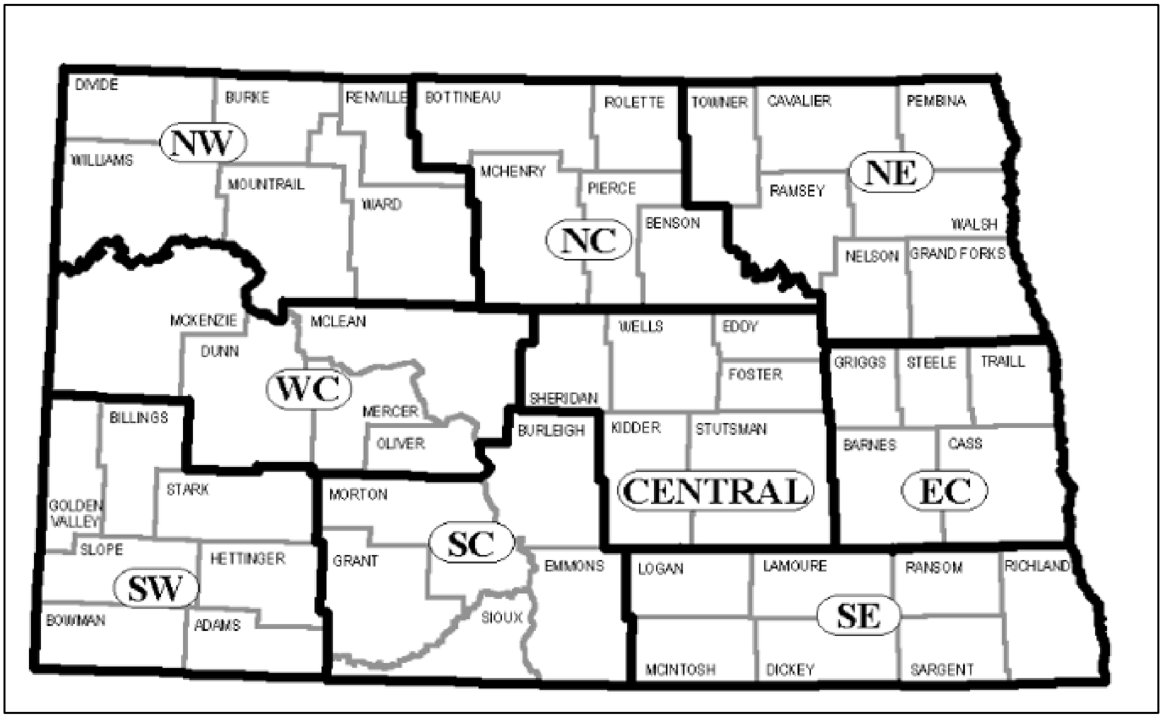

Figure 1. Map of North Dakota Crop Reporting Districts (CRDs).

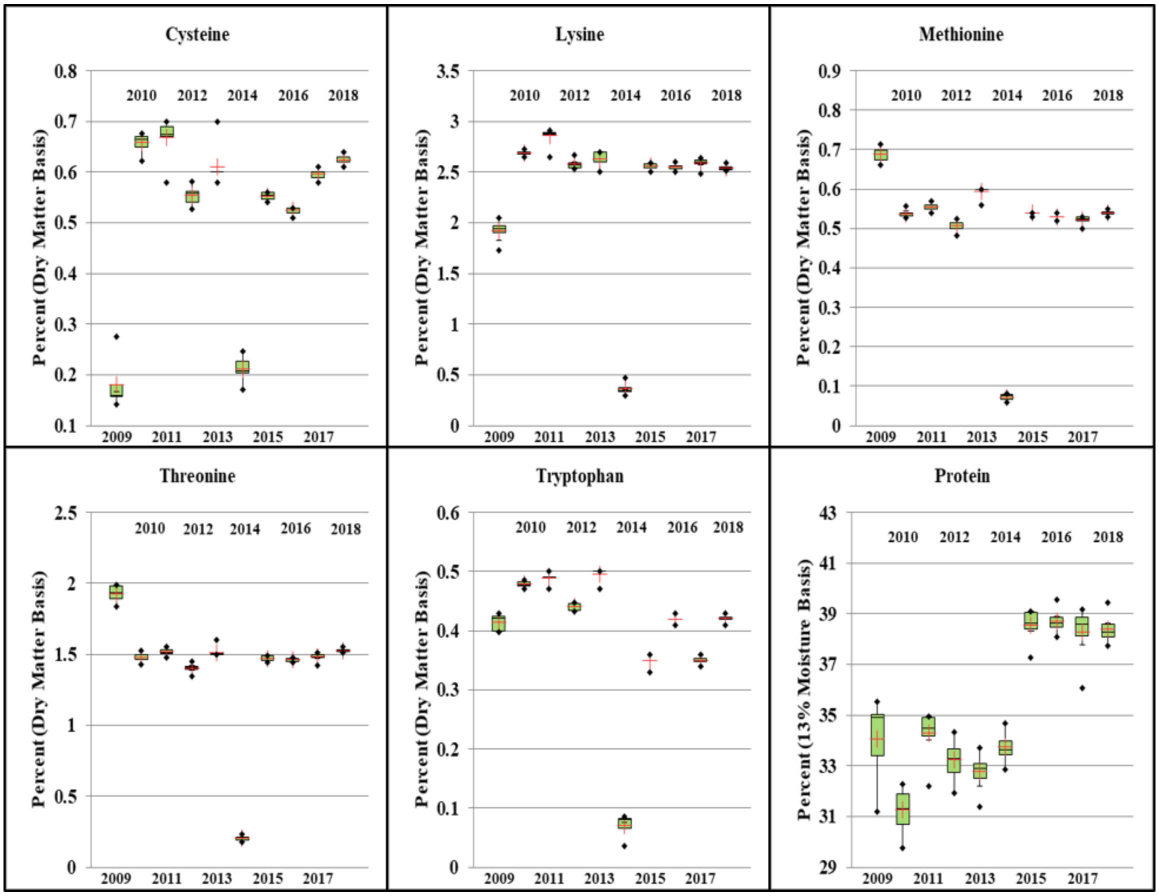

Figure 2. Time Series Boxplot of Soybean Quality Variables Across North Dakota CRDs.

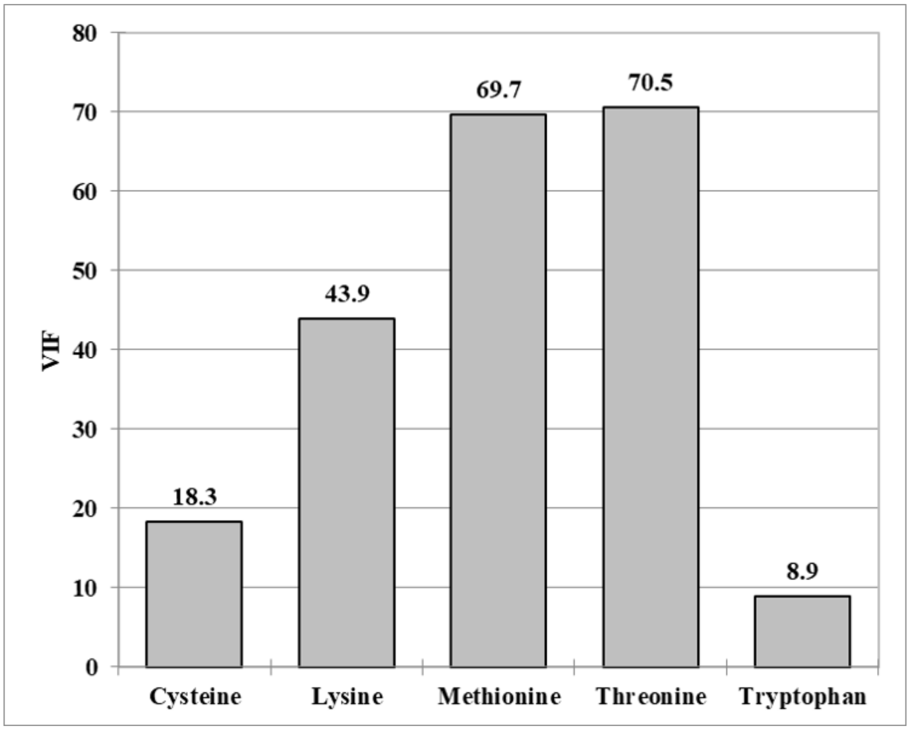

Figure 3. Variance Inflation Factors (VIF) for the Amino Acid Variables (80 observations).

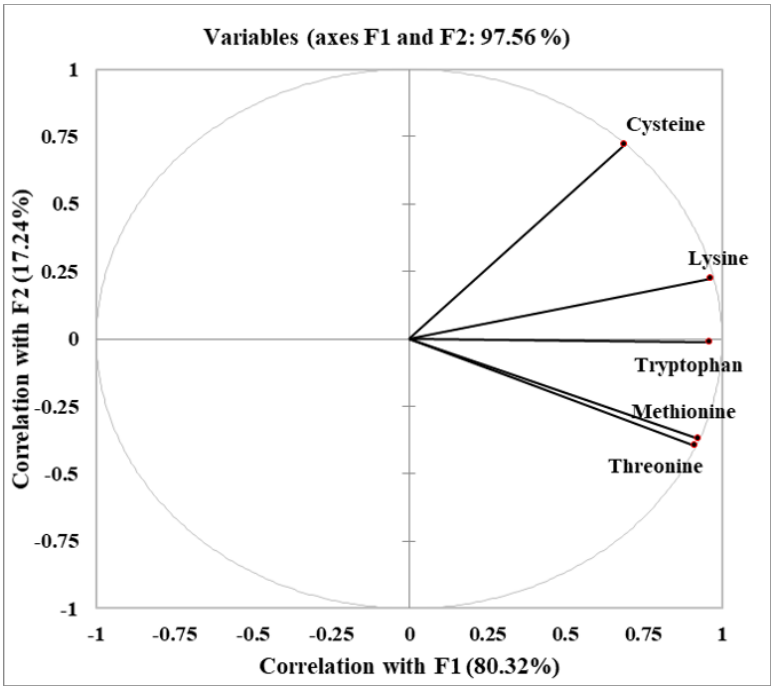

Figure 4. Pearson Correlation of Amino Acid Variables with Factor Variables

Appendix B: Tables

Table 1. List of Origin Elevators for Basis Data by Crop Reporting District[a]

Crop Reporting District	Origin Elevator for Basis Data
Northeast	Cenex Harvest States, Edmore, ND
North Central	BTR Farmers Cooperative, Leeds, ND
Northwest	Berthold Farmers Elevator, Berthold, ND
East Central	Arthur Companies, Page, ND
Central	Fessenden Cooperative, Fessenden, ND
West Central	Wilton Farmers Union, Wilton, ND
South Central	South-Central Grain, Hague, ND
Southeast	Farmers Elevator, Hankinson, ND

[a] Southwest region excluded due to lack of available quality data.

Table 2. Spearman Rank-Order Correlation of Amino Acid Contents with Protein Level by CRD

	North Dakota Crop Reporting District (CRD)							
Amino Acid	CENT	EC	NC	NW	NE	SC	SE	WC
Cysteine	-0.1763	-0.1763	-0.3091	-0.2892	-0.1879	0.1758	-0.1763	-0.1515
Lysine	-0.1879	-0.2121	-0.4012	-0.3780	-0.2485	0.2242	-0.0788	-0.1636
Methionine	-0.1641	-0.0675	0.1785	-0.4771	0.2067	-0.1829	-0.0305	-0.2012
Threonine	-0.0488	0.1398	0.1398	-0.0854	0.1758	-0.0608	0.0667	0.0732
Tryptophan	-0.5915	-0.6403	-0.5671	-0.6177	-0.5106	-0.2857	**-0.6565**	-0.4788

[a] Values in **bold font** are significant at the 95% confidence level.

Table 3. Principal Components Analysis (PCA) Summary for Amino Acid Variables

	Eigenvectors		Correlation		Contribution (%)		Squared Cosines	
Amino Acid	F1	F2	F1	F2	F1	F2	F1	F2
Cysteine	0.344	0.776	0.689	0.721	11.81	60.26	0.474	0.519
Lysine	0.482	0.242	0.965	0.224	23.20	5.84	0.932	0.050
Methionine	0.461	-0.398	0.924	-0.370	21.28	15.85	0.855	0.137
Threonine	0.455	-0.425	0.911	-0.394	20.69	18.03	0.831	0.155
Tryptophan	0.480	-0.013	0.962	-0.012	23.02	0.02	0.925	0.000

Table 4. Summary of the Panel Regression Identification Tests

Test	Model				
	MY	Q1	Q2	Q3	Q4
Groupwise Heteroskedasticity[a]	420.06***	174.77***	39.63***	38.85***	210.12***
Breusch and Pagan (1980)[b]	152.24***	97.10***	161.72***	120.32***	75.48***
Mundlak (1978)[c]	26.12***	32.12***	69.42***	74.21***	9.76**

[a] Modified Wald test for groupwise heteroskedasticity (H0: $\sigma_i^2 = \sigma^2$ for all $i = 1,\ldots,8$ groups). Distributed as $\chi^2(8)$.

[b] Breusch and Pagan lagrange multiplier test (H0: Pooled OLS is correct model). Distributed as chi-bar(1).

[c] Mundlak augmented regression test (H0: Random Effects is correct model). Distributed as $\chi^2(7)$.

Table 5. Fixed Effects Panel Regression Results with Origin Basis as Dependent Variable (t-statistics in parentheses)

Explanatory Variable	Model				
	MY	Q1	Q2	Q3	Q4
Constant	-2.4237***	-0.4399	-2.3416***	-4.3504***	-2.322**
	(-8.59)	(-0.85)	(-5.82)	(-6.26)	(-3.02)
F1	0.0250***	0.0145	0.0438***	0.0381***	-0.0115
	(5.22)	(1.02)	(9.49)	(4.57)	(-1.42)
F2	-0.0640**	0.0721**	-0.0858***	-0.0908***	-0.0927*
	(-3.08)	(3.43)	(-4.49)	(-4.10)	(-2.06)
Protein	0.0118	0.0184*	0.0090*	0.0147	0.0165
	(1.46)	(2.20)	(1.91)	(1.63)	(1.07)
Tariff_FSC	0.0002	-0.0004***	0.0001	0.0004*	-0.0001
	(0.30)	(-4.22)	(0.33)	(2.15)	(-0.11)
PNW_Basis	0.9250***	1.0180***	0.9633***	1.0033***	0.8008***
	(8.05)	(6.84)	(5.38)	(12.59)	(3.53)
Sec_Car_Value	-0.0002***	-0.0002***	-0.0001***	-0.0003***	-0.0003***
	(-6.42)	(-7.08)	(-5.42)	(-7.29)	(-3.58)
NB_Fut_Sprd	0.2785	-0.3842	0.6294	0.3757*	0.1331
	(1.29)	(-0.82)	(1.31)	(2.12)	(0.35)
R^2 (within groups)	0.626	0.705	0.861	0.808	0.167
R^2 (between groups)	0.459	0.001	0.632	0.660	0.213
R^2 (overall)	0.351	0.507	0.648	0.617	0.096
rho[a]	0.674	0.551	0.693	0.607	0.496

[a] Fraction of variance explained by group (location) effects.

Table 6. Fixed Effects Regression Results with Factor Variable Eigenvector Conversions
(t-statistics in parentheses)

Explanatory Variable	Model				
	MY	**Q1**	**Q2**	**Q3**	**Q4**
Cysteine	-0.0411** (-2.53)	0.0609*** (3.57)	-0.0516*** (-3.45)	-0.0574** (-3.29)	-0.0759* (-2.16)
Lysine	-0.0034 (-0.62)	0.0244** (2.86)	0.0004 (0.07)	-0.0036 (-0.54)	-0.0279** (-2.41)
Methionine	0.0370*** (4.32)	-0.0220* (-2.07)	0.0544*** (6.87)	0.0537*** (5.58)	0.0316 (1.72)
Threonine	0.0386*** (4.24)	-0.0240* (-2.18)	0.0564*** (6.72)	0.0559*** (5.51)	0.0341 (1.75)
Tryptophan	0.0129*** (5.55)	0.0060 (0.88)	0.0221*** (9.93)	0.0195*** (4.85)	-0.0043 (-1.09)