Introduction

Stability in the farming sector provides stability in rural economies, with a varying but large portion of employment in rural communities across the nation directly related to agriculture or to the agricultural services and processing industries. Instability in the agricultural sector can send ripple effects throughout the economy through increased food and fiber prices. Additionally, there has been a movement towards land investment by equity firms. As rent is the primary source of revenue, understanding movements in rent is useful for mitigating risk and understanding the market.

Farmland comprises the vast majority of the asset value on the agricultural sector’s balance sheet; totaling about 85%\(^1\), therefore when there is volatility, or rapid movement in one direction, there is cause for concern about the overall farm economy. National farmland prices have been increasing, overall. The Midwest or “Corn Belt” region has seen the largest growth in farmland prices, increasing 243% from 2000 to 2014, on average. In recent years, row crop prices have been declining. The national average corn price in 2012 was $6.89 and has since declined to $3.61.\(^2\) USDA reports have confirmed large corn and soybean crops for 2016. Large harvest volume will drive the price of soybeans and corn down even further leading to low returns in 2017 and likely causing operational losses in 2018. The depressed crop prices put strain on the operators’ ability to breakeven as rent values have risen following the rapid increase in land prices.

Forecasting cash rents can help producers and landowners prepare for changes in the market. Recent and expected volatility in the market for row crops makes forecasting prices

\(^1\) USDA Economic Research Service- Farm Income and Wealth Statistics
\(^2\) USDA National Agricultural Statistic Service Data
timely and necessary to producers and investors alike. Little research regarding forecasting of land values has occurred since the historic farmland bubble in the late 1970’s, and little academic research on forecasting cropland cash rents exists in the literature.

The purpose of the following research is to address the deficit in recent forecast literature pertaining to land and cash rent prices and to identify the best methodology for forecasting. Tested methodologies include a holt-winters naive forecast, a structural model with lagged rent, farmland prices and crop prices as explanatory variables, an error correction model (ECM), an autoregressive integrated moving average (ARIMA) forecast model, and a composite forecast. Each model is evaluated using mean absolute percent error (MAPE) and root mean squared error (RMSE).

**Literature Review**

Rent is a representation of the return to land from production or return for the use land can be put towards. It is also considered the underlying source of value and a major determinant of farmland prices by previous literature. Phipps (1984) proves the fundamental relationship of rent and farmland by showing rents Granger-cause land prices and that past rents and land prices explain the dynamics of future price changes. A substantial amount of literature is available on the determinants of farmland price and rents with the primary focus being on the determinants of farmland values and how rent affects the changes in the market price for farmland. Klinefelter (1973), Castle and Hoch (1982), Falk (1991), Lloyd (1994), and more provide possible determinants of farmland price with some discussion of rent. An overall understanding and ability to forecast rent is not largely considered but intuitively provides insight for both the income from the land and the land price determined in the market.
Rent is the payment of a non-owner operator to a landowner in exchange for the use of that owner’s land for production. There are two primary lease agreements for farmland rent—cash and share. Cash rent is a pre-specified cash payment paid to the landlord. The tenant assumes all risk and return associated with the production of the land. Cash rent is typically agreed upon prior to, and remains constant over the growing season. A share rent agreement, is an arrangement between the tenant and landowner where both parties share a portion of the risk and return of production. Contracts can vary but the landowner to some extent pays for a portion of expenses to production and receives a portion of the revenue from production. Share rent agreements inherently consider a premium needed to assume a portion of risk.

Rent agreements can also vary in terms of specified periods. Cash rent agreements and land prices in general, vary across states and regions. As shown by Robinson, Lins and VenKatarman in “Cash Rent and Land Value in U.S. Agriculture,” studies of rent and land must take place on a state or region-wide basis as there are significant differences across local land markets and even due to local non-agricultural demand for land.

Klinefelter (1973) states that as the returns to farmland increase from technological advances in production equipment, better management practices, and greater capital inputs any excess income above current costs will be capitalized into the value of the non-reproducible land resource following the Ricardian theory of rents. As the value of farmland increases from the capitalization process, the cash rental rate charged by the landowner will increase to compensate them for the annual use of the land. Excluding the 1980s, cash rental rates have historically aligned with the returns to farming operations and in the long run and one would expect cash rental rates and returns to farming operations to continue to align, in the long run. As the goal of studying the rental rate is to understand the relationship between income to the land and long
term farmland values, using cash rent is the most appropriate. Additionally, rent values will be addressed on a state-to-state basis to account for regional differences.

Forecasting rent more efficiently has become even more important in recent years. A primary concern for the relationship between cash rent and farmland prices is how they move together and how much of farmland prices can be explained by rent prices. Falk (1991) considers rent the primary determinant of farmland prices noting that in the short term, land prices will likely be more volatile than cash rents. Alston (1986) considers net rental income as the driver for most of the real growth in U.S. land prices during the twenty years prior to 1982. Others such as Klinfelter (1973) consider a number of explanatory variables in addition to rent such as average farm size, capital gains and the number of voluntary transfers of farmland. Melichar (1979) places emphasis on forward looking trends in rent focusing on expectations of changes in rent values. Ibendahl and Griffin (2013), show there is an asymmetric relationship between changes in rent and land values created by the lessee choosing to share varying amounts of information during productive years and more difficult years. The asymmetric relationship between lessee and landowner creates the need for the use of lags in determining land prices. Just and Mirankowski (1993) found that they could successfully predict land prices using current and lagged rent of up to 5-8 years. They also point out that while lagging rent by those amounts proves effective, it lacks substance in terms of “defensible economic rationale.” Featherstone and Baker (1998), and Falk (1991) agree that these techniques, while proving strong in terms of an r-squared, do not prove applicable.

Studied even more often than whether rent is a determinant of land prices, is testing the relationship between the capitalized value of land and the price seen in the market. In theory, rent, being the income to land, capitalized by a market interest rate should yield the value of the
farmland and therefore should equal the price determined in the market. Several studies have focused on the ability of the present value model to hold given the market prices for rent and farmland and the interest rate over time. Lloyd (1994) states two reasons the present value model is attractive. Firstly, the present value model is the quintessential example of the economic theory that in the long run the income and value of a good should not diverge. Secondly, the return from a good should be reflected by its price, in the long run.

Overall, rent is seen as an essential part of forecasting changes in farmland prices but, rent, specifically, is not largely the focus of the forecast or estimation literature. In some cases, articles reference the estimation of rent. Burt (1986) includes input prices, farm commodities and technological changes as primary explanatory variables. He also models rent as a second order rational multiplicative distributed lag to determine changes in land price movements.

During the period preceding the 1980s considerable attention was paid to estimating and forecasting land values using multiple modeling methodologies. During the period preceding the 1980s land values were increasing at a rate which outpaced multiple series thought to be an indicator or driver of farmland value. The rapid increase in value led to the research interest and proliferation of literature pertaining to the topic of forecasting land value. Following the “collapse” of farmland value in the mid to late 1980s the majority of the literature pertaining to farmland value transitioned into “post-mortem” ideology attempting to explain the rapid decrease in farmland value. Since the late 1980s and early 1990s little forecasting research has been added to the literature.

The bulk of the agricultural balance sheet’s assets are in land. Recent volatility in commodity price and returns to production has rapidly increased land value across multiple regions, impacting the entire agricultural sector. The increase in farmland value coupled with the
recent drop in crop prices warrants a return to research addressing forecasting land values and its determinants.

Data

Historical statewide rent values are the forecast object. Data from Iowa is used due to its contiguous series of crops grown throughout the state. Additionally, the variation in crop choices—largely corn and soybeans—are correlated in terms of pricing (the correlation coefficient of the two price series is greater than .96) therefore, studying one crop provides intuition about the other as well. Iowa State University Extension & Outreach provides historical rents which are reported and consistent with the USDA NASS reports of rent values for the state.

Rent contracts are made on a yearly basis. The data series used extends from 1975 to 2015. A rolling forecast is performed with the initial model estimation period from 1975 to 2009 allowing for six years of out-of-sample forecasts. The model is updated five times with an additional year of data added to the estimation until only the 2015 value is forecasted using data from 1975 to 2014. Six forecasts for each model are performed, resulting in 30 sets of forecasts.

For the structural and ECM model, two variables explain the variation in the rent values—lagged farmland prices and the settlement price on the December corn futures contract. All values are deflated using the same method implemented by the USDA NASS service. The deflator is the GDP chain-type price index published by the St. Louis Federal Reserve Economic Data service. All data is deflated to comparable 2009 levels.

The settlement price for the December futures price of corn is an indicator of the expectation of profitability of the upcoming production year. Since, rent values are determined
at the onset of the marketing year December futures price of corn is a good representation of the sentiment of the producers when they are negotiating rent contracts.

**Methodology**

There are essentially three overarching structures of forecasting—intuitive modeling, econometric or structural modeling and time-series modeling. Models from each of these categories are considered. From the intuitive class of models, a solely univariate modeling technique, the Holt-Winters model, is used. In the time-series category, there are both univariate and multivariate considerations. For these specifications, the ARIMA process and ECM processes are developed for each, respectively. Structural models take into consideration both multiple explanatory variables as well as multiple equations in estimation of the variables. Finally, a composite of all models considered is developed. As noted before, a rolling forecast beginning in 2010 is developed for all models. A total of six models and forecasts for each method are developed.³

**Holt-Winters**

The Holt-Winters model is a type of intuitive model, which are often used as baselines to compare more “complex” models against, and is an industry standard for determining the value added from using other forecasting methods. Holt-Winters is a linear trend model with an adjusting intercept and slope that employs a univariate structure with more weight being placed on the more recent period. The univariate structure only requires the use of historic rents. Upon graphing historic rent there was no evidence of seasonality or of trend, and so no augmentation for those factors was required.

$$ \hat{y}_{T+h,T} = \bar{y}_T + h F_T $$

³ In order to conserve space, the individual equations for each period are included in Appendix A.
Structural

Rent is a proxy for return to land therefore, measures of revenue and costs would be intuitive explanatory variables for rent itself. Three explanatory variables were found to explain 96.7% of the variation in rent. Both forward and backward looking variables were included. To measure the past changes in income to cropland, a lagged farmland price was included. A moving average specification of the variable was also considered but did not perform as well. The forward looking consideration for rent prices was the average trading price of the December futures corn contract. Additionally, as expected agents will consider the previous rent levels when deciding next year’s value, therefore, a lagged rent values was also included as an independent variable. All variables were significant at the 5% level.

\[ (2) \text{Rent}_t = \beta_0 + \beta_1 \text{Rent}_{t-1} + \beta_2 F L_{t-1} + \beta_3 F_C + \beta_4 MA(3) \]

Autocorrelation is present without the inclusion of the lagged rent value and an additional moving average process of order 3 was included to ensure the residuals are white noise. Tests for multicollinearity, heteroscedasticity in residuals and autocorrelation all show no signs of an issue in the final model.

Autoregressive Iterative Moving Average

The ARIMA modeling is again a univariate process where only the forecast objet itself, rent, is considered in the model development. There are three separate processes considered in the estimation process to best fit and forecast the data. In ARIMA model estimation, the data series must satisfy an invertibility condition and a stationarity condition. The invertibility condition is satisfied when all roots of the equation are larger than one. The condition ensures that the forecast object can be expressed in terms of the residuals and a weighted some of lagged values of itself.
The second condition of the stationarity is satisfied when the mean and variance of the series are constant throughout time, which is not likely for most data series. The mean gradually changes over time showing either a positive or negative trend. Non-stationary may also occur due to seasonality, a repeated pattern over a yearly time frame. Removing the trend and seasonal patterns before proceeding to the auto regressive (AR) process or moving average (MA) process determinants is necessary for the success of the estimation process.

The integration of the ARIMA modeling defines the degree of differencing required for series to achieve stationarity. The Augmented-Dickey Fuller Test (ADF) and the Phillips-Perron Tests (PPT) are considered when addressing the stationarity of the rent series. The ADF test regresses a lagged value of the forecast object on the first difference of the series. If the coefficient of the lagged value is statistically significant, then the series is stationary.

Following the integration process four steps are performed: identify preliminary models, estimate and refine the preliminary model, check for white noise residuals and finally forecasting. The MA process is a trend elimination process where error terms of the forecast object regressed on a time trend are included as explanatory variables. The AR process is a series of lags of the forecast object included as explanatory variables.

The correlogram is examined to determine the order of both the AR and MA processes. When the appropriate MA process has been identified the autocorrelation function (ACF), which considers the correlation of the forecast object over time with itself. The p-values of the ACF will progressively fall to zero following the optimal lag of the identified MA process. Alternatively, the partial autocorrelation function (PACF) is used to identify the AR process. The PACF considers the correlation of the forecast object with lags of itself and nets out the intermittent lags. In the case of the AR process, the p-values of the PACF will
immediately drop to zero following the optimal lag of the AR process. The final process will have the following structure where d, r and q represent the orders of the integration, AR, and MA processes, respectively:

\[(3) \Delta^d \text{Rent}_t = \alpha_1 \text{Rent}_{t-1} + \cdots + \alpha_r \text{Rent}_{t-r} + \epsilon_t + \beta_1 \epsilon_{t-1} + \cdots + \beta_q \epsilon_{t-q} \]

When all included explanatory variables are identified and prove statistically significant, the residuals of the model must be white noise; otherwise, the model is not properly specified. Following the estimation of the model, forecasting of out-of-sample observations are performed.

*Vector Autoregression/Error Correction Model*

Vector autoregression (VAR) models are multivariate processes like the structural process, however all the variables included in the estimation are endogenous. A VAR model is a system of equations in which each endogenous variable is a function of its own historic series and the historic series of the other included variables. The purpose of employing a VAR methodology is taking into account the interdependence between several time series, which is the case with rent, farmland prices and the December futures contract on corn. When employing a VAR process there are several issues that must be accounted for prior to estimation, including investigating the stationarity of the series, accounting for co-integration, the determination of the appropriate lag length of the data, and determining whether or not there is Granger-causality for ordering purposes. Equation 4. is a theoretical two equation VAR:

\[(4) V_1 = \beta_0 + \beta_1 V_{1t-1} + \cdots + \beta_s V_{1(t-n)} + \alpha_0 V_2(t-1) + \cdots + \alpha_m V_2(t-q) + \epsilon_{1t} \]

\[V_2 = \beta_0' + \beta_1' V_{1t-1} + \cdots + \beta_s' V_{1(t-n')} + \alpha_1' V_2(t-1) + \cdots + \alpha_m' V_2(t-q') + \epsilon_{2t} \]

The process to determine stationarity is the same process outlined in the ARIMA methodology. The ADF test and/or the PPT can be used as suitable measures for determining
whether or not the rent series and associated variables are stationary. If the series is not stationary, which is common particularly with a time series that is extensive in length, using a first or second differencing of the data will provide a stationary series.

Following the determination, and if necessary construction of a stationary series, it should be determined whether or not there exists a co-integrating relationship between the variables employed in the VAR. The existence of co-integration implies that there is a long run equilibrium relationship between a pair of variables within the VAR. With three variables included in the VAR, there can be up to two co-integrating relationships. While there is not necessarily any economic meaning to a co-integrating relationship, it is possible to draw economic ramifications.

The Johansen test procedure, which is composed of a trace test and maximal eigenvalue test, is used to discern whether or not there exists a co-integrating relationship between any pairwise combinations of the variables included in a VAR. In the presence of a co-integrating relationship an ECM should be used in place of a VAR in order to discern the short run dynamics between the time series. The ECM weights the exogenous variable within each VAR equation by the one-period lag of the residuals from the co-integrated relationship between the two co-integrated variables. Equation 5. is a theoretical two equation ECM:

\[
E_1 = \rho_1 z_{t-1} + \beta_0 + \beta_1 V_{1t-1} + \cdots + \beta_s V_{1(t-n)} + \alpha_0 V_{2(t-1)} + \cdots + \alpha_m V_{2(t-q)} + \epsilon_{1t}
\]

\[
E_2 = \rho_2 z_{t-1} + \beta_0 + \beta_1 V_{1t-1} + \cdots + \beta_s V_{1(t-n)} + \alpha_0 V_{2(t-1)} + \cdots + \alpha_m V_{2(t-q)} + \epsilon_{2t}
\]

Determining the appropriate amount of lags is critical whenever constructing a VAR or ECM model. Estimating the model with a different amount of lags and considering the Schwarz Information Criterion (SIC) and the Akaike Information Criterion (AIC) will yield the appropriate amount of lags to include. If the two criteria yield inconsistent results, the SIC is
considered more appropriate for large data sets while the AIC is more appropriate for small data sets, or the Hannan-Quinn information criterion can be used as a tiebreaker.

Testing for Granger-causality can provide insight into the direction of an economic relationship between two variables despite the fact that in a VAR all variables are considered endogenous. The Granger-causality test is a hypothesis test which regresses one variable upon the lags of itself and the lags of the other series. If the coefficients of the second series lagged variables equal zero, the second series fails to Granger-cause the first. The Granger-causality test dictates the ordering of estimation of the equations.

**Composite**

Following the estimation of the previous models, a composite forecast will combine the output of all models by averaging their forecasts. In doing so, the model collects and combines varying input from models that take into consideration varying facets of the forecast object, rent.

**Benchmark**

In order to determine whether any of the selected forecasting methods add value to a standard forecasting method, a benchmark was established for comparison. Published rent forecasts for Iowa were not found therefore a three-year and five-year moving average were constructed using the same rolling forecast periods the other five models operated under.

**Forecast Evaluation Measures**

Following the forecasting of these models two measures of forecast accuracy are considered to determine the strength of the models’ ability to forecast out-of-sample observations of the forecast object in terms of both magnitude and direction. Minimum values and ranges of both the RMSE and MAPE are considered in addressing the strength of forecasting the magnitude of the forecast object. Following the determination of these models, a test of co-
integration and Granger-causality of rent and farmland values is performed as a formal test of the importance of the relationship between the movement in farmland prices and rent.

**Root Mean Squared Error**

The RMSE evaluates the average unit amount the forecast is off over the forecast period. Equation 6 provides the calculation used to determine the RMSE:

\[ RMSE = \left( \frac{1}{M} \sum_{t=1}^{M} (F_t - A_t)^2 \right)^{1/2} \]

M is the number of periods included in the forecast; Ft is the forecasted value and At is the actual value. With the rolling forecast periods employed, there are six RMSE values provided for each model and the benchmark.

**Mean Absolute Percent Error**

The MAPE is a measure of the average percentage error of the out-of-sample forecast from the actual observation. Equation 7 defines the equation for MAPE.

\[ MAPE = \frac{1}{M} \sum_{t=1}^{M} \left| \frac{F_t - A_t}{A_t} \right| * 100 \]

It states, on average, the forecast is incorrect by a certain percentage. One issue to note is that each observation is weighted equally. In addition to the average of these errors, a range of percent differences for each forecast is considered to get an understanding of how many of the observations are outside a particular percentage range. There are also six MAPE values provided for each model and the moving average forecasts.

**Results**

Overall, the structural model performs the best in terms of minimizing the MAPE and RMSE consistently over the rolling out-of-sample forecast periods. The structural model minimizes the MAPE and RMSE in the 2012-2015 and 2013-2015 forecast periods. In the other four periods, the structural model was the second best forecast. The structural model forecasts were also
consistent across forecast periods. The MAPE range was only 7.0% with a minimum MAPE of 3.8% and a maximum value of 10.9%.

In the earlier estimation periods, 1975-2009 and 1975-2010, the Holt Winters estimation minimizes the MAPE and RMSE. In the later periods, 1975-2014 and 1975-2015, the ARIMA performs the best. Both the Holt Winters and ARIMA perform poorly outside these estimation ranges. Forecast MAPEs ranges for the Holt Winters and ARIMA are 48.8% and 35.5%, respectively.

The ECM and composite model estimation both perform the worst. The ECM’s average MAPE across the rolling forecasts was 31.88% with a range of 53.48% (minimum MAPE, 8.418% and maximum MAPE, 61.9%). Overall, the ECM was never competitive nor was it consistent in forecasting rent values. The composite also performed poorly in terms of the MAPE and RMSE.

Comparing to the 3-year moving average, the minimized forecasts all beat the benchmark and the structural model only comes second to the benchmark in two cases, the first forecast period of 2010 – 2015 and the last period 2015. The MAPE of the structural model is only 0.2 percentage points and 1.3 percentage points greater than the benchmark in each case. The ECM never beats the benchmark and the other forecasts beat the benchmark in at least two periods.

In Table 2 a distribution of the calculated average percent errors is calculated so as to provide intuition on consistency and range of the forecast models. Overall, the structural model remains in the 5-10% area. The Holt Winters range tends to be large in the larger sample periods whereas the ARIMA’s range is larger in the earlier samples.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Holt Winters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAPE</td>
<td>9.803</td>
<td>1.491</td>
<td>17.050</td>
<td>50.268</td>
<td>39.966</td>
<td>34.071</td>
</tr>
<tr>
<td>RMSE</td>
<td>24.014</td>
<td>3.753</td>
<td>40.777</td>
<td>119.920</td>
<td>92.596</td>
<td>76.289</td>
</tr>
<tr>
<td><strong>Structural</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAPE</td>
<td>10.846</td>
<td>8.620</td>
<td>4.744</td>
<td>3.813</td>
<td>4.094</td>
<td>5.627</td>
</tr>
<tr>
<td>RMSE</td>
<td>27.228</td>
<td>23.874</td>
<td>14.926</td>
<td>10.443</td>
<td>11.377</td>
<td>12.600</td>
</tr>
<tr>
<td><strong>ARIMA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAPE</td>
<td>34.429</td>
<td>34.030</td>
<td>19.564</td>
<td>1.494</td>
<td>0.906</td>
<td>1.546</td>
</tr>
<tr>
<td>RMSE</td>
<td>65.308</td>
<td>65.726</td>
<td>38.091</td>
<td>2.724</td>
<td>1.481</td>
<td>1.996</td>
</tr>
<tr>
<td><strong>ECM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAPE</td>
<td>61.897</td>
<td>44.835</td>
<td>14.094</td>
<td>8.418</td>
<td>27.493</td>
<td>12.908</td>
</tr>
<tr>
<td>RMSE</td>
<td>121.276</td>
<td>87.382</td>
<td>29.312</td>
<td>16.214</td>
<td>44.203</td>
<td>16.673</td>
</tr>
<tr>
<td><strong>Composite</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAPE</td>
<td>5.975</td>
<td>4.579</td>
<td>11.016</td>
<td>31.016</td>
<td>27.542</td>
<td>23.160</td>
</tr>
<tr>
<td>RMSE</td>
<td>12.598</td>
<td>10.603</td>
<td>26.647</td>
<td>72.747</td>
<td>63.686</td>
<td>51.857</td>
</tr>
<tr>
<td><strong>Rolling Forecast 3 Year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(Benchmark)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSE</td>
<td>27.839</td>
<td>30.332</td>
<td>32.917</td>
<td>28.369</td>
<td>17.039</td>
<td>9.760</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>------</td>
</tr>
<tr>
<td><strong>Holt Winters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1%</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1%-5%</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5%-10%</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10%-20%</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;20%</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>Structural</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1%</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1%-5%</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5%-10%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10%-20%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;20%</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>ARIMA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1%-5%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5%-10%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10%-20%</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;20%</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>ECM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1%-5%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5%-10%</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10%-20%</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>&gt;20%</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Composite</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1%-5%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>5%-10%</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10%-20%</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;20%</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Rolling Forecast 3 Year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1%-5%</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5%-10%</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10%-20%</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;20%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Conclusion

While the forecasts evaluation show a range of optimal forecasting in various time periods, the structural model overall provided both a best fit and consistent ability to forecast well throughout the updates in sample. The structural model also beats the benchmark forecast in almost all periods with a marginal difference in the periods it is not optimal.

The univariate models, overall, only seem to do well in either the beginning or end of the sample periods indicating a possible structural break in the data series making it difficult for the univariate models to forecast either before or after that major change. Testing for the presence of a structural change, using a test such as the Chow Test, would allow us to understand if and when a structural change may have occurred. Additionally, adding a structural change variable to the structural model may also provide better forecasting.

Additional forecast evaluations such as the Henricksson-Merton test for directional changes and the Granger-Newbold test for statistical difference in outcomes may also provide more intuition on the efficiency of these models in predicting not only magnitude but also directional change.
References


Castle, Emery N. and Irving Hoch. 1982. “Farm Real Estate Price Components.” *American Journal of Agricultural Economics. 64.1: 8-18*


Duncan, Marvin. 1977. “Farm Real Estate Values—Some Important Determinants.” *Economic Review. Mar: 3-12*


Appendix A

1. Holt Winters
   a. $Rent_{1975-2009} = 175.000 + 5.904 T$
   b. $Rent_{1975-2010} = 173.868 + 1.132 T$
   c. $Rent_{1975-2011} = 189.711 + 9.420 T$
   d. $Rent_{1975-2012} = 223.342 + 33.630 T$
   e. $Rent_{1975-2013} = 238.463 + 18.364 T$
   f. $Rent_{1975-2014} = 239.204 + 10.953 T$

2. Structural Model
   a. $Rent_{1976-2009} = 23.742 + 0.484Rent_{t-1} + 0.007F_{t-1} + 11.016P_{F,C} - 0.692MA_{3} + \varepsilon$
   b. $Rent_{1976-2010} = 24.411 + 0.502Rent_{t-1} + 0.006F_{t-1} + 10.5349P_{F,C} - 0.634MA_{3} + \varepsilon$
   c. $Rent_{1976-2011} = 24.195 + 0.512Rent_{t-1} + 0.006F_{t-1} + 10.274P_{F,C} - 0.651MA_{3} + \varepsilon$
   d. $Rent_{1976-2012} = 19.907 + 0.562Rent_{t-1} + 0.007F_{t-1} + 9.290P_{F,C} + 0.484MA_{1} + \varepsilon$
   e. $Rent_{1976-2013} = 24.089 + 0.538Rent_{t-1} + 0.006F_{t-1} + 9.551P_{F,C} + 0.422MA_{1} + \varepsilon$
   f. $Rent_{1976-2014} = 22.828 + 0.551Rent_{t-1} + 0.006F_{t-1} + 9.328P_{F,C} - 0.406A_{1} + \varepsilon$

3. ARIMA
   a. $\Delta^{1}Rent_{t} = \alpha_{t}\epsilon_{t-1} + \epsilon_{t}$

4. ECM
   a. 1975-2009
      i. $\Delta F_{1975-2009} = C_{1} \cdot (F_{t-1} - 39.599 \cdot Rent_{t-1} + 1188.026 \cdot P_{F,C(t-1)} - 295.069) + C_{2} \cdot \Delta(F_{t-1}) + C_{3} \cdot \Delta(F_{t-2}) + C_{4} \cdot \Delta(F_{t-3}) + C_{5} \cdot \Delta(Rent_{t-1}) + C_{6} \cdot \Delta(Rent_{t-2}) + C_{8} \cdot \Delta(P_{F,C(t-1)}) + C_{11}$
      ii. $\Delta Rent_{1975-2009} = C_{12} \cdot (F_{t-1} - 39.599 \cdot Rent_{t-1} + 1188.026 \cdot P_{F,C(t-1)} - 295.069) + C_{13} \cdot \Delta(F_{t-1}) + C_{14} \cdot \Delta(F_{t-2}) + C_{15} \cdot \Delta(F_{t-3}) + C_{17} \cdot \Delta(Rent_{t-1}) + C_{21} \cdot \Delta(Rent_{t-2}) + C_{22}$
      iii. $\Delta P_{F,C(1975-2009)} = C_{23} \cdot (F_{t-1} - 39.599 \cdot Rent_{t-1} + 1188.026 \cdot P_{F,C(t-1)} - 295.069) + C_{24} \cdot \Delta(F_{t-1}) + C_{25} \cdot \Delta(F_{t-2}) + C_{26} \cdot \Delta(F_{t-3}) + C_{27} \cdot \Delta(Rent_{t-1}) + C_{29} \cdot \Delta(Rent_{t-2}) + C_{32} \cdot \Delta(P_{F,C(t-3)}) + C_{33}$
   b. 1975-2010
      i. $\Delta F_{1975-2010} = C_{1} \cdot (F_{t-1} + 910.231 \cdot P_{F,C(t-1)} - 5489.282) + C_{4} \cdot \Delta(F_{t-2}) + C_{12}$
      ii. $\Delta Rent_{1975-2010} = C_{14} \cdot (Rent_{t-1} - 13.344 \cdot P_{F,C(t-1)} - 104.461) + C_{15} \cdot \Delta(F_{t-1}) + C_{16} \cdot \Delta(F_{t-2}) + C_{17} \cdot \Delta(F_{t-3}) + C_{21} \cdot \Delta(P_{F,C(t-1)}) + C_{22} \cdot \Delta(P_{F,C(t-2)}) + C_{4}$
      iii. $\Delta P_{F,C(1975-2010)} = C_{25} \cdot (F_{t-1} - 39.599 \cdot Rent_{t-1} + 1188.026 \cdot P_{F,C(t-1)} - 295.069) + C_{26} \cdot \Delta(Rent_{t-1}) - 13.344 \cdot P_{F,C(t-1)} - 104.461) + C_{28} \cdot \Delta(Rent_{t-2}) + C_{31} \cdot \Delta(Rent_{t-3}) + C_{33} \cdot \Delta(P_{F,C(t-1)}) + C_{36}$
   c. 1975-2011
i. \[ \Delta F_{1975-2011} = C_4 \cdot \left(F_t - 191.685 \cdot Rent_{t-1} + 3022.083 \cdot P_{F,C(t-1)} + 16309.971 \right) + C_4 \cdot \Delta(F_{t-3}) + C_{11} \]

ii. \[ \Delta Rent_{1975-2011} = C_{12} \cdot \left(F_t - 191685 \cdot Rent_{t-1} + 3022.083 \cdot P_{F,C(t-1)} + 16309.971 \right) + C_{13} \cdot \Delta(F_{t-1}) + C_{14} \cdot \Delta(F_{t-2}) + C_{15} \cdot \Delta(F_{t-3}) + C_{18} \cdot Rent_{t-3} + C_{19} \cdot \Delta(P_{F,C(t-1)}) + C_{20} \cdot \Delta(P_{F,C(t-2)}) + C_{22} \]

iii. \[ \Delta^d P_{F,C(1975-2011)} = C_{33} \]

d. 1975-2012

i. \[ \Delta F_{1975-2012} = C_1 \cdot \left(F_t - 108.156 \cdot Rent_{t-1} + 2035.882 \cdot P_{F,C(t-1)} + 6956.786 \right) + C_2 \cdot \Delta(F_{t-1}) + C_7 \cdot \Delta(Rent_{t-2}) + C_{11} \]

ii. \[ \Delta Rent_{1975-2012} = C_{12} \cdot \left(F_t - 108.156 \cdot Rent_{t-1} + 2035.882 \cdot P_{F,C(t-1)} + 6956.786 \right) + C_{13} \cdot \Delta(F_{t-1}) + C_{14} \cdot \Delta(F_{t-2}) + C_{19} \cdot \Delta(F_{t-3}) + C_{20} \cdot \Delta(P_{F,C(t-2)}) + C_{22} \]

iii. \[ \Delta^d P_{F,C(1975-2012)} = C_{33} \]

e. 1975-2013

i. \[ \Delta F_{1975-2013} = C_1 \cdot \left(F_t - 618.104 \cdot P_{F,C(t-1)} + 446.648 \right) + C_3 \cdot \Delta(F_{t-1}) + C_9 \]

ii. \[ \Delta Rent_{1975-2013} = C_{10} \cdot \left(F_t - 618.104 \cdot P_{F,C(t-1)} + 446.648 \right) + C_{11} \cdot \Delta(P_{F,C(t-1)}) - 55.069 + C_{12} \cdot \Delta(F_{t-1}) + C_{17} \cdot \Delta(F_{t-2}) + C_{17} \cdot \Delta^d(P_{F,C(t-2)}) + C_{18} \]

iii. \[ \Delta P_{F,C(1975-2013)} = C_{22} \cdot \left(Rent_{t-1} - 25.790 \cdot P_{F,C(t-1)} - 55.069 \right) + C_{25} \cdot P_{F,C(t-1)} + C_{27} \]

f. 1975-2014

i. \[ \Delta F_{1975-2014} = C_2 \cdot \Delta(F_{t-1}) + C_8 \]

ii. \[ \Delta Rent_{1975-2014} = C_9 \cdot \left(F_t - 108.128 \cdot Rent_{t-1} + 2175.242 \cdot P_{F,C(t-1)} + C_{13} \cdot \Delta(Rent_{t-2}) + C_{14} \cdot \Delta(P_{F,C(t-1)}) \right) + C_{15} \cdot \Delta(P_{F,C(t-2)}) + C_{16} \]

iii. \[ \Delta P_{F,C(1975-2014)} = C_{17} \cdot \left(F_t - 108.128 \cdot Rent_{t-1} + 2175.242 \cdot P_{F,C(t-1)} + 6344.005 + C_{20} \Delta Rent_{t-1} + C_{21} \Delta Rent_{t-2} + C_{24} \right) \]