Informing Long Term Lumber Buying: A Decision-Making Criteria for Wood Buyers Using a Simple Algorithm

Thanos Gentimis
Louisiana State University AgCenter

Shaun Tanger
Louisiana State University AgCenter

Maria Bampasidou
Louisiana State University AgCenter

Lumber purchasing decisions rely mostly on historic lumber prices, or the corresponding price of their futures. Various, lumber specific, supply and demand conditions lead to an unpredictable weekly price if one employs basic economic and statistics tools. We propose a prediction model that builds on the historical change of the price from week to week to provide purchasing recommendations yearly. We tested our model using data from the Random Lengths weekly price catalog for six lumber types (1995-2014), comparing it to other purchasing strategies created from the spot and futures prices showing that it is better almost always.

INTRODUCTION

Lumber in the United States is a heavily used commodity in manufacturing and housing construction. Thus, the corresponding lumber markets are well established entities in the world of trade, and the various lumber prices are the object of speculation from sellers and buyers alike. However, due to complex supply and demand conditions, these prices regularly exhibit extreme volatility in the short-run constituting forecasting a difficult task (Oliveira et al., 1977). These conditions can be linked to stumpage costs, railroad strikes or railcar shortages, residential construction, and prices of substitutes like steel and aluminum. In addition, the lumber market can be affected by macroeconomic conditions; the housing bubble of 2006 followed by the great recession in 2009 are recent examples.

Traditionally, forecasting models for purchasing recommendations in lumber depend on historic price series and futures prices. So far, studies have employed moving average analysis and regression analysis to forecast lumber price series and address the volatility these series exhibit, by incorporating other factors, like variables for demand, price movement and volume traded. The complexity of forecasting lumber price series intensifies when we consider different production and trading regions, different species and grades, lumber uses and respective dimensions needed. Considering an aggregated lumber price series for grades, regions, and species implying that the law of one price holds for the lumber market can lead to misspecifications of the forecasting model (Yin and Baek, 2005). This issue can also be encountered in studies using futures prices and renders some of these models impractical in extended periods of time.

In this paper, we propose a new prediction model that builds on the movement of prices instead of their actual values (spot or futures prices). The proposed model complements the existing literature, utilizing only the aforementioned time series. Specifically, we use data from the Random Lengths weekly price catalog for six lumber types for the years 1995 to 2014 and no other information. We do not aggregate our data but instead examine stationarity, correlations between these six series, and seasonality patterns separately for each price series. Keeping the data series disaggregated allows for a more robust empirical analysis and targeted purchasing recommendations tied to the respective lumber type.

As indicated, successful buying and selling of lumber is dependent on traders accurately predicting prices that they will face in the near future (Buongiorno et al, 1984). Many lumber market participants engage in price discovery through following and/or transacting in the lumber futures market (Deneckere et al., 1986; Hasan and Hoffman-MacDonald, 2012). Though thinly traded, the trading volume threshold, which is necessary to facilitate efficient price discovery, is very low, the common price discovery measures suggest that futures markets are dominant in the price discovery process (Admmer et al., 2016). Calculations of the referenced authors indicates that only 42 percent of the total production in covered by the futures market and find that only 1 of the 6 largest lumber trading companies took active positions in the futures market despite other risk hedging behavior. However, many lumber buyers do not engage this mechanism of price discovery, relying more heavily on recent prices and experience of trading networks developed over time. Linkedin correspondence with a lumber trader indicated not all traders use the futures market to determine buying and selling strategy. While some use it as a basis for contracting, those that trade in the open markets will look at futures to decide if they should buy or sell. Others simply buy based on inventory needs and demand that they face in the present. While recent findings in the lumber price literature suggest that engaging the futures markets is efficiency increasing over not using it, consensus does not exist on the subject (Parajuli and Zhang, 2016). Given lumber's thinly traded futures, several researchers have found that spot and futures price series are not co-integrated, implying very little to no role of futures in the price discovery process for lumber spot prices[1].

However, using novel statistical methodology the two most recent publications on the subject, find that indeed futures do aid traders in the price discovery process (Parajuli and Zhang, 2016; Mehrotra and Carter, 2017). Two questions arise repeatedly in the literature:

1. Are futures prices merely following the same process as the spot prices due to end use demand? In other words, are the same expectations and supply and demand conditions that influence spot prices and futures prices and thus are redundant in the information they provide (Mehrotra and Carter, 2017), therefore showing co-integration when its spurious? This question, while addressed remains unanswered in the futures literature.
2. Are lumber futures a strong predictor for specific species that are not included in the futures contract? For example, can futures be used to predict prices for southern yellow pine? "The species mix in the futures contracts is a significant component of North American lumber production, but it does not represent the majority of the softwood lumber produced. We estimate that 20-25 percent of the softwood lumber produced in 2009 was of species that could qualify for the contracts."(Lutz, 2012, p.3). Lumber futures are comprised of 2x4's (8' to 20'), graded at #1 and #2 of western SPF, Hem-fir, Engelmann Spruce, and Lodgepole pine. This question has been examined empirically, but the evidence is mixed. However, if there are strong correlations among lumber species, seemingly, the answer should be yes.

Yin and Baek (2005), undertake the most ambitious analysis on the subject of co-integration among lumber species in the North American market. They find evidence supporting the law of one price for the entire United States softwood lumber market (in other words, co-integrated), but unlike the futures and spot prices literature there is far more agreement on the matter (Uri and Boyd, 1990; Jung and Doroodian, 1994; Shahi et al., 2006; Shahi and Kant, 2009). They do mention that this relationship is not unanimous among price series relationships. However, depending on the test groupings were co-integrated approximately 90 percent of the time, no matter the test.

Unfortunately, the same statistical question plagues this literature on the issue of co-integration as that mentioned by Mehrotra in the futures prices literature. Namely, that species may be co-integrated, but not due to substitutability, instead owing to the fact that all these co-integrating relationships are caused by common-demand side factors (Shook et al., 2009). Softwood lumber is largely traded through wholesalers, who can take speculative positions on a number of species through storage of the commodity and hold long positions on a number of lumber species. Thus, prices for various species may show co-integration due to liquidity constraints faced by the wholesalers, who trade multiple species concurrently (Pindyck and Rotemberg, 1988; Shook et al., 2009). Regardless of a causal relationship, this would indicate that prices for various lumber species are highly correlated and thus the futures should adequately predict any lumber type prices including the southern yellow pine lumber we examine in this paper especially the #2 series.

On the other hand, there are several researchers who have examined the relationship of other specific spot prices and the futures market and have not found this to be true (e.g. Hasan and Hoffman-MacDonald, 2012; He and Holt, 2004; French, 1986; Fama and French, 2016). Mehrotra and Carter (2017) believe this may be either due to the species differences in the lumber futures and spot prices, since the specific lumber commodity traded on the futures market is not traded in any spot market, or due to a statistical averaging issue common in time series (using monthly or quarterly averages). They use the price of the expiring contract to serve as a spot price (thus removing the species mix issue). However, Parajuli and Zhang (2016) using the same specific species series as Manfredo and Sanders (2008) do find co-integration between the two series. This would indicate that the lack of integration is due to some other factor(s). In fact, it may be a specification issue with how the authors are constructing their futures price series, specifically the expiration date and when the price rolls over to the subsequent futures series. Parajuli and Zhang (2016) use a continuous weekly series similar to that developed by Quandl. Interestingly, the implications of these more recent findings, is that futures can be used as a hedging mechanism against market volatility and as advance information of lumber markets.

While many researchers have devised various models and methods to examine the relationships of lumber prices, those that employ futures or otherwise, do not reach the same conclusions. It may in fact be the case that the series are exhibiting long memory (Niquidet and Sun, 2012), which may explain why at times, two or more price series may co-integrate, but at other times they do not (a common issue in the futures literature). Niquidet and Sun (2012) examine several lumber and pulp products and find that after price shocks, despite the shock dissipating within 50 months, the effects can linger for 30 years. Among lumber markets, Sun and Ning (2014) find that the southern markets tend to achieve equilibrium more quickly than those of other North American lumber markets after price shocks. Assuming the co-integration is true, one could use the delayed futures price to predict the movement of the actual price. Unfortunately, our model neither confirmed, nor denied the existence of co-integration in a definite way, as we will explain in section results.

In terms of forecasting lumber prices, the papers by Oliveira et al., (1977), Buongiorno et al., (1984), and Deneckere et al., (1986) are the most relevant to our analysis. Using a series of models from ARIMA, Oliveira et al. (1977), found these simple time series models to be relatively accurate for short-run predictions of lumber prices (within 10 percent of actual prices up to 4 weeks for SYP). Buongiorno et al. (1984), compared a relatively complex econometric model, a futures model, and a lagged cash price model in terms of predictive power. In the shorter-run (one quarter), FORSIM and futures models compared favorably, but for longer term forecasts (2-3 quarters) the FORSIM model was superior. Both outperformed the lagged cash price model. Further, findings in Deneckere et al. (1986), indicate that futures are an effective hedge to blunt the variance of cash positions taken by the trader, making them effective for risk averse traders.

The only other literature we were able to find on the issue of a decision framework for buying and selling lumber was Kingslien (1975). They give an explicit methodology for producers to sell futures once agreeing to contract to produce lumber, they also state that wholesalers and end purchasers can use the tables with a simple formula to determine desired futures selling prices (i.e. Desired profit margin + Cost of production - Adjustment factor for item(s) = Futures price). Then using their tables to arrive at the adjustment factor, they offer producers guidelines for trading futures, given relationships that they establish between non-contract and contract grade lumber prices and how they are related to expiring futures prices.

In this paper, we take a slightly different approach to advance the literature by adding this interesting applied wrinkle: Given historical information in the respective weekly price series (futures or spot), what action should be taken by lumber buyers (buy or don't buy)? Specifically, we offer purchasing recommendations based off probabilities of price movements rather on actual prices. We compare this "long-history" model against other possible strategies to determine the best strategy to use over a given period (2000-2015). Our findings indicate that among model alternatives, our model outperforms all other strategies, including one that uses futures prices with up to six months of lags. We also examine these recommendations in the context of warehouse space that the buyer has available. This has implications both for the literature on futures and spot lumber prices and their relationships as well as real world implications for buyers who wish to employ this method as a buying strategy for purchasing lumber throughout the entire year.

Our paper is structured as follows. The data section, discusses the lumber price series used in the analysis and examines some of their properties, namely potential trends, correlations and seasonal patterns. The following section, introduces the mathematical underpinnings of our model's estimation algorithm, formulates the conceptual model and describes the computations. In the subsequent section, we compare our model to others for various parameters and report our results. In the last two sections, we discuss our findings in terms of main economic conditions and conclude this paper by presenting future venues for this analysis.

DATA

Dataset Description

Our dataset comprises of weekly lumber prices provided by the independent price recording company "Random Lengths" for two grades and six types of softwood lumber described in the table below (See Table 1).

TABLE 1
LUMBER PRICE SERIES

Price Series	Pubdate	Description	Price	Year	Month	Week	Issue
LAGD	06-Jan-95	KD Southern Pine (Eastside) #2 2x4 random Prices Net f.o.b. Mill	420.00	1995	1	1	1
LAGD	13-Jan-95	KD Southern Pine (Eastside) #2 2x4 random Prices Net f.o.b. Mill	425.00	1995	1	2	2
LAGD	20-Jan-95	KD Southern Pine (Eastside) #2 2x4 random Prices Net f.o.b. Mill	420.00	1995	1	3	3
LAGD	27-Jan-95	KD Southern Pine (Eastside) #2 2x4 random Prices Net f.o.b. Mill	408.00	1995	1	4	4

Lumber products come in different grades, which are often related to the quality of the sawlog, originally harvested and converted to the corresponding wood product. Typically, these are some form of structural lumber (i.e. 2x4, 2x6) or plywood depending on the size and quality of the original sawlog. Softwoods (or pine) such as these under examination in this manuscript, come in four grades based on either their strength and/or appearance. Knots and other defects result in a lower grade. Most two-inch thick softwood lumber (the 2 in 2x4) is graded for its strength rather than appearance.

The common grades found at your local lumberyard from best to worst are:
#1 Construction grade, #2 Standard grade, #3 Utility grade and #4 Economy grade. This study originated from a collaborative project with a wood pallet producer in the Southeastern United States. Typically, pallets used for shipping and hauling products utilize #4 grade quality lumber. However, many chain stores (e.g. Walmart, Lowes) use #2 grade for floor displays of merchandise. This particular pallet producer creates pallets for both of these uses, utilizing the lumber dimensions listed earlier in the text and thus our analysis focuses on these particular lumber products. While we cannot make any generalizable claims about other lumber species and how well our model performs relative to other predictive models, we see no reason why this analysis could not be done to estimate purchasing options for other lumber types or other popular wood products, such as plywood and engineered wood products.
The dataset covers the period of 1995-2014 and it was pre-processed to fit a standard 52-week calendar with appropriate interpolations. In the original dataset, all months had 5 weeks, but some of them had no entries. We found the weeks with the fewest entries throughout our dataset for each lumber type, and removed them until a 52 week calendar was created. This may result to a small error in computations, but since the comparisons are averaged out over many years, we do not anticipate that to be a major flaw. A typical entry (row) of our dataset is presented in the table below (See Table 2) and our dataset has 6,240 rows. We use the dataset to perform both the training and testing of the model in this paper, but we plan to expand our methods to other types in the future.

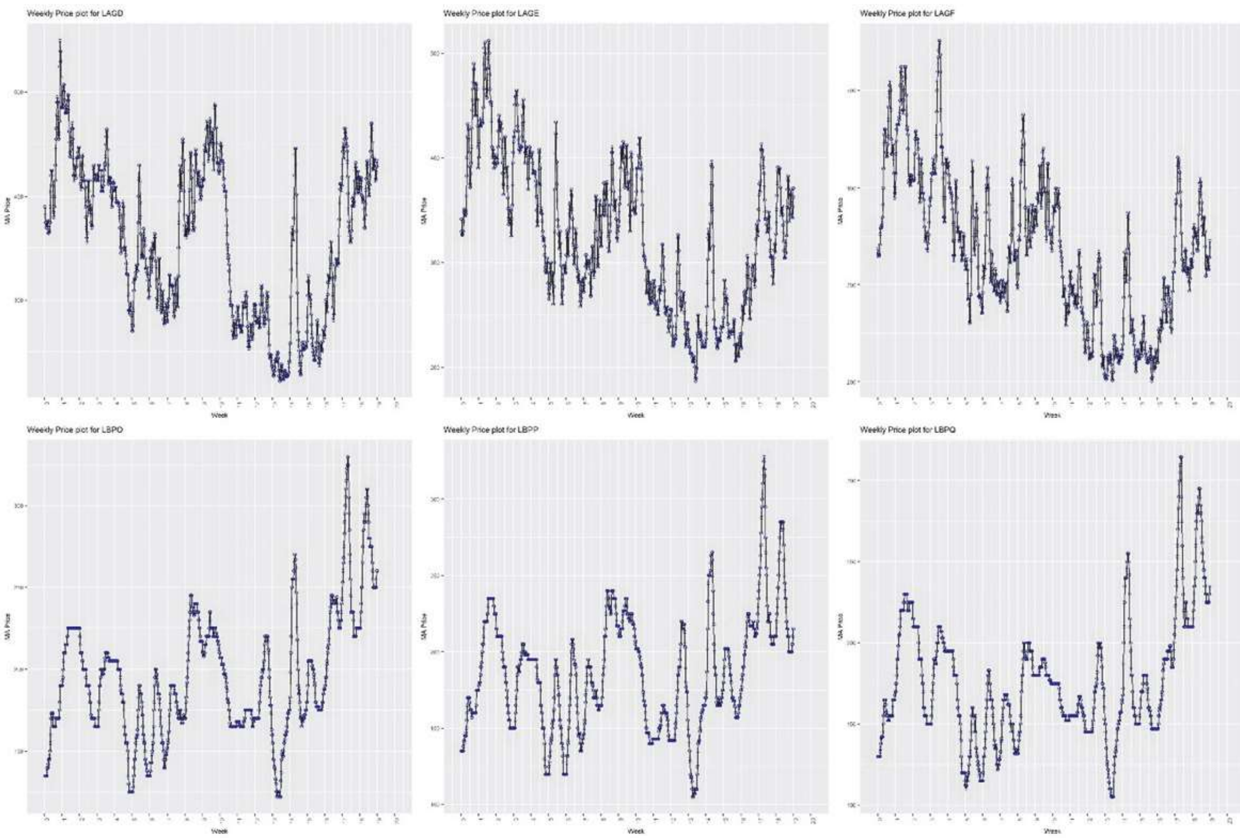
TABLE 2
TYPICAL ENTRY OF DATASET

Price Series	Description	Year	Week	Price
LAGE	KD Southern Pine (Eastside) #2 2x4 random Prices Net f.o.b. Mill	1995	1	395

Trends and Graphs

In this subsection, we present some simple descriptive statistics of our dataset and the associated graphs. We viewed each lumber type as a separate time-series thus creating the graphs in Figure 1.

FIGURE 1
PRICES FOR VARIOUS LUMBER TYPES FOR THE YEARS 1995-2014



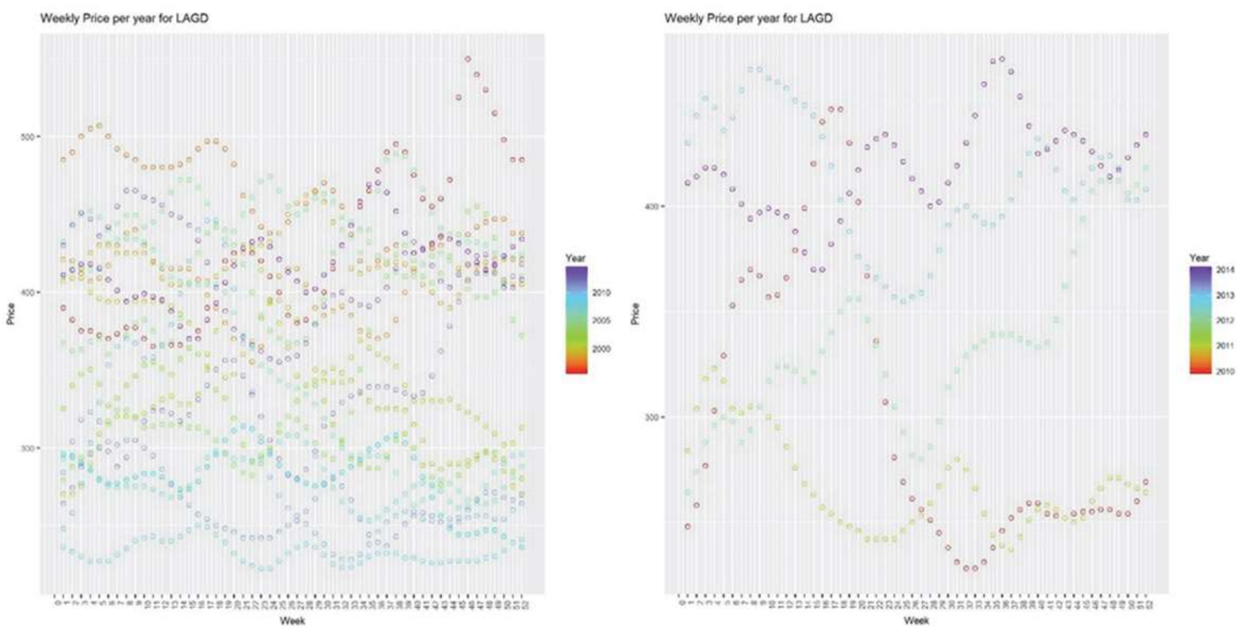
To check for stationarity and explore the attributes of our dataset we examined each price series separately. The first three panels in Figure 1 (top row) are for grade #2 2x4, 2x6 and 2x8, respectively, whereas the next three panels (bottom row) are for grade #4 2x4, 2x6 and 2x8, respectively. As we see in Figure 1, there are no clear periodic behaviors overall but a further analysis of each time series separately revealed some seasonality which we comment on below. Our analysis of the one week lagged price differences showed that the week by week differences are not (statistically) significantly different than zero throughout the years. Table 3 gives the relevant statistics.

TABLE 3
STATISTICAL ANALYSIS OF THE ONE WEEK DIFFERENCE IN PRICE PER LUMBER PRICE

Price Series	Mean	St. Dev.	SE	Tstat	Min	Max
LAGD	0.013	8.509	0.264	200.871	-35	53
LAGE	-0.023	9.228	0.286	139.792	-36	40
LAGF	-0.043	9.588	0.297	134.535	-40	40
LBPO	0.082	3.323	0.103	174.709	-20	18
LBPP	0.038	3.691	0.114	157.272	-20	18
LBPQ	0.059	3.485	0.108	138.808	-25	15

We then tried to identify common peaks and lows for the price of each individual lumber type over the span of a year. In Figure 2, left panel for example, we plot the weekly prices for type LAGD for the period 1995-2014. No clear trend can be discerned. This holds for the other five lumber types. Even when we focus on the last five years of our study period, trends are not easy to see on a yearly basis as shown in Figure 2 right panel.

FIGURE 2
LAGD WEEKLY PRICES FOR PERIODS (1995-2014) & (2012-2014)



We used the autocorrelation function both on the reported prices and on the one-week price differences but no immediate periodic behavior was apparent. As one can see in Figure 3 and Figure 4, the levels rarely exceed the $0.2 mark independent of how many lags we considered.

FIGURE 3
AUTOCORRELATION OF THE WEEKLY DIFFERENCE IN PRICES BY LUMBER TYPE AND VARIOUS LAGS, PERIOD 1995-2014

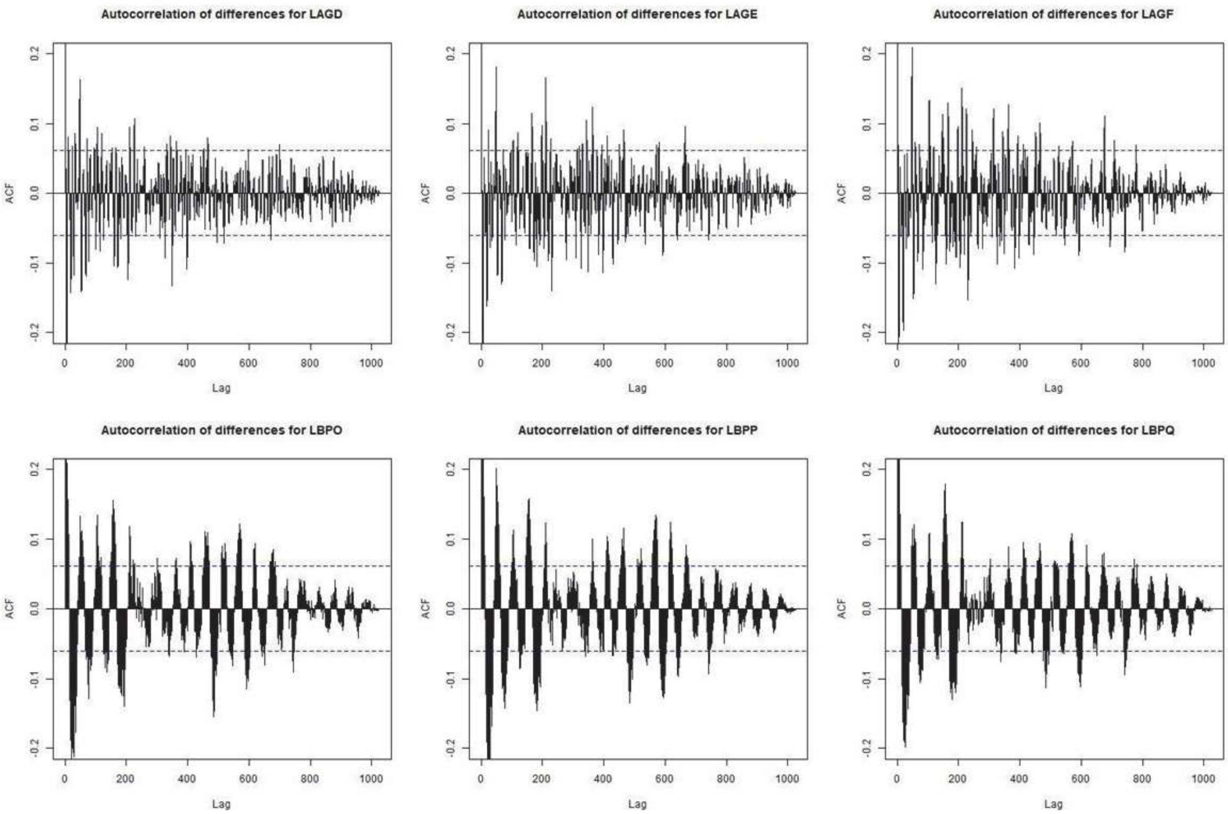
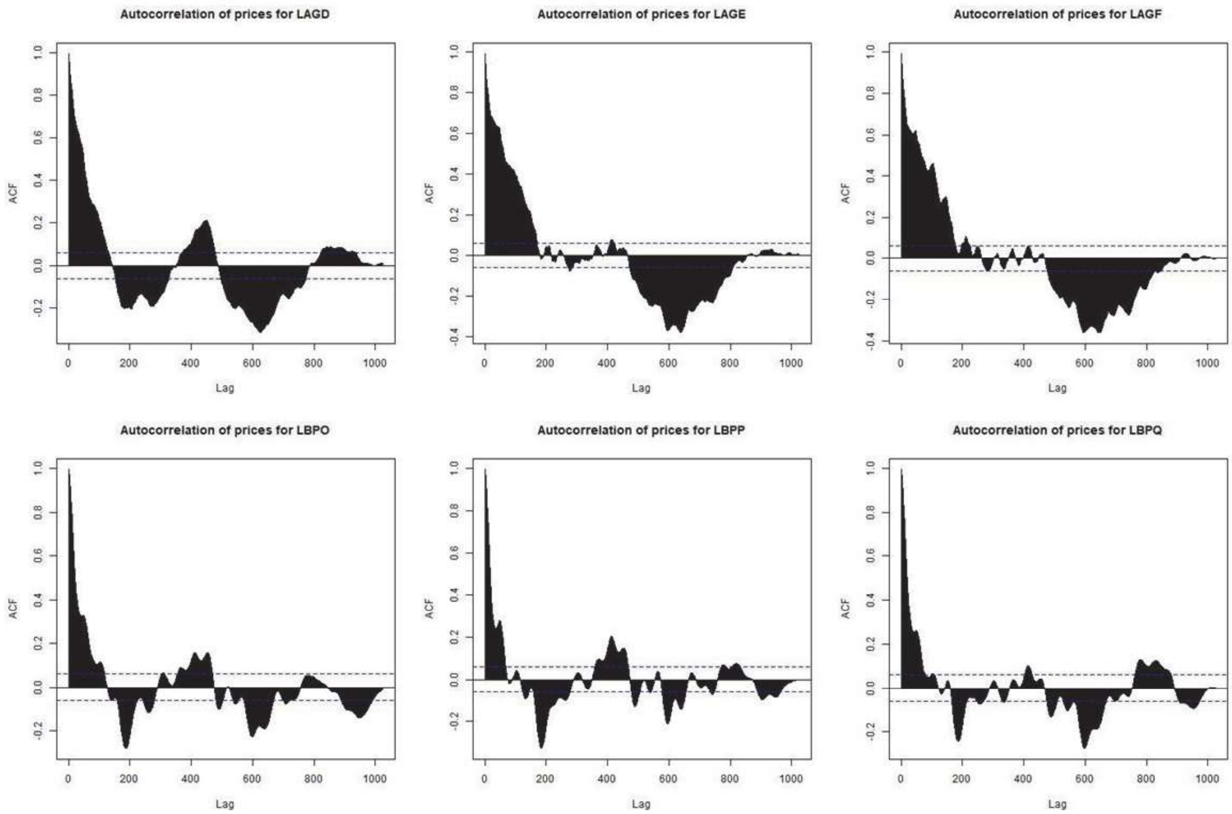
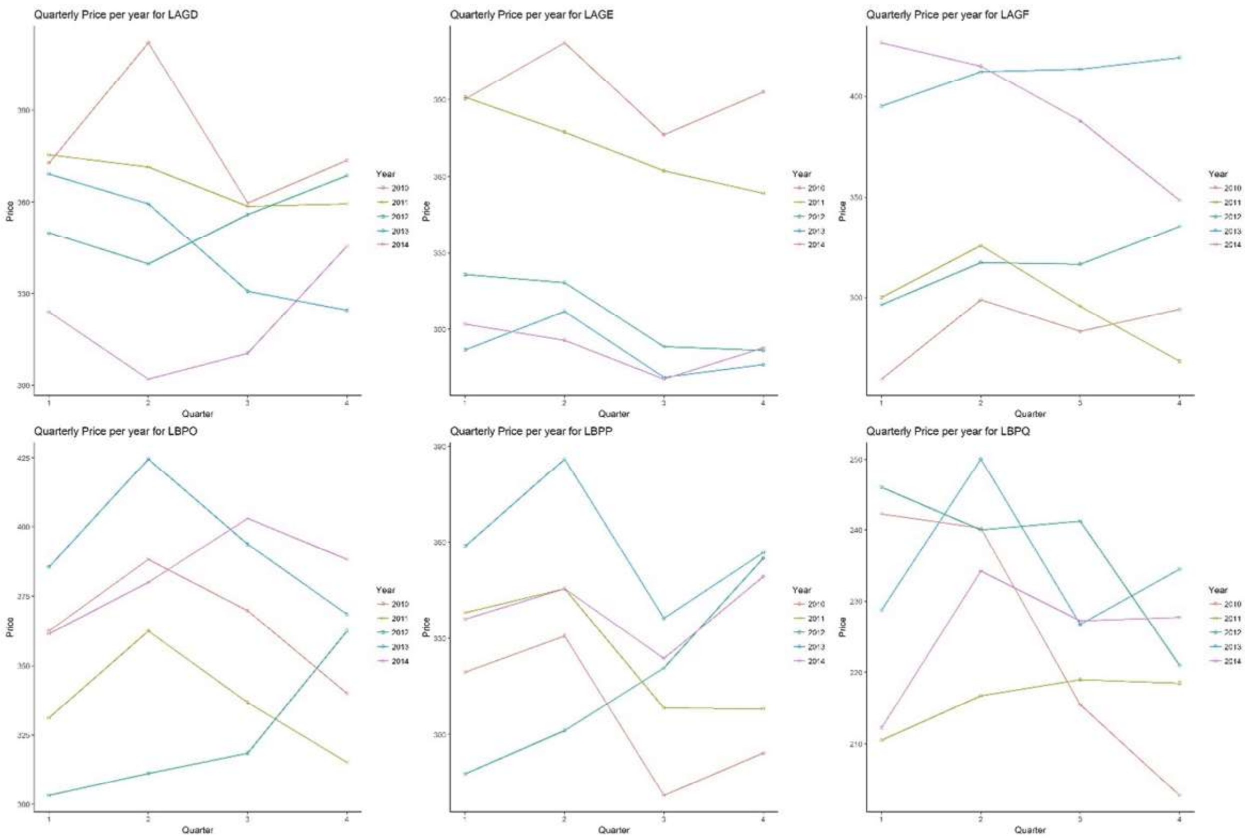


FIGURE 4
AUTOCORRELATIONS OF THE PRICES BY LUMBER TYPE AND VARIOUS LAGS, PERIOD 1995-2014



Our analysis revealed some seasonality in the quarterly price over the last five years as Figure 5 suggests, but again nothing definitive. For example, we see a spike on the price for the second quarter followed by an average dip on the third quarter. Notice again some variations among years not only in the pattern the prices for each lumber type exhibit but also at the levels. The normalization to 52 weeks had minimal effect on the trend since we are averaging over a 13-week period.

FIGURE 5
QUARTERLY PRICES FOR ALL LUMBER TYPES FOR THE YEARS 2010-2014



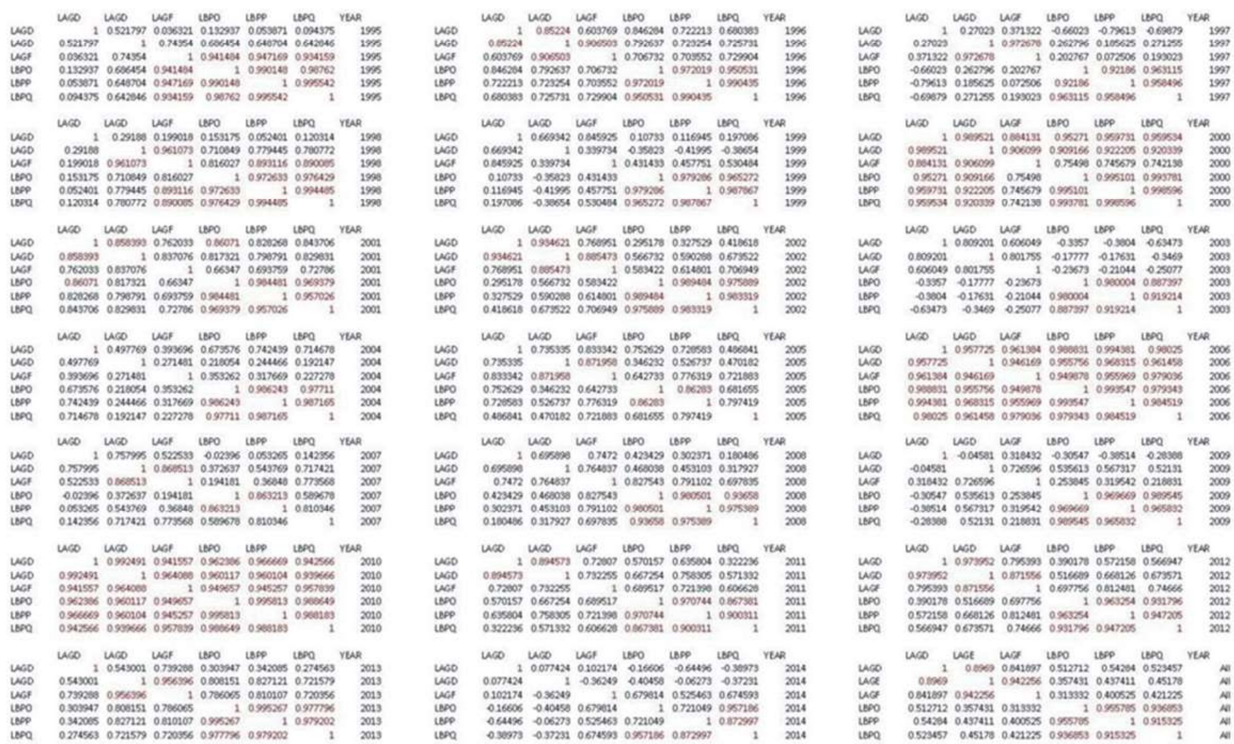
One can thus suggest that the data, in its entirety, indicate stationarity, and no seasonality is discernible. As an appendix, we included an analysis of our time series using the Dickey Fuller test and we discuss our findings there, which conform to analyzing the lags.

Correlations

In this subsection, we shift our attention to the correlations between the prices of different types on a yearly basis. Our observations hint at distinct trends among lumber types based on their number valuation. Figure 6 presents our findings, where colored in red are all correlations above 90 percent.

We can see that various strong correlations appear among the six lumber types for most of the years; for example in 2006 all prices are correlated among themselves. This would give credit to the “law of one price” described in the literature. However, there are years like 2014 that only a few very strong correlations can be found, which puts the global power of this law into question. In Figure 6, the last sub-table shows that if the prices are treated as a long time series (1995-2014) the correlations are stronger among the types that share a number valuation rather than between grade valuations (#2 vs. #4). This is something worth exploring in future publications. Notice also that in general many correlation coefficients are large, close to 0.9. This correlation can be explained if one thinks about the common uses of these types of lumber as it was described in subsection dataset description above. In Figure 7 we provide more information regarding the correlations between different lumber prices for all years.

FIGURE 6
CORRELATIONS BETWEEN DIFFERENT LUMBER TYPES FOR THE YEARS 1995-2014



	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.521797	0.036321	0.132937	0.053871	0.094375	1995
LAGD	0.521797	1	0.74354	0.686454	0.648704	0.642846	1995
LAGF	0.036321	0.74354	1	0.941484	0.947169	0.934159	1995
LBPO	0.132937	0.686454	0.941484	1	0.990148	0.98762	1995
LBPP	0.053871	0.648704	0.947169	0.990148	1	0.995542	1995
LBPQ	0.094375	0.642846	0.934159	0.98762	0.995542	1	1995

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.85224	0.603769	0.846284	0.722213	0.680383	1996
LAGD	0.85224	1	0.906503	0.792637	0.723254	0.725731	1996
LAGF	0.603769	0.906503	1	0.706732	0.703552	0.729904	1996
LBPO	0.846284	0.792637	0.706732	1	0.972019	0.950531	1996
LBPP	0.722213	0.723254	0.703552	0.972019	1	0.990435	1996
LBPQ	0.680383	0.725731	0.729904	0.950531	0.990435	1	1996

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.27023	0.371322	-0.66023	-0.79613	-0.69879	1997
LAGD	0.27023	1	0.972678	0.262796	0.185625	0.271255	1997
LAGF	0.371322	0.972678	1	0.202767	0.072506	0.193023	1997
LBPO	-0.66023	0.262796	0.202767	1	0.92186	0.963115	1997
LBPP	-0.79613	0.185625	0.072506	0.92186	1	0.958496	1997
LBPQ	-0.69879	0.271255	0.193023	0.963115	0.958496	1	1997

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.29188	0.199018	0.153175	0.052401	0.120314	1998
LAGD	0.29188	1	0.961073	0.710849	0.779445	0.780772	1998
LAGF	0.199018	0.961073	1	0.816027	0.893116	0.890085	1998
LBPO	0.153175	0.710849	0.816027	1	0.972633	0.976429	1998
LBPP	0.052401	0.779445	0.893116	0.972633	1	0.994485	1998
LBPQ	0.120314	0.780772	0.890085	0.976429	0.994485	1	1998

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.669342	0.845925	0.10733	0.116945	0.197086	1999
LAGD	0.669342	1	0.339734	-0.35823	-0.41995	-0.38654	1999
LAGF	0.845925	0.339734	1	0.431433	0.457751	0.530484	1999
LBPO	0.10733	-0.35823	0.431433	1	0.979286	0.965272	1999
LBPP	0.116945	-0.41995	0.457751	0.979286	1	0.987867	1999
LBPQ	0.197086	-0.38654	0.530484	0.965272	0.987867	1	1999

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.989521	0.884131	0.95271	0.959731	0.959534	2000
LAGD	0.989521	1	0.906099	0.909166	0.922205	0.920339	2000
LAGF	0.884131	0.906099	1	0.75498	0.745679	0.742138	2000
LBPO	0.95271	0.909166	0.75498	1	0.995101	0.993781	2000
LBPP	0.959731	0.922205	0.745679	0.995101	1	0.998596	2000
LBPQ	0.959534	0.920339	0.742138	0.993781	0.998596	1	2000

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.858393	0.762033	0.86071	0.828268	0.843706	2001
LAGD	0.858393	1	0.837076	0.817321	0.798791	0.829831	2001
LAGF	0.762033	0.837076	1	0.66347	0.693759	0.72786	2001
LBPO	0.86071	0.817321	0.66347	1	0.984481	0.969379	2001
LBPP	0.828268	0.798791	0.693759	0.984481	1	0.957026	2001
LBPQ	0.843706	0.829831	0.72786	0.969379	0.957026	1	2001

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.934621	0.768951	0.295178	0.327529	0.418618	2002
LAGD	0.934621	1	0.885473	0.566732	0.590288	0.673522	2002
LAGF	0.768951	0.885473	1	0.583422	0.614801	0.706949	2002
LBPO	0.295178	0.566732	0.583422	1	0.989484	0.975889	2002
LBPP	0.327529	0.590288	0.614801	0.989484	1	0.983319	2002
LBPQ	0.418618	0.673522	0.706949	0.975889	0.983319	1	2002

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.809201	0.606049	-0.3357	-0.3804	-0.63473	2003
LAGD	0.809201	1	0.801755	-0.17777	-0.17631	-0.3469	2003
LAGF	0.606049	0.801755	1	-0.23673	-0.21044	-0.25077	2003
LBPO	-0.3357	-0.17777	-0.23673	1	0.980004	0.887397	2003
LBPP	-0.3804	-0.17631	-0.21044	0.980004	1	0.919214	2003
LBPQ	-0.63473	-0.3469	-0.25077	0.887397	0.919214	1	2003

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.497769	0.393696	0.673576	0.742439	0.714678	2004
LAGD	0.497769	1	0.271481	0.218054	0.244466	0.192147	2004
LAGF	0.393696	0.271481	1	0.353262	0.317669	0.227278	2004
LBPO	0.673576	0.218054	0.353262	1	0.986243	0.97711	2004
LBPP	0.742439	0.244466	0.317669	0.986243	1	0.987165	2004
LBPQ	0.714678	0.192147	0.227278	0.97711	0.987165	1	2004

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.735335	0.833342	0.752629	0.728583	0.486841	2005
LAGD	0.735335	1	0.871958	0.346232	0.526737	0.470182	2005
LAGF	0.833342	0.871958	1	0.642733	0.776319	0.721883	2005
LBPO	0.752629	0.346232	0.642733	1	0.86283	0.681655	2005
LBPP	0.728583	0.526737	0.776319	0.86283	1	0.797419	2005
LBPQ	0.486841	0.470182	0.721883	0.681655	0.797419	1	2005

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.957725	0.961384	0.988831	0.994381	0.98025	2006
LAGD	0.957725	1	0.946169	0.955756	0.968315	0.961458	2006
LAGF	0.961384	0.946169	1	0.949878	0.955969	0.979036	2006
LBPO	0.988831	0.955756	0.949878	1	0.993547	0.979343	2006
LBPP	0.994381	0.968315	0.955969	0.993547	1	0.984519	2006
LBPQ	0.98025	0.961458	0.979036	0.979343	0.984519	1	2006

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.757995	0.522533	-0.02396	0.053265	0.142356	2007
LAGD	0.757995	1	0.868513	0.372637	0.543769	0.717421	2007
LAGF	0.522533	0.868513	1	0.194181	0.36848	0.773568	2007
LBPO	-0.02396	0.372637	0.194181	1	0.863213	0.589678	2007
LBPP	0.053265	0.543769	0.36848	0.863213	1	0.810346	2007
LBPQ	0.142356	0.717421	0.773568	0.589678	0.810346	1	2007

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.695898	0.7472	0.423429	0.302371	0.180486	2008
LAGD	0.695898	1	0.764837	0.468038	0.453103	0.317927	2008
LAGF	0.7472	0.764837	1	0.827543	0.791102	0.697835	2008
LBPO	0.423429	0.468038	0.827543	1	0.980501	0.93658	2008
LBPP	0.302371	0.453103	0.791102	0.980501	1	0.975389	2008
LBPQ	0.180486	0.317927	0.697835	0.93658	0.975389	1	2008

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	-0.04581	0.318432	-0.30547	-0.38514	-0.28388	2009
LAGD	-0.04581	1	0.726596	0.535613	0.567317	0.52131	2009
LAGF	0.318432	0.726596	1	0.253845	0.319542	0.218831	2009
LBPO	-0.30547	0.535613	0.253845	1	0.969669	0.989545	2009
LBPP	-0.38514	0.567317	0.319542	0.969669	1	0.965832	2009
LBPQ	-0.28388	0.52131	0.218831	0.989545	0.965832	1	2009

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.992491	0.941557	0.962386	0.966669	0.942566	2010
LAGD	0.992491	1	0.964088	0.960117	0.960104	0.939666	2010
LAGF	0.941557	0.964088	1	0.949657	0.945257	0.957839	2010
LBPO	0.962386	0.960117	0.949657	1	0.995813	0.988649	2010
LBPP	0.966669	0.960104	0.945257	0.995813	1	0.988183	2010
LBPQ	0.942566	0.939666	0.957839	0.988649	0.988183	1	2010

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.894573	0.72807	0.570157	0.635804	0.322236	2011
LAGD	0.894573	1	0.732255	0.667254	0.758305	0.571332	2011
LAGF	0.72807	0.732255	1	0.689517	0.721398	0.606628	2011
LBPO	0.570157	0.667254	0.689517	1	0.970744	0.867381	2011
LBPP	0.635804	0.758305	0.721398	0.970744	1	0.900311	2011
LBPQ	0.322236	0.571332	0.606628	0.867381	0.900311	1	2011

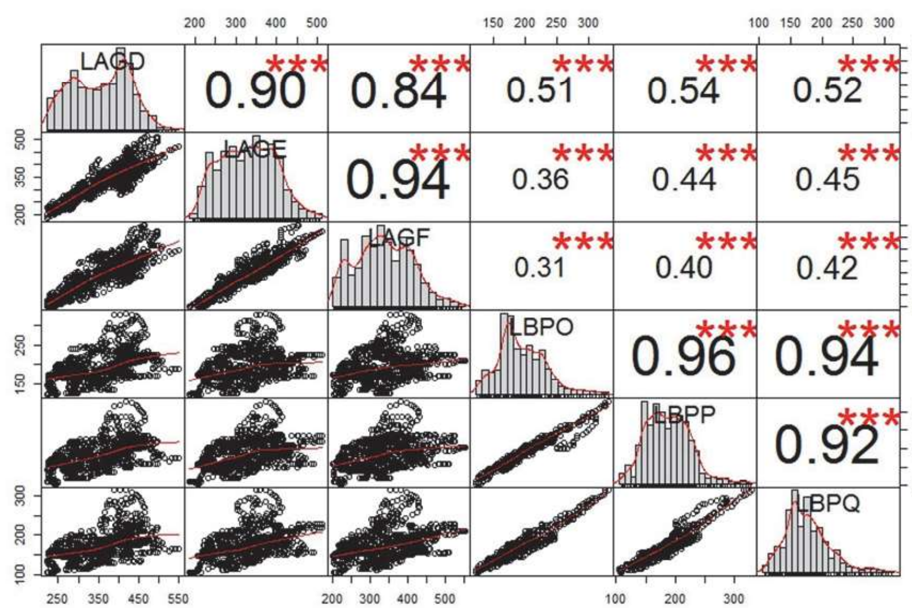
	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.973952	0.795393	0.390178	0.572158	0.566947	2012
LAGD	0.973952	1	0.871556	0.516689	0.668126	0.673571	2012
LAGF	0.795393	0.871556	1	0.697756	0.812481	0.74666	2012
LBPO	0.390178	0.516689	0.697756	1	0.963254	0.931796	2012
LBPP	0.572158	0.668126	0.812481	0.963254	1	0.947205	2012
LBPQ	0.566947	0.673571	0.74666	0.931796	0.947205	1	2012

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.543001	0.739288	0.303947	0.342085	0.274563	2013
LAGD	0.543001	1	0.956396	0.808151	0.827121	0.721579	2013
LAGF	0.739288	0.956396	1	0.786065	0.810107	0.720356	2013
LBPO	0.303947	0.808151	0.786065	1	0.995267	0.977796	2013
LBPP	0.342085	0.827121	0.810107	0.995267	1	0.979202	2013
LBPQ	0.274563	0.721579	0.720356	0.977796	0.979202	1	2013

	LAGD	LAGD	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.077424	0.102174	-0.16606	-0.64496	-0.38973	2014
LAGD	0.077424	1	-0.36249	-0.40458	-0.06273	-0.37231	2014
LAGF	0.102174	-0.36249	1	0.679814	0.525463	0.674593	2014
LBPO	-0.16606	-0.40458	0.679814	1	0.721049	0.957186	2014
LBPP	-0.64496	-0.06273	0.525463	0.721049	1	0.872997	2014
LBPQ	-0.38973	-0.37231	0.674593	0.957186	0.872997	1	2014

	LAGD	LAGE	LAGF	LBPO	LBPP	LBPQ	YEAR
LAGD	1	0.8969	0.841897	0.512712	0.54284	0.523457	All
LAGE	0.8969	1	0.942256	0.357431	0.437411	0.45178	All
LAGF	0.841897	0.942256	1	0.313332	0.400525	0.421225	All
LBPO	0.512712	0.357431	0.313332	1	0.955785	0.936853	All
LBPP	0.54284	0.437411	0.400525	0.955785	1	0.915325	All
LBPQ	0.523457	0.45178	0.421225	0.936853	0.915325	1	All

FIGURE 7
CORRELATIONS BETWEEN DIFFERENT LUMBER TYPES FOR ALL YEARS



Once more, we would like to comment that lumber price series are characterized by high volatility especially in the short-run. In addition, lumber price series are sensitive to market fluctuations including the lumber market (use of lumber, regions, etc.) and the housing market as well as macroeconomic factors like interest rates (e.g. Karali, 2011). These factors could help explain the price variation we observe throughout the years and suggest that no simple model like a regression or moving average would fit the data series well. Although our model did not exploit the connections we identified, we will surely be pursuing that venue in future publications.

MODEL

The analysis above indicated that a clear forecast of the weekly price is intractable. In this section, we present a conceptual model that relates buying strategies to predicted movement of prices. Our models aim at identifying the turning points, which seemed more stable through the years.

Conceptual Model

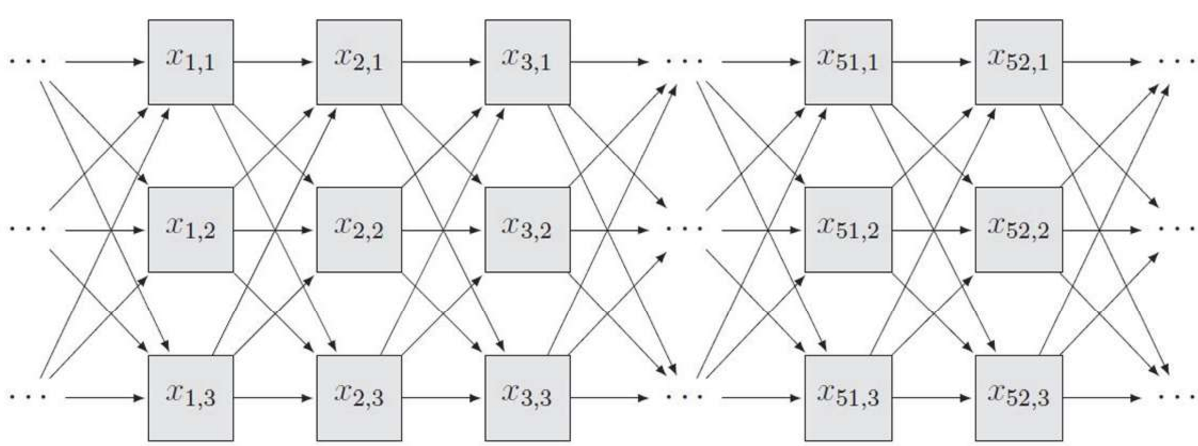
We consider an agent, a lumber buyer, who has historic data on lumber prices (spot or futures can be considered). Each week, the agent makes purchasing decisions based on the direction he or she anticipates the prices will move based on the forecast and disregards the spot price. This setting allows us to discretize the price in the following sense. Every week there are three possible scenarios: "The price will go up from the previous week, the price will go down, or the price will stay the same."

Let $X_1, X_2, \ldots, X_{52}$ be random variables corresponding to the purchasing decisions, one for each week. For each of the $X_i's$ we have three possible directions:

1) $x_{i,1}$ =The price goes up from the previous week.
2) $x_{i,2}$ =The price stays the same as the last week.
3) $x_{i,3}$ =The price goes down from last week.

Due to the cyclic nature of the yearly calendar, week 1 uses week 52 as a previous week with minor adjustments.

FIGURE 8
MATHEMATICAL DESCRIPTION OF THE STATES



We can compute the conditional probabilities $P(X_{i+1} = x_{i+1,j} | X_i = x_{i,k})$ for all possible cases on the assumption that the variables X_i and X_{i+1} are not independent. The purchasing decision for week i may affect the purchasing decision for week $i+1$. To make predictions stronger, we also compute the two-step conditional probabilities, i.e. $P(X_{i+1} = x_{i+1,j} | X_i = x_{i,k}, X_{i-1} = x_{i-1,l})$. Clearly, these probabilities change throughout the year and seasonality is easier to uncover using aggregate historical data in this setting.

We propose three weekly purchasing recommendations for our buyer, namely "Don't Buy (DB), Buy (B+), and Fill the inventory (B++)." The (DB) and (B++) options are self-explanatory. The (B+) strategy allows the buyer to satisfy immediate demand and store some extra units for future applications. In our implementation, the (B+) strategy corresponds to buying twice as much as the needed quantity (2 units) but that can be changed at will. The aggressive purchasing strategy (B++) is not uncommon when there is lack of information or uncertainty and in our implementation is driven by two consecutive indications of increased lumber prices for the weeks that follow although again that can be extended to a bigger time period for other examples. The aggressive buying strategy could be justified with the buyer making sure they capitalize on a forecasted big increase in prices.

Predictive Model

Our method uses a greater than five year history period (training period) to identify the movements in lumber prices for the selected period. Additional inputs include, the starting year for the predictions, the lumber type; the storage capacity and how long of a history should be included in the prediction. In our implementation, all previous history is included up to the selected year.

Two new variables are created; one that contains the prices forwarded by one week (FOW1) and another one where prices are forwarded by two weeks (FOW2). The direction is captured by the difference between the current price and the two forward prices (FOW1 and FOW2) and turned into an indicator variable with values, +1 if the price goes up, -1 if the price goes down and 0 if it is unchanged. For the last two weeks (week 51 and week 52) prices from the next year are utilized, and when those are not available, they are extrapolated using past prices from these weeks.

For each week, the model computes the probability of the price change in the next two weeks using the selected years as follows:

$$a_1 = P(FOW1 > Price)\,, a_2 = P(FOW1 < Price), a_3 = P(FOW1 = Price) \quad (1)$$

$$b_1 = P(FOW2 > Price)\,, b_2 = P(FOW2 < Price), b_3 = P(FOW2 = Price) \quad (2)$$

Obviously $a_1 + a_2 + a_3 = 1$ and $b_1 + b_2 + b_3 = 1$. Finally, the model creates purchasing recommendations based on the following three cases:

1) $a_1 - a_2 > \alpha$ and $b_1 - b_2 > \alpha$, where $\alpha = 0.2$ is a threshold chosen through experimentation. To compute this parameter, we tested various alphas with increments of 0.05, on the first 10 years of the dataset and we chose the alpha with the lowest yearly cost for all lumber types on average. In this case, we anticipate that on average the price will increase dramatically for both the next two weeks. Thus, the algorithm suggests an aggressive buy (B++).
2) $a_1 - a_2 \leq 0$ and $b_1 - b_2 \leq 0$. In this case, we assume that on average the price will go down or stay the same the next two weeks. The model suggests a halt on purchasing (DB).
3) Everything else, which means that the prices do not follow a clear trend the next two weeks, but at least one of the two is on average a little bit greater or equal to the current week's price. The model recommends a moderate buy (B+).

Model Implementation

Our dataset contains information on six types of lumber and respective prices for the years 1995-2014. In order for the model to stabilize, we require five years of price series data, so our smallest starting year is 2000. To allow for the implementation of the predictive algorithm and the three recommendations (B+, B++, and DB) we impose a lower bound on the storage capacity to be four units. That is the minimal warehouse size since we require one operational unit and a possible purchase up to three more units. Note that the variable "units" is an arbitrary measure of quantity and it can be adjusted to the operational schema of any lumber purchasing entity. Also, note that a smaller storage capacity would disregard the three purchasing choices. For example, at storage capacity of three units the recommendations (B++) and (B+) will lead to the same purchasing strategy of purchasing two units.

Although theoretically there is no upper bound to the capacity of the warehouse, we decided to stop it at nine, presenting us with six different cases to test in our experiment. Thus, following the rule of thumb we have more than 30 comparisons per year (36 in our case) which leads to safe and easily interpretable statistical results. Furthermore, when we multiply by the number of years (15) we end up with 540 different possible test cases.

As mentioned earlier, our algorithm provides information for the forward prices (FOW1, and FOW2), and the direction of the price change with respect to the current price for the two weeks that follow, (D_{FOW1}, and D_{FOW2}, respectively). A snapshot of the information can be found in Table 4.

TABLE 4
INTERMEDIATE TABLE SHOWING THE ALGORITHM AT WORK

Price Series	Description	Year	Week	Price	FOW1	FOW2	D_{FOW1}	D_{FOW2}
LAGE	KD Southern…	1995	1	395	400	380	+1	-1
LAGE	KD Southern…	1995	2	400	380	380	-1	-1
LAGE	KD Southern…	1995	3	380	380	…	0	…
LAGE	KD Southern…	1995	4	380	…	…	…	…

Using the logic statements above the algorithm outputs a list of recommendations for each week of that year. A representation of the information we have (after cleaning up the dummy columns) is shown in Table 5. As a reminder, the recommendation Buy (B+) recommends the purchase of units that will satisfy immediate demand and some additional units, in our experimentation set arbitrarily to two. One could alter this to match other purchasing schemes if needed.

TABLE 5
WEEKLY PRICES FOR YEAR 2002 FOR LUMBER TYPE LAGE INCLUDING THE MODEL'S RECOMMENDATIONS

Price Series	Description	Year	Week	Price	Recommendation
LAGE	KD Southern…	2002	1	395	B++
LAGE	KD Southern…	2002	2	400	B+
…	…	…	…	…	…
LAGE	KD Southern…	2002	52	430	DB

EXPERIMENTAL RESULTS

To test the predictive capabilities of our model we created a purchasing strategy for each year and each lumber type. We do not allow arbitrage to take place; our agent is only allowed to purchase lumber units and not sell, hence not taking advantage of potential price differences. We impose two restricting assumptions: we demand one lumber unit to operate and the warehouse size is limited. The first assumption is important to satisfy immediate demand for the operation; the second assumption is important for the implementation of the predictive algorithm and is a realistic restriction. Lastly, our model requires having the same amount of units (s) in store at the start and the end of each year.

The purchasing strategy, based on the suggestion tables of our model proceeds week by week as follows: Based on our recommendation table we fill the warehouse if the suggestion is (B++). We buy two units if the suggestion is (B+), in anticipation of a moderate price increase. Finally, if the recommendation is (DB) we halt purchases unless the warehouse is empty in which case we buy one unit to cover the operational needs of the week. In addition, as we get closer to the end of the year our purchases are modified in such a way so that we do not exceed the amount of units (s) set as our end of the year goal in the warehouse.

Table 6 shows the proposed purchases, stored quantity and the actual cost per week for the lumber type "KD Southern Pine (Eastside) #2 2x8 random Prices Net f.o.b. Mill (LAGF)," during the year 2010. The total cost for that year is computed and compared to the other suggestion models described below. Notice that although our method does not always predict the correct movement of the spot prices, if one looks at the overall output for the year then on average the prediction is informative and in most cases leads to a better yearly purchasing strategy.

TABLE 6
A PURCHASING STRATEGY FOR LAGF IN 2010

Number	TAG	Year	Week	Price	Recommendation	Purchases	Stored Quantity	Cost
2861	LAGF	2010	1	239	B++	9	1	2151
2862	LAGF	2010	2	247	B+	1	9	247
2863	LAGF	2010	3	264	B+	1	9	264
2864	LAGF	2010	4	287	B+	1	9	287
2865	LAGF	2010	5	308	B+	1	9	308
2866	LAGF	2010	6	326	B++	1	9	326
2867	LAGF	2010	7	333	B++	1	9	333
2868	LAGF	2010	8	327	B++	1	9	327
2869	LAGF	2010	9	317	B+	1	9	317
2870	LAGF	2010	10	304	DB	0	9	0
2871	LAGF	2010	11	302	DB	0	8	0
2872	LAGF	2010	12	310	DB	0	7	0
2873	LAGF	2010	13	320	B+	2	6	640
2874	LAGF	2010	14	337	B++	3	7	1011
2875	LAGF	2010	15	357	B++	1	9	357
2861	LAGF	2010	16	371	B+	1	9	371
2862	LAGF	2010	17	375	DB	0	9	0
2863	LAGF	2010	18	369	DB	0	8	0
…	…	…	…	…	…	…	…	…
2904	LAGF	2010	44	229	B+	2	3	458
2905	LAGF	2010	45	232	B+	2	4	464
2906	LAGF	2010	46	233	B+	2	5	466
2907	LAGF	2010	47	233	B+	0	6	0
2908	LAGF	2010	48	228	DB	0	5	0
2909	LAGF	2010	49	224	DB	0	4	0
2910	LAGF	2010	50	223	B+	0	3	0
2911	LAGF	2010	51	224	B+	0	2	0
2912	LAGF	2010	52	227	B++	0	1	0

In the following subsections, we compare our "long-history" method with other strategies and recommendation methods, using different warehouse sizes.

Naive Method

The simplest purchasing strategy that will satisfy all the assumptions of our experiment is buying one unit every week. We call this the "naive method", and we use it as a first benchmark to prove the predictive power of our method. Figure 9 shows the difference in yearly cost of our long-history method vs the naïve method, for all years and all lumber types for various warehouse sizes.

FIGURE 9
LONG-HISTORY METHOD VS NAÏVE METHOD

Price Series	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Size
LAGD	83	348	355	500	428	474	219	325	264	249	465	325	676	455	410	4
LAGE	42	193	384	382	420	355	271	251	200	201	464	242	544	438	394	4
LAGF	91	179	421	265	614	168	324	180	193	256	446	282	588	463	291	4
LBPO	87	256	295	251	356	215	147	175	174	120	330	227	373	354	430	4
LBPP	90	250	330	257	346	222	99	155	165	107	342	186	359	295	375	4
LBPQ	85	196	256	205	311	181	144	143	151	89	245	172	315	277	390	4
LAGD	25	407	356	509	409	502	161	356	272	245	566	319	718	439	373	5
LAGE	-39	229	427	419	407	337	243	262	239	174	533	282	675	494	449	5
LAGF	-39	236	509	242	781	89	329	169	260	268	552	306	735	510	299	5
LBPO	67	321	357	288	431	228	129	178	199	90	387	237	428	396	515	5
LBPP	75	325	415	297	416	240	69	164	199	72	399	186	414	320	445	5
LBPQ	70	244	301	235	379	189	134	143	171	44	268	179	358	294	465	5
LAGD	8	372	388	502	388	502	117	384	235	251	576	321	762	438	317	6
LAGE	-97	336	453	448	388	340	190	256	293	137	570	310	799	541	509	6
LAGF	-146	334	585	245	954	42	326	159	320	272	640	312	836	520	336	6
LBPO	47	386	414	320	501	248	109	180	224	57	429	239	483	428	600	6
LBPP	60	400	493	331	481	263	35	176	242	37	444	181	464	335	500	6
LBPQ	55	292	343	262	442	199	124	145	196	-1	281	184	401	306	540	6
LAGD	-30	344	408	522	365	508	65	386	183	249	598	321	821	467	254	7
LAGE	-135	481	476	475	379	360	100	251	376	100	655	304	890	552	489	7
LAGF	-243	477	644	236	1109	44	289	127	379	269	707	307	920	508	381	7
LBPO	27	451	468	350	569	273	84	182	252	24	464	239	538	445	685	7
LBPP	45	475	568	365	546	286	1	188	288	2	476	171	514	345	545	7
LBPQ	40	340	383	286	502	209	109	147	221	-46	274	184	444	288	610	7
LAGD	-94	400	435	537	351	492	-11	370	140	251	641	277	918	401	227	8
LAGE	-158	575	482	489	366	382	15	245	455	60	716	273	967	555	472	8
LAGF	-300	583	687	227	1199	94	270	94	432	263	756	296	997	484	447	8
LBPO	2	511	520	380	637	298	59	189	282	-9	486	234	593	457	765	8
LBPP	30	545	640	399	611	304	-36	205	336	-33	503	156	564	340	580	8
LBPQ	25	383	423	310	562	219	94	154	249	-91	257	179	487	270	675	8
LAGD	-145	472	400	545	404	486	-96	380	100	237	715	225	1041	324	216	9
LAGE	-171	567	475	493	361	387	-64	246	466	23	704	238	1040	537	487	9
LAGF	-337	680	717	225	1281	156	250	69	482	259	785	285	1067	457	525	9
LBPO5	-23	571	567	410	702	318	32	196	312	-42	503	224	643	459	840	9
LBPP	15	615	712	431	674	317	-73	222	384	-68	520	141	609	320	600	9
LBPQ	10	426	463	334	620	229	79	161	277	-136	233	174	530	262	717	9

Note: In red, we report the losses of the long history methods vs the naive one, for all years, all lumber types and various warehouse sizes with initial quantity 1.

Our long-history method outperforms the naive one more than 95 percent of the times assuming the starting lumber quantity in the warehouse is one. It is interesting that the naive method performs better than ours mostly in year 2000 and only for large warehouses during the years 2006 and 2009. For year 2000, an explanation could be that the 5-year history is not enough for our prediction algorithm to compute the correct yearly movements. Another reason could be the mini recession that hit the lumber market during that period. A thorough analysis for the years of the recession including 2009 is presented in the section discussion that follows. If we confine ourselves to a warehouse of size 4 (moderate or average size warehouse) we win constantly independent of the starting and ending quantity.

Random Method

In order to prove the predicting capabilities of our long-history method independent of the purchasing schema, we created the "random" purchasing method as follows. For each year and each lumber type a new recommendation table was created using the same labels (B++, B+ and DB) drawn randomly from a distribution whose probabilities are equal to the probabilities of (B++, B and DB) in the recommendation tables of our long-history method.

The algorithm again fills the warehouse when the recommendation is (B++), buys two units when the recommendation is (B+) and halts purchases if there recommendation is (DB). Once more, we make sure the demand of one unit is covered weekly and the purchases are adjusted so that at the end of the year we have the same quantity s in our inventory as when we started.

Our strategy is better than the “random” one 80 percent of the times for warehouse sizes of four to nine units and a starting quantity of one. We note here that we almost always lose in 2009 against the “random” strategy as Figure 10 suggests. We discuss this in depth in section Recession.

FIGURE 10
LONG-HISTORY METHOD VS RANDOM METHOD

Price Series	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Size
LAGD	1	-21	50	56	11	78	90	-16	42	-18	330	138	-73	39	28	4
LAGE	-110	-119	104	-30	160	15	122	-3	69	-72	226	128	188	109	20	4
LAGF	-50	31	99	-101	47	-8	151	-84	135	11	319	166	185	166	6	4
LBPO	66	132	114	49	72	-14	44	-15	-22	-64	118	17	79	134	115	4
LBPP	29	109	129	79	43	23	45	26	-6	-113	128	65	82	177	125	4
LBPQ	64	58	73	70	94	29	26	18	-14	-72	90	58	29	24	107	4
LAGD	-40	-17	160	-121	-99	150	148	-72	94	-41	363	41	66	45	-147	5
LAGE	-135	-178	59	-64	-116	22	238	143	105	19	234	145	155	257	320	5
LAGF	-173	-4	483	-160	238	-79	92	12	434	42	349	153	113	285	92	5
LBPO	125	152	247	147	182	10	3	21	50	-110	255	83	108	194	190	5
LBPP	118	178	278	100	142	25	19	9	26	-151	191	19	106	133	215	5
LBPQ	80	170	236	98	68	-3	23	20	9	-177	-90	66	40	104	247	5
LAGD	163	312	318	-200	7	81	385	63	115	-7	2	165	24	229	-185	6
LAGE	-223	418	371	-36	-131	153	409	-56	212	-133	431	143	231	254	205	6
LAGF	-263	169	483	-71	567	-397	141	-23	-152	48	656	-40	210	485	153	6
LBPO	194	293	252	100	158	27	24	-32	38	-171	228	93	85	376	270	6
LBPP	207	280	303	79	243	35	1	16	29	-254	366	19	97	305	267	6
LBPQ	182	125	191	148	127	35	33	54	-30	-220	286	80	69	189	230	6
LAGD	90	-55	303	-493	-348	127	283	34	255	-50	630	220	215	106	-193	7
LAGE	-201	143	169	41	78	33	263	290	200	-22	577	316	26	498	22	7
LAGF	-202	131	590	-651	644	-136	286	134	168	-49	818	169	289	448	142	7
LBPO	255	277	258	128	161	4	60	-12	97	-215	300	112	143	284	360	7
LBPP	185	524	392	114	245	122	14	63	202	-355	284	99	152	311	404	7
LBPQ	220	239	254	139	123	48	55	30	5	-285	119	46	24	-53	312	7
LAGD	326	-376	427	-640	-290	51	588	52	174	-15	233	74	-195	275	-199	8
LAGE	141	180	442	9	-105	336	127	199	279	-190	655	233	258	508	184	8
LAGF	-217	350	570	-183	1147	-75	467	236	298	-108	594	114	109	542	167	8
LBPO	239	345	352	89	443	72	83	41	120	-313	321	159	207	502	390	8
LBPP	272	245	395	246	294	161	28	71	-128	-341	424	108	84	235	430	8
LBPQ	255	216	234	170	162	89	61	89	-58	-369	256	115	147	161	460	8
LAGD	301	-179	392	-737	315	132	622	109	85	35	-160	286	93	-65	-179	9
LAGE	120	32	424	-412	128	16	370	218	373	-154	1278	306	133	584	230	9
LAGF	-193	634	447	-110	814	-130	458	315	467	-93	388	136	230	411	232	9
LBPO	319	421	421	157	468	105	80	53	87	-341	864	132	282	266	370	9
LBPP	230	611	474	116	133	93	88	85	162	-490	224	105	167	804	571	9
LBPQ	340	184	340	238	320	92	33	83	46	-413	167	160	46	320	556	9

Note: In red, we report the losses of the long history methods vs the random one, for all years, all lumber types and various warehouse sizes with initial quantity 1.

Short-term Method

We also implemented a classic “short term” prediction strategy as follows. For each week, we computed the movement based only on the previous week. If the price went up significantly, we recommend (B++) in anticipation of a price hike. If it went up by a moderate amount, we recommend (B+) in anticipation of a moderate price increase. Finally, if the prices went down we recommend (DB) as we believe prices will keep falling. Our purchase suggestions again follow the rules of meeting demand (i.e. have at least one unit a week) and adjust so that by the end of the year we have the same starting quantity in storage. To distinguish big increases vs short increases we computed the mean of the positive increases for each lumber type throughout the years up to the starting year and used that as a cutoff point.

Once again, the “long-history” method is better than the “short-term” method 78 percent of the times as can be seen in Figure 11. Again, we see a failure of our method to correctly identify the price movements in 2009 compared to this one.

FIGURE 11
LONG-HISTORY METHOD VS SHORT-TERM METHOD

Price Series	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Size
LAGD	13	-24	217	38	13	206	86	78	111	151	-42	143	78	169	-27	4
LAGE	-4	6	301	120	235	90	158	102	22	83	26	242	185	167	42	4
LAGF	-77	-148	202	89	69	50	241	70	-3	177	166	234	195	165	2	4
LBPO	135	80	54	99	37	153	123	126	83	39	83	104	66	53	60	4
LBPP	125	79	110	133	145	148	111	142	60	17	99	102	98	138	105	4
LBPQ	140	135	116	119	110	143	140	135	65	31	52	155	86	64	120	4
LAGD	-34	-134	263	-60	-50	256	74	154	130	185	-121	139	-16	125	-87	5
LAGE	-23	98	353	171	256	114	220	106	52	56	47	319	239	180	58	5
LAGF	-207	-77	231	147	101	35	328	110	32	208	147	309	222	157	-22	5
LBPO	130	35	29	71	-13	173	113	123	56	-27	53	96	21	18	30	5
LBPP	112	59	126	125	150	168	94	149	28	-44	101	85	88	170	93	5
LBPQ	140	148	106	100	93	148	140	135	42	-27	27	165	64	57	110	5
LAGD	-54	-209	326	-204	-41	282	71	232	140	223	-180	155	-106	100	-128	6
LAGE	-26	257	413	222	248	184	237	91	109	6	51	349	271	199	74	6
LAGF	-322	47	279	177	174	32	354	140	73	230	97	315	196	93	17	6
LBPO	120	-5	16	48	-59	185	103	120	34	-86	3	91	-14	-2	0	6
LBPP	97	57	138	120	165	188	72	156	6	-97	55	68	106	205	86	6
LBPQ	140	153	121	82	46	153	140	135	7	-85	-3	175	39	35	60	6
LAGD	-92	-230	373	-359	-27	311	59	247	193	237	-218	194	-175	52	-177	7
LAGE	-41	445	486	268	251	290	201	86	250	-48	129	365	307	205	20	7
LAGF	-419	237	344	188	281	106	310	151	146	246	73	327	185	31	96	7
LBPO	105	-45	21	30	-102	197	91	117	14	-142	-34	91	-44	-7	-40	7
LBPP	82	59	151	107	170	208	46	165	-9	-145	22	51	123	205	63	7
LBPQ	140	166	144	66	26	158	140	135	-23	-140	-30	180	27	33	15	7
LAGD	-122	-254	452	-526	-51	323	26	258	251	270	-226	252	-195	0	-265	8
LAGE	-19	642	553	241	253	399	165	90	384	-99	203	383	316	216	-47	8
LAGF	-466	389	393	140	303	238	306	180	232	261	49	341	191	-16	204	8
LBPO	87	-67	6	15	-137	209	79	114	4	-192	-64	91	-71	-7	-70	8
LBPP	67	56	146	84	170	228	18	174	-9	-198	-19	31	134	180	25	8
LBPQ	135	186	139	50	6	163	140	135	-43	-189	-45	180	20	36	-5	8
LAGD	-181	-282	527	-695	-72	335	-18	274	311	272	-222	302	-194	-42	-347	9
LAGE	18	695	609	206	248	482	103	107	395	-144	218	408	323	244	-89	9
LAGF	-488	523	436	125	315	387	300	219	302	271	5	356	215	-76	300	9
LBPO	72	-69	1	5	-167	209	68	117	-13	-233	-79	91	-98	3	-100	9
LBPP	52	29	103	51	167	248	-14	186	4	-251	-54	9	132	215	-18	9
LBPQ	125	196	139	34	-14	168	140	135	-53	-241	-45	180	15	49	-20	9

Note: In red, we report the losses of the long history methods vs the short-term one, for all years, all lumber types and various warehouse sizes with initial quantity 1.

Futures Method

Finally, we wanted to analyze the predictive capabilities of the futures price series. We also wanted to test the assumption that is prevalent in the literature, that the futures and spot price are co-integrated. This would imply that using delayed versions of the futures price time series instead of the reported price would lead to a better price forecasting. In our first attempt at creating recommendations from the futures prices, we repeated our method but replaced the price series with various delayed versions of the futures price up to 26 weeks (roughly half a year).

FIGURE 12
LONG-HISTORY METHOD VS FUTURES METHOD

Price Series	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Size
LAGD	82	29	133	5	-141	-12	76	82	22	23	236	45	-16	-29	-54	4
LAGE	62	-76	105	-51	-58	-84	89	102	-99	8	210	39	-34	8	-152	4
LAGF	35	-42	48	-83	-1	-161	47	28	-119	19	176	86	37	14	-44	4
LBPO	57	128	94	44	32	16	28	9	-3	-44	160	-5	58	59	145	4
LBPP	70	155	138	41	0	21	18	17	-42	-37	165	-27	60	2	156	4
LBPQ	60	88	90	40	29	18	2	11	-44	-48	115	21	41	-55	155	4
LAGD	170	71	173	-77	-278	-10	96	110	-13	17	350	34	-141	-64	-129	5
LAGE	146	-19	155	-116	-147	-149	101	136	-161	9	296	75	-53	28	-171	5
LAGF	53	29	78	-181	19	-224	42	57	-126	28	306	109	50	16	-48	5
LBPO	90	206	126	45	40	31	32	9	5	-64	232	-18	69	64	175	5
LBPP	108	256	191	39	-2	44	15	25	-27	-62	224	-44	69	-11	181	5
LBPQ	93	161	109	40	29	38	-3	20	-37	-75	155	26	34	-105	185	5
LAGD	277	21	245	-160	-413	-30	183	119	-69	13	367	48	-233	-56	-230	6
LAGE	220	129	196	-173	-224	-198	132	145	-162	-3	346	109	-48	58	-171	6
LAGF	68	170	107	-236	42	-255	89	92	-102	21	423	117	38	3	-27	6
LBPO	115	284	140	38	38	53	46	6	13	-87	286	-42	70	84	190	6
LBPP	138	355	225	26	-16	69	25	33	-8	-95	268	-61	66	-14	191	6
LBPQ	123	234	110	39	19	60	4	32	-25	-103	185	27	17	-135	200	6
LAGD	383	-41	350	-298	-504	-80	301	102	-129	13	390	96	-268	-2	-307	7
LAGE	332	343	278	-250	-211	-242	153	140	-102	-6	442	139	-34	71	-194	7
LAGF	108	382	159	-320	50	-296	151	105	-97	18	517	128	49	-10	26	7
LBPO	170	352	172	29	47	70	63	6	-1	-90	330	-68	70	79	235	7
LBPP	198	448	272	14	-18	74	48	33	-35	-100	302	-81	60	-24	241	7
LBPQ	183	295	125	33	21	75	9	35	-52	-109	195	18	-1	-208	245	7
LAGD	458	-11	477	-435	-573	-154	400	75	-170	27	432	112	-298	-61	-349	8
LAGE	444	514	360	-274	-198	-282	183	108	-29	-6	514	147	-59	82	-230	8
LAGF	175	576	211	-327	8	-293	238	67	-73	18	597	135	49	-34	102	8
LBPO	220	420	207	34	58	82	80	6	-15	-85	367	-88	83	79	290	8
LBPP	258	536	322	22	-18	69	73	33	-55	-100	336	-93	75	-39	291	8
LBPQ	243	353	148	50	28	85	17	35	-73	-109	200	9	4	-266	295	8
LAGD	526	36	500	-585	-555	-217	481	103	-195	22	511	126	-305	-125	-380	9
LAGE	546	579	409	-322	-218	-334	206	134	-6	-31	518	153	-88	94	-242	9
LAGF	232	783	257	-380	23	-268	314	92	-25	-3	653	145	46	-40	183	9
LBPO	275	488	252	49	73	87	95	6	-27	-80	397	-110	96	99	340	9
LBPP	318	625	386	38	-10	64	98	35	-70	-100	363	-103	83	-34	326	9
LBPQ	303	411	186	72	38	95	25	42	-92	-109	198	0	9	-279	322	9

Note: In red, we report the losses of the long history methods vs the futures one, for all years, all lumber types and various warehouse sizes with initial quantity 1.

We then computed 26 recommendation tables and found the one that yielded the smallest yearly costs on average for all lumber types. We then compared that to our own method. The following table (See Table 7) presents the results of the best futures-based recommendation system against ours. It turns out that the best performance happens when we chose the non-lagged version.

TABLE 7
PERCENTAGE OF WINNINGS OF THE LONG-HISTORY METHOD VS FUTURES METHOD FOR VARIOUS LAGS

Lag	Percentage
None	67.22%
1 Week	69.63%
2 Weeks	74.26%
3 Weeks	74.81%
4 Weeks	72.81%
5 Weeks	74.81%
6 Weeks	75%
7 Weeks	79.26%
8 Weeks	83.7%
9 Weeks	84.44%

As one can see our long-history strategy is better than the futures recommendation in the majority of cases (67 percent). Again, in 2008 and 2009 our method is clearly outperformed by the futures one.

DISCUSSION

Our analysis revealed some interesting results for the period of the recent recession of 2006 to 2010.[2] We found that during this time both lumber prices and housing starts had been dropping since 2006. Lumber traders recognized the phenomena in late 2005 as volumes of lumber futures reached levels not seen since 1985, which not surprisingly coincided with a large drop in housing starts-although lumber prices generally moved in a positive direction. As housing starts, lumber prices, and futures prices collapsed, the volume of futures traded continued to rise, with abnormally high volumes until the trough in housing starts and lumber prices subsided in April of 2009. Again, futures prices seemed to indicate a recovery as their numbers began to improve for the first time in January of that year.

Table 8 presents the performance of the methods examined for the six different lumber types and warehouse sizes four through nine (for a total of 36 cases) for the years 2007-2009.

TABLE 8
METHOD COMPARISONS FOR YEARS 2007-2009

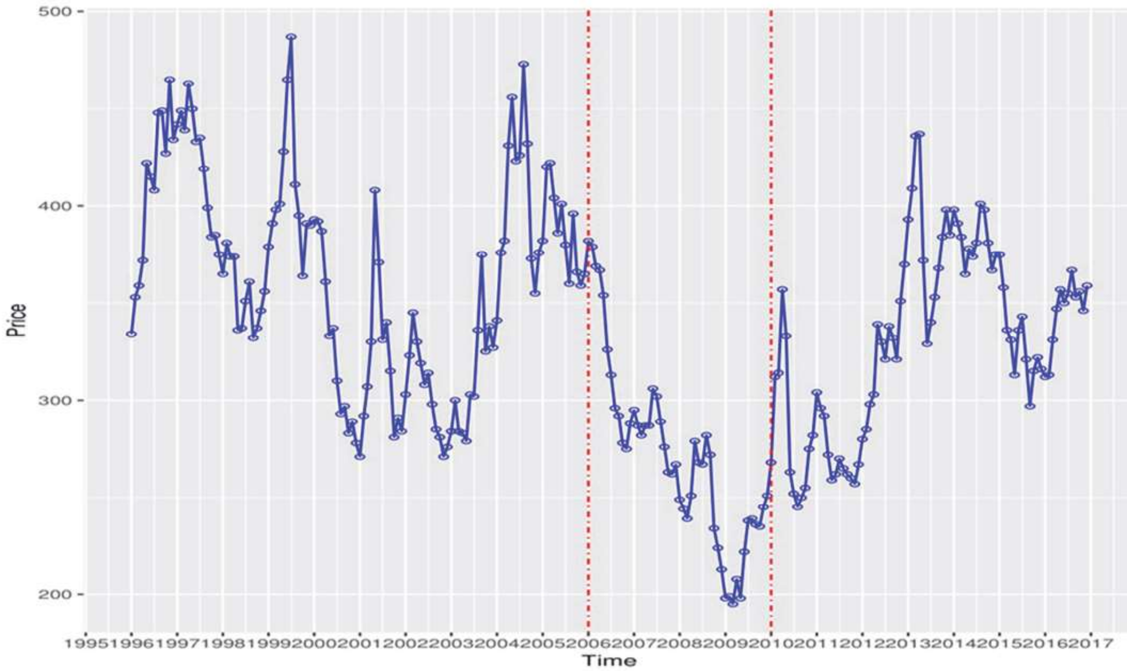
Methods	2007	2008	2009
Long-history	27/36	3/36	5/36
Naïve	0/36	0/36	0/36
Random	9/36	1/36	31/36
Futures	0/36	32/36	0/36
Short-term	0/36	0/36	0/36

It is clear that in 2007, traders believed that markets would rally and they did slightly, as such the “normal” price histories still do a good job of predicting behavior, however the “random” model performs second best as credit markets begin to short circuit. By the start of 2008, we are in the designated time of the credit crisis and lumber traders continue to hedge in large volume futures trading. At this point (January of 2008) housing starts are at their lowest since 1991. While lumber prices rally slightly in the middle of the year, housing starts continue to slide. This can be attributed to the steepest lumber production decline on record (Random Lengths). It is clear that traders are now driving the markets in futures trading as the futures model dominates the other two. The randomness has been driven from the model, largely everyone's belief is that the market is contracting and will continue to do so.

Our price histories no longer perform as well, as such a cataclysmic event is not included in their history. Lastly, and probably the most strangely is the 2009 outcomes, where the “random” model performs more strongly than either the “long-history" model or the “futures” model. Interestingly, this is right when the housing market starts to make its recovery, although the recovery is muted as another panic hits in late summer of 2009 as housing starts start to decline in August and continue down until October, before rallying again.

This U-shaped market shown in Figure 13 seems to cause problems for both the “futures” model as well as the “long-history” model. Here it is probable that all methods based on historic price series fail since the market is in full disarray and thus the fact that the “random” model’s recommendations give the best results for this year should not come as a surprise. It is important here to note that the “naïve” approach loses all the time to some other model. So even during the years of crisis employing some of the other models will lead to better results than employing no model at all.

FIGURE 13
SPOT LUMBER PRICES FOR THE YEARS 1995-2017



CONCLUSIONS-FUTURE WORK

Our long-history data predicting model, although simple in its nature, manages to capture effectively the changes in the price of these six lumber types on a yearly basis. Without the use of other external information, the purchasing strategies produced almost always exceed the "naïve" and "random" approach. Once again, in a future endeavor associations between the various lumber prices will be analyzed, since according to common practice, these types of lumber are used in conjunction with each other for various constructions and applications.

In this version of our algorithm, we are not concerned about the magnitude of the price change but only for its direction. In a later publication, the magnitude of price changes will be used to get a better understanding of the phenomenon. Also, during our analysis, we discovered inconsistencies in the creation and dissemination of the "futures" time-series. Thus in a future effort we will try to identify the roots of some of the problems we spotted in the creation of a continuous futures times series, and utilize some general statistics-driven ideas to create a more robust version of it.

Furthermore, reflecting on the failure of our method to predict well during the times of crisis, we will be exploring the creation of a switching mechanism to other models, utilizing shorter histories or even randomness, when various market indices cross specific critical thresholds. Exclusion criteria for various turbulent years will be implemented, since the time series information from those years could affect negatively the predicting capabilities of our algorithm when the market resumes its natural cycles.

Finally, we note that due to the simple nature of the information needed, the model and its extensions can be employed to forecast prices series for other forest products and agricultural commodities. Our codes where written mostly in the statistical language R and some of the time series analysis utilized the package "forecast" (Hyndman and Khandakar, 2008). The codes can be made available after a communication with the authors.

ENDNOTES

1. Information can be found at http://www.cmegroup.com/trading/agricultural/lumber-and-pulp/random-length-lumber_quotes_volume_voi.html. Trading volume was at 1,244 as of this writing.
2. Officially, the subprime mortgage crisis occurred from December of 2007 and lasted until June of 2009. Housing starts peaked in January 2006, at 2.273 million seasonally adjusted starts. The lumber futures composite reached its low in January of 2009 and housing starts reached their bottom approximately four months later, which interestingly enough is the time lag of two futures contracts. Perhaps futures traders saw evidence of the recovery in their business.

REFERENCES

Admmer, P., Bohl, M. T., & Gross, C. (2016). Price discovery in thinly traded futures markets: How thin is too thin? *Journal of Futures Markets*, 36(9), 851-869.

Buongiorno, J., Huang, F.M., & Spelter, H. (1984). Forecasting the price of lumber and plywood: econometric model versus futures markets. *Forest Products Journal*. 34(7), 13-18.

Deneckere, R., Buongiorno, J., & Il Bark, S. (1986). Optimal hedging in lumber futures markets. *Forest Science*, 32(3), 634-642.

Fama, E. F., & French, K. R. (2016). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. In: *The World Scientific Handbook of Futures Markets. World Scientific*, 79-102.

French, K. R. (1986). Detecting spot price forecasts in futures prices. *Journal of Business*, 59(2), S39-S54.

Hasan, S., & Hoffman-MacDonald, J. (2012). Price convergence in the lumber futures market. *Journal of Global Business Management*, 8(2), 126-133.

He, D., & Holt, M. (2004). Efficiency of forest commodity futures markets. In: Meetings of the American Agricultural Economics Association Selected Paper.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3), 1-22.

Jung, C. & Doroodian, K. (1994). The law of one price for us softwood lumber: a multivariate cointegration test. *Forest Science*, 40(4), 595-600.

Karali, B. (2011). What drives daily volatility in lumber futures markets? *Forest Science*, 57(5), 379-392.

Kingslien, H. K. (1975). *A decision framework for trading lumber futures.* Corvallis: Oregon State University, School of Business.

Lutz, J. (2012). There are no futures in timber. *Forest Research Notes,* 9(4), 4th Quarter.

Manfredo, M. R., & Sanders, D. R. (2008). Price discovery in a private cash forward market for lumber. *Journal of Forest Economics*, 14(1), 73-89.

Mehrotra, S. N., & Carter, D. R. (2017). Forecasting performance of lumber futures prices. *Economics Research International*, (Vol. 2017) (Article ID 1650363), 8 pages, doi:10.1155/2017/1650363.

Niquidet, K., & Sun, L. (2012). Do forest products prices display long memory? *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 60(2), 239-261.

Oliveira, R. A., Buongiorno, J., & Kmiotek, A. M. (1977). Time series forecasting models of lumber cash, futures, and basis prices. *Forest Science*, 23(2), 268-280.

Parajuli, R., & Zhang, D. (2016). Price linkages between spot and futures markets for softwood lumber. *Forest Science*, 62(5), 1-8.

Pindyck, R. S. & Rotemberg, J. J. (1988). *The excess co-movement of commodity prices*.

Shahi, C., Kant, S., & Yang, F. (2006). The law of one price in the North American softwood lumber markets. *Forest Science*, 52(4), 353-366.

Shahi, C. K., & Kant, S. (2009). Cointegrating relationship and the degree of market integration among the North American softwood lumber product markets. *Canadian journal of forest research*, 39(11), 2129-2137.

Shook, S.R., Plesha, N., & Nalle, D. J. (2009). Does cointegration of prices of North American softwood lumber species imply nearly perfectly substitutable products? *Canadian Journal of Forest Research*, 39(3), 553-565.
Sun, C. & Ning, Z. (2014). Timber restrictions, financial crisis, and price transmission in North American softwood lumber markets. *Land Economics*, 90(2), 306-323.
Uri, N. D., & Boyd, R. (1990). Considerations on modeling the market for softwood lumber in the United States. *Forest Science*, 36(3), 680-692.
Yin, R., & Baek, J. (2005). Is there a single national lumber market in the United States? *Forest Science*, 51(2), 155-164.

APPENDIX

DICKEY-FULLER TESTS

The Dickey-Fuller test assumes as a null hypothesis that the differences between consecutive elements of the time series do not depend on the previous values, but are just random errors. We say then that a unit root is present in our autoregressive model. We tested our time series for each lumber type on the whole history of our dataset and on the last five years using the specific Dickey-Fuller t-distribution (computations in R). Table A.1 presents the test results for the whole history and Table A.2 for the last five years.

TABLE A.1
DICKEY-FULLER TEST FOR EACH LUMBER PRICE (1995-2014)

	LAGD	LAGE	LAGF	LBPO	LBPP	LBPQ
p-value	0.007	0.011	0.01	0.01	0.01	0.014

TABLE A.2

DICKEY-FULLER TEST FOR EACH LUMBER PRICE (2010-2014)

	LAGD	LAGE	LAGF	LBPO	LBPP	LBPQ
p-value	0.341	0.353	0.421	0.133	0.038	0.142

According to table A.1 we should reject the null hypothesis, i.e. that the time-series has a unit root, for all types if we view it throughout the years. On the other hand, if we focus only on the last five years we have a completely different picture, namely according to table A.2 we fail to reject the null hypothesis for all types except LBPP.