# Bordeaux 2016 Abstract Submission

## Title
FORECASTING BORDEAUX AOC WINE PRICES USING STATE SPACE METHODS

## I want to submit an abstract for:
Conference Presentation

## Corresponding Author
Stephen Bazen

## E-Mail
stephen.BAZEN@univ-amu.fr

## Affiliation
Aix-Marseille University (AMSE), CNRS and EHESS ; Bordeaux Wine Economics

## Co-Author/s

<table>
<thead>
<tr>
<th>Name</th>
<th>E-Mail</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jean-Marie Cardebat</td>
<td><a href="mailto:jean-marie.cardebat@u-bordeaux.fr">jean-marie.cardebat@u-bordeaux.fr</a></td>
<td>LAREFI, Université de Bordeaux and INSEEC ; Bordeaux Wine Economics</td>
</tr>
</tbody>
</table>

## Keywords
Wine prices, state space methods, forecasting

## Research Question
How relevant is the state space method to forecast Bordeaux wine prices?

## Methods
Econometrics, state space methods

## Results
The state space methods are really efficient for short term (week and month) predictions.

## Abstract
Bordeaux wine prices have risen substantially over the last ten years. While most attention has been directed at the top end of the market, the price of generic AOC wine has also increased since the mid 2000s, following a period of decline and stagnation which began in 1997. Demand has expanded as a consequence of overseas demand from the United States and China and the price is now higher that in 2000. The current paper is concerned with forecasting the price of generic Bordeaux wine using state space methods with weekly and monthly data.

The functioning of this particular market is characterised by the risk aversion of producers who hold limited stocks and decide on how much wine to put on the market while seeking to avoid being caught short as a result a bad harvest in the following year. The market functions as follows. A buyer, say a British supermarket, wishes to purchase a quantity of claret to be sold under its own label and approaches a merchant. The offer is opened...
up to producers who will sell some of their existing stock depending on the price and the producer’s need for cash revenue. However it will also depend on the extent of remaining stock, since the producer will always desire to hold some wine back in case the next harvest is not a good one.

The next harvest will depend on climatic factors which affect the survival of the grape on the vine once the latter flowers. Through the spring frost may kill the grape flower and over the following months hail storms may damage the grapes. General precipitation and lack thereof at particular points of time through the cultivation period will also be relevant factors, and certain forms of disease may occur. Taken together the decision to sell part of existing stocks will be more likely to occur once the producer has less uncertainty concerning the size of forthcoming harvest. Thus there will be more sales in volume terms after the period of risk of frost is over, and then again when the nature of the summer has been ascertained. There is also the matter of freeing up space for stocking the new harvest.

There are a number of approaches available in order forecast the future path of a variable using time series data. A recent paper used machine-learning methods to predict returns on fine wines (Yeo et al, 2015). Here we are concerned by the price rather than the return and the series that we use are weekly and monthly average prices per barrel for the period late 1999 to late 2015. The data were provided by the Bordeaux wine trade professional organisation.

II THE SERIES USED

The data cover the period week 32 of 1999 to week 43 of 2015. The weekly price series (Figure 1) show a trend decline between 2000 and 2005, then rises slightly before falling back to it 2005 in 2010. Thereafter the wine price of rises sharply almost exponentially to a record price in early 2014 before falling back slightly over the remainder of 2014 and early 2015. In the weekly series there is a blatant outlier in week 28 of 2015. The data provider confirms that this was indeed the average price for that week. The monthly series “irons out” this outlier (Figure 2) and provides a clearer picture of the trend in prices. It is interesting that the volumes of sales do not vary is the same way (see Figure 3) and on their own cannot explain why the price increases dramatically after 2010. The role of the grape harvest will be taken into account since there are at least two disappointing harvest (in 2008 and 2013, see Figure 4).

These series are analysed using a state space model which is estimated for the purposes of predicting wine prices. This kind of model considers that a series can be decomposed into a trend, cycle, seasonal and irregular components. The different components can be replaced by artificial scales (linear trends and dummy variables) associated to parameters to be estimated, resulting in a deterministic representation of these components. In the state space approach, see Harvey (1989), and Commandeur and Koopman (2007), these components are treated as unobservable and stochastic. The deterministic formulation is in fact a special case in these models. From a forecasting point of view, the flexibility of the state space approach allows previous forecast errors to be rapidly taken into account in the modelling of the different components so that future forecasts will be more accurate.

III UNIVARIATE MODELS
In this first approach, a univariate approach is adopted using the price series in weekly and monthly frequencies. Such an approach is both a useful benchmark for evaluating richer models and also has the advantage of requiring no more data than on the series to be predicted. The estimates of the variances of the different stochastic terms and the predictions presented here were obtained using the STAMP Version 8 software package (Koopman et al, 2007).

(i) A univariate model for the weekly series

Weekly data are more numerous, and therefore provide for more information than series with a lower data frequency over a given period. On the other hand, they are likely to be noisier than say monthly or quarterly data which are usually averages or sums of the weekly series. One step forecasts of a weekly series are also likely to be less accurate on an observation-by-observation basis, although in a state space framework, the operation of the Kalman filter will usually correct the one-step forecast errors quite quickly.

The model is specified as above, with lagged price terms and no cyclical component. A dummy variable is included for the period running from week 40 of 2013 to week 39 of 2014 to capture the effect of the poor 2013 harvest. The model is estimated for the period 1999 week 32 to 2015 week 43 in order to identify the different components of the price series. Then the model is estimated up to 2014 week 44, and the remaining observations used to evaluate the predictive performance of the model for this 52 week period. The estimates of the variances of the different stochastic components (or hyper parameters) suggest that while the trend is stochastic, its slope and the seasonal component are both deterministic as their variances are close to zero (see Figure 5). The latter means that other things equal wine price movements in particular weeks are the same from one year to the next. For example, according to the estimated seasonal component the price rises in late February and early October, independently of other factors.

The predictions (one step ahead) over the post-sample period track the actual series reasonably well (Figure 6). The CUSUM plot is close to the horizontal axis. However there are four occasions on which the actual price lies above the upper bound of the 90% confidence interval, one of which is the outlier mentioned above. However in each case the prediction error reassuringly reverts to zero after these sizeable under-estimates. This is one of the major advantages of using state space methods for forecasting purposes.

(ii) A univariate model for the monthly series

The monthly average price irons out some of the noisy fluctuations observed in the weekly series. The maximum monthly average price occurs in April 2014. The model is first estimated as before with lagged price terms, a dummy variable for the 2013 harvest and trend, slope and seasonal components plus an irregular term (again there is no role for a cyclical component). While each component initially treated as being potentially stochastic, the slope component has a zero variance and is treated as deterministic. The stochastic seasonal component indicates that the seasonal variation of prices is lower after 2008 (Figure 7). The trend is stochastic and tracks the underlying evolution of the price series, while the stochastic seasonal accounts for much of the variation around this smooth trend. This model is used to predict the price of wine over the twelve month period from October 2014 - see Figure 8. The price predictions are more accurate than those obtained for the weekly series and the actual price is outside the 90% confidence interval on only one occasion. The CUSUM plot remains close to the horizontal axis and reverts to zero at the end of the period covered.

These two univariate models have slightly different stochastic properties regarding seasonal variation. However, both have a satisfactory out-of-sample predictive performance in statistical terms, in the sense that the prediction errors are generally inside the 90% confidence interval and the CUSUM plot reverts to the horizontal
axis. On the occasions when the predicted prices significantly diverge from the actual ones they quickly come back onto track. This underlines the efficiency of state space methods for forecasting purposes.

Figure 7 Monthly price series – model components

Figure 8 Predicted monthly price and CUSUM plot

IV A BIVARIATE MODEL

From a forecasting point of view, univariate models have the advantage of standing alone. However additional variables provide supplementary information, and can improve the precision of predictions obtained. The volume of wine sold in each period is now included in the model, and predictions are made as before for the final twelve months of the data run. In this bivariate framework the model is specified as a vector autoregression (VAR), in which each series has its own unobserved trend, seasonal, cyclical and irregular components. The components for the two series could be correlated. The full specification is a generalisation of the univariate framework, in which variance-covariance matrices replace the scalar component variances of the univariate model (see Harvey, 1989, for more details).

We estimate a model of monthly prices and sales volumes and then proceed to make out of sample predictions of the price of wine for the same period as before. The model components are quite different for the two series. As before, the stochastic trend for the price is evolves in a manner similar to the actual series, whereas for the volume of sales, the trend component follows a smooth downward trend (Figure 9). The latter series is dominated by a stochastic seasonal component (Figure 10), which indicates that seasonal variation is more pronounced toward the end of the period (after 2008). The seasonal variation in prices moves in the opposite direction, and is much smaller toward the end of the period.

The out of sample predictions for prices are similar but slightly better than those obtained from the univariate model (Figures 11 and 12). There is still one case where the actual price is outside the 90% confidence interval. The predictions of the bivariate model are slightly more accurate for the first half of 2015.