Modeling crops returns to secure the food supply chain

ECO511 Applied econometrics project

Côme Cheritel, Théo Guichaoua,
Raphaël Lafrogne-Joussier, Vincent Rollet

Abstract

Large retail firms face risks due to unexpected variation of commodity prices. Pierre Picard and Alexis Louaas argue in one of their working paper that the financial hedging strategy used by some retail firms (mainly consisting of using futures contracts) can be dominated by another hedging strategy: hedging by the use of catastrophe bonds and commodity bundling. The aim of our project is to provide empirical results to complement the theoretical results of Picard and Louaas. We use time series data on commodities traded in the London stock market between 1988 and 2018, and ARMA, VAR, and GARCH methods to forecast the prices of five crops. We then use these forecasts to test different insurance strategies, using cat-bonds.

Contents

1 Introduction

1.1 Context of our study

1.1.1 Securitizing the supply-chain

Supply-chain management has become an increasingly important activity for large firms in a globalized world. Some firms like supermarkets and other retail firms have a business model heavily relying on an efficient procurement and transport of goods. When there is a negative supply shock for a retail firm (i.e. when a supplier increases its commodity prices), the retail firm faces an unexpected cost, as it either has to remain with its current set of suppliers and endure the increased costs without always having the possibility to increase prices down the line, or change its set of suppliers. In the latter case, the retail firm also endures increased costs as establishing a procurement relationship with a supplier is costly, and reorganizing the supply-chain may also come along with significant costs.

1.1.2 Risk hedging

Operational and financial hedging There are several ways a firm can reduce the risks it is exposed to, but we can divide the possible approaches in two distinct categoriesm according to Treanor and al. ([10]): financial hedging and operational hedging. Operational hedging consists in diversification of risks through the firm operational structure. For example, a firm might set factories in various countries to protect itself from various risks (natural catastrophes, political instability, new regulation, etc.) Financial hedging consists in buying financial instruments that limit risks (futures, call and put options).

Retail firms often rely on futures markets to secure their future prices. In the working paper here at stake ([7]), Alexis Louaas and Pierre Picard argue that this strategy is dominated by a use of catastrophe bonds (henceforth cat-bonds).

Insight on cat-bonds Insurance companies can limit their risk through reinsurance. However, in the mid-1990s, another financial instrument was created to limit the risk of insurance companies: cat-bonds. These bonds are characterized by a maturity, a principal, a coupon and a trigger. The trigger describes a risky event, such as a hurricane, an earthquake or a wildfire with precise characteristics (wind speed for a hurricane, sum lost by the company due to the catastrophe...) and a recognized agency determines whether the trigger event has occurred or not.

When a cat-bond is issued, an investor invests the principal to buy it. If the bond is triggered before maturity, the principal is transferred to the issuer who can use the proceeds to pay their claim-holders. However, if the bond is not triggered, the principal is returned to the investor with a coupon. The issuing of cat-bonds by insurers can therefore be used as a complement or substitute to reinsurance.

Cat-bonds attract investors as their returns are generally uncorrelated with other market outcomes, and therefore hedge-funds and other financial actors can use them as a source of diversification. They can also be useful for firms as financial hedging using futures is often imperfect. This is because futures contracts do not exist for all commodities, and the firms specific constraints are not reflected in the future market prices. For this reason, tailored cat-bonds may help to limit the base risk[1] of a firm more than futures contracts.

[1] The base risk of a firm is the risk that it is exposed to due to imperfect hedging.

1.1.3 Optimal hedging with risk pooling and cat-bonds

Risk pooling It is quite easy to show that the price risk of a bundle of goods is lower than the sum of the price risks of all goods of the bundle , as those prices are not perfectly correlated[2]. The less risk averse the insurer, and the less correlated the risk lines are, the less the cost of bundled insurance is compared to unbundled insurance. As firms that acquire many goods can reduce their risk in-house before transferring it to the market, they can use this property to reduce their insurance costs. Using data from the London stock market, Louaas and Picard find that the cost of procurement risk is lowered by 5 to 14% annually by using their proposed bundling strategy rather than the classic line-by-line hedging strategy (using futures contracts).

Using cat-bonds to reduce risks Even though bundled insurance is more efficient than line-by-line insurance, implementing this strategy is not direct: one usually cannot buy bundled futures contracts. A retail firm could buy a set of futures contracts from different investors, but could not achieve risk reduction through this manner. Cat-bonds, however, can be used to bundle risks. Two types of cat-bonds may be issued: cat-bonds with indemnity triggers and cat-bonds with parametric triggers.

- Indemnity triggers are defined by a level of losses incurred by the issuer of the bond (in our case, the retail firm). Firms can issue a cat-bond to secure the price of a bundle of goods instead of individual goods: this is why cat-bonds permit the bundling strategy that the future markets cannot provide. Indemnity triggers are attractive for issuers because they leave them with no basis risk. However, they are less attractive for investors because of information asymmetry: the issuer exactly knows what losses she has endured, whereas the investor cannot always verify the issuer's claims. In case of a dispute, the investor might have to wait long periods for a settlement to be reached and part of his investment to be reimbursed. Furthermore, the fact that the issuer has no basis risk leaves the possibility of moral hazard. In our case, the retail company may buy more goods from areas exposed to high risk: these goods would be cheaper but expose the investor to higher losses. This higher risk level for the investor is associated with a risk premium and higher prices.
- Parametric triggers are defined by the occurrence of well-defined events (for example, a wind speed higher than 120 km/hr in a specified location or an earthquake higher than 7.5 on the Richter scale). These parametric triggers can be bundled together (for example, an issuer might define a bundle of events that lead to the loss of the principal) which enable in-house pooling of risks. However, these triggers leave a certain basis risk to the issuer of the bond: it is for example difficult to estimate precisely the impact of a hurricane in Malaysia on rice prices. Contrary to indemnity triggers, they do not expose the investor to information asymmetry or moral hazard. Cat-bonds with parametric triggers are therefore less costly than cat-bonds with indemnity triggers.

Retail firms can use cat-bonds to bundle their risks in-house before transferring them to markets. However, they still have to solve the trade-off between base risk and cat-bond issuing prices. A low base risk involves higher bond prices, and vice-versa.

[2] The proof of this result can be found in [7]

1.1.4 Forecasting prices to design optimal bundles

We have now defined the broad scope of our study. Our goal is to tackle one of the numerous issues brought forth by Louaas and Picard: the issue of bundling. As we have mentioned, a firm can reduce risks by bundling different products in a hedging strategy, and this reduction is made possible because prices of different goods are not perfectly correlated. It is therefore evident that optimal bundling heavily relies on the correlations between the prices of different goods: the main goal of our project is to determine the correlations between the prices of five crops that we would like to bundle together. Just finding the variance-covariance matrices of the prices of these five crops is insufficient. In fact, part of the evolution of these prices can be forecasted, and the risk that retail firms have to face comes precisely from the unforecastable evolution of prices. For this reason, our project's aim is to provide a forecasting model for the prices of these five crops, and to determine the correlations between the unforecastable variations of these prices.

1.2 A first analysis of the data

1.2.1 The data

We will work on five series of crop prices : maize, wheat, rice, barley and soy, traded on the London stock market between 1988 and 2018. Figure 1 plots the prices over time and Table 1 summarizes the time averages and standard deviations of prices. All prices are expressed in GBP per ton. Rice is the most expensive crop yet the most volatile, followed by wheat, corn, sorghum and rice. Table 2 displays the correlation matrix of the time series. All crops have pairwise positive correlation; corn and sorghum prices are the most correlated lines. This high correlation invites us to investigate the determinants of the global prices trends over the past decades.

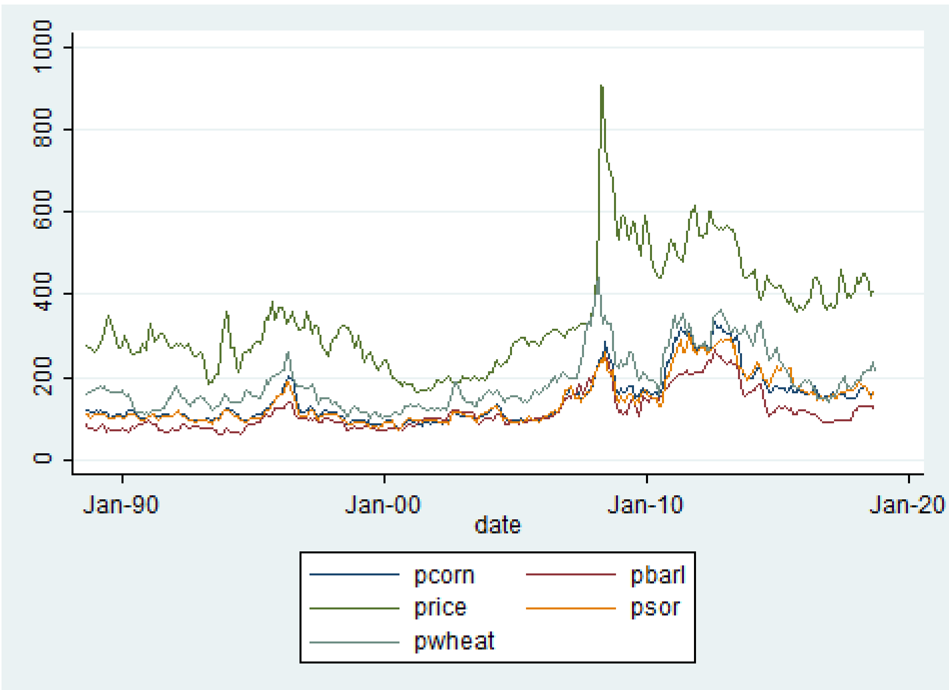

Figure 1: Crop prices 1988-2018

	corn	rice	wheat	sorghum	barley
μ	145.69	345.35	189.67	141.87	117.48
σ	60.06	128.29	68.70	55.82	47.57

Table 1: Input prices mean and std dev. Sep 1988-Sep 2018

	corn	rice	wheat	sorghum	barley
corn	1.0000				
rice	0.8331	1.0000			
wheat	0.8941	0.7779	1.0000		
sorghum	0.9699	0.8107	0.8889	1.0000	
barley	0.9191	0.7652	0.8991	0.8909	1.0000

Table 2: Correlation matrix of gross time series

1.2.2 From 1988 to 2000

Wheat prices varied little from 1988 until January of 1995, at the time when prices began to increase dramatically to reach a all-time high in 1996. The Bureau of Labor Statistics of the United States ([1]) provides some explanation about this global increase throughout the world. Extremely low total ozone over Arctic in 1995 led to the winter 1995-1996 being of the coldest winter of all time. As a consequence, winter wheat crop had a hard time growing, leading to a sharp increase in prices. Concerning corn, it appears that production in the United States, the largest corn producer in the world, went down from 10.1 million bushels in 1994-1995 to 7.4 million in 1995-1996, due to a severe drought in the Midwest and West. Jointly to this natural decrease, a pessimistic forecast of demand and a rather inflexible foreign demand led to an additional rise in prices, demand being much more important than expected.

1.2.3 From 2000 to 2008

The period between the beginning of the 2000s until 2008 is characterized by a constant rise in all food crops prices. For example, one ton of corn was sold $83.31 in June 2001, while seven years later, in June 2008, the same amount of corn was sold $287.81. After that, prices dramatically decreased: in two years, corn lost 46% of its value ($152.75 in June 2010).

Crop	**Rise between 2000 and 2008's peak**	**Fall between 2008's peak and 2010**
Corn	245%	46%
Barley	247%	58%
Rice	454%	51%
Sorghum	256%	50%
Wheat	330%	64%

Table 3: Rise and drop of food crops prices in the 2000s. Main trend.

What could be the reasons of such a scheme ? Many institutions and states came with their own story to explain those trends. The first story to come in mind is the important growth of resource intensive food demand. This growth mainly comes from India and China, whose middle class grew dramatically in the 2000s. Both countries' population grew at high rates until the beginning of the 1990's for China and until the beginning of 2000's in India. Combined with higher life standards, this increase in population led to a much larger increase in resource intensive food demand [3]: generalized consummation of meat for example. One could also argue, following the International Food Policy Research Institute in 2011, that the major driver of the food crisis in 2008 was energy prices (including oil), similarly to the food crisis that occurred in 1972-1974. Because petroleum prices are high, and as natural gas is a substitute to petroleum, the price of fertilizers made up of natural gas increased. Price of transportation has also increased in the meantime. Aside of this, bio-fuel crops took a larger part in world crop production. Even if the effect of bio-fuels on food prices through the reallocation of production, does not have a large effect on food crops prices (according to the World Bank and the OECD), energy prices may have led to higher crops prices. Finally, an additional effect due to natural disasters reducing the harvest may also have pushed prices upward. A six years drought in Australia, the second-largest exporter of wheat, have decreased the production of wheat and rice by 98%, according to the New York Times [4]. A heat wave in California in 2006 and out-of-season rainfalls in India in 2008 as well contributed to a reduced supply of wheat and rice during those years.

Along with those explanations, specific crops have their own stories. Let's focus on wheat prices, on figure 2. Although food production grew faster than population growth, allegedly sufficient to secure food supply, wheat production during 2006 and 2007 was 4% lower than that in 2004 and 2005, according to the FAO. Rising demand in both India and Egypt helped to ramp up demand for American wheat during the bull market in August 2007. The price of wheat reached record highs after Kazakhstan began to limit supplies being sold overseas in early 2008, but had slowed down by late 2008. Concerning rice, its dramatic increase in price, out of the global crises, is the result of various export restrictions by India and Vietnam mainly. Because rice price was already rising, those countries secured their own demand for rice by banning partly rice exports, thus pushing foreign rice price upward, especially in London.

In December 2008, the global economic slowdown, decreasing oil prices, and speculation of decreased demand for commodities worldwide brought a natural sharp decreases in crops' prices.

1.2.4 Since 2008

Food prices again started to rise in 2010. Focusing on corn, as a result of a very dry summer in the United States and Europe, prices reached all-time highs in July 2012 and prices remained high throughout 2012.

One other reason for the increase in food prices may again be the increase in oil prices at the same time, leading to increased demand for bio-fuels. And the use of corn for ethanol fuel production rose from 15% of total U.S. maize production in 2006 to 40% in 2012, according to the Council of Foreign Relations of the United States ([8]). It is then tempting to link oil prices variations to corn prices variations. An alleged piece of evidence is provided by figures 3 and 4: corn and oil prices are undoubtedly correlated; they at least follow the same trend.

[3] In 2002,China's middle class was only 4% of its population, while a decade later, it was 31%, that is 420 million people, according to the CSIS China Power Project

[4] "A Drought in Australia, a Global Shortage of Rice", April 17, 2008

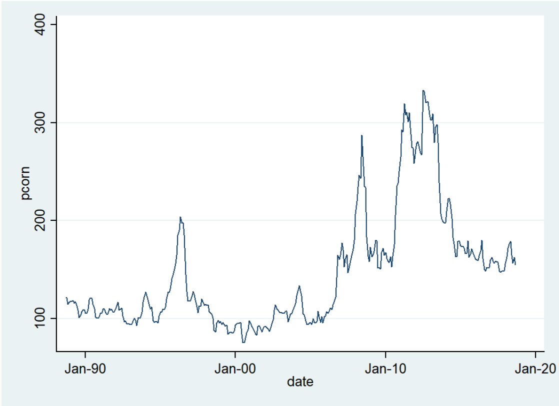

Figure 2: Wheat prices from 1988 to 2018

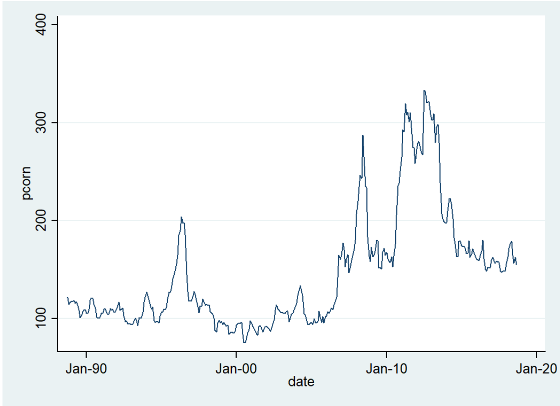

Figure 3: Corn prices from 1988 to 2018

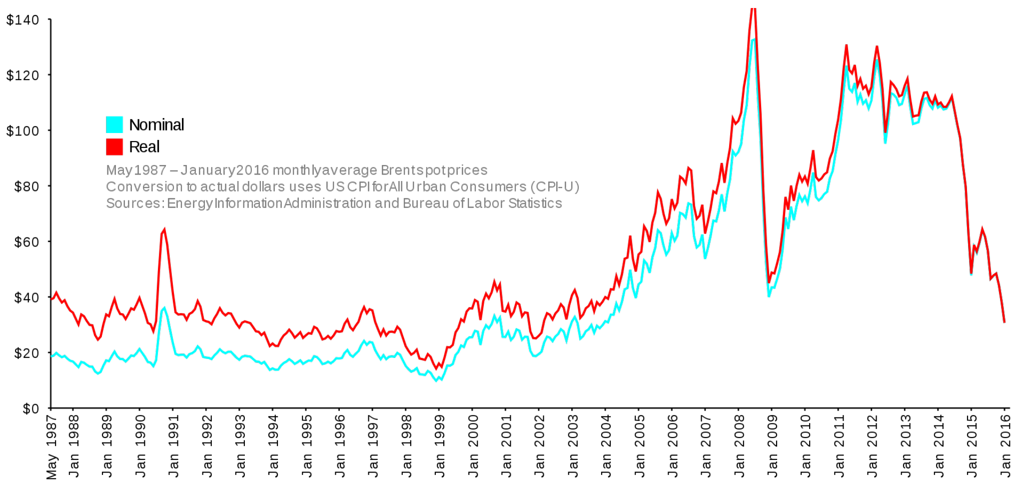

Figure 4: Monthly average oil spot prices from 1987 to 2018. Source: Wikipedia

1.3 Our approach

Being aware of the details mentioned in part 1.1, the core of our work will be to compare different forecasting models and provide the best possible forecasts of crop returns.

The first step of our study will consist in applying the simple methods of ARMA forecasting to our time-series. This will provide us baseline forecasts for each of our five crops. We start the next section by providing a few theoretical models that motivate our use of ARMA models. In an attempt to improve our first ARMA forecasts, we will then explore a more general method of forecasting, using all of the time series of our data set to predict the price evolution of each of our five crops. This technique is known as vector auto regression (VAR). Finally, as crop prices may be subject to time-varying volatility, we use a forecasting technique which is widely used in finance : the ARCH and GARCH models. Comparing the results we obtain with these three models (ARMA, VAR and GARCH), we select for each crop the model that provides the best forecasts. We will then finally be able to estimate the correlations between the unpredictable components of each time series.

At the end of the process, we will use our forecasts in a simple insurance model.

2 Models of forecasting

2.1 Before applying empirical methods: a theoretical approach and the cobweb model

2.1.1 Setting of the cobweb model

In order to start our analysis of price forecasting, we introduce a first model which can explain the evolution of prices in a market: the cobweb model, using the work of González ([5]). This model assumes classical demand and supply functions for consumers and suppliers of the economy, which are respectively decreasing and increasing with price. The equilibrium price is found when these two functions cross and the market is cleared. Let's assume the following linear demand and supply functions:

$$q^D = \alpha_1 - \beta_1 p$$

$$q^S = \alpha_2 + \beta_2 p$$

Where $\beta_1, \beta_2 > 0$ and $\alpha_1 > \alpha_2$ (implying that both curves cross at one point). The equilibrium price is found at market clearing:

$$p^* = \frac{\alpha_1 - \alpha_2}{\beta_1 + \beta_2}$$

However, this equilibrium price is not systematically reached, especially when there are shocks: participants in the market take time to adjust their behaviour to the new information contained in prices. We now assume that there are shocks on the supply side of the market, and no demand shocks. This assumption is realistic in the market for agricultural commodities in which the supply side is often subject to shocks (because of bad weather conditions or catastrophic events for example). On the contrary, the demand side of agricultural markets can be considered less volatile.

Within this framework, let's consider that producers make their production decisions for time t at time t-1 by seeing price evolution as a Markovian process: they expect that tomorrow's price will be today's price. Consumers buy the quantity produced at a price determined by the demand curve, clearing the market.

2.1.2 Effect of a single supply shock

Let's now study the effect of a single supply shock on this market. At time 0, both prices and quantities are at equilibrium value. At time 1, there is a supply shock that reduces the quantity that producers can actually supply: they supply a quantity q_1 which is below the equilibrium q^*. As the quantity available in the market is below q^*, the price at time 1 is $p_1 = p^* + \epsilon_1 > p^*$. As producers anticipate the price of time 2 to be p_1, they adjust their production to a value over q^*. The supply function yields:

$$q_2^S = \alpha_2 + \beta_2 p_1 = \alpha_2 + \beta_2 (p^* + \epsilon_1)$$

Consumers buy this produced quantity at a price p_2, given by the demand function:

$$p_2 = \frac{\alpha_1 - q_2^S}{\beta_1} = p^* - \frac{\beta_2}{\beta_1}\epsilon_1$$

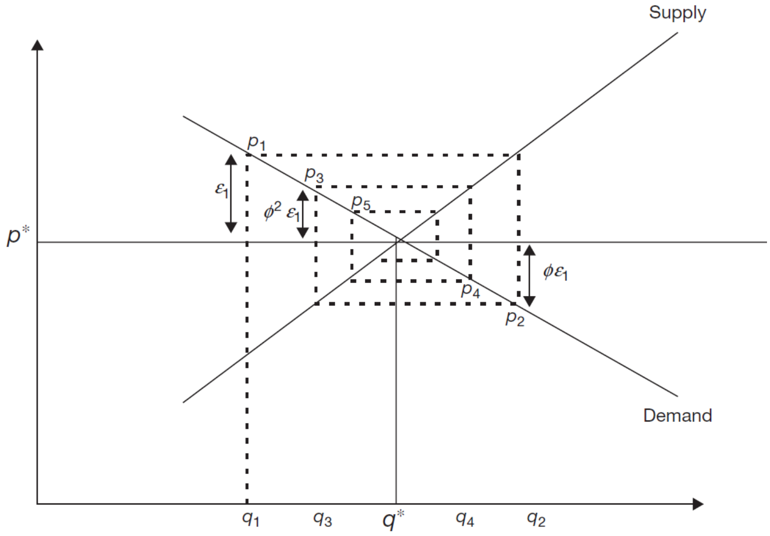

Figure 5: The cobweb model

At this point, the producers realize that they have overproduced at time 2, leading to a price p_2 under the equilibrium price. They therefore decide to produce less at time 3, and estimating that p_3 will be equal to p_2, they produce below the equilibrium quantity, at $q_3^S = \alpha_2 + \beta_2 p^* - \frac{\beta_2^2}{\beta_1}\epsilon_1 = q^* - \frac{\beta_2^2}{\beta_1}\epsilon_1$.
At time 3, consumers that can only buy quantity $q_3^S < q^*$ clear the market by buying the good at price:

$$p_3 = \frac{\alpha_1 - q_3^S}{\beta_1} = p^* + \left(\frac{\beta_2}{\beta_1}\right)^2 \epsilon_1$$

If we set $\phi = \frac{\beta_2}{\beta_1}$, we can show by induction that in this model, the initial shock persists in the economy long after it occurred:

$$p_t = p^* + \phi^{t-1}\epsilon_1$$

If $\phi < 1$, i.e. if the supply curve is steeper than the demand curve, then the shock gradually disappears over time. In the other case, it never disappears. However, any solution with $\phi > 1$ is implausible as it leads to infinitely high prices (and negative prices).

2.1.3 Effect of multiple shocks

In the plausible case, shocks thus fade away at rate ϕ over time. In a realistic model, there would also not be one single shock, but shocks at every time period and of different magnitudes. We consider a series of shocks (ϵ_t), and it is simple to prove that in the case of multiple shocks, the price at each period is:

$$p_t = p^* + \sum_{s=1}^{t} \phi^{t-s}\epsilon_s$$

(ϵ_t) is a white noise process, as shocks cannot be forecasted. This representation of current prices as a function of past shocks is a **moving average (MA)** representation. We can also note that:

$$p_{t+1} = p^*(1-\phi) + \phi p_t + \epsilon_t \Longleftrightarrow (p_t - p^*) = \phi(p_{t-1} - p^*) + \epsilon_t$$

This representation of future prices (or future deviations from equilibrium prices as a function of past prices (as well as an error term) is the **autoregressive (AR)** representation. In this simple case, the two representations are equivalent, and in the AR representation, the fact that only the one-time-lagged observation is used, we have a AR(1) process.
The cobweb model shows that exploiting AR, MA or ARMA models can be a good way to forecast prices in agricultural markets. The AR(1) model that the theory predicts here is obviously too simple: the assumption that producers see price evolution as a simple Markovian process and thus have a very myopic behaviour needs to be relaxed by using models of higher order.
In the following subsections, we analyze our data using ARMA models as hinted by the cobweb model.

2.2 Modeling price evolutions with an ARMA process

2.2.1 The ARMA framework

To show how we model price evolutions using an ARMA model, we first focus on one of our time series: the price of corn. As can be seen in Figure 6, corn prices have varied very significantly in thirty years. In this time series, we can observe that the process at hand is not covariance-stationary. The mean corn price in our sample is of \$146, but this mean has not been constant over time: the mean of corn prices between 1988 and 2000 is of \$111, while the mean price between 2001 and 2018 is of \$170. There is an upward trend in prices that makes the process non-covariance-stationary. Furthermore, the variance of the time series varies also over time: in some time periods, such as the period 2015-2018, prices do not vary much. On the contrary, there are periods in which prices vary a lot (for example between 2007 and 2013).

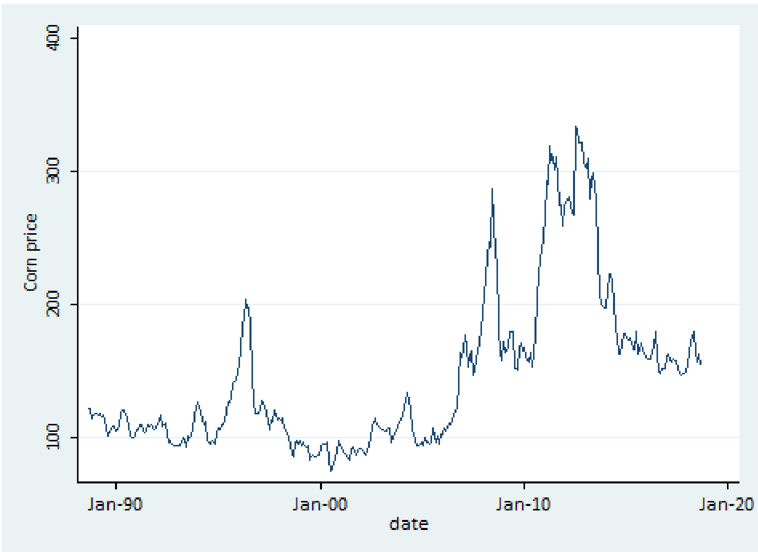

Figure 6: Evolution of corn prices

When modeling processes with ARMA models, covariance-stationarity is an important prerequisite as it is a necessary condition for the application of the Wold representation theorem, which states that covariance-stationary processes can be always modeled by a linear representation: $Y_t = V_t + \psi(L)\epsilon_t$ [5]. In ARMA models, we approximate the infinite polynomial $\psi(L)$ by

[5] where Y_t is the covariance-stationary process we want to model, V_t is a deterministic component, (ϵ_t) is a

a rational function of finite orders.
To proceed with our modeling, we compute the series of monthly returns (which can be obtained by taking the first difference of the logged time series). This new time series is displayed in Figure 7. Its mean value of 0.00227 is very close to zero.

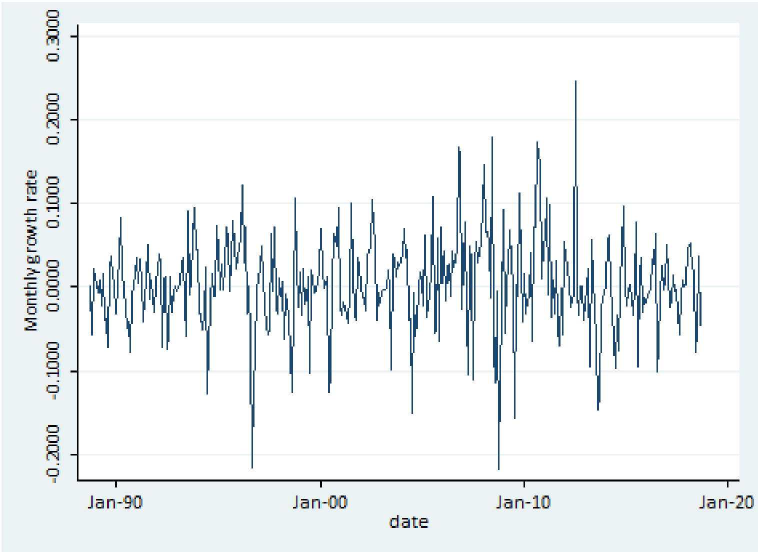

Figure 7: Evolution of monthly returns for corn (1988-2018)

This new time series resembles more to a covariance-stationary time series. In Figure 8, we observe the monthly returns for corn for years 2000 to 2002. This subsample exhibits substantial autocorrelation: if a return rate is negative a given month, it has a high probability to remain negative the next month (and vice-versa for positive returns).

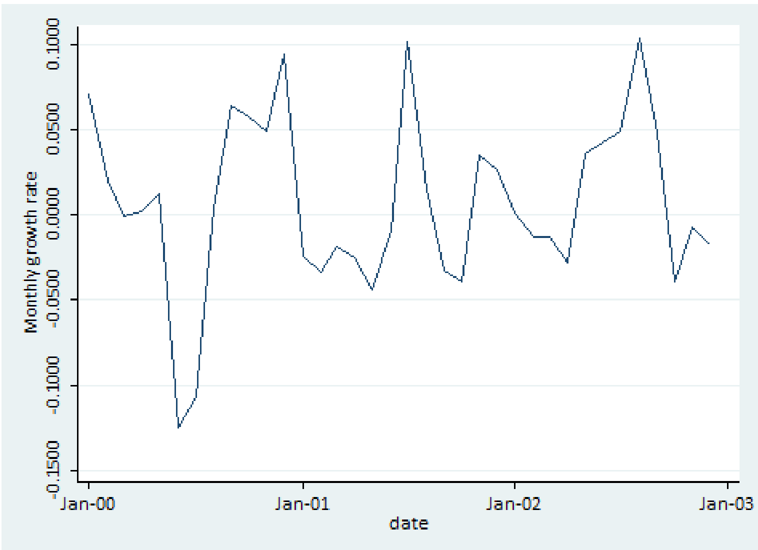

Figure 8: Evolution of monthly returns for corn (2000-2002)

To evaluate the autocorrelation of the time series, we compute the autocorrelation function (ACF) and partial autocorrelation function (PACF). They are visible in Figure 9. Performing

white noise process and $\psi(L)$ is a possibly infinite polynomial of the lag function L

the Ljung–Box Q-test enables us to reject the no-autocorrelation hypothesis for all reasonable confidence intervals.

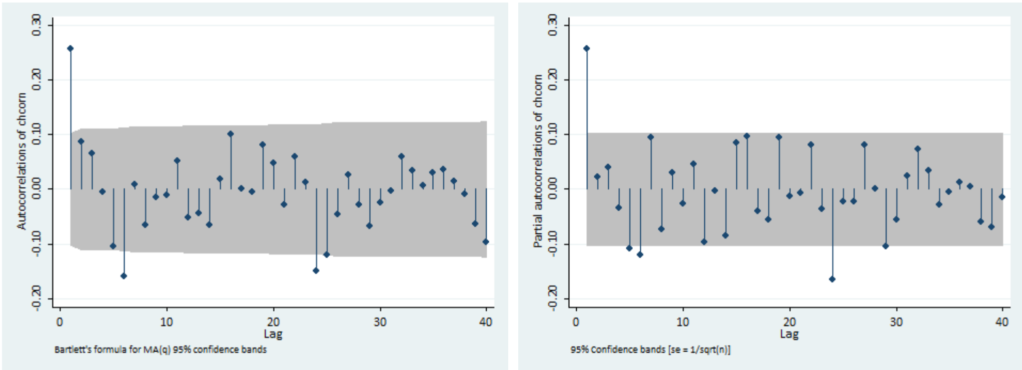

Figure 9: Autocorrelation and partial autocorrelation functions for the return rate of corn

Does the cobweb model still hold when using returns instead of prices ? We justified the use of an ARMA estimation by introducing the cobweb model. According to this very simple model, in a market subject to supply shocks and where suppliers form adaptative expectations, the series of prices can be modelled either by an AR(1) or an MA(∞) process. Seeing that the prices did not follow a covariance-stationary process, we have decided to focus on the evolution of returns. In this case, can we still justify the use of an ARMA model with the cobweb model? Unfortunately, it can be shown that the cobweb model does not predict that returns will follow an ARMA process (Appendix A.1). However, in an extension of the cobweb model that we developed (the store-and-sell model), returns do follow an ARMA process (Appendix A.2).

2.2.2 Model selection

We use classical model selection techniques to find the best-suited ARMA parameters for our five crops. The selection process is explained for one crop in Appendix B, and leads to the following results when replicated for all five crops :

Crop	**Selected model**
Corn	ARMA(2,2)
Barley	ARMA(2,1)
Rice	ARMA(1,2)
Sorghum	ARMA(2,5)
Wheat	AR(1)

Table 4: Selected ARMA models for all 5 crops

2.3 Modeling price evolutions with Vector Autoregression (VAR)

2.3.1 The VAR framework

When we look at the evolution of crop prices over time, it is clear that the times series are not independent. An example in which this is especially striking is the relation between prices of sorghum and corn. Figure 10 shows the almost perfect correlation between prices of sorghum and corn (R^2=0.94), and the strong correlation associated to their respective returns. This very high correlation may be due to the fact that these two crops are both used intensively in animal husbandry (70% of the total corn production in Europe is used for fodder). Breeders may buy on markets the crop that happens to be the cheapest, and it is the activity of arbitrageurs that might lead to such a high correlation.

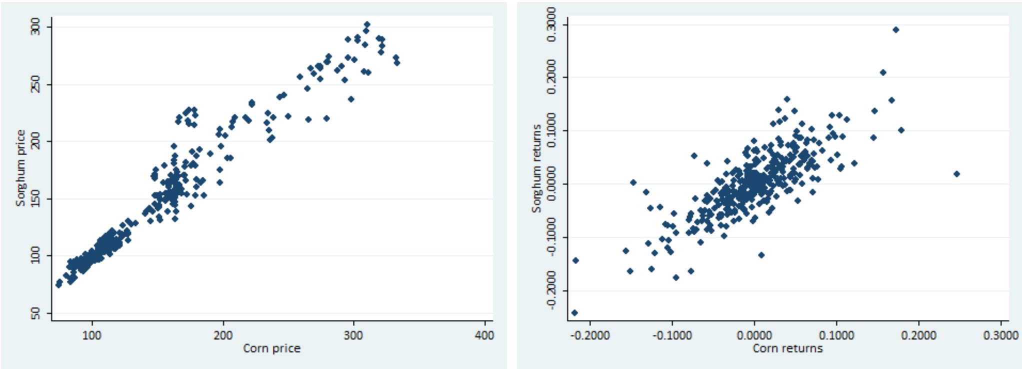

Figure 10: Scatter plots of the prices of sorghum and corn, and of the revenues associated with these two crops

Seeing the strong relationship between various crop prices leads us to study the following question: do changes in the price of one crop help us forecast the changes in prices of other crops?
In order to answer this question, we use another important method of forecasting theory: **vector autoregression models**.

2.3.2 Model selection

We perform a VAR on the five price series of our data set. In order to choose the best order for the VAR process, we compute VARs for orders 1, 2 and 3 and choose the model which yields the lowest information criteria (AIC and BIC)[6]. Performing these three regressions yield the following results :

Model	AIC	BIC
VAR(1)	**-5686.117**	**-5569.617**
VAR(2)	-5682.926	-5469.496
VAR(3)	-5656.545	-5346.326

Table 5: Selecting a VAR model with information criteria

Both the AIC and BIC suggest that we should use a VAR(1) to model the evolution of crop

[6]Akaike's Information Criterion and Bayesian Information Criterion

prices. Formally, this amounts to performing the following regressions:

$$\begin{cases} B_t = \Gamma_1 + \alpha_1 B_{t-1} + \beta_1 R_{t-1} + \gamma_1 W_{t-1} + \delta_1 S_{t-1} + \zeta_1 C_{t-1} + \epsilon_{1t} \\ R_t = \Gamma_2 + \alpha_2 B_{t-1} + \beta_2 R_{t-1} + \gamma_2 W_{t-1} + \delta_2 S_{t-1} + \zeta_2 C_{t-1} + \epsilon_{2t} \\ W_t = \Gamma_3 + \alpha_3 B_{t-1} + \beta_3 R_{t-1} + \gamma_3 W_{t-1} + \delta_3 S_{t-1} + \zeta_3 C_{t-1} + \epsilon_{3t} \\ S_t = \Gamma_4 + \alpha_4 B_{t-1} + \beta_4 R_{t-1} + \gamma_4 W_{t-1} + \delta_4 S_{t-1} + \zeta_4 C_{t-1} + \epsilon_{4t} \\ C_t = \Gamma_5 + \alpha_5 B_{t-1} + \beta_5 R_{t-1} + \gamma_5 W_{t-1} + \delta_5 S_{t-1} + \zeta_5 C_{t-1} + \epsilon_{5t} \end{cases}$$

Where B_t, R_t, W_t, S_t, C_t are the prices of our five crops, $(\Gamma_i)_{i\in\{1,2,3,4,5\}}$ are constants and $(\epsilon_i)_{i\in\{1,2,3,4,5\}}$ are error terms, assumed to have a normal distribution, and that are not required to be uncorrelated.

Using Granger-causality tests, we compare the VAR and ARMA approaches (the tests are described in Appendix C).

This leads us to select the models of table 6.

Crop	**Selected model**
Corn	VAR(1)
Barley	VAR(1)
Rice	ARMA(1,2)
Sorghum	ARMA(2,5)
Wheat	AR(1)

Table 6: Selected ARMA models for all 5 crops

2.4 Modeling price evolution with a GARCH process

2.4.1 The GARCH framework

The Wold decomposition, as mentioned in [3], makes clear that every covariance stationary series may be viewed as ultimately driven by an underlying white noise ϵ_t. In many cases, it does not matter whether ϵ_t is independent (that is, independent of its past), or serially uncorrelated. However, the Wold decomposition only requires serial uncorrelation of ϵ_t, and not independence. Thus, one could possibly imagine a process where ϵ_t is serially uncorrelated but dependent on its own past. That is what we do in a general autoregressive conditional heteroskedasticity (GARCH) model: residuals are driven by their past values through their variance. Hence,

$$\epsilon_t | \Omega_{t-1} \sim (0, \sigma_t)$$

where σ_t^2 is the conditional variance. In this class of models, the conditional variance varies as Ω_{t-1} evolves, which stress out the possibility and the seek for of time-varying volatility. Specifically, a pure $GARCH(p, q)$ process would be defined as

$$\epsilon_t | \Omega_{t-1} \sim \mathcal{N}(0, \sigma_t)$$

$$\sigma_t^2 = \omega + \alpha(L)\epsilon_t^2 + \beta(L)\sigma_t^2$$

where $\alpha(L)$ and $\beta(L)$ are lags polynomial of order p and q. When q is zero, we talk about ARCH process. Then, ARCH is to conditional variance dynamics what AR is to conditional mean dynamics, and GARCH is to ARCH what ARMA is to AR. A GARCH model is about a white noise series which variance is different at each period, and those changes are fully determined by the past values of the series (ARCH) and of the variance itself (GARCH). In our framework, as we pointed out that residuals ϵ_t after ARMAs were not variance stationary, we will try to fit those residuals by a GARCH model. Implementation of the GARCH methodology is explained in

Appendix D.1 focuses on the methodology used when searching for a GARCH fit of a time series.

2.4.2 Model selection

Using the GARCH parameter selection method as detailed in Appendix D.2, we select the following models :

Crop	Selected model
Corn	GARCH(1,2)
Barley	GARCH(1,2)
Rice	ARCH(2)
Sorghum	GARCH(1,2)
Wheat	GARCH(1,2)

Table 7: Selected GARCH models for all 5 crops

2.5 Finding the best forecast : comparison of ARMA, VAR and GARCH approaches

We have already made a comparison between the ARMA and VAR models for the five crops of our dataset. We can now compare the models we have selected previously with the new forecasts we have obtained with the ARCH framework. As previously, we will compare our different forecasts by comparing the sum of the squares of the forecast errors (SSE) for the time series. Our results can be found in Table 8.

Table 8: Comparing ARMA, VAR and GARCH forecasts

	ARMA	**VAR**	**GARCH**
Sorghum	**1.163**	1.186	1.165
Barley	1.137	**1.079**	1.111
Corn	1.044	**1.018**	1.046
Rice	**1.109**	1.125	1.139
Wheat	**1.319**	1.337	1.321

What we find in this table is not surprising : for none of the five crops of our dataset do GARCH models provide better forecasts than their ARMA and VAR counterparts. This is not surprising given the results of the Engle test performed (see Appendix D.2), which gave a weak evidence of ARCH effects.

The models that we finally select as best suited for the analysis of our five time series are therefore those of Table 6.

From the forecasts for returns that we obtain with the aforementioned methods, we can forecast the evolution of the prices of our five crops. Figure 11 shows the one-step ahead forecasts of prices for the period 2000-2002, as well as the realized prices: we chose to use those simple forecasts to challenge at first our model in a simple way. They make it clear that there already is a discrepancy between the one-step ahead forecast and the effective realisation of the price. Since the returns are not strongly serially correlated, this is not a surprise. This is a common finding when studying financial markets (see González [5]).

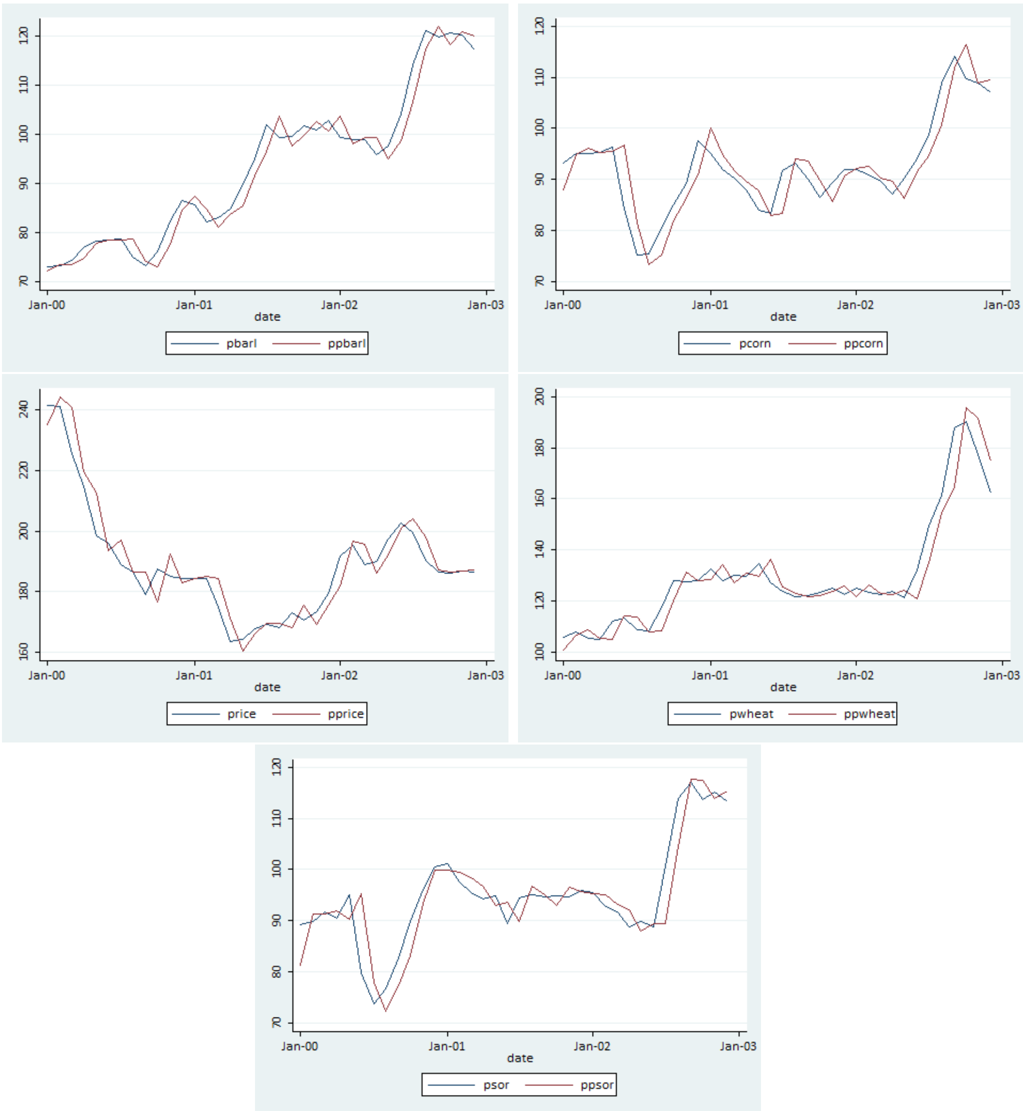

Figure 11: Predicted prices (in red) vs. realized prices (in blue)

2.6 Towards insurance theory, and the returns from bundling : correlations between the unpredictable components of the time series

2.6.1 Empirical results

Now that we have estimated forecasts for our five time series, we have been able to decompose each one into predictable and unpredictable components. This decomposition is essential to tackle the problem at hand.

Bundling of different risk lines can only be efficient if the unpredictable components of the risk lines are uncorrelated. For example, if each unpredictable shock affects all of the five crops we study in the same way (i.e. if the unpredictable components of the time series are highly correlated), then bundling the five crops will not substantially reduce risk. However, if the unpredictable components of the risk lines do not show a substantial correlation, bundling the risk lines enables risk-reduction. In Table 9, we show the correlation matrix between the unpredictable components of the time series. It it clear that the correlations between these unpredictable components are generally low. The only exception is the high correlation between the unpredictable components of the returns on corn and sorghum, which is easily understandable given the very high correlation between the prices of these two crops, as highlighted above on figure 10.

Table 9: Correlation matrix of the forecast errors, i.e. the unpredictable components of the time series

	Sorghum	**Rice**	**Barley**	**Corn**	**Wheat**
Sorghum	1				
Rice	0.0825	1			
Barley	0.3164	0.1176	1		
Corn	0.7324	0.0949	0.4132	1	
Wheat	0.4420	-0.0311	0.3184	0.4665	1

2.6.2 One last forecasting question : how should we forecast the evolution of the price of a bundle of goods ?

Before moving on to the analysis of the insurance problem, we introduce a bundle of goods that will be used as a reference. This bundle contains one unit of each good we have studied so far (i.e. the bundle consists of a ton of wheat, a ton of barley, a ton of corn, a ton of sorghum and a ton of rice). To end this section on forecasting, we analyze the problem of the forecast of the price of this bundle of goods. There are two methods to estimate such a forecast:

1. Forecast the price of each of the five crops of our bundle, and sum the forecasted prices;

2. Use an ARMA model to estimate directly a forecast of the price of the bundle.

The first strategy is easy to implement given the previous work. To implement the second one, we check that the time series of the returns on the bundle is covariance-stationary, study the AC and PAC functions of the time series, and test different ARMA models. This process selects an ARMA(2,4).

With the predicted returns for the bundle, we can estimate a one-step ahead forecast of the price of the bundle. Our results are visible in Figure 12 for the period 2000-2002.

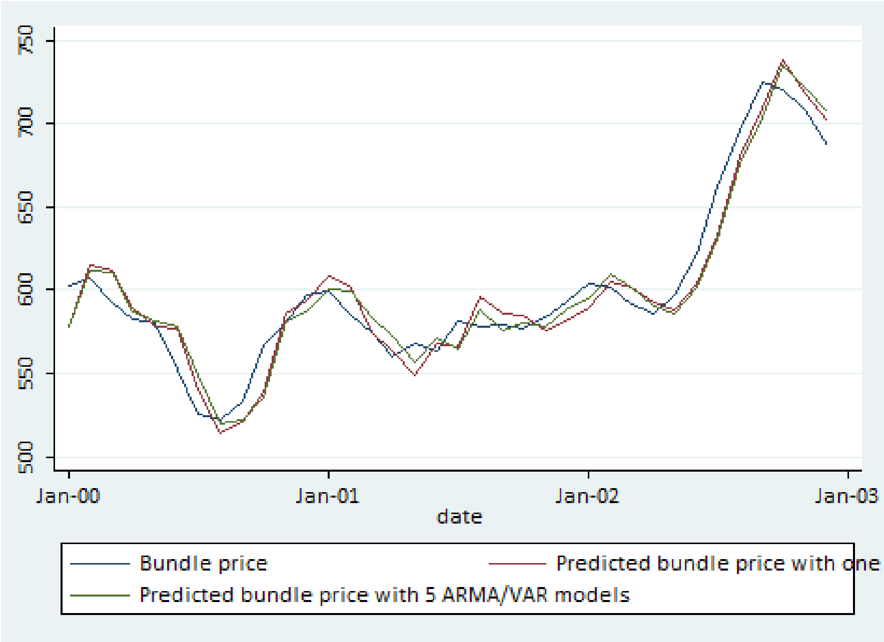

Figure 12: Predicted prices with the previous ARMA and VAR models (strategy 1, in green), with an ARMA(2,4) model (strategy 2, in red) and realized prices (in blue)

How can we compare the two strategies? As before, a relevant method is to compare the mean square of the one-step ahead forecast error. This comparison method gives us the following results :

	Mean square one-step ahead forecast error
Multiple ARMA and VAR forecasts (strategy 1)	1508.67
ARMA(2,4) forecast (strategy 2)	1685.62

It appears that using a single ARMA model provides better forecasts for the price of the bundle than using ARMA or VAR forecasts for each of the components of the bundle. This can be explained by the low or negative correlation between the unpredictable components of the returns on our five crops. In order to predict the price of the bundle, we will from now on rely on predictions of its components rather than of its own price.

3 Using forecasts for insurance

Now that we have estimated forecasts for the prices of five crops and of their bundle, we can test different insurance strategies with cat-bonds, and estimate the returns from insuring a bundle of goods vs. insuring each risk line separately.

3.1 Theoretical setting

We consider a firm which can issue cat-bonds, and an insurer who buys the bonds and sets the price at which the bonds are purchased. At each period, the firm wishes to purchase one unit of the previously defined bundle. The firm plans its purchases one period ahead: at the end of each time period, it saves the amount of money it considers necessary to buy the unit of the bundle at the next period. If the amount saved is lower that the amount actually needed to purchase the bundle, the firm faces a loss.

3.1.1 The cat-bond

For a product i, we define a cat-bond issued at time t as $I_{i,t} = (I_{i,t}(x; a, b, \theta), P_{i,t})$ where $I_{i,t}(x; a, b, \theta)$ is the indemnity received in case of a loss x, given parameters a, b, θ and $P_{i,t}$ is the price at which the cat-bond is exchanged. The loss x is defined in a contract between the buyer and seller of the bond. In our model, we consider that the loss is the difference between the price of product i at time $t+1$ and the expected price at t if this difference is positive. I.e.:

$$x = \max\{0, P_{t+1} - \mathbb{E}_t[P_{t+1}]\}$$

For example, consider a firm that plans at time t to buy a ton of rice at a price of £100 at time $t+1$. If the price of a ton of rice at time $t+1$ actually ends up being £102, the firm must face a loss of £2. However, if a ton of rice at time $t+1$ only costs £98, then the firm does not face any loss in our framework. The indemnity $I_{i,t}(x; a, b, \theta), P_{i,t})$ obeys to the following rules:

$$I_{it}(x; a, b, \theta) = \begin{cases} 0 \text{ if } x < a \\ \theta(x-a) \text{ if } a < x < b \\ \theta(b-a) \text{ if } x > b \end{cases}$$

A full insurance corresponds to buying a cat-bond with indemnity $I_{it}(x; 0, +\infty, 1) = x$ for all positive x, 0 otherwise.

3.1.2 The insurer

We consider that at the end of each period t, the insurer can buy for all $a, b, \theta \in \mathbb{R}^+ \times \overline{\mathbb{R}^+} \times [0, 1]$ cat-bonds for period $t+1$ associated with an indemnity $I_{i,t}(x; a, b, \theta)$, at price $P_{i,t}$. At the end of time period $t+1$, the insurer has a net gain of $P_{i,t} - I_{i,t}(x; a, b, \theta)$.

The price of the cat-bond with indemnity $I_{i,t}(x; a, b, \theta)$ is set using the following formula, also used in our reference paper [7]:

$$P_{i,t} = (1+\lambda)\chi_i + \frac{c}{2}\xi_i^2 \tag{1}$$

Where λ represents a loading associated with administrative and handling costs for the insurer. In this subsection, we will consider $\lambda = 0$ for simplicity. χ_i is the expected value of $I_i(x; a, b, \theta)$, and ξ_i is its expected standard deviation. c is a parameter that reflects the degree of risk aversion of the insurer. A standard value for c is 0.05, and we will use this value in numerical estimates.

How are χ_i and ξ_i estimated by the insurer ? In order to estimate χ_i and ξ_i, the insurer uses the same forecasting methods as we did before. He estimates a one-step ahead forecast of P_{t+1}, and considers at time t that:

$$P_{t+1} \sim \mathbb{E}_t[P_{t+1}] + \epsilon_{t+1} \tag{2}$$

Where $\mathbb{E}_t[P_{t+1}]$ is the one-step ahead forecast and ϵ_{t+1} forecast error, which follows a distribution that we need to determine. Using this distribution of P_{t+1}, the insurer sets the price of the bond to:

$$P_{i,t} = \mathbb{E}_t(I_{i,t}(P_{t+1} - \mathbb{E}_t[P_{t+1}]; a, b, \theta)) + \frac{c}{2}\sigma_I^2 = \mathbb{E}_t(I_{i,t}(\epsilon_{t+1}; a, b, \theta)) + \frac{c}{2}\sigma_I^2 \tag{3}$$

Where $\epsilon_{t+1} = x$ is the loss faced by the firm, and σ_I^2 is the variance of the indemnity.

Are the forecast errors normally distributed ? In order to assign a price to the cat-bonds, we need to know the distribution of the forecast errors ϵ_{t+1}. This distribution will determine the expectation of the indemnities and their standard deviation. At first glance, the distribution of forecast errors seems to be normal. But estimating its density *via* its kernel density shows us the discrepancy between the two hypothesis (see figure 13). The kurtosis of a normal distribution is 3 and its skewness zero (as it is a symmetric distribution). Here, we test[7] for nonnormality due either to skewness or to excess kurtosis as in figure 14: p-values of 13% and 0% respectively, suggesting that at reasonable confidence levels, skewness is not statistically different from zero, whereas excess kurtosis is undoubtuous.

Table 10: Kurtosis of errors of price forecasting

	Sorghum	**Rice**	**Barley**	**Corn**	**Wheat**	**Bundle**
Kurtosis	9.534236	42.54344	10.01047	12.18014	9.982071	9.438219

Rather than a normal distribution which is described by its two first moments, it seems like the series' distribution has to be described by at least three parameters, if not four: one for the variance, one for the kurtosis, one for normalization, and one for the mean as well. However, the shape of the estimated distribution makes us think about a Laplace distribution, although it only allows for two degrees of freedom. Nevertheless, we use this simple distribution as a first approximation. The probability density function of the Laplace is the following, with parameters μ and b describing the mean and standard deviation of the distribution:

$$f(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}$$

We use the sample mean and variance of the forecast errors to estimate μ and b. In order to determine how good the distribution fits the data, we use the Kolmogorov-Smirnov test (KS-test), as detailed by Shorack and Wellner ([9]), where the null hypothesis is "both distributions

[7] this test is presented by D'Agostino and Belanger in [2], where the null hypothesis is normality

are equal". The test gives two statistical aggregates: the KS-statistic is the maximum distance between the empirical cumulative distribution function and the one we draw by estimating μ and b, and the corresponding p-value is the standard statistic allowing for decision about the relevance of the model. If the p-value is less than 0.1, we reject the hypothesis that the distribution is well described by a Laplace distribution, at the 10% significance level. Results are shown in table 11.

Choosing a Laplace distribution calibrated on the two first moments seems to lead to better results than choosing a normal one. However, we still do not have a perfect fit, as shown in Table 11. Here, we do not focus on finding a better distribution, and will compute empirical estimates using a Laplace distribution. We find that a Laplace distribution is actually a good fit for the forecasting errors on the bundle.

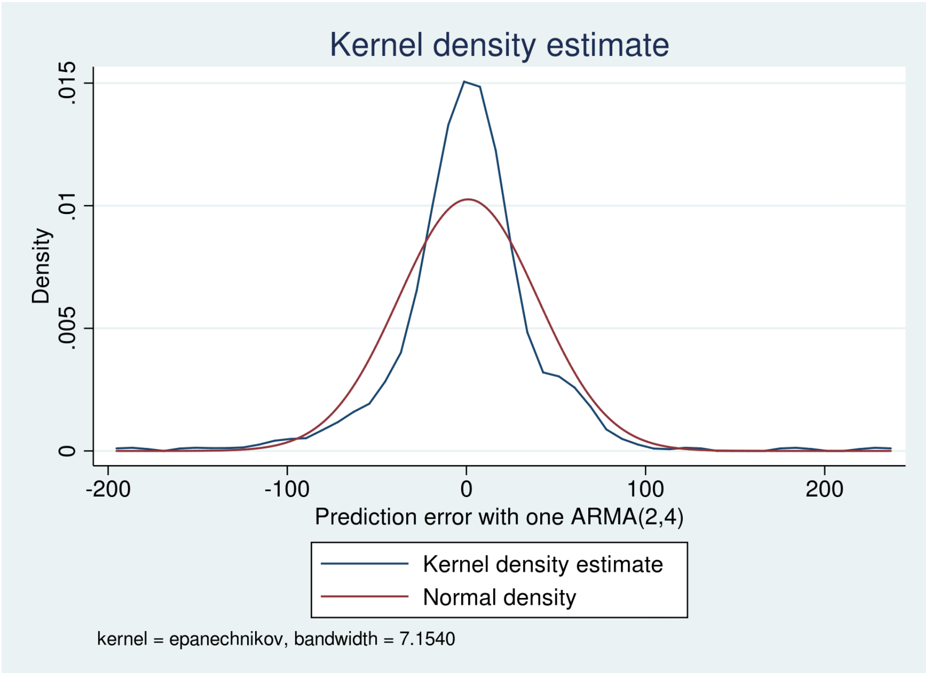

Figure 13: Inferred density of forecast errors and the corresponding normal density, or what it would have been if the series was normally distributed, computed on the bundle's price forecast

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	joint adj chi2(2)	joint Prob>chi2
eppbundle1	360	0.1303	0.0000	46.13	0.0000

Figure 14: Skewness and kurtosis tests on the distribution of errors of forecasting on the bundle

3.1.3 The firm

At each time period, the firm decides to issue cat-bonds with parameters a, b, θ, and announcing its predicted price $\mathbb{E}_t[P_{t+1}]$. For simplicity, we will here assume that the firm uses the same

Table 11: Goodness of fit of a mean and variance calibrated Laplace distribution for forecast errors (KS test). We also show KS statistics against a calibrated normal distribution. Comparing KS-statistics, i.e. distances between distributions, makes clear that the Laplace distribution is a better approximation than the normal.

	Sorghum	**Rice**	**Barley**	**Corn**	**Wheat**	**Bundle**
Laplace KS statistic	0.0530	0.0828	0.0748	0.0684	0.048	0.0369
p-value	0.2546	0.0135	0.0337	0.0656	0.3673	0.7111
Gaussian KS statistic	0.1113	0.142	0.1353	0.1219	0.0975	0.0955
p-value	0.000	0.000	0.000	0.000	0.002	0.002

forecasting methods as the insurer. Thus, both the firm and the insurer have the same expectations concerning the losses of the firm. At period $t+1$, the firm makes a payment of $P_{i,t}$ to the insurer, and receives a payment $I_{i,t}(\epsilon_{t+1}; a, b, \theta)$. We define the gain of the firm with a mean-variance preference:

$$W_t = \sum_i G_{i,t} + \frac{\kappa}{2} Var(\sum_i G_{i,t})$$

Where:

$$G_{i,t} = I_{i,t}(\epsilon_{t+1}; a, b, \theta) - P_{i,t} - \epsilon_{t+1}$$

and the i are the different risk lines of the firm.

If the firm uses bundling, its gain is:

$$W_t = G_t + \frac{\kappa}{2} Var(G_t)$$

where:

$$G_t = I_{i,t}(\sum_i \epsilon_{i,t+1}; a, b, \theta) - P_{bundle,t} - \sum_i \epsilon_{i,t+1}$$

We consider two strategies for the firm:

1. At each period, issue 5 cat-bonds (one for each crop of the bundle);
2. At each period, issue only one cat-bond for the entire bundle.

Here, we wish to show that strategy 2 dominates strategy 1 in terms of the gains for the firm.

3.2 A comparison of hedging strategies: theoretical and empirical results

3.2.1 Theoretical results

Problem setup. With the framework of the previous subsection, we try to find theoretically an optimal contract (i.e. an indemnity and a contract price) both for the firm and the insurer in the case of a bundle contract as well as in the case of an Line-By-Line contract. The best indemnity function will be the one maximising the gain of the firm as well as minimising the uncertainty of its loss. We consider a mean-variance preference for the insured firm like in [7], with which the best indemnity is found (with reasoning close to the one in [6]).

Line-By-Line strategy In order to compare the bundle contract with an Line-By-Line one we examine a simple case:

$$Bundle = \frac{1}{n}\sum_{i=1}^{n} Crop_i$$

So we have the following equality:

$$\epsilon_{t+1} = \frac{1}{n}\sum_{i=1}^{n} \epsilon_{i,t+1} \tag{4}$$

In order to have a very symmetrical problem, we suppose that the $\epsilon_{i,t+1}$ are identically distributed so:

$$\mathbb{E}_t(\epsilon_{t+1}) = \mathbb{E}_t(\epsilon_{i,t+1})$$
$$\mathbb{V}_t(\epsilon_{t+1}) = \mathbb{V}_t(\epsilon_{i,t+1})$$

and we suppose that:

$$\forall i,j, i \neq j, \mathbb{E}_t((\epsilon_{i,t+1} - \mathbb{E}_t(\epsilon_{i,t+1}))(\epsilon_{j,t+1} - \mathbb{E}_t(\epsilon_{j,t+1}))) = \sigma_{ij}^2 = \sigma^2$$

In this case the price of the contract on one crop is:

$$P_{i,t} = \mathbb{E}_t(I_{i,t}(P_{i,t+1} - \mathbb{E}_t[P_{i,t+1}])) + \frac{c}{2}\sigma_I^2 = \mathbb{E}_t(I_{i,t}(\epsilon_{i,t+1})) + \frac{c}{2}\sigma_I^2 \tag{5}$$

The total price of the cat-bond is:

$$P_{tot,t} = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{E}_t(I_{i,t}(\epsilon_{i,t+1})) + \frac{c}{2}\sigma_I^2\right) \tag{6}$$

The firm wants to maximize the following objective function:

$$V(I,P) = \frac{1}{n}\sum_{i=1}^{n}(I_{it}(\epsilon_{it+1}) - P_{it} - \epsilon_{it+1}) - \frac{\kappa}{2}(I_{it}(\frac{1}{n}\sum_{i=1}^{n}(\epsilon_{it+1}) - \epsilon_{it+1}))^2 \tag{7}$$

Subject to:

$$P_{it} = \mathbb{E}_t(I_{it}(\epsilon_{it+1})) + \frac{c}{2}\sigma_I^2 \tag{8}$$

Because of symmetry among crops, we have a symmetric solution, that is the same indemnity function for each crop and the same price as well (see Appendix E.1). The indemnity $I(x)$ is solution of the following equation:

$$\left(c + \frac{\kappa}{n}\right) I(x) + \frac{\kappa(n-1)}{n}\mathbb{E}(I(y)|x) = c\mathbb{E}(x) + \frac{\kappa}{n}x + \frac{\kappa(n-1)}{n}\mathbb{E}(y|x) \tag{9}$$

For a case $n = 2$, we have for example:

$$cI(x) + \frac{\kappa}{2}I(x) + \frac{\kappa}{2}\mathbb{E}(I(y)|x) = c\mathbb{E}(x) + \frac{\kappa}{2}x + \frac{\kappa}{2}\mathbb{E}(y|x) \tag{10}$$

This illustrates well a trade off between the risk aversion of the insurer (denoted by c) and by the one of the firm (denoted by κ). Moreover, it also shows another arbitrage that the firm takes into account between the crop price variation and the conditionally expected variation of other crops prices. However, we were not able to find a closed-form solution for this equation: we thus revert to empirical estimations in the following subsection.

Bundle case. We consider a bundle with an insurance price:

$$P_{bundle,t} = \mathbb{E}_t(I_{bundle,t}(\epsilon_{t+1})) + \frac{c}{2}\sigma_I^2 \tag{11}$$

The firm wants to maximize the expectation of the following objective function:

$$V(I) = I_{bundle,t}(\epsilon_{t+1}) - P_{bundle,t} - \epsilon_{t+1} - \frac{\kappa}{2}(I_{i,t}(\epsilon_{t+1}) - \epsilon_{t+1})^2 \tag{12}$$

The firm wants to maximize its expected gain, and wishes to avoid volatility. We obtain the following results (see Appendix E.2):

$$I(x) = \frac{\kappa}{\kappa + c}x + \frac{c}{\kappa + c}\mathbb{E}_t(\epsilon_{t+1}) \tag{13}$$

and price:

$$P_{bundle,t} = \mathbb{E}_t(\epsilon_{t+1}) + \frac{c}{2}\left(\frac{\kappa}{\kappa + c}\right)^2 \sigma_\epsilon^2 \tag{14}$$

3.2.2 Empirical results

Since the theoretical approach does not yield closed-form solutions, we revert to an empirical investigation. We suppose that the firm decides to issue cat-bonds every month, for the next month. In this subsection, we will consider the price of 5 cereals from 1990 to 2018. The firm wants to get insurance for a bundle of weights $\alpha_i = 0.2$, for each cereal i. We keep the notations of[7], and suppose that $\lambda = 0$.

We compute for the line-by-line and bundle strategies:

- The price of the cat-bonds;
- The gains of the firm, as defined previously.

As stated before, we use the classical value of $c = 0.05$. In order to have a more risk-averse firm than the insurer, we choose $\kappa = 0.1$.

We can compare the gains of the firm for both strategies, in full insurance ($a = 0$, $b = +\infty$, $\theta = 1$).

The gain of a risk-neutral firm ($\kappa = 0$) is displayed in Appendix F in figure 20.

In order to compute the gains of the firm, we also need to compute the variance of the gains of the firm, see figure 21.

The variance for the in-line strategy is almost always greater than for the bundling strategy. This hints to a first advantage of the bundling strategy.

It is difficult to visualize on figure 20 which strategy maximizes the utility of the firm. Yet, we find that :

The in-line strategy brings an average gain of -6.2, and is optimal to use in 78 periods of our sample;

1. Bundling brings an average gain of -5.0, and is optimal 263 times

This supports the claim that bundling brings is optimal for the firm. We want to test this

Table 12: Utilities of the firm in in-line and pooling strategies, depending on values of a, b and θ

a	b	θ	U_{line}	U_{pool}	best strategy
0	∞	0.5	-1.58	-1.32	pool
0	∞	1	-1.23	-0.84	pool
5	∞	0.5	-1.61	-1.40	pool
5	∞	1	-1.27	-0.93	pool
0	10	0.5	-2.19	-2.05	pool
0	10	1	-2.11	-1.86	pool
5	10	0.5	-2.22	-2.15	pool
5	10	1	-2.18	-2.05	pool

result out of the full insurance case, for different a, b, θ. To compare the two strategies, we will assume that in the in-line insurance, each cereal i is insured with parameters:

$$a_i = a/5 \tag{15}$$

$$b_i = b/5 \tag{16}$$

$$\theta_i = \theta \tag{17}$$

where a, b, and θ are the parameters for the bundle.

We can see that in the range of values of a, b and θ chosen, the best strategy is most often to insure the bundle.

4 Conclusion

In our study, we have used various forecasting methods to model the evolution of crop prices. These techniques are essential to price insurance contracts, but as with most commodities, forecasting crop prices is quite difficult. We show that the use of ARMA methods provide the best results in most cases, and that VAR regressions can improve forecasts of highly correlated price lines. Finally, we show that modeling volatility within the GARCH framework does not provide significant improvement to ARMA estimates.

Using the chosen forecasting models, we challenged the theoretical hypothesis formulated by Picard and Louaas. Is it optimal for firms to bundle risk lines using cat-bonds than to use classical hedging techniques? We show that under full insurance or simple cat-bond contracts, pooling may benefit to the insured firm.

Further steps in research include the definition of optimal bundles, and parameters for the cat-bond contracts. As closed-form solutions are likely not to be found to solve the problem, numerical approximations should be computed, using classical optimization techniques.

A The store-and-sell model

A.1 The cobweb model does not hold when dealing with returns

As proved in the previous subsection, prices in the cobweb model follow an AR(1) process, with :

$$(p_t - p^*) = \phi(p_{t-1} - p^*) + \epsilon_t$$

Defining returns as $r_t = \frac{p_t - p_{t-1}}{p_{t-1}}$, we have :

$$\forall t, p_t = p_0 \prod_{i=1}^{t}(1 + r_i)$$

We can attempt to separate r_t in an AR and MA component :

$$\begin{aligned}
r_t &= \frac{p_t - p_{t-1}}{p_{t-1}} \\
p_{t-1} r_t &= (\phi - 1)p_{t-1} + (1-\phi)p^* + \epsilon_t \\
p_0 \left(\prod_{i=1}^{t-1}(1 + r_i) \right) r_t &= (\phi - 1)p_{t-1} + (1-\phi)p^* + \epsilon_t \\
p_0 \left(\prod_{i=1}^{t-1}(1 + r_i) \right) r_t &= (\phi - 1)\left(p^* + \sum_{s=1}^{t-1} \phi^{t-s}\epsilon_s \right) + (1-\phi)p^* + \epsilon_t
\end{aligned}$$

While the right-hand side of the relation exhibits a clear MA structure, the left-hand side does not have an AR structure. The ARMA model for prices fails to extend completely to returns. However, it is still possible to theoretically justify an ARMA analysis of returns by using another model which is a more complex version of the cobweb model : we call this model the store-and-sell model.

A.2 The store-and-sell model

We describe here the setting of the store-and-sell model. In this model, prices are denoted at each time period by p_t, returns by r_t.

Supply side : at each point in time, suppliers produce one unit of the good. At each time period t, they take three actions :

1. They sell a quantity $q_t^M \geq 0$ of the good on the spot market;
2. They store a quantity $q_t^S \geq 0$ of the good for one time period, to sell it on the spot market at the next period;
3. They can enact futures contracts for a quantity $q_t^F \geq 0$ of the good : at time $t+1$, the producer will need to give q_t^F to the other party of the contract, against a payment of q_t^F.

At each time period, we thus have :

$$q_t^M + q_t^S + q_t^F = 1$$

Demand side : on the spot market, at each time period, there is a quantity $q_t = q_t^M + q_{t-1}^S$ of the good provided by the suppliers. Part of the demand has already been satisfied by the futures contracts enacted at time $t-1$. We have the following linear demand function :

$$q_t = 2 - q_{t-1}^F - p_t$$

Market clearing : when the market clears, we have the following price function:

$$p_t = 1 + (q_t^S - q_{t-1}^S) + (q_t^F - q_{t-1}^F)$$

Storage, contract cost and expectations : finally, to make the model complete, we need to set costs for storage, contracts and expectations. We assume that agents make adaptative expectations on returns : $\mathbb{E}[r_t] = r_{t-1}$. We set quadratic costs for both storage and enacting contracts. This means that when the supplier wishes to store quantity q_t^S, he must pay a storage cost of $c(q_t^S) = \frac{(q_t^S)^2}{2\gamma}$, and when he enacts a futures contract quantity q_t^F, he must pay a enactment cost of $c(q_t^F) = \frac{(q_t^F)^2}{2\gamma'}$. If we assume that $p_t \sim 1$ (which is verified when the amounts of good stored and exchanged on futures markets do not vary much over time), that $\gamma = \gamma'$ (for convenience in computations) and that the interest rate is very small compared to one, then we find that :

- When $\mathbb{E}[r_t] \geq 0$, $q_t^S = \gamma\mathbb{E}[r_t]$ and $q_t^F = 0$;
- When $\mathbb{E}[r_t] \leq 0$, $q_t^F = \gamma\mathbb{E}[r_t]$ and $q_t^S = 0$

We then have for all t : $q_t^S + q_t^F = \gamma\mathbb{E}[r_t] = \gamma r_{t-1}$, and since $p_t = 1 + (q_t^S + q_t^F) - (q_{t-1}^S + q_{t-1}^F)$, we have :

$$p_t = 1 + \gamma r_{t-1} - \gamma r_{t-2}$$

And, finally, for returns, as $p_t \approx 1$, $r_t = ln(p_t) - ln(p_{t-1}) \approx p_t - p_{t-1}$, we have :

$$r_{t+1} = \gamma r_t - 2\gamma r_{t-1} + \gamma r_{t-2}$$

With this new store-and-sell model, returns have an AR(3) structure. We will now leave from theory to continue our empirical workd and estimate whether an ARMA model is appropriate to analyse our dataset.

B Finding the right ARMA model

We need to use the ACF and PACF to decide which model should be used. There is no clear cutoff in the PACF (which would have hinted towards a AR(p) process) or in the ACF (which would have suggested a MA(q) process). The best model for this time series is therefore probably an ARMA. The first peak of the ACF and PACF functions is statistically significant, but not the 2nd, 3rd and 4th. In the ACF, the 6th peak is statistically different from zero at a 95% level, and in the PACF, this is the case for the 5th and 6th peak.
Keeping these observations in mind, we are going to test the following models: AR(1), ARMA(1,1), ARMA(1,2), ARMA(2,1), ARMA(2,2), ARMA(5,6) and ARMA(6,6). To choose between these different models, we report the **Akaike information criteria (AIC)** and the **Bayesian information criteria (BIC)** of these different models in the following table:

Model	AIC	BIC
AR(1)	-1068.644	**-1056.985**
ARMA(1,1)	-1066.893	-1051.349
ARMA(2,1)	-1064.923	-1045.492
ARMA(1,2)	-1064.966	-1045.536
ARMA(2,2)	-1069.492	-1046.176
ARMA(5,6)	-1073.137	-1022.617
ARMA(6,6)	**-1073.983**	-1023.464

Table 13: Comparing different ARMA models with information criteria

The lowest AIC coefficient is found for the ARMA(6,6) and the lowest BIC is found for an AR(1) model. This reflects the fact that the BIC penalizes more models with higher orders than the AIC. For simplicity, we choose to first model the time series with an AR(1):

$$Y_t = c + \phi Y_{t-1} + \epsilon_t$$

AR models can be estimated by OLS, which yields the following results:

chcorn	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
chcorn						
_cons	.0022204	.0038624	0.57	0.565	-.0053498	.0097906
ARMA						
ar						
L1.	.2563972	.0458494	5.59	0.000	.166534	.3462604
/sigma	.0543896	.0013937	39.02	0.000	.051658	.0571212

Figure 15: AR(1) estimation for the returns on corn

The next step to validate the model is to study whether the residuals of our estimation are a white noise or not. If the residual is not a white noise (e.g. if there is autocorrelation), then the selected model is unsatisfactory: the residuals still contain information ! We therefore compute the residuals of the AR(1) predictions, and find in the correlogramm that some AC and PAC coefficients are statistically different from 0, and the Ljung–Box test rejects the hypothesis of no autocorrelation. We must thus change our model. Testing different possibilities leads to an ARMA(2,2) model: even though it has not been chosen by any of the two information criteria, it constitutes a good compromise, as will be shown. We estimate the following model:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \epsilon_t$$

In the case of MA and ARMA models, as the terms ϵ_t are not directly observable, we cannot use ordinary least squares to estimate this model, and need to use nonlinear least squares: this estimation can be easily made with the help of statistical software. We find the following results:

In this new estimation, the estimated coefficients are jointly different from zero,[8] and the p-values associated to the Q-statistics 1 to 23 are all over 0.05. Therefore, there is barely any

[8] Here, we observe that the second MA coefficient is not statistically different from zero. does this mean

chcorn	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf. Interval]	
chcorn						
_cons	.0022121	.0040692	0.54	0.587	-.0057634	.0101875
ARMA						
ar						
L1.	-.4294545	.2360495	-1.82	0.069	-.892103	.033194
L2.	.3977911	.1834267	2.17	0.030	.0382813	.7573009
ma						
L1.	.7026966	.2464802	2.85	0.004	.2196043	1.185789
L2.	-.2339617	.2215603	-1.06	0.291	-.6682119	.2002885
/sigma	.0538587	.0014081	38.25	0.000	.0510989	.0566185

Figure 16: ARMA(2,2) estimation for the returns on corn

difference between our residuals and a white noise.
Would an ARIMA model add some precision to our forecasts? The answer is no: as we performed an estimation on the returns series and not directly on the price series, we do not need to integrate our data.

C Granger-causality tests

VAR enables us to use all of the information available to make predictions. However, some variables in our enlarged information set may be more useful than others in forecasting. To see which variables are important in the VAR we use a **Granger-causality test**. A variable X is said to Granger-cause Y if forecasts that use past values of X and Y yield lower forecast errors than forecasts that only use past values of Y. However, establishing a Granger-causality of X on Y does not mean that X truly causes Y.
In a general five variable context, the Granger-causality tests consists in the following :

1. For each crop Y (and noting the four other crops X_1, X_2, X_3, X_4), we estimate 6 alternative VAR models: the first one is a regression of Y on the lagged values of Y and X_1, X_2, X_3, X_4 (it is one of the 5 regressions of our previous VAR), the next four are the regressions of values of Y on lagged values of Y and of the lagged values of 3 of the 4 variables in the set $\{X_1, X_2, X_3, X_4\}$. The last regression is simply an AR(1) regression for Y. The five last regressions are restrictions of the first one, with respectively (1,1,1,1,4) restricted variables.

2. Using the unrestricted model and the five restricted model, we can then perform F-tests or, as STATA performs, chi-squared tests, in order to estimate whether the coefficients associated with each crop X_1, X_2, X_3, X_4 are jointly significant or not, and whether the coefficients associated to all other crops X_1, X_2, X_3, X_4 are jointly significant or not.

that the model is too complex and we should revert to an ARMA(2,1)? The answer is no: if we estimate an ARMA(2,1) model, the first MA coefficient becomes statistically indifferent from zero. Including a second lag is therefore of interest in this estimation.

3. Using the computed test statistics, we can test the following null hypotheses:

$$\forall i \in \{1,2,3,4\}, H_{0i} : X_i \text{ does not Granger-cause } Y$$

$$H_0^{'} : X_1, X_2, X_3, X_4 \text{ do not Granger-cause } Y$$

In our simple VAR(1) context, testing the H_{0i} is quite easy, as the F-test needed to test them only make one restriction: they are therefore equivalent to t-tests. To test the H_{0i}, we can simply look at the results of the VAR(1) estimation, and reject the H_{0i} hypothesis when the coefficient associated with the once-lagged value of X_i is statistically different from 0. To test $H_0^{'}$ however, we still need to compute an F-test. Our results are found in Table 14.

Crop X/Y	**Corn**	**Barley**	**Rice**	**Sorghum**	**Wheat**
Corn		NO	NO	NO	NO
Barley	NO		NO	NO	NO
Rice	NO	NO		YES*	NO
Sorghum	YES**	NO	NO		NO
Wheat	NO	NO	YES*	YES*	
ALL	YES**	YES**	NO	YES*	NO

Table 14: Results for the Granger-causality tests : for each crop Y (each column, we can find whether the hypotheses H_{0i} and $H_0^{'}$ can be rejected or not at a 5% confidence level. For example, the cell in row "Barley" and column "Corn" tells us that we cannot reject the hypothesis that the changes in prices of barley **do not** Granger-cause changes of prices of corn. A * indicates that we can reject the null at a 5% confidence level, a ** indicates that we can do so at a 1% confidence level.

What can we conclude from these results? First, it appears that when predicting the price changes of rice and wheat, estimating a VAR does not substantially increase the performance of the forecast, compared to a AR(1). Since we have shown previously that the AR(1) model was one of the most efficient in the ARMA family (according to AIC and BIC) to forecast the returns on wheat, and that the ARMA(1,2) dominated the AR(1) to forecast the returns on rice, we drop the VAR(1) model to predict the evolution of returns of these two crops, and return to the previously selected AR and ARMA models.

What about the three other crops? For sorghum, barley and corn, how to choose between the VAR and ARMA models? To distinguish between these two types of model which performs better in forecasting, we decided to compare the sum of squares of forecast errors for the two estimates. For each of the three crops and the two types of models, we compute the SSE and report them below :

	ARMA	**VAR**
Sorghum	**1.163**	1.186
Barley	1.137	**1.079**
Corn	1.044	**1.018**

D GARCH model selection

D.1 Methodology

The way to find whether a process is ARCH or not can be separated within four stages. **First**, we identify the ARMA parameters of the series at stake. **Second**, we compute the residuals after the ARMA modelisation. We then want to know if these residuals can actually be described by a GARCH model. That is, we check for heteroskedasticity in squares of residuals, as hinted by the structure of this model. Engle in [4] developed a Lagrange multiplier test to answer this question: it checks the hypothesis that ϵ_t is an independant identically distributed (IID) white noise against the alternative that ϵ_t is an $ARCH(q)$ process. **Third**, if a GARCH effect is detected, we estimate the number of necessary lags in the GARCH, via criteria such as AIC and BIC, both compounding different criteria of model match to the data. According to them, we choose the adequate number of lags and finally run the estimation of the GARCH parameters using a statistical software. Estimation is often conducted via maximum likelihood estimation. Finally the outcome of the algorithm must be tested: is there any effects we failed to model ? One simple criterion to (un)validate the GARCH fit of the series is to compute the ACF and PACF of the squared residuals after the GARCH. We use Q-statistic to determine whether those residuals are correlated or not. In the latter case, the GARCH model is a good tool to model the series at stake.

D.2 Finding the right GARCH model

To show how we model price evolutions here using a GARCH model, we again first focus on one of our time series: the returns of corn. As GARCH is used to model the behavior of residuals obtained after an ARMA model, we will take the predictions of the previous subsection as granted and work on their residuals. Then, we compute the ACF and PACF of squares of residuals, as shown in figure 17. They point out that a GARCH model of at least nine lags for both parameters is relevant, which can be hard to calculate. We also compute Engle's LM-test (see figure 18). LM's test clearly shows the need to use at least an ARCH on these residuals. Nevertheless, for our own interest, we compute the Akaike's information criterion and the Bayes' information criterion to decide which orders we will choose, as in table 15. We stop the calculus at GARCH(2,2), because the method of optimization we used couldn't estimate the parameters beyond. The analysis of table 15 provides us with an optimal GARCH model for residuals: we choose to run a GARCH(1,2), as both AIC and BIC are at their lowest level here. In order to be sure of our choice, we again compute the ACF and PACF of squares of residuals obtained after GARCH. Results are shown in figure 19.

Model	**AIC**	**BIC**
ARCH(1)	-2665.251	-2653.592
ARCH(2)	-2665.703	-2650.159
ARCH(3)	-2666.845	-2647.414
GARCH(1,1)	-2663.803	-2648.259
GARCH(2,1)	-2663.711	-2644.281
GARCH(1,2)	**-2707.154**	**-2687.724**

Table 15: Comparing different GARCH models for corn with information criteria

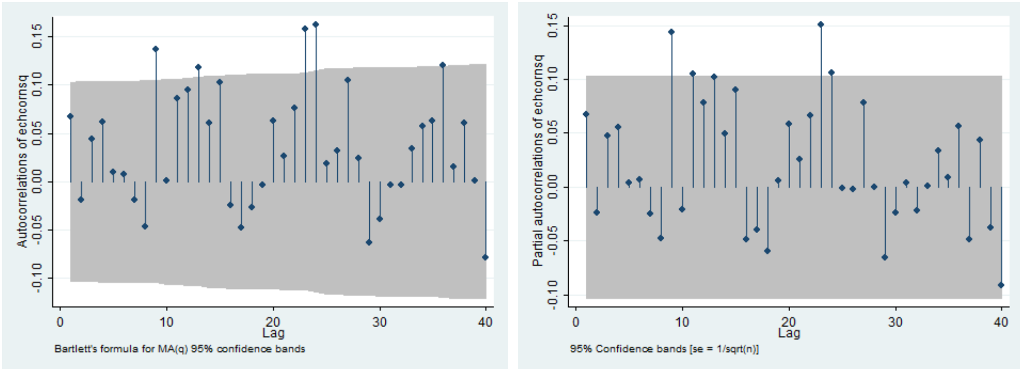

Figure 17: Auto-correlation function of square residuals after ARMA on corn returns.

```
. estat archlm, lags(1/5)
LM test for autoregressive conditional heteroskedasticity (ARCH)
```

lags(p)	chi2	df	Prob > chi2
1	1.641	1	0.2002
2	1.849	2	0.3967
3	2.657	3	0.4476
4	3.685	4	0.4503
5	3.639	5	0.6024

```
H0: no ARCH effects      vs.  H1: ARCH(p) disturbance
```

Figure 18: Engle's test on after ARMA residuals for corn. The null hypothesis being "there's no ARCH effect in this series" cannot be rejected at any significance level below 20% for one single lag, below 39% for two lags, etc.

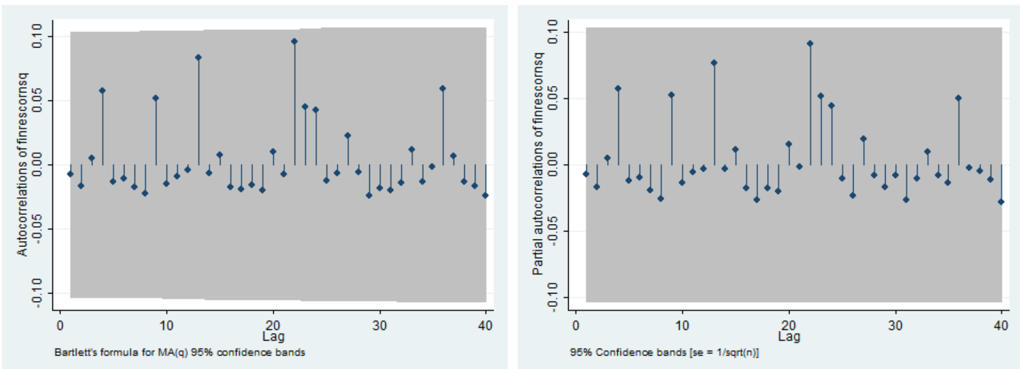

Figure 19: Auto-correlation function of square of residuals after GARCH on corn returns. All correlations are within the two standard error bands.

E Analysis of the best indemnity function

E.1 Best contract for a bundle contract

During this whole proof we only consider functions that admit an expectation and a variance. We want to get the indemnity that maximises the expectation of the following objective function :

$$V(I,P) = I_{bundlet}(\epsilon_{t+1}) - P_{bundlet} - \epsilon_{t+1} - \frac{\kappa}{2}(I_{bundlet}(\epsilon_{t+1}) - \epsilon_{t+1})^2 \tag{18}$$

Knowing :

$$P_{bundlet} = \mathbb{E}_t(I_{bundlet}(\epsilon_{t+1})) + \frac{c}{2}\sigma_I^2 \tag{19}$$

The problem is to find, considering the expectation conditionally to the fact that $\epsilon_{t+1} \geq 0$: the insured firm is interested in getting an indemnity only when she faces a loss. So, the (unknown) probability density we consider is the density of ϵ_{t+1} restricted to the positive value :

$$\sup_I \mathbb{E}_t(I_{bundle,t}(\epsilon_{t+1}) - P_{bundle,t} - \epsilon_{t+1} - \frac{\kappa}{2}(I_{bundle,t}(\epsilon_{t+1}) - \epsilon_{t+1})^2) \tag{20}$$

equal to

$$\sup_I \mathbb{E}_t(I_{bundle,t}(\epsilon_{t+1}) - (\mathbb{E}_t(I_{bundle,t}(\epsilon_{t+1}) + \frac{c}{2}\sigma_I^2) - \epsilon_{t+1} - \frac{\kappa}{2}(I_{i,t}(\epsilon_{t+1}) - \epsilon_{t+1})^2) \tag{21}$$

We define the application

$$U : I \longrightarrow U(I)$$

that we are going to maximise, with simplified notations :

$$U(I) = \mathbb{E}(-\frac{c}{2}\sigma_I^2 - \frac{\kappa}{2}(I(\epsilon) - \epsilon)^2) = -\frac{c}{2}\mathbb{E}((I(\epsilon) - \mathbb{E}(I(\epsilon)))^2) - \frac{\kappa}{2}\mathbb{E}((I(\epsilon) - \epsilon)^2) \tag{22}$$

Let us compute the differential of U :

$$U(I+i) - U(I) = -c\mathbb{E}\left((I(\epsilon) - \mathbb{E}(I(\epsilon)))\, i(\epsilon)\right) - \kappa\mathbb{E}((I(\epsilon) - \epsilon)i(\epsilon)) \tag{23}$$

Which depends linearly on i. So, $U(I+i) - U(I)$ is the differential of U. Hence, we want to find I such as :

$$\forall i, -c\mathbb{E}\left((I(\epsilon) - \mathbb{E}(I(\epsilon)))\, i(\epsilon)\right) - \kappa\mathbb{E}((I(\epsilon) - \epsilon)i(\epsilon)) = 0 \tag{24}$$

Which implies the following relation (we suppose that the density of ϵ is non-zero almost everywhere):

$$\forall x \in [0, +\infty], c(I(x) - \mathbb{E}(I)) + \kappa(I(x) - x) = 0 \tag{25}$$

Taking the expectation we get :

$$\mathbb{E}(I) = \mathbb{E}(\epsilon) \tag{26}$$

Thus :

$$I(x) = \frac{\kappa}{\kappa + c}x + \frac{c}{\kappa + c}\mathbb{E}(\epsilon) \tag{27}$$

E.2 Best contract in the case of a line-by-line strategy

We suppose the same hypothesis in this subsection as in the previous one. The firm wants to maximise the following objective function :

$$V(I,P) = \frac{1}{n}\sum_{i=1}^{n}(I_{i,t}(\epsilon_{i,t+1}) - P_{i,t} - \epsilon_{i,t+1}) - \frac{\kappa}{2}(\frac{1}{n}\sum_{i=1}^{n}(I_{i,t}\,(\epsilon_{i,t+1}) - \epsilon_{i,t+1}))^2 \tag{28}$$

Knowing :

$$P_{i,t} = \mathbb{E}_t(I_{i,t}(\epsilon_{i,t+1})) + \frac{c}{2}\sigma_I^2 \tag{29}$$

The problem is to find, considering the expectation conditionally to the fact that $\forall i, \epsilon_{i,t+1} \geq 0$: the insured firm is interested in getting an indemnity only when the price have raised. So, the (unknown) probability density we consider is the joint density of the $\epsilon_{i,t+1}$ restricted to the positive values :

$$\sup_I \mathbb{E}_t(\frac{1}{n}\sum_{i=1}^{n}(I_{i,t}(\epsilon_{i,t+1}) - P_{i,t} - \epsilon_{i,t+1}) - \frac{\kappa}{2}\frac{1}{n}(\sum_{i=1}^{n}(I_{i,t}(\epsilon_{i,t+1}) - \epsilon_{i,t+1}))^2 \tag{30}$$

equal to

$$\sup_I \mathbb{E}_t(\frac{1}{n}\sum_{i=1}^{n}\left(-\frac{c}{2}\sigma_I^2 - \epsilon_{it+1}\right) - \frac{\kappa}{2}(\frac{1}{n}\sum_{i=1}^{n}((I_{i,t}(\epsilon_{t+1}) - \epsilon_{t+1}))^2) \tag{31}$$

Because of the symmetry of all the ϵ_i, the solution must be symmetric as well, hence all the indemnity are equal and we find the following function to maximise (with simplified notations) :

$$U(I) = -\frac{c}{2n}\sum_{i=1}^{n}\mathbb{E}((I(\epsilon_i) - \mathbb{E}(I(\epsilon_i)))^2) - \frac{\kappa}{2n^2}\mathbb{E}((\sum_{i=1}^{n}(I(\epsilon_i) - \epsilon_i))^2) \tag{32}$$

Which differential is, with $i : \mathbb{R} \longrightarrow \mathbb{R}$:

$$\delta U_I(i) = -\mathbb{E}(\frac{c}{n}\sum_{i=1}^{n}(I(\epsilon_i) - \mathbb{E}(I(\epsilon_i)))i(\epsilon_i)) + \frac{\kappa}{2n^2}\sum_{i=1}^{n}2(I(\epsilon_i)-\epsilon_i)i(\epsilon_i) + \sum_{n,n}^{i\neq j}((I(\epsilon_i)-\epsilon_i)i(\epsilon_j) + (I(\epsilon_j)-\epsilon_j)i(\epsilon_i))) \tag{33}$$

from which we get the following equality :

$$\left(c + \frac{\kappa}{n}\right)I(x) + \frac{\kappa(n-1)}{n}\mathbb{E}(I(y)|x) = c\mathbb{E}(x) + \frac{\kappa}{n}x + \frac{\kappa(n-1)}{n}\mathbb{E}(y|x) \tag{34}$$

Where x is the loss on a crop and y is a random variable representing the impact of the correlation between crops.

F Empirical comparison

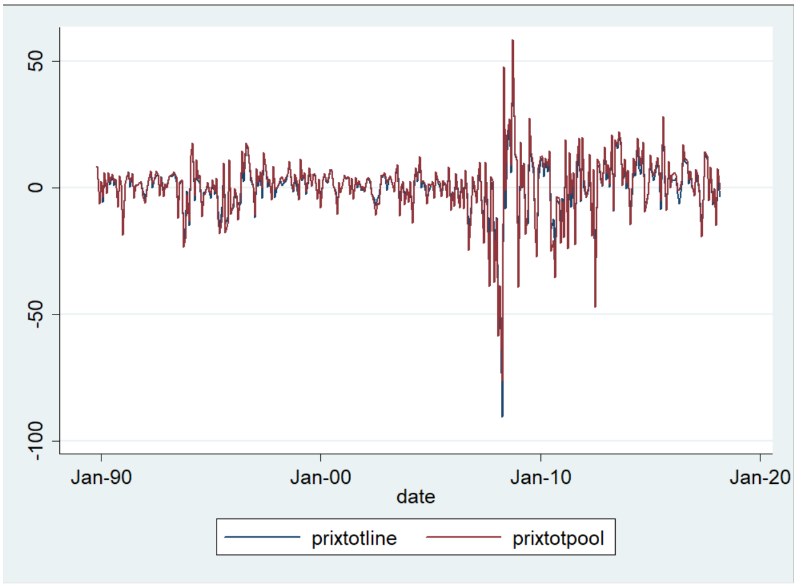

Figure 20: Gain of the firm for the in-line (blue) and pooling (red) strategies, with full insurance.

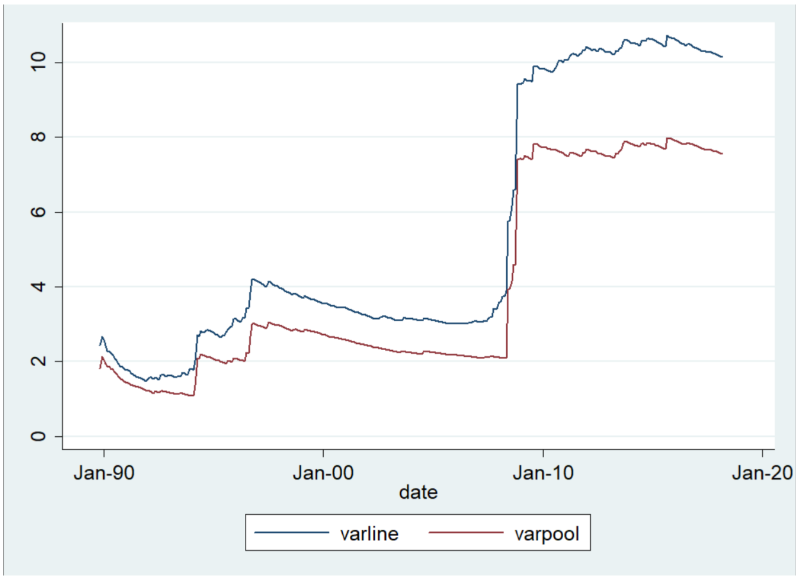

Figure 21: $\frac{\kappa}{2}Var(G)$, where G is the gain of a risk-neutral firm. for the in-line (blue) and pooling (red) strategies. At each t, the quantity is evaluated from the previous data available. Results obtained in full insurance

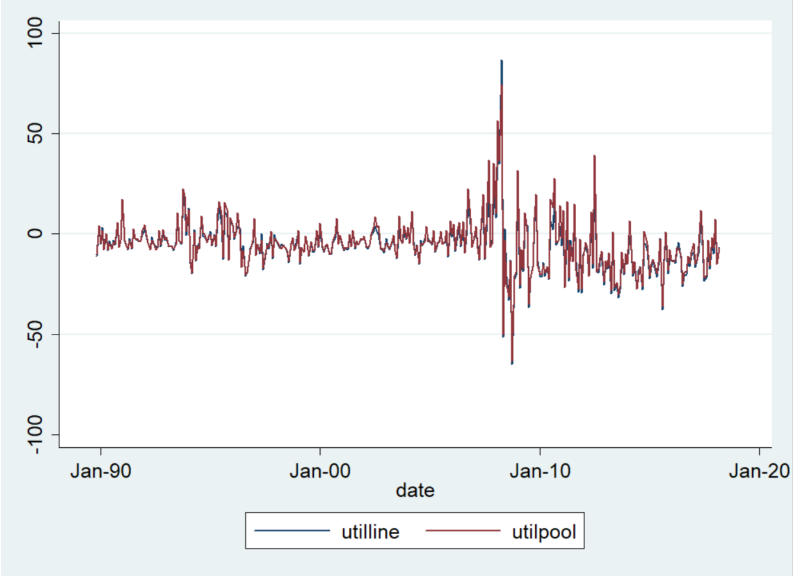

Figure 22: Gain of the firm, each month, for each strategy, in full insurance

References

[1] U.S. Department of Labor Bureau of Labor Statistics. Three factors led to 1996 grain price shock. *The Economics Daily*, 1998.

[2] Ralph B D'agostino, Albert Belanger, and Ralph B D'Agostino Jr. A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44(4):316–321, 1990.

[3] Francis X Diebold. *Elements of forecasting*. Citeseer, 1998.

[4] Robert F Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

[5] Gloria González-Rivera. *Forecasting for Economics and Business: Instructor's Review Copy*. Routledge, 2016.

[6] Chengguo Weng Jingong Zhang, Ken Seng Tan. Index insurance design.

[7] Picard P. Louaas A. Securitizing the supply chain risk. 2018.

[8] Council on Foreign Relations. Food price volatility and insecurity. 2013.

[9] Galen R Shorack and Jon A Wellner. *Empirical processes with applications to statistics*, volume 59. Siam, 2009.

[10] Stephen D Treanor, David A Carter, Daniel A Rogers, and Betty J Simkins. Operational and financial hedging: friend or foe? evidence from the us airline industry. *Journal of Accounting and Finance*, 13(6):64–91, 2013.