

# Les Cahiers de la Chaire / N°54

A simple equilibrium model for a commodity market with spot trades and futures contracts

Ivar Ekeland, Delphine Lautier & Bertrand Villeneuve



# A simple equilibrium model for a commodity market with spot trades and futures contracts\*

Ivar Ekeland

Université Paris-Dauphine (CEREMADE)

Delphine Lautier

Université Paris-Dauphine (DRM)

Bertrand Villeneuve

Université Paris-Dauphine (LEDA)

December 21, 2012

## Abstract

We propose a simple model which offers a unified theoretical framework for the analysis of price and quantity relationships in commodity markets. We study the simultaneous equilibrium in the physical and futures markets. We demonstrate the existence and uniqueness of this equilibrium and we provide explicit expressions. We provide insights into the hedging function of the futures market and the informational role of prices. The model is particularly efficient for precise qualitative and quantitative comparative statics. Among other possibilities, we compare equilibrium variables with and without futures markets and we show that the level and volatility of spot prices increases with the number of speculators. We also provide neat predictions on the political economy of potential reforms of market structure.

## 1 Introduction

In the literature on commodity derivatives, the analysis of price relationships is split into two strands: one focuses on the cost of storage of the underlying asset; the other is centered on the risk premium. Although they are complementary, until now, these two strands remained apart. In this paper, we propose a model of commodity markets which offers a unified framework.

In this simple (perhaps the simplest possible) model of commodity trading, the financial market interacts with the physical market. There are two periods, a single commodity, a numéraire, and two markets: the spot market at time  $t = 1$  and  $t = 2$ , and the futures market, which is open at  $t = 1$  and settled at  $t = 2$ . The spot market is physical (no shorting is allowed: there is a non negativity constraint on inventories), while the futures market is financial (both long and short positions are allowed). There are three types

---

\*The authors acknowledge conversations with Larry Karp and Eugenio Bobenrieth in 2011, and comments from audiences at Paris-Dauphine (FIME Lab) and Zurich (ETH) and Montreal (IAES) during 2012. This article is based upon work supported by the Chaire Finance et Développement Durable.

of traders: inventory holders and industrial processors of the commodity, both of which operate on the two markets, and money managers, who operate on the futures market only. All of them are utility maximizers, and have mean-variance utility.<sup>1</sup> There is also a background demand (or supply), attributed to spot traders, which helps clear the spot markets. The sources of uncertainty are the amount of commodity produced and the demand of the spot traders at  $t = 2$ : their realization is unknown at  $t = 1$ , but their law is common knowledge. As only the difference between these two quantities matters, there is only one source of uncertainty. All decisions are taken at  $t = 1$ , conditionally on expectations about  $t = 2$ .

Our main contributions are three: qualitative, quantitative and normative. They are the consequences of the tractability of the model.

Qualitatively, we give necessary and sufficient conditions on the fundamentals of this economy for a rational expectations equilibrium to exist, and we show that it is unique. We characterize the four possible regimes in equilibrium, given the non-negativity constraints on physical positions. While each of these four regimes is simple to understand on economical grounds, we believe that our model is the first to allow them in a unified framework and to give explicit conditions on the fundamentals of the economy determining which one will actually prevail in equilibrium. In each of the regimes, we give explicit formulas for the equilibrium prices. This enables us to characterize regimes in detail and to perform complete and novel comparative statics. For instance, whether there is a contango (the “current basis”, defined as the difference between the futures price and the *current* spot price, is positive) or a backwardation (the current basis is negative). We can also compare the futures price with the expected spot price and ask whether or not there is a bias in the futures price (we define the “expected basis” as the difference between the futures price and the *expected* spot price). The sign and the level of the bias depend directly on which regime prevails. In the third one, for instance, there is no bias; in the first regime, there are two sub-regimes, one where the futures price is higher than the expected spot price, and one where it is lower. Here the model depicts the way futures market are used to reallocate risk between operators, the price to pay for such a transfer, and thus provides insights into the other main economic function of derivative markets: hedging against price fluctuations.

Quantitatively, we show that, as the importance of money managers in the futures market increases, for example because access to the market is relaxed, the volatility of prices goes up. This effect has empirical support, and it may sound inefficient. Our interpretation is that speculation increases the informativeness of prices: volatility brings more efficiency. The mechanism is quite simple. As the number of speculators increases, the cost of hedging decreases and demand for futures grows along with physical positions. Smaller hedging costs make storers and processors amplify the differences in their positions in response to different pieces of information, implying that their market impact increases. This increases in turn the volatility of prices.

Normatively, we use our model to perform a welfare analysis. For instance, we can show that there is an optimal number of speculators for speculators themselves. Storers and

---

<sup>1</sup>This assumption is very powerful for calculability, as we shall show. We don't think that rationalizing this assumption (as other authors do) with CARA utility functions and normal distributions is satisfactory. We rather see mean-variance utility maximization as a decent approximation of plausible behavior, and it is flexible enough to allow any distribution of random variables, comprised those with bounded support.

processors have opposite views on the desirability of speculators. Indeed, speculators are worthless when the positions of storers and processors match exactly; but when one type of agents has more needs than the other type can supply, then the former wants more (the latter wants less) speculators because this reduces the cost of hedging for those who need it most.

**Short literature review** The latter effect had not been investigated before. There is, of course, a vast amount of literature on the first two subjects, the questions we have raised have been investigated in other contexts. Contrary to what is done in this paper, the literature on commodity prices indeed usually separates the question of the links between the spot and the futures prices and that of the bias in the futures price. The latter has been investigated first by Keynes (1930) through the theory of normal backwardation (today known as the hedging pressure theory) whereas the former is usually associated to the theory of storage, initiated by Kaldor (1940), Brennan (1958) and Working (1949). The same is true for the equilibrium models developed so far.

An important number of equilibrium models of commodity prices focuses on the bias in the futures price and the risk transfer function of the derivative market. This is the case, for example, of Anderson and Danthine (1983a), Anderson and Danthine (1983b), Hirshleifer (1988), Hirshleifer (1989), Guesnerie and Rochet (1993), and Acharya et al. (2011).

The model proposed by Anderson and Danthine (1983a) is an important source. Their main question is whether or not the futures price coincides with the expected value of the spot price for the date when the futures contract matures. To answer it, they examine the relationships between fundamental economic structures (producers, processors and storage companies) and the bias reflected in equilibrium futures prices. They also discuss the partial equilibrium impact of uncertainties on the quantities produced and processed, as well as that of some rigidities in the production process. Compared with this work, our model is more simple (for example the producers are not directly modeled) and completely specified, whereas Anderson and Danthine (1983a) stay at a more general level. This gives us the possibility to obtain explicit formulas for the equilibrium prices and to investigate further economics issues.

The models developed by Hirshleifer (1988) and Hirshleifer (1989) are also inspired by Anderson and Danthine (1983a): they focus on the bias in the futures price and they also propose a framework in which there is a joint equilibrium in spot and futures markets. In these papers, Hirshleifer analyzes two points which are of particular interest for our model: first, the simultaneous existence of futures and forward markets; second, the role of the spot traders. Hirshleifer (1989) studies in particular whether or not vertical integration and futures trading can be substitute means of diversifying risk. We focus instead, in the comparative statics, on the impact type of agent by type of agent, with a rich variety of cases.

Let us also mention that, contrary to Anderson and Danthine (1983b), Hirshleifer (1989) and Routledge et al. (2000), we do not undertake an inter-temporal analysis in the present version of the model. The article of Anderson and Danthine (1983b) is a natural extension of Anderson and Danthine (1983a). To obtain results while keeping tractable equations, the authors however must simplify their model so that only one category of hedger remains in the inter-temporal version. When equilibrium analysis stands at the heart of

all concerns (which is our case), this is a strong limitation. Routledge et al. (2000) give another interesting example of intertemporal analysis. Their model is indeed compatible with the analysis of the whole term structure and it makes predictions about volatilities of futures prices at different horizons. In this sense, the model tends towards realism in the description of prices processes, but is not adapted to normative analysis.

Beyond the question of the risk premium, equilibrium models in the same spirit than Anderson and Danthine (1983a) have also been used in order to examine the possible destabilizing effect of the presence of a futures market. Through the study of mental (“eductive”) coordination strategies, Guesnerie and Rochet (1993) show that speculation on futures markets reduces the likelihood of occurrence of a stable rational expectations equilibrium: although the equilibrium price is less volatile after the futures market is opened (which is usually viewed as a stabilizing effect), traders may find it more difficult or even impossible to coordinate their expectations in order to implement this equilibrium. Apart from this original contribution, the model is classic and more stripped down than ours. Our explicit formulas for equilibrium prices allows for interesting comparisons depending on the presence or absence of a futures market, as in Newbery (1987).

Another strand of the literature on equilibrium models focuses on the *current* spot price and the role of inventories in the behavior of commodity prices, as well as on the relationship between the spot and the futures prices. With their rational equilibrium model, Deaton and Laroque (1992) pay particular attention on an essential characteristics of commodity markets: the non negativity constraint on inventories. They apply their model to thirteen commodity markets and they use it to show how one could explain (and obtain) well-known price properties of commodity prices, like skewness, volatility jumps and volatility clustering. The model developed by Chambers and Bailey (1996) is conceived and employed in the same spirit as Deaton and Laroque (1992), with important differences in the stochastic specification. In these models, there are no futures market; there is in fact a single type of representative agent, which prevents examining risk allocation and the political economy of structural change.

Apart from the specific behavior of prices, the non-negativity constraint on inventories raises another issue. Empirical facts indeed testify that there is more than a non-negativity constraint in commodity markets: the level of inventories never falls to zero, leaving thus unexploited some supposedly profitable arbitrage opportunities. The concept of a convenience yield associated with inventories, initially developed by Kaldor (1940) and Brennan (1958) is generally used to explain such a phenomenon, which has been regularly confirmed, on an empirical point of view, since Working (1949). In their model, Routledge et al. (2000) introduce a convenience yield in the form of an embedded timing option associated with physical stocks. Contrary to these authors, we do not take into account the presence of a convenience yield in our analysis. While this would probably constitute an interesting improvement of our work, it is hardly compatible with a two-period model.

Recent attempts to test equilibrium models must also be mentioned, as they are rare. Acharya et al. (2011) build an equilibrium model of commodity markets in which speculators are capital constrained. The limits to arbitrage induced by such constraints is at the core of their analysis. On the basis of this model, they show that limits to financial arbitrage generate limits to hedging by producers and affect good prices. To test their model, they employ data on spot and futures prices for heating oil, gasoline and natural gas over the period 1980 to 2006. Interestingly, they propose proxies for changes in producer’s risk aversion and hedging demand: they indeed use movements in the default

risk of commodity producing firms. Contrary to these authors, we do not examine the limits to arbitrage for speculators and their possible consequences. Neither do we perform empirical tests. However, while not totally operable in our context, the tests undertaken by Acharya et al. (2011) could be used as an interesting source of inspiration for further developments.

Our contributions to the extant literature on commodity prices in the presence of a derivative market are the following: first, we offer a unified theoretical framework, which comprises the hedging pressure theory as well as the storage theory, where the questions of the hedging and the informational functions performed by the derivative market can be simultaneously assessed; second, the study of the simultaneous equilibrium in the physical and in the futures markets leads us to demonstrate the existence and the uniqueness of this equilibrium. We also give an analysis of this equilibrium in commodity in other settings: without storage and without futures markets. Finally, we perform a welfare analysis.

## 2 The model

This is a two-period model. There is one commodity, a numéraire, and two markets: the spot markets at times  $t = 1$  and  $t = 2$ , and a futures market, which is open at  $t = 1$  and such that contracts are settled at time  $t = 2$ . It is important to note that short positions are allowed on the futures market. When an agent sells (resp. buys) futures contracts, his position is short (resp. long), and the amount  $f$  he holds is negative (resp. positive). On the spot market, such positions are not allowed: you can't sell what you don't hold. In other words, the futures is a financial market, while the spot is a physical market.

There are three types of traders.

- *Industrial users*, or processors, who use the commodity to produce other goods which they sell to consumers. Because of the inertia of their own production process, and/or because all their production is sold forward, they decide at  $t = 1$  how much to produce at  $t = 2$ . They cannot store the commodity, so they have to buy all of their input on the spot market at  $t = 2$ . They also trade on the futures market.
- *Inventory holders*, who have storage capacity, and who can use it to buy the commodity at  $t = 1$  and release it at  $t = 2$ . They trade on the spot market at  $t = 1$ , where they buy, and at  $t = 2$ , where they sell. They also operate on the futures market.
- *Money managers*, or speculators, who use the commodity price as a source of risk, to make a profit on the basis of their positions in futures contracts. They do not trade on the spot market.

In addition, we think of these markets as operating in a partial equilibrium framework: in the background, there are other users of the commodity, and producers as well. These additional agents will be referred to as *spot traders*, and their global effect will be described by a demand function. At time  $t = 1$ , the demand is  $\mu_1 - mP_1$ , and it is  $\tilde{\mu}_2 - m\tilde{P}_2$  at time  $t = 2$ .  $P_t$  is the spot price at time  $t$  and the demand can be either positive or negative; the sign  $\sim$  indicates a random variable.

All decisions are taken at time  $t = 1$ , conditional on the information available for  $t = 2$ . The timing is as follows:

- for  $t = 1$ , the commodity is in total supply  $\omega_1$ , the spot market and the futures market open. On the spot market, there are spot traders and storers on the demand side, the price is  $P_1$ . On the futures markets, the processors, the storers and the speculators all initiate a position, and the price is  $P_F$ . Note that the storers have to decide simultaneously how much to buy on the spot market and what position to take on the futures market.
- for  $t = 2$ , the commodity is in total supply  $\tilde{\omega}_2$ , to which one has to add the inventory which the storers carry from  $t = 1$ , and the spot market opens. The processors and the spot traders are on the demand side, and the price is  $\tilde{P}_2$ . The futures contracts are then settled at that price, meaning that every contract brings a financial result of  $\tilde{P}_2 - P_F$ .

There are  $N_S$  speculators,  $N_P$  processors,  $N_I$  storage companies ( $I$  for inventories). We assume that all agents (except the spot traders) are risk averse inter-temporal utility maximizers. To take their decisions at time  $t = 1$ , they need to know the distribution of the spot price  $\tilde{P}_2$  at  $t = 2$ . We will show that, under mean-variance specifications of the utilities, there is a unique price system  $(P_1, P_F, \tilde{P}_2)$  such that all three markets clear.

Uncertainty is modeled by a probability space  $(\Omega, \mathcal{A}, P)$ . Both  $\tilde{\omega}_2$ ,  $\tilde{\mu}_2$  and  $\tilde{P}_2$  are random variables on  $(\Omega, \mathcal{A}, P)$ . At time  $t = 1$ , their realizations are unknown, but their distributions are common knowledge.

Before we proceed, some clarifications are in order.

- Production of the commodity is inelastic: the quantities  $\omega_1$  and  $\tilde{\omega}_2$  which reach the spot markets at times  $t = 1$  and  $t = 2$  are exogenous to the model. Traders know  $\omega_1$  and  $\mu_1$ , and share the same priors as to  $\tilde{\omega}_2$  and  $\tilde{\mu}_2$ .
- This said, a negative spot demand can be understood as extra spot supply: if for instance  $P_1 > \mu_1/m$ , then the spot price at time  $t = 1$  is so high that additional means of production become profitable, and the global economy provides additional quantities to the spot market. The coefficient  $m$  is the elasticity of demand (or production) with respect to prices. The number  $\mu_1$  (demand when  $P_1 = 0$ ) is the level at which the economy saturates: to induce spot traders to demand quantities larger than  $\mu_1$ , one would have to pay them, that is, offer negative price  $P_1 < 0$  for the commodity. The same remark applies to time  $t = 2$ .
- We separate the roles of the industrial user and the inventory holder, whereas in reality industrial users may also hold inventory. It will be apparent in the sequel that this separation need not be as strict, and that the model would accommodate agents of mixed type. In all cases, agents who trade on the physical markets would also trade on the financial market for two separate purposes: hedging their risk, and making additional profits. In the sequel, we will see how their positions reflect this dual purpose.
- Note also that the speculators would typically use their position on the futures market as part of a diversified portfolio; our model does not take this into account.

- We also suppose that there is a perfect convergence of the basis at the expiration of the futures contract. Thus, at time  $t = 2$ , the position on the futures markets is settled at the price  $\tilde{P}_2$  then prevailing on the spot market.
- For the sake of simplicity, we set the risk-free interest rate to 0.

In what follows, as we examine an REE (rational expectation equilibrium), we look at two necessary conditions for such an equilibrium to appear: the maximization of the agent's utility, conditionally on their price expectations, and market clearing.

## 3 Markets

### 3.1 Utility maximization

All agents have mean-variance utilities. For all of them, a profit  $\tilde{\pi}$  brings utility:

$$E[\tilde{\pi}] - \frac{1}{2}\alpha_i \text{Var}[\tilde{\pi}].$$

where  $\alpha_i$  is the risk aversion parameter of a type  $i$  individual.

**Speculator** For the speculator, the profit resulting from a position in the futures market  $f_S$  is the r.v.:

$$\pi_S(f_S) = f_S(\tilde{P}_2 - P_F),$$

and the optimal position is:

$$f_S^* = \frac{E[\tilde{P}_2] - P_F}{\alpha_S \text{Var}[\tilde{P}_2]}. \quad (1)$$

This position is purely speculative. It depends mainly on the level and on the sign of the bias in the futures price. The speculator goes long whenever he thinks that the expected spot price is higher than the futures price. Otherwise he goes short. Finally, he is all the more inclined to take a position as his risk aversion and volatility of the underlying asset are low.

**Storer** The storer can hold any non-negative inventory. However, storage is costly: holding a quantity  $x$  between  $t = 1$  and  $t = 2$  costs  $\frac{1}{2}Cx^2$ . Parameters  $C$  (cost of storage) and  $\alpha_I$  (risk aversion) characterize the storer. He has to decide how much inventory to buy at  $t = 1$ , if any, and what position to take in the futures market, if any.

If he buys  $x \geq 0$  on the spot market at  $t = 1$ , resells it on the spot market at  $t = 2$ , and takes a position  $f_I$  on the futures market, the resulting profit is the r.v.:

$$\pi_I(x, f_I) = x(\tilde{P}_2 - P_1) + f_I(\tilde{P}_2 - P_F) - \frac{1}{2}Cx^2.$$

The optimal position on the physical market is:

$$x^* = \frac{1}{C} \max\{P_F - P_1, 0\}. \quad (2)$$

The storer holds inventories if the futures price is higher than the current spot price. This position is the only one, in the model, that directly links the spot and the futures prices. This is consistent with the theory of storage and, more precisely, its analysis of contango and the informational role of futures prices.

The optimal position on the futures market is:

$$f_I^* = \frac{E[\tilde{P}_2] - P_F}{\alpha_I \text{Var}[\tilde{P}_2]} - x^*. \quad (3)$$

This position can be decomposed into two elements. First, a negative position  $-x^*$ , which simply hedges the physical position: the storer sells futures contracts in order to protect himself against a decrease in the spot price. Second, a speculative position, structurally identical to that of the speculator, which reflects the storer's risk aversion and his expectations about the relative level of the futures and the expected spot prices.

**Processor** The processor decides at time  $t = 1$  how much input  $y$  to buy at  $t = 2$ , and which position  $f_P$  to take on the futures market. The revenue from sales at date  $t = 2$  is  $(y - \frac{\beta}{2}y^2)P$ , where  $P$  is our convention for the forward price of the output, and the other factor reflects decreasing marginal revenue. Due to these forward sales of the production, this revenue is known at time  $t = 1$ . The resulting profit is the r.v.:

$$\pi_P(y, f_P) = \left( y - \frac{\beta}{2}y^2 \right) P - y\tilde{P}_2 + f_P(\tilde{P}_2 - P_F).$$

An easy computation then gives his optimal decisions, namely:

$$y^* = \frac{1}{\beta P} \max\{P - P_F, 0\}, \quad (4)$$

$$f_P^* = \frac{E[\tilde{P}_2] - P_F}{\alpha_P \text{Var}[\tilde{P}_2]} + y^*. \quad (5)$$

The futures market is also used by the processor to plan his production, all the more so if the price of his input  $P_F$  is below that of his output  $P$ . The position on the futures market, again, can be decomposed into two elements. First, a positive position  $y^*$ , which hedges the position on the physical market: the processor goes long on futures contracts in order to protect himself against an increase in the spot price. Then, a speculative position reflecting the processor's risk aversion and his expectations about the level of the expected basis.

**Remarks on optimal positions** In this framework, all agents have the possibility to undertake speculative operations. After having hedged 100 percent of their physical positions, they adjust this position according to their expectations. The separation of the physical and the futures decisions was derived by [Danthine \(1978\)](#). As shown by [Anderson and Danthine \(1983a\)](#), this property does not hold if the final good price is stochastic, unless a second futures market for the final good is introduced. As we shall see, this separation result is very convenient for equilibrium analysis. This is one of the reasons why we choose, for the processor, not to introduce uncertainty on the output price and/or on the quantities produced.

## 3.2 Market clearing

Although we assume that all individuals are identical in each category of agents, more subtle assumptions could be retained without much complication. For example, remark that if the storers had different technologies, say, storer  $i$  with  $i = 1, \dots, N_I$  had technology  $C_i$ , then, instead of  $\frac{N_I}{C} \max\{P_F - P_1, 0\}$ , total inventories would be  $(\sum_i 1/C_i) \max\{P_F - P_1, 0\}$ . In other words, storers are easily aggregated. In the following, when relevant, we shall use the index  $n_I$  representing a synthetic number of storage units, and per-unit inventories  $X^*$  defined by:

$$n_I = \begin{cases} N_I/C & \text{if storers are identical,} \\ \sum_i 1/C_i & \text{otherwise,} \end{cases}$$

$$X^* = \max\{P_F - P_1, 0\}.$$

Similarly, if producers had different technologies, say, producer  $i$  with  $i = 1, \dots, N_P$  had technology  $\beta_i$ , then total input demand would be  $\sum_i 1/(\beta_i P) \cdot \max\{P - P_F, 0\}$  instead of  $\frac{N_P}{\beta P} \max\{P - P_F, 0\}$ . Thus, when relevant, we shall use the index  $n_P$  representing a synthetic number of production units, and per-unit demand  $Y^*$  defined by:

$$n_P = \begin{cases} \frac{N_P}{\beta P} & \text{if producers are identical,} \\ \frac{1}{P} \sum_i \frac{1}{\beta_i} & \text{otherwise,} \end{cases}$$

$$Y^* = \max\{P - P_F, 0\}.$$

**The spot market at time 1** On the supply side we have the harvest  $\omega_1$ . On the other side we have the inventory  $n_I X^*$  bought by the storers, and the demand of the spot traders. Market clearing requires:

$$\omega_1 = n_I X^* + \mu_1 - mP_1,$$

hence:

$$P_1 = \frac{1}{m} (\mu_1 - \omega_1 + n_I X^*). \quad (6)$$

**The spot market at time 2** We have, on the supply side, the harvest  $\tilde{\omega}_2$ , and the inventory  $n_I X^*$  sold by the storers; on the other side, the input  $n_P Y^*$  bought by the processors and the demand of the spot traders. The market clearing condition is:

$$\tilde{\omega}_2 + n_I X^* = n_P Y^* + \tilde{\mu}_2 - m\tilde{P}_2,$$

with  $X^*$  and  $Y^*$  as above. We get:

$$\tilde{P}_2 = \frac{1}{m} (\tilde{\mu}_2 - \tilde{\omega}_2 - n_I X^* + n_P Y^*). \quad (7)$$

**The futures market** Market clearing requires:

$$N_S f_S^* + N_P f_P^* + N_I f_I^* = 0.$$

Replacing the  $f_i^*$  by their values, we get:

$$E[\tilde{P}_2] - P_F = \frac{\text{Var}[\tilde{P}_2]}{\frac{N_P}{\alpha_P} + \frac{N_I}{\alpha_I} + \frac{N_S}{\alpha_S}} (n_I X^* - n_P Y^*). \quad (8)$$

Remark that if, say, different storers had different risk aversions  $\alpha_{Ij}$  (for  $j = 1, \dots, N_I$ ), then we would see  $\sum_j 1/\alpha_{Ij}$  instead of  $N_I/\alpha_I$  in equation (8). This is an illustration of a more general fact: we sum up the inverse of the risk aversions of all agents to represent the inverse of the overall (or market) risk aversion.

Equation (8) gives a formal expression for the bias in the futures price, which confirms the findings of [Anderson and Danthine \(1983a\)](#). It shows indeed that the bias depends primarily on fundamental economic structures (the characteristics of the storage and production functions, which are embedded in  $X^*$  and  $Y^*$ ) and the number of operators, secondarily on subjective parameters (the risk aversion of the operators), and thirdly on the volatility of the underlying asset. Note also that the sign of the bias depends only on the sign of  $(n_I X^* - n_P Y^*)$ . As the risk aversion of the operators only influences the speculative part of the futures position, it does not impact this sign. Finally, when  $n_I X^* = n_P Y^*$ , there is no bias in the futures price, and the risk transfer function is entirely undertaken by the hedgers, provided that their positions on the futures market are the exact opposite of each others. Thus the absence of bias is not exclusively the consequence of risk neutrality but may have other structural causes.

## 4 Existence and uniqueness

The equations characterizing the equilibrium are the optimal choices on the physical market (equations (2) and (4)), the clearing of the spot market at dates 1 and 2 (equations (6) and (7)), as well as the clearing of the futures market (8):

$$\begin{cases} X^* = \max\{P_F - P_1, 0\} & (2) \\ Y^* = \max\{P - P_F, 0\} & (4) \\ P_1 = \frac{1}{m} (\mu_1 - \omega_1 + n_I X^*) & (6) \\ \tilde{P}_2 = \frac{1}{m} (\tilde{\mu}_2 - \tilde{\omega}_2 - n_I X^* + n_P Y^*) & (7) \\ P_F = E[\tilde{P}_2] + \frac{\text{Var}[\tilde{P}_2]}{\frac{N_P}{\alpha_P} + \frac{N_I}{\alpha_I} + \frac{N_S}{\alpha_S}} (n_P Y^* - n_I X^*) & (8) \end{cases}$$

Let us also remind that the distribution of  $\tilde{\mu}_2 - \tilde{\omega}_2$  is common knowledge. We introduce the following notations:

$$\begin{aligned} \xi_1 &:= \mu_1 - \omega_1, \\ \tilde{\xi}_2 &:= \tilde{\mu}_2 - \tilde{\omega}_2, \\ \xi_2 &:= E[\tilde{\mu}_2 - \tilde{\omega}_2], \\ \rho &:= 1 + m \frac{\text{Var}[\tilde{\mu}_2 - \tilde{\omega}_2]}{\frac{N_P}{\alpha_P} + \frac{N_I}{\alpha_I} + \frac{N_S}{\alpha_S}} \end{aligned}$$

where  $m$  is the elasticity of demand.

From (7), we can derive useful moments:

$$\mathbb{E}[\tilde{P}_2] = \frac{1}{m} (\xi_2 - n_I X^* + n_P Y^*), \quad (7E)$$

$$\text{Var}[\tilde{P}_2] = \frac{\text{Var}[\xi_2]}{m^2}. \quad (7V)$$

We assume  $\text{Var}[\tilde{\xi}_2] > 0$ , so there is uncertainty on the future availability of the commodity. It is the only source of uncertainty in the model. Likewise, we assume (for the time being) that  $\alpha_P, \alpha_I$  and  $\alpha_S$  all are non-zero numbers. These restrictions will be lifted later on.

## 4.1 Definitions

**Definition 1.** An *equilibrium* is a family  $(X^*, Y^*, P_1, P_F, \tilde{P}_2)$  such that all prices are non-negative, processors, storers and speculators act as price-takers, and all markets clear.

Technically speaking,  $(X^*, Y^*, P_1, P_F, \tilde{P}_2)$  is an equilibrium if equations (2), (4), (6), (7), and (8) are satisfied, with  $X^* \geq 0, Y^* \geq 0, P_1 \geq 0, P_F \geq 0$  and  $\tilde{P}_2(\omega) \geq 0$  for all  $\omega \in \Omega$ . Note that the latter condition depends on the realization of the random variable  $\tilde{P}_2$ , which can be observed only at  $t = 2$ , while the first four can be checked at time  $t = 1$ . This leads us to the following:

**Definition 2.** A *quasi-equilibrium* is a family  $(X^*, Y^*, P_1, P_F, \tilde{P}_2)$  such that all prices except possibly  $\tilde{P}_2$  are non-negative, processors, storers and speculators act as price-takers and all markets clear.

Technically speaking, a quasi-equilibrium is a family  $(X^*, Y^*, P_1, P_F, \tilde{P}_2) \in \mathbb{R}_+^4 \times L^0(\Omega, \mathcal{A}, P)$  such that equations (2), (4), (6), (7) and (8) are satisfied.

We now give two existence and uniqueness results, the first one for quasi-equilibria and the second one for equilibria.

## 4.2 Quasi-equilibrium

**Theorem 1.** *There is a quasi-equilibrium if and only if  $(\xi_1, \xi_2)$  belongs to the region:*

$$\xi_2 \geq -\rho n_P P \quad \text{if } \xi_1 \geq 0, \quad (9)$$

$$\xi_2 \geq -\rho n_P P - ((m + \rho n_P)/n_I + \rho) \xi_1 \quad \text{if } -n_I P \leq \xi_1 \leq 0, \quad (10)$$

$$\xi_2 \geq -(m/n_I + \rho) \xi_1 \quad \text{if } \xi_1 \leq -n_I P, \quad (11)$$

and then it is unique.

*Proof.* To prove this theorem, we begin by substituting equation (7E) in equation (8). We get:

$$mP_F - \rho(n_P Y^* - n_I X^*) = \xi_2. \quad (12)$$

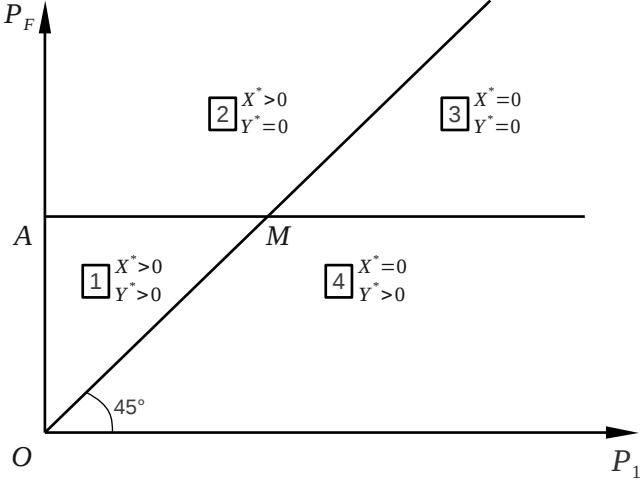


Figure 1: Phase diagram of physical and financial decisions in space  $(P_1, P_F)$ .

We now have two equations, (6) and (12) for  $P_1$  and  $P_F$ . Replacing  $X^*$  and  $Y^*$  by their values, given by (2) and (4), we get a system of two nonlinear equations in two variables:

$$mP_1 - n_I \max \{P_F - P_1, 0\} = \xi_1, \quad (13)$$

$$mP_F + \rho(n_I \max \{P_F - P_1, 0\} - n_P \max \{P - P_F, 0\}) = \xi_2. \quad (14)$$

Remark that if we can solve this system with  $P_1 > 0$  and  $P_F > 0$ , we get  $\tilde{P}_2$  from (7). So the problem is reduced to solving (14) and (13). Consider the mapping  $F : \mathbb{R}_+^2 \rightarrow \mathbb{R}^2$  defined by:

$$F(P_1, P_F) = \begin{pmatrix} mP_1 - n_I \max \{P_F - P_1, 0\} \\ mP_F + \rho(n_I \max \{P_F - P_1, 0\} - n_P \max \{P - P_F, 0\}) \end{pmatrix}.$$

In  $\mathbb{R}_+^2$ , take  $P_1$  as the horizontal coordinate and  $P_F$  as the vertical one, as depicted by Figure 1. There are four regions, separated by the straight lines  $P_F = P_1$  and  $P_F = P$ :

- Region 1, where  $P_F > P_1$  and  $P_F < P$ . In this region, both  $X^*$  and  $Y^*$  are positive.
- Region 2, where  $P_F > P_1$  and  $P_F > P$ . In this region,  $X^* > 0$  and  $Y^* = 0$ .
- Region 3, where  $P_F < P_1$  and  $P_F > P$ . In this region,  $X^* = 0$  and  $Y^* = 0$ .
- Region 4, where  $P_F < P_1$  and  $P_F < P$ . In this region,  $X^* = 0$  and  $Y^* > 0$

Moreover, in the regions where  $X^* > 0$ , we have  $X^* = P_F - P_1$  and in the regions where  $Y^* > 0$ , we have  $Y^* = P - P_F$ . So, in each region, the mapping is linear, and it is obviously continuous across the boundaries. Denote by  $O$  the origin in  $\mathbb{R}_+^2$ , by  $A$  the point  $P_1 = 0, P_F = P$ , and by  $M$  the point  $P_1 = P_F = P$  (so, for instance, region 1 is the triangle  $OAM$ ). In region 1, we have:

$$F(P_1, P_F) = \begin{pmatrix} mP_1 - n_I (P_F - P_1) \\ mP_F + \rho(n_I (P_F - P_1) - n_P (P - P_F)) \end{pmatrix}.$$

The images  $F(O)$ ,  $F(A)$ , and  $F(M)$  are easily computed:

$$\begin{aligned} F(O) &= (0, -\rho n_P P), \\ F(A) &= P(-n_I, m + \rho n_I), \\ F(M) &= m P(1, 1). \end{aligned}$$

From this, one can find the images of all four regions (see Figure 2). The image of region 1 is the triangle  $F(O)F(A)F(M)$ . The image of region 2 is bounded by the segment  $F(A)F(M)$  and by two infinite half-lines, one of which is the image of  $\{P_1 = 0, P_F \geq P\}$ , the other being the image of  $\{P_1 = P_F, P_F \geq P\}$ . In region 2, we have:

$$F(P_1, P_F) = \begin{pmatrix} mP_1 - n_I(P_F - P_1) \\ mP_F + \rho n_I(P_F - P_1) \end{pmatrix}.$$

The first half-line emanates from  $F(A)$  and is carried by the vector  $(-n_I, m + \rho n_I)$ . The second half-lines emanates from  $F(M)$  and is carried by the vector  $(1, 1)$ . Both of them (if extended in the negative direction) go through the origin. The image of region 4 is bounded by the segment  $F(O)F(M)$  and by two infinite half-lines, one of which is the image of  $\{P_F = 0\}$ , the other being the image of  $\{P_1 \geq P, P_F = P\}$ . In region 4, we have:

$$F(P_1, P_F) = \begin{pmatrix} mP_1 \\ mP_F - \rho n_P(P - P_F) \end{pmatrix},$$

so the first half-line emanates from  $F(O)$  and is horizontal, with vertical coordinate  $-\rho n_P P$ , and the second emanates from  $F(M)$  and is horizontal. The image of region 3 is entirely contained in  $\mathbb{R}_+^2$ , where it is the remainder of the three images we described. To prove the theorem, we have to show that the system (14) and (13) has a unique solution. It can be rewritten as:

$$F(P_1, P_F) = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix},$$

and it has a unique solution if and only if the right-hand side belongs to the image of  $F$ , which we have just described. This leads to the conclusion of the proof: based on the previous remark summarized in Figure 2, we easily find the expressions of the theorem.  $\square$

### 4.3 Equilibrium

To get an equilibrium instead of a quasi-equilibrium, we need the further condition, calculated last,  $\tilde{P}_2 \geq 0$ . By equation (7), this is equivalent to:

$$\inf \{\tilde{\mu}_2 - \tilde{\omega}_2\} \geq n_I X^\star - n_P Y^\star. \quad (15)$$

This amount to

$$P_F \leq \frac{\inf \{\tilde{\mu}_2 - \tilde{\omega}_2\} + n_I P_1 + n_P P}{n_I + n_P} \quad \text{in region 1,} \quad (16)$$

$$P_F \leq P_1 + \frac{\inf \{\tilde{\mu}_2 - \tilde{\omega}_2\}}{n_I} \quad \text{in region 2,} \quad (17)$$

$$0 \leq \inf \{\tilde{\mu}_2 - \tilde{\omega}_2\} \quad \text{in region 3,} \quad (18)$$

$$P_F \leq P + \frac{\inf \{\tilde{\mu}_2 - \tilde{\omega}_2\}}{n_P} \quad \text{in region 4.} \quad (19)$$

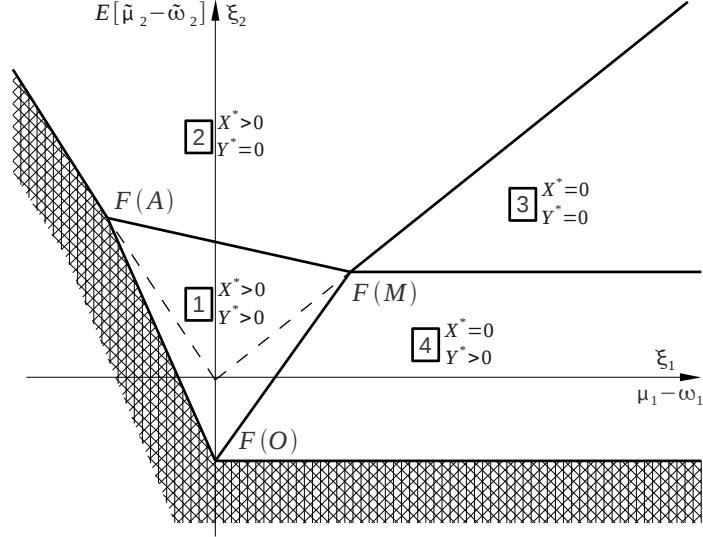


Figure 2: Phase diagram of physical and financial decisions.

**Theorem 2.** Let  $(\xi_1, \xi_2)$  belong to the region (9), (10), (11), so there exists a unique quasi-equilibrium. It is an equilibrium if and only if  $\tilde{\mu}_2 - \tilde{\omega}_2$  satisfies an additional condition, namely:

$$\begin{aligned} \inf \{\tilde{\mu}_2 - \tilde{\omega}_2\} &\geq \frac{n_P(m + n_I)(\xi_2 - mP) + mn_I(\xi_2 - \xi_1)}{n_P(m + n_I)\rho + m(m + (1 + \rho)n_I)} \text{ in region 1;} \\ \inf \{\tilde{\mu}_2 - \tilde{\omega}_2\} &\geq \frac{n_I(\xi_2 - \xi_1)}{m + (1 + \rho)n_I} \text{ in region 2;} \\ \inf \{\tilde{\mu}_2 - \tilde{\omega}_2\} &\geq 0 \text{ in region 3;} \\ \inf \{\tilde{\mu}_2 - \tilde{\omega}_2\} &\geq -\frac{mn_P(P - \frac{\xi_2}{m})}{m + \rho n_P} \text{ in region 4.} \end{aligned}$$

*Proof.* The proof for region 1 comes from applying  $F$  on equation (16). For region 2, a direct application of  $F$  shows that equation (17) implies

$$\xi_2 - \xi_1 \leq \frac{m + (1 + \rho)n_I}{n_I} \inf \{\tilde{\mu}_2 - \tilde{\omega}_2\},$$

which must be read directly as a restriction on  $\inf \{\tilde{\mu}_2 - \tilde{\omega}_2\}$  given  $\xi_2$ . For region 3, the theorem is directly derived from equation (15), since  $X^* = 0$  and  $Y^* = 0$ . For region 4, a direct application of  $F$  shows that equation (19) gives the condition. Note that  $\inf \{\tilde{\mu}_2 - \tilde{\omega}_2\} \geq 0$  is a sufficient condition for an equilibrium to exist in region 4.  $\square$

Remark that the condition for region 1 is general in the following sense. Take  $n_P = 0$ , you get the condition for region 2; take  $n_I = 0$ , you get the condition for region 4; take now  $n_I = n_P = 0$ , you get the condition for region 3. This simple shortcut works for other analytical results.

## 5 Equilibrium analysis

In this section we analyze the equilibrium in two steps. Firstly, we examine the four regimes depicted in Figure 1. They correspond to very different types of decisions undertaken in the physical and the financial markets. Secondly, we turn to Figure 2 and enrich the discussion with the analysis of the net scarcity of the commodity, both immediate and expected.

### 5.1 Prices, physical and financial positions

A first general comment on Figure 1 is that in regimes 1 and 2 where  $X^* > 0$ , the futures market is in contango:  $P_F > P_1$ . Inventories are positive and they can be used for inter-temporal arbitrages. In regimes 3 and 4, there is no inventory ( $X^* = 0$ ) and the market is in backwardation:  $P_F < P_1$ . These configurations are fully consistent with the theory of storage. The other meaningful comparison concerns  $P_F$  and  $E[\tilde{P}_2]$ . From Equation (8), we know that  $n_I X^* - n_P Y^*$  gives the sign and magnitude of  $E[\tilde{P}_2] - P_F$ , i.e. the way risk is transferred between the operators on the futures market.

The analysis of the four possible regimes, with a focus on regime 1 (it is the only one where all operators are active and it gathers two important subcases), enables us to unfold the reasons for the classical conjecture: backwardation on the expected basis, i.e.  $P_F < E[\tilde{P}_2]$ . More interestingly, we show why the reverse inequality is also plausible, as mentioned by several empirical studies.<sup>2</sup>

The equation  $n_I X^* - n_P Y^* = 0$  cuts regime 1 into two parts, 1U and 1L. It passes through  $M$  as can be seen in Figure 3. This frontier can be rewritten as:

$$\Delta : \quad P_F = \frac{n_I}{n_I + n_P} P_1 + \frac{n_P}{n_I + n_P} P. \quad (20)$$

- Along the line  $\Delta$ , there is no bias in the futures price, and the risk remains entirely in the hands of the hedgers (storers and producers have perfectly matching positions).
- Above  $\Delta$ ,  $n_I X^* > n_P Y^*$  and  $P_F < E[\tilde{P}_2]$ . This concerns the upper part of regime 1 (regime 1U) and regime 2.
- Below  $\Delta$ ,  $n_I X^* < n_P Y^*$  and  $P_F > E[\tilde{P}_2]$ . This concerns the lower part of regime 1 (regime 1L) and regime 4.

When  $n_I X^* > n_P Y^*$ , the net hedging position is short and speculators in long position are indispensable to the clearing of the futures market. In order to induce their participation, there must be a profitable bias between the futures price and the expected spot price: the bias  $E[\tilde{P}_2] - P_F$  is positive. This backwardation on the expected basis corresponds to the situation depicted by Keynes (1930) as the normal backwardation theory. On the contrary, when  $n_I X^* < n_P Y^*$ , the net hedging position is necessarily long and the speculators must be short. The expected spot price must be lower than the futures price, and the bias  $E[\tilde{P}_2] - P_F$  is negative.

---

<sup>2</sup>For extensive analyses of the bias in a large number of commodity markets, see for example Fama and French (1987), Kat and Oomen (2006) and Gorton et al. (2012).

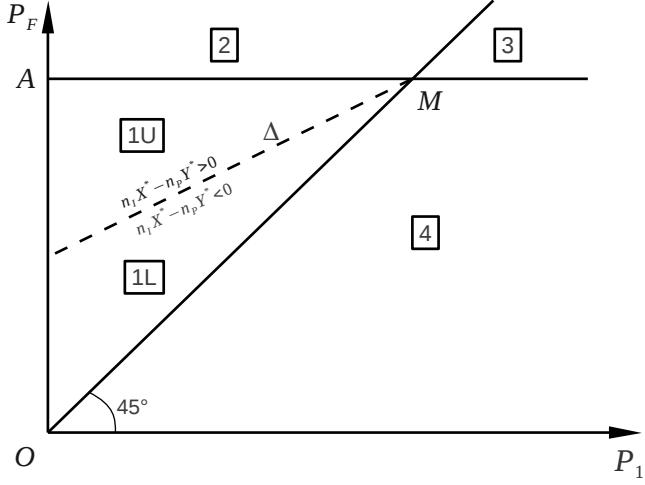


Figure 3: Phase diagram of physical and financial decisions in space  $(P_1, P_F)$  (zoom on Regime 1).

Table 1 summarizes for each regime the relationships between the prices and the physical and financial positions. An attentive scrutiny of the table shows that the regimes are very contrasted.

For example, in regime 2, we have simultaneously a contango on the current basis and a backwardation on the expected basis (or a positive bias). In short,  $P_1 < P_F < E[\tilde{P}_2]$ . In regime 3, in the absence of hedging of any sort, the futures market is dormant, and this is no bias on the expected basis. Regime 4 is the opposite of regime 2: the market is in backwardation and, as  $X^* = 0$ , the net hedging position is long, the net speculative position is short and the bias is negative. In short,  $P_1 > P_F > E[\tilde{P}_2]$ .

## 5.2 Supply shocks

To exploit usefully Figure 2, one must bear in mind that the horizontal and vertical variables measure scarcity, not abundance:  $\xi_1 = \mu_1 - \omega_1$  is the extent to which current production  $\omega_1$  fails short of the demand of spot traders, and  $\xi_2 = E[\tilde{\mu}_2 - \tilde{\omega}_2]$  is the (expected) extent to which future production falls short of the demand of spot traders.

Assume that no markets are open before  $\xi_1$  is realized and assume that  $\xi_1$  brings no news (or revision) about  $\xi_2$ . We can fix  $\xi_2$ , and see what happens on equilibrium variables, depending on  $\xi_1$ . To fix ideas suppose that the expected situation at date 2 is a moderate scarcity, situated at  $\xi_2 = \bar{\xi}_2$ . The level of  $\bar{\xi}_2$  is common knowledge for the operators. Take it as drawn in Figure 4. In the case of a low  $\xi_1$  (abundance in period 1), we are in regime 1U. If  $\xi_1$  is bigger, we are in regime 1L, and if  $\xi_1$  is even bigger, the equilibrium is in regime 4.

The interpretation is straightforward. If period 1 experiences abundance (regime 1U), there is massive storage (the current price is low and expected profits are attractive, since a future scarcity is expected). Storers need more hedging than processors, first because inventories are high, second because the expected release of stocks reduces the needs of the

|          |             |                        |           |
|----------|-------------|------------------------|-----------|
| 1U       | $P_1 < P_F$ | $P_F < E[\tilde{P}_2]$ | $P_F < P$ |
|          | $X^* > 0$   | $f_S > 0$              | $Y^* > 0$ |
| $\Delta$ | $P_1 < P_F$ | $P_F = E[\tilde{P}_2]$ | $P_F < P$ |
|          | $X^* > 0$   | $f_S = 0$              | $Y^* > 0$ |
| 1L       | $P_1 < P_F$ | $P_F > E[\tilde{P}_2]$ | $P_F < P$ |
|          | $X^* > 0$   | $f_S < 0$              | $Y^* > 0$ |
| 2        | $P_1 < P_F$ | $P_F < E[\tilde{P}_2]$ | $P_F > P$ |
|          | $X^* > 0$   | $f_S > 0$              | $Y^* = 0$ |
| 3        | $P_1 > P_F$ | $P_F = E[\tilde{P}_2]$ | $P_F > P$ |
|          | $X^* = 0$   | $f_S = 0$              | $Y^* = 0$ |
| 4        | $P_1 > P_F$ | $P_F > E[\tilde{P}_2]$ | $P_F < P$ |
|          | $X^* = 0$   | $f_S < 0$              | $Y^* > 0$ |

Table 1: Relationships between prices, physical and financial positions.

processors. Thus, there is a positive bias in the futures price and speculators have a buy position. For a less marked abundance (regime 1L), storage is more limited. The hedging needs of the storers diminishes while those of the processors increase. So the net hedging position is long, the bias in the futures price becomes negative and the speculators have a sell position. If the commodity is even scarcer (regime 4), there is no storage, only the processors are active and they hedge their positions.

The combination of the exogenous variables of the model (i.e. current or expected scarcity) with the activities on the physical market makes it possible to create a link between the storage and the normal backwardation theories. For example, it explains why, when there is a contango on the current basis in regime 1, we can have either an expected backwardation or an expected contango.

## 6 Welfare analysis

In this section, we shall express the indirect utilities of the various agents in equilibrium, and compute their sensitivities with respect to the parameters. We proceed in two steps. First, we compute the indirect utilities of the agents in equilibrium, as functions of equilibrium prices  $P_1$  and  $P_F$ . Second, we compute the elasticities of  $P_1$  and  $P_F$  to deduce the elasticities of the indirect utilities. We restrict ourselves to the richer case, *i.e.* Regime 1, where all agents are active. Recall that then we have:

$$P_F < P \text{ and } P_1 < P_F; \quad (21)$$

$$\tilde{P}_2 = \frac{1}{m} (\tilde{\mu}_2 - \tilde{\omega}_2 - n_I (P_F - P_1) + n_P (P - P_F)); \quad (22)$$

$$mP_1 - n_I (P_F - P_1) = \xi_1; \quad (23)$$

$$mP_F + \rho (n_I (P_F - P_1) - n_P (P - P_F)) = \xi_2. \quad (24)$$

As above, we shall set  $\xi_2 := E[\tilde{\mu}_2 - \tilde{\omega}_2]$  and  $\xi_1 := \mu_1 - \omega_1$ .

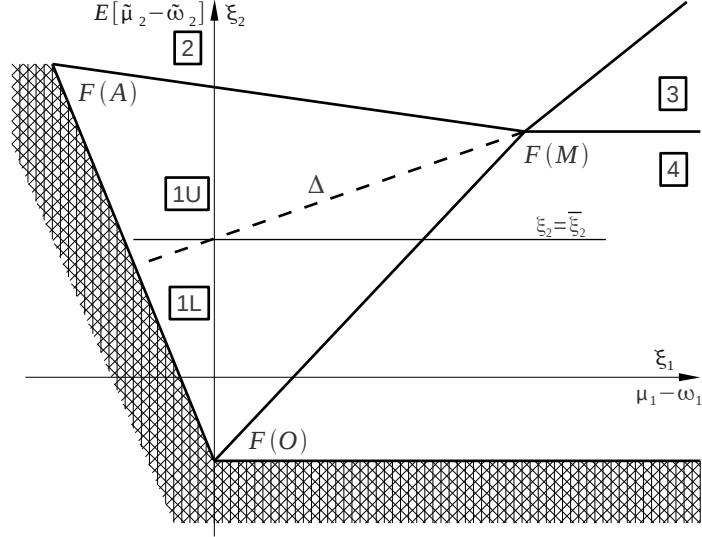


Figure 4: Phase diagram of physical and financial decisions in space (zoom on Regime 1).

## 6.1 Indirect utilities

The indirect utility of the speculators is given by:

$$U_S = f_S^*(E[\tilde{P}_2] - P_F) - \frac{1}{2}\alpha_S f_S^{*2} \text{Var}[\tilde{P}_2],$$

where we have to substitute the value of  $f_S^*$ , which leads to:

$$U_S = \frac{(E[\tilde{P}_2] - P_F)^2}{2\alpha_S \text{Var}[\tilde{P}_2]}. \quad (25)$$

Let us now turn to the storers. Their indirect utility is given by:

$$U_I = (x^* + f_I^*)E[\tilde{P}_2] - x^*P_1 - f_I^*P_F - \frac{1}{2}Cx^{*2} - \frac{1}{2}\alpha_I(x + f_I^*)^2 \text{Var}[\tilde{P}_2],$$

where we substitute the values of  $f_I^*$ ,  $x^*$  and  $y^*$ :

$$U_I = \frac{(E[\tilde{P}_2] - P_F)^2}{2\alpha_I \text{Var}[\tilde{P}_2]} + \frac{(P_F - P_1)^2}{2C}. \quad (26)$$

For the processors we have, in a similar fashion:

$$U_P = \frac{(E[\tilde{P}_2] - P_F)^2}{2\alpha_P \text{Var}[\tilde{P}_2]} + \frac{(P_F - P)^2}{2\beta P}. \quad (27)$$

We thus obtain, for all categories of agents, a clear separation between two additive components of the indirect utilities. The first is associated with the level of the expected basis and is clearly linked with speculation. The second is associated with the level of the

current basis or the futures prices and is linked with the hedged activity on the physical market. We shall name  $U_{Si}$  this first component for the category of agent  $i$ , and  $U_{Hi}$  the second one.

Quite intuitively, for all operators,  $U_{Si}$  is all the more important as the futures market is biased, whatever the sign of the bias; it decreases with respect to risk aversion and to the variance of the expected spot price.  $U_{Hi}$  changes with the category of agent under consideration. For the storers, it is positively correlated to the current basis and diminishes with storage costs. For the processors, it rises with the margin on the processing activity and decreases with the production costs.

We will now particularize formulas (25), (26) and (27) to the case when the markets are in equilibrium. In that case,  $\tilde{P}_2$  becomes a function of  $(P_1, P_F)$ , and the formulas become (after replacing the  $n_i$  by their values in terms of the  $N_i$ ):

$$U_S = \frac{\text{Var}[\tilde{\xi}_2]}{2m^2\alpha_S \left(\sum \frac{N_i}{\alpha_i}\right)^2} \left( \frac{N_I}{C} (P_F - P_1) - \frac{N_P}{\beta P} (P - P_F) \right)^2; \quad (28)$$

$$U_I = \frac{\text{Var}[\tilde{\xi}_2]}{2m^2\alpha_I \left(\sum \frac{N_i}{\alpha_i}\right)^2} \left( \frac{N_I}{C} (P_F - P_1) - \frac{N_P}{\beta P} (P - P_F) \right)^2 + \frac{(P_F - P_1)^2}{2C}; \quad (29)$$

$$U_P = \frac{\text{Var}[\tilde{\xi}_2]}{2m^2\alpha_S \left(\sum \frac{N_i}{\alpha_i}\right)^2} \left( \frac{N_I}{C} (P_F - P_1) - \frac{N_P}{\beta P} (P - P_F) \right)^2 + \frac{(P_F - P)^2}{2\beta P}. \quad (30)$$

Note for future use that these are indirect utilities *per head*: for instance, there are  $N_I$  storers, they are all identical, and  $U_I$  is the indirect utility of each one of them. This will enable us to do a welfare analysis in the next subsection.

## 6.2 The impact of speculators on the welfare of others

Formulas (28), (29) and (30) give us the indirect utilities of the agents at equilibrium in terms of the equilibrium prices  $P_1$  and  $P_F$ . These can in turn be expressed in terms of the fundamentals of the economy, namely  $\xi_1$  and  $\tilde{\xi}_2$  (see Appendix A): substituting formulas (43), (44) and (45), we get new expressions, which can be differentiated to give the sensitivities of the indirect utilities with respect to the parameters in the model.

However, it is better to work directly with formulas (28), (29) and (30). We will then need the sensitivities of  $P_1$  and  $P_F$  with respect to the varying parameter, but these can be derived from the system (23)-(24) by the implicit function theorem. To see how it is done, let us compute the sensitivities with respect to  $N_S$ , the number of speculators. In other words, we will investigate whether an increase in the number of speculators increases or decreases the welfare of speculators, of inventory holders, and of industry processors.

**Sensitivities of prices** We first compute the sensitivities  $\frac{dP_1}{dN_S}$  and  $\frac{dP_F}{dN_S}$ . We get them by differentiating (23)-(24):

$$m \frac{dP_1}{dN_S} - n_I \left( \frac{dP_F}{dN_S} - \frac{dP_1}{dN_S} \right) = 0,$$

$$m \frac{dP_F}{dN_S} + \rho \left( n_I \left( \frac{dP_F}{dN_S} - \frac{dP_1}{dN_S} \right) + n_P \frac{dP_F}{dN_S} \right) = - \frac{d\rho}{dN_S} (n_I (P_F - P_1) - n_P (P - P_F)),$$

which yields:

$$\frac{dP_F}{dN_S} = \left( \frac{m}{n_I} + 1 \right) \frac{dP_1}{dN_S}, \quad (31)$$

$$\begin{aligned} \frac{dP_1}{dN_S} &= \frac{d\rho}{dN_S} \frac{n_P (P - P_F) - n_I (P_F - P_1)}{\left( \frac{m}{n_I} + 1 \right) (m + \rho n_I + \rho n_P) - \rho n_I} \\ &= - \frac{m}{\alpha_S} \frac{\text{Var}[\tilde{\xi}_2]}{\left( \sum \frac{N_i}{\alpha_i} \right)^2} \frac{n_P (P - P_F) - n_I (P_F - P_1)}{(m + \rho n_I + \rho n_P) - \rho n_I}. \end{aligned} \quad (32)$$

**Sensitivity of  $U_S$**  Differentiating formula (28) yields:

$$\begin{aligned} \frac{dU_S}{dN_S} &= - \frac{\text{Var}[\tilde{\xi}_2]}{m^2 \alpha_S \left( \sum \frac{N_i}{\alpha_i} \right)^2} (n_I (P_F - P_1) - n_P (P - P_F)) \left( m + n_P \left( 1 + \frac{m}{n_I} \right) \right) \frac{dP_1}{dN_S} \\ &\quad - \frac{\text{Var}[\tilde{\xi}_2]}{m^2 \alpha_S^2 \left( \sum \frac{N_i}{\alpha_i} \right)^3} (n_I (P_F - P_1) - n_P (P - P_F))^2 \\ &= - \frac{\text{Var}[\tilde{\xi}_2]}{m^2 \alpha_S^2 \left( \sum \frac{N_i}{\alpha_i} \right)^3} \left( \frac{\text{Var}[\tilde{\xi}_2]}{\sum \frac{N_i}{\alpha_i}} \frac{m \left( m + n_P \left( 1 + \frac{m}{n_I} \right) \right)}{\left( \frac{m}{n_I} + 1 \right) (m + \rho n_I + \rho n_P) - \rho n_I} - 1 \right) \\ &\quad \times (n_I (P_F - P_1) - n_P (P - P_F))^2. \end{aligned} \quad (33)$$

The sign of  $\frac{dU_S}{dN_S}$  is constant in region 1: it is the sign of the middle term. Given that the variance is hidden in  $\rho$ , it is positive if

$$\text{Var}[\tilde{\xi}_2] > \frac{(m + n_I)(m + n_P) + m}{m^2 n_I} \sum \frac{N_i}{\alpha_i}.$$

Remark that there is an optimal number of speculators for speculators themselves: above a certain  $N_S$ , adding speculators ceases to be profitable to incumbent speculators.

Careful examination of the equations above shows that increasing the number of speculators have two opposite effects. First, it decreases the margin on hedging, since the overall risk tolerance is bigger; this effect is negative. Second, the lower price of hedging increases demand thereof; this effect is positive on speculators welfare. The second effect dominates in situations where the second-period shock has a relatively high volatility, as the inequality above shows, because this means that there is lot to gain to risk sharing.

**Sensitivity of  $U_I$**  Differentiating formula (29) yields:

$$\frac{dU_I}{dN_S} = - \frac{\text{Var}[\tilde{\xi}_2]}{m^2 \alpha_I \left( \sum \frac{N_i}{\alpha_i} \right)^2} (n_I (P_F - P_1) - n_P (P - P_F)) \left( m + n_P \left( 1 + \frac{m}{n_I} \right) \right) \frac{dP_1}{dN_S}$$

$$\begin{aligned}
& - \frac{\text{Var}[\tilde{\xi}_2]}{m^2 \alpha_I^2 \left( \sum \frac{N_i}{\alpha_i} \right)^3} (n_I(P_F - P_1) - n_P(P - P_F))^2 + \frac{P_F - P_1}{C} \left( \frac{dP_F}{dN_S} - \frac{dP_1}{dN_S} \right) \\
& = \frac{\text{Var}[\tilde{\xi}_2]}{m^2 \alpha_I^2 \left( \sum \frac{N_i}{\alpha_i} \right)^3} \left( \frac{\text{Var}[\tilde{\xi}_2]}{\sum \frac{N_i}{\alpha_i}} \frac{m \left( m + n_P \left( 1 + \frac{m}{n_I} \right) \right)}{\left( \frac{m}{n_I} + 1 \right) (m + \rho n_I + \rho n_P) - \rho n_I} - 1 \right) \\
& \quad \times (n_I(P_F - P_1) - n_P(P - P_F))^2 \\
& + \frac{P_F - P_1}{C} \frac{m}{n_I} \frac{m}{\alpha_S} \frac{\text{Var}[\tilde{\xi}_2]}{\left( \sum \frac{N_i}{\alpha_i} \right)^2} \frac{n_I(P_F - P_1) - n_P(P - P_F)}{\left( \frac{m}{n_I} + 1 \right) (m + \rho n_I + \rho n_P) - \rho n_I}.
\end{aligned} \tag{34}$$

We will not pursue the calculations further, noting simply that  $(n_I(P_F - P_1) - n_P(P - P_F))$  factors, so that the result is of the form:

$$\frac{dU_I}{dN_S} = A(n_I(P_F - P_1) - n_P(P - P_F))(K_1(P_F - P_1) + K_2(P - P_F)),$$

for suitable constants  $A$ ,  $K_1$ , and  $K_2$ . This means that the sign changes across

- the line  $\Delta$ , already encountered, defined by  $n_I(P_F - P_1) + n_P(P - P_F) = 0$ ;
- the line  $D$ , defined by the equation  $K_1(P_F - P_1) + K_2(P - P_F) = 0$ .

Both  $\Delta$  and  $D$  go through the point  $M$  where  $P_1 = P_F = P$ . If  $K_2/K_1 < 0$ , the line  $D$  enters region 1, if  $K_2/K_1 > 0$ , it does not. So, if  $K_2/K_1 > 0$ , region 1 is divided in three subregions by the lines  $D$  and  $\Delta$ , and the sign changes when one crosses from one to the other. If  $K_2/K_1 > 0$ , region 1 is divided in two subregions by the line  $\Delta$ , and the sign changes across  $\Delta$ . In all cases, the response of inventory holders to an increase in the number of speculators will depend on the equilibrium.

**Sensitivity of  $U_P$**  Differentiating formula (30) yields:

$$\begin{aligned}
\frac{dU_P}{dN_S} & = \frac{\text{Var}[\tilde{\xi}_2]}{m^2 \alpha_P \left( \sum \frac{N_i}{\alpha_i} \right)^2} (n_I(P_F - P_1) - n_P(P - P_F)) \left( m + n_P \left( 1 + \frac{m}{n_I} \right) \right) \frac{dP_1}{dN_S} \\
& \quad - \frac{\text{Var}[\tilde{\xi}_2]}{m^2 \alpha_P^2 \left( \sum \frac{N_i}{\alpha_i} \right)^3} (n_I(P_F - P_1) - n_P(P - P_F))^2 + \frac{P_F - P}{\beta P} \frac{dP_F}{dN_S} \\
& = \frac{\text{Var}[\tilde{\xi}_2]}{m^2 \alpha_P^2 \left( \sum \frac{N_i}{\alpha_i} \right)^3} \left( \frac{\text{Var}[\tilde{\xi}_2]}{\sum \frac{N_i}{\alpha_i}} \frac{m \left( m + n_P \left( 1 + \frac{m}{n_I} \right) \right)}{\left( \frac{m}{n_I} + 1 \right) (m + \rho n_I + \rho n_P) - \rho n_I} - 1 \right) \\
& \quad \times (n_I(P_F - P_1) - n_P(P - P_F))^2 \\
& + \frac{P_F - P}{\beta P} \left( \frac{m + n_I}{n_I} \right) \frac{m}{\alpha_S} \frac{\text{Var}[\tilde{\xi}_2]}{\left( \sum \frac{N_i}{\alpha_i} \right)^2} \frac{n_I(P_F - P_1) - n_P(P - P_F)}{\left( \frac{m}{n_I} + 1 \right) (m + \rho n_I + \rho n_P) - \rho n_I}.
\end{aligned} \tag{35}$$

Again, we will not pursue the calculations further, noting simply that  $n_I(P_F - P_1) - n_P(P - P_F)$  factors again, so that:

$$\frac{dU_P}{dN_S} = A^*(n_I(P_F - P_1) - n_P(P - P_F))(K_1^*(P_F - P_1) + K_2^*(P - P_F))$$

As in the preceding case, there will be a line  $D^*$  (different from  $D$ ), which enters region 1 if  $K_1^*/K_2^* < 0$  and does not if  $K_1^*/K_2^* > 0$ . In the first case, region 1 is divided into three subregions by  $D$  and  $\Delta^*$ , in the second it is divided into two subregions by  $\Delta$ , and the sign of  $\frac{dU_I}{dN_P}$  changes when one crosses the frontiers.

**Speculation and welfare in summary** Remark that all agents are speculators in their ways. This activity gives the sign of the first term of the derivative of welfare with respect to  $N_S$  (the speculation term): if the speculators gain from being more, then all agents gain as far as only speculation is concerned. This said, remark that the second term in the derivative of welfare concerns only the storers and the processors (the hedging term). They go in opposite direction in regime 1: in subcase 1U, if the number of speculators increases, the hedging term is positive for storers and negative for processors. It is the other way around in subcase 1L.

In terms of political economy (in the sense that economic interests may determine political positions), we can simplify the message as follows. Note that in the neighborhood of  $\Delta$ , the speculation term is of second order with respect to the hedging term. Therefore, the interests of storers and processors are systematically opposed. Storers are in favor of (processors are against) an increase in the number of speculators if they demand more futures (in absolute value) than processors can offer (subcase 1U). The opposite positions are taken if processors are demanding more futures in absolute value (subcase 1L).

### 6.3 The impact of speculators on prices

Formula (45) gives  $\tilde{P}_2$  conditional on  $\xi_1$ :

$$\tilde{P}_2 = \frac{\tilde{\xi}_2}{m} + \frac{n_I \frac{\xi_1}{m} - ((1 + \frac{n_I}{m})n_P + n_I) \frac{\xi_2}{m} + (1 + \frac{n_I}{m})n_P P}{(n_I + n_P + \frac{1}{m}n_I n_P) \rho + (m + n_I)}, \quad (36)$$

where

$$\xi_2 = E[\tilde{\xi}_2] \text{ and } \rho = 1 + m \frac{\text{Var}[\tilde{\xi}_2]}{\sum \frac{N_i}{\alpha_i}}.$$

$\tilde{P}_2$  is clearly a decreasing function of  $\rho$ , which in turn is a decreasing function of  $N_S$ . So  $\tilde{P}_2$  is an increasing function of  $N_S$ . On the other hand

$$\text{Var}[\tilde{P}_2 | P_1] = \frac{\text{Var}[\tilde{\xi}_2]}{m^2}, \quad (37)$$

which depends only on the fundamentals of the economy, not on  $N_I$ ,  $N_P$  nor  $N_S$ . So the number of speculators does not influence the conditional volatility of  $\tilde{P}_2$ .

The unconditional expectation and volatility of  $\tilde{P}_2$  is a different matter. Let us assume, for instance, that  $\tilde{\xi}_1$  and  $\tilde{\xi}_2$  are independent samples from a random variable  $\tilde{\xi}$ . Substituting in (45) gives:

$$E[\tilde{P}_2] = \frac{E[\tilde{\xi}]}{m} + \frac{(1 + \frac{n_I}{m})n_P \left( P - \frac{E[\tilde{\xi}]}{m} \right)}{(n_I + n_P + \frac{n_I n_P}{m})\rho + m + n_I}, \quad (38)$$

$$\text{Var}[\tilde{P}_2] = \left( 1 + \left( \frac{n_I}{(n_I + n_P + \frac{1}{m}n_I n_P)\rho + (m + n_I)} \right)^2 \right) \frac{\text{Var}[\tilde{\xi}]}{m^2}. \quad (39)$$

Let us also investigate  $P_1$  as a random variable, assuming again that  $\tilde{\xi}_1$  and  $\tilde{\xi}_2$  are independent samples from a random variable  $\tilde{\xi}$ . Transforming equation (43), we get:

$$P_1 = \frac{(m + (n_I + n_P)\rho)\frac{\xi_1}{m} + n_I\frac{\xi_2}{m} + n_I n_P \rho m^{-1} P}{m + (n_I + n_P)\rho + n_I + n_I n_P \rho m^{-1}} = Q_0 + Q_\rho, \quad (40)$$

where

$$Q_0 = \frac{(n_I + n_P)\xi_1 + n_I n_P P}{m(n_I + n_P) + n_I n_P},$$

$$Q_\rho = \frac{\left( \xi_1 + n_I \frac{\xi_2}{m} \right) - \frac{(n_I + n_P)\frac{\xi_1}{m} + n_I n_P m^{-1} P}{n_I + n_P + n_I n_P m^{-1}} (m + n_I)}{m + n_I + (n_I + n_P + n_I n_P m^{-1})\rho}.$$

Note that  $Q_0$  is independent of  $\rho$  while  $Q_\rho$  contains  $\rho$  in its denominator only. We find

$$E[P_1] = E[Q_0] + E[Q_\rho], \quad (41)$$

$$E[Q_0] = \frac{(n_I + n_P)E[\xi] + n_I n_P P}{m(n_I + n_P) + n_I n_P},$$

$$E[Q_\rho] = \frac{\left( 1 + \frac{n_I}{m} \right) E[\xi] - \frac{(n_I + n_P)\frac{E[\xi]}{m} + n_I n_P m^{-1} P}{n_I + n_P + n_I n_P m^{-1}} (m + n_I)}{m + n_I + (n_I + n_P + n_I n_P m^{-1})\rho}.$$

It is easily checked that the numerator of  $E[Q_\rho]$  is positive provided  $P < \frac{E[\xi]}{m}$ .

We calculate directly

$$\text{Var}[P_1] = \text{Var} \left[ \frac{(m + (n_I + n_P)\rho)\frac{\xi_1}{m}}{m + (n_I + n_P)\rho + n_I + n_I n_P \rho m^{-1}} \right] \quad (42)$$

$$= \left( \frac{m + (n_I + n_P)\rho}{m + (n_I + n_P)\rho + n_I + n_I n_P \rho m^{-1}} \right)^2 \frac{\text{Var}[\xi]}{m^2}.$$

**Speculation and prices in summary** Increasing the number of speculators  $N_S$

- doesn't influence  $E[\tilde{P}_2 | P_1]$  and  $\text{Var}[\tilde{P}_2 | P_1]$ ;
- increases  $E[\tilde{P}_2]$ ;
- increases  $E[P_1]$  if  $P < \frac{E[\xi]}{m}$ ;

- decreases  $E[P_1]$  if  $P > \frac{E[\xi]}{m}$ ;
- increases  $\text{Var}[P_1]$  and  $\text{Var}[\tilde{P}_2]$ .

Note that Appendix C completes this analysis with comparisons between the basic scenario and another in which the futures market is closed.

The interpretation is the following. As the number of speculators increases, the overall risk appetite increases, which decreases the cost of hedging. This stimulates risk-taking in the physical sphere, pushing storers to store more (other things being equal) and processors to program more production. More precisely, storers respond more intensely to shocks, implying that their purchases have a higher level on average and a higher variability. The latter effect causes higher first-period prices as well as higher volatility when there are more speculators. In the second period, the release of inventories also have higher market impact, which explains the higher (unconditional) volatility of the price. Note that these effects are more systematic (i.e. less contrasted) than those on welfare, which confirms a well-known theoretical fact: volatilities alone are poor indicators of welfare.

## 7 Conclusion

Our model, although extremely simple, perhaps the simple possible which allows the interaction between a physical spot markets and a financial futures market, exhibits a surprising variety of behaviors. In equilibrium, there may be contango and there may be backwardation, the futures price may be higher or lower than the expected spot price, inventory holders may or may not hold stocks, industrial processors may or may not buy forward, adding speculators may increase or decrease the indirect utilities of inventory holders and of industrial processors. All depends, in a way we determine, on market fundamentals and the realization of the shocks. This rich variety of behaviors can be found in commodities markets as they go, and we have not found in the literature a model which encompasses them all. We hope that the present paper fills that gap.

Of course, our model is too simple to capture some important effects; for instance, we would like to understand the so-called convenience yield, which is usually explained as the option value of holding stock. This cannot be understood within a two-period model. So developing a multi-period model, perhaps even an infinite-horizon one, is the next step in our approach. It would be interesting to see how the conclusions of the two-period model (for instance, that speculators increase the volatility of spot prices) fare in the multi-period or the infinite-horizon model. This will be the topic of further work.

## References

- ACHARYA, V. V., LOCHSTOER, L. A., AND RAMADORAI, T. 2011. Limits to Arbitrage and Hedging: Evidence from Commodity Markets. Technical Report 16875, NBER.
- ANDERSON, R. W. AND DANHINE, J.-P. 1983a. Hedger Diversity in Futures Markets. *Economic Journal* 93:370–389.
- ANDERSON, R. W. AND DANHINE, J.-P. 1983b. The Time Pattern of Hedging and the Volatility of Futures Prices. *Review of Economic Studies* 50:249–266.

- BRENNAN, M. J. 1958. The Supply of Storage. *American Economic Review* 48:50–72.
- CHAMBERS, M. J. AND BAILEY, R. E. 1996. A Theory of Commodity Price Fluctuations. *Journal of Political Economy* 104:924–957.
- DANTHINE, J.-P. 1978. Futures Markets and Stabilizing Speculation. *Journal of Economic Theory* 17:79–98.
- DEATON, A. AND LAROQUE, G. 1992. On the Behaviour of Commodity Prices. *Review of Economic Studies* 59:1–23.
- FAMA, E. F. AND FRENCH, K. R. 1987. Commodity Futures Prices: Some Evidence on Forecast Power, Premiums and the Theory of Storage. *Journal of Business* 60:55–73.
- GORTON, G. B., HAYASHI, F., AND ROUWENHORST, K. G. 2012. The Fundamentals of commodity futures returns. *Yale ICF working paper* 0708.
- GUESNERIE, R. AND ROCHE, J.-C. 1993. (De)stabilizing Speculation On Futures Markets: An Alternative View Point. *European Economic Review* 37:1046–1063.
- HIRSHLEIFER, D. 1988. Risk, Futures Pricing and the Organization of Production in Commodity Markets. *Journal of Political Economy* 96:1206–1220.
- HIRSHLEIFER, D. A. 1989. Futures, Trading, Storage and the Division of Risk: A Multiperiod Analysis. *Economic Journal* 99:700–719.
- KALDOR, N. 1940. A Note on the theory of the forward market. *Review of Economic Studies* 7:196–201.
- KAT, H. M. AND OOMEN, R. C. A. 2006. What Every Investor Should Know About Commodities, Part I: Univariate Return Analysis. *Cass Business School Research Paper* 29.
- KEYNES, J. M. 1930. A Treatise on Money, volume 2. Macmillan.
- NEWBERY, D. M. 1987. When do Futures Destabilize Spot Prices? *International Economic Review* 28:291–297.
- ROUTLEDGE, B. R., SEPPI, D. J., AND SPATT, C. S. 2000. Equilibrium Forward Curves for Commodities. *Journal of Finance* 55:1297–1338.
- WORKING, H. 1949. The Theory of the Price of Storage. *American Economic Review* 31:1254–1262.

## A Prices: explicit expressions

Note that  $\xi_1 := \mu_1 - \omega_1$ ,  $\tilde{\xi}_2 := \tilde{\mu}_2 - \tilde{\omega}_2$ ,  $\xi_2 := E[\tilde{\mu}_2 - \tilde{\omega}_2]$ ,  $n_I := N_I/C$  and  $n_P := \frac{N_P}{\beta P}$ .  $(\xi_1, \xi_2)$  determine the regime, and the final expressions of equilibrium prices are as follows.

Regime 1:

$$P_1 = \frac{(m + (n_I + n_P)\rho)\frac{\xi_1}{m} + n_I\frac{\xi_2}{m} + n_I n_P \rho m^{-1} P}{m + (n_I + n_P)\rho + n_I + n_I n_P \rho m^{-1}}, \quad (43)$$

$$P_F = \frac{n_I \rho \frac{\xi_1}{m} + (m + n_I) \frac{\xi_2}{m} + (m + n_I) n_P \rho m^{-1} P}{n_I \rho + (m + n_I) + (m + n_I) n_P \rho m^{-1}}, \quad (44)$$

$$\tilde{P}_2 = \frac{\tilde{\xi}_2 + \frac{n_I \xi_1}{m} - ((m + n_I) n_P m^{-1} + n_I) \frac{\xi_2}{m} + (m + n_I) n_P m^{-1} P}{n_I \rho + (m + n_I) + (m + n_I) n_P \rho m^{-1}}, \quad (45)$$

$$X^* = \frac{-(m + n_P \rho) \frac{\xi_1}{m} + m \frac{\xi_2}{m} + n_P \rho P}{n_I \rho + (m + n_I) + (m + n_I) n_P \rho m^{-1}}, \quad (46)$$

$$Y^* = \frac{-n_I \rho \frac{\xi_1}{m} - (m + n_I) \frac{\xi_2}{m} + (m + (1 + \rho) n_I) P}{n_I \rho + (m + n_I) + (m + n_I) n_P \rho m^{-1}}. \quad (47)$$

Remark that all denominators are equal. They are written in different ways only to show that  $P_1$  and  $P_F$  are convex combinations of  $\frac{\xi_1}{m}$ ,  $\frac{\xi_2}{m}$  and  $P$ .

Note that starting from regime 1, setting  $n_I$  or  $n_P$  to 0 in the expressions to get the prices for any other region works perfectly. For example, the prices for regime 2 can be directly retrieved by posing  $n_P = 0$  in equations (43)-(47).

Regime 2:

$$P_1 = \frac{(m + n_I \rho) \frac{\xi_1}{m} + n_I \frac{\xi_2}{m}}{m + n_I (1 + \rho)}; \quad P_F = \frac{n_I \rho \frac{\xi_1}{m} + (m + n_I) \frac{\xi_2}{m}}{m + n_I (1 + \rho)}; \quad \tilde{P}_2 = \frac{\tilde{\xi}_2 + \frac{n_I (\xi_1 - \xi_2)}{m}}{m + n_I (1 + \rho)};$$

$$X^* = \frac{-m (\frac{\xi_1}{m} - \frac{\xi_2}{m})}{m + n_I (1 + \rho)}; \quad Y^* = 0.$$

Regime 3:

$$P_1 = \frac{\xi_1}{m}; \quad P_F = \frac{\xi_2}{m}; \quad \tilde{P}_2 = \frac{\tilde{\xi}_2}{m}; \quad X^* = 0; \quad Y^* = 0.$$

Regime 4:

$$P_1 = \frac{\xi_1}{m}; \quad P_F = \frac{m \frac{\xi_2}{m} + n_P \rho P}{m + n_P \rho}; \quad \tilde{P}_2 = \frac{\tilde{\xi}_2 + \frac{n_P (P - \frac{\xi_2}{m})}{m}}{m + n_P \rho};$$

$$X^* = 0; \quad Y^* = \frac{m (P - \frac{\xi_2}{m})}{m + n_P \rho}.$$

## B Existence: qualitative comparative statics on the basic scenario

We claim in this section that, for a given distribution of  $(\xi_1, \xi_2)$ , if risk or risk aversions increase, regime 1, in which all markets and all operators are active, becomes more likely.

Beside  $\mu_1 - \omega_1$  and  $E[\tilde{\mu}_2 - \tilde{\omega}_2]$ , the model parameters are  $\text{Var}[\tilde{\mu}_2 - \tilde{\omega}_2]$ , the forward price  $P$  at which the processor can sell his product, the numbers  $N_P$ ,  $N_I$  and  $N_S$ , the risk aversion

parameters  $\alpha_P, \alpha_I$  and  $\alpha_S$ , the production coefficient  $\beta$ , the cost of storage  $C$  and the elasticity of spot traders' demand  $m$ .

Before we proceed with the comparative statics, note that  $\text{Var}[\tilde{\mu}_2 - \tilde{\omega}_2]$  and  $\alpha_P, \alpha_I$  and  $\alpha_S$  appear only through the single parameter:

$$\rho = 1 + m \frac{\text{Var}[\tilde{\mu}_2 - \tilde{\omega}_2]}{\frac{N_P}{\alpha_P} + \frac{N_I}{\alpha_I} + \frac{N_S}{\alpha_S}}.$$

We will thus concentrate our attention on this parameter.

Of particular interest is the case  $\rho = 1$ . This happens when one of the following conditions is verified:

- $\text{Var}[\tilde{\mu}_2 - \tilde{\omega}_2] = 0$  (the future is deterministic and known);
- at least one agent is risk neutral (it is worth emphasizing that our model does not rely on risk aversion to function: even in the risk-neutral case, the four market configurations appear);
- at least one sector is extremely competitive  $N_i = +\infty$  for some  $i$ ;
- $m$ , the elasticity of the demand to the price is zero.

In all other cases,  $\rho$  is higher than 1. It increases with the elasticity of the demand, with the uncertainty on the future availability of the commodity, with risk aversion, and it diminishes with the number of operators.

In the case  $\rho = 1$ , we have:

$$\begin{aligned} F(O) &= (0, -n_P P), \\ F(A) &= P(-n_I, m + n_I), \\ F(M) &= mP(1, 1). \end{aligned}$$

and the upper boundary of region 3 has slope 1, so that region 3 has the minimum possible size. This means that the conditions are the most favorable to active physical markets.

Figure 5 illustrates what happens as  $\rho$  increases from 1 to  $+\infty$  due to a change of  $\text{Var}[\tilde{\mu}_2 - \tilde{\omega}_2]$  or of one of the  $\alpha_i$ . The point  $F(M)$  remains fixed, while  $F(O)$  and  $F(A)$  move vertically, the first one downwards and the second one upwards. So region 1 is enlarged, but the effects are ambiguous for region 4. Region 3 also enlarges, as the slope of the half-line emanating from  $F(M)$  increases towards the vertical. Region 2 is the only one to be unambiguously reduced.

## C Comparison with the no-futures scenario (NF)

When there is no futures markets (scenario NF), speculators are inactive. There remains only three kinds of operators: storers, processors, and spot traders.

In what follows, we can economize on calculations by applying the parameter transposition table 2 on the previous results. The validity of the method is formally established in the proof of Theorem 3.

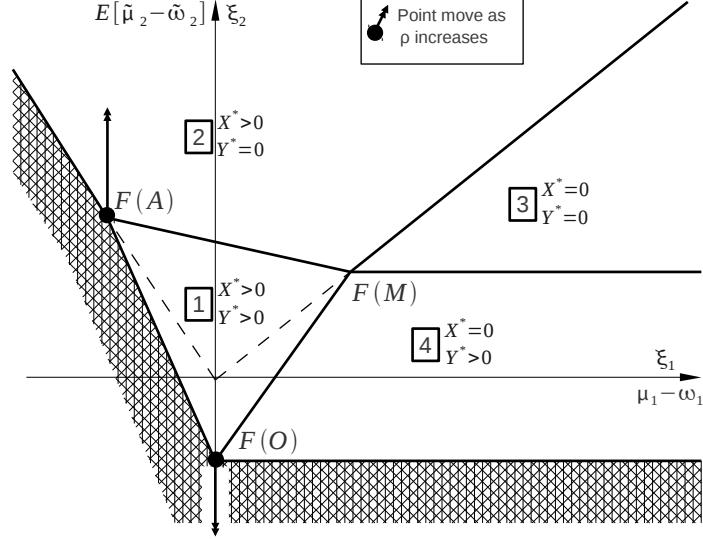


Figure 5: Phase diagram of physical and financial decisions (effect of  $\rho$ ).

|                 |                  |        |   |   |
|-----------------|------------------|--------|---|---|
| Basic scenario: | $P_F$            | $\rho$ | $n_I$   | $n_P$   |
| NF scenario:    | $E[\tilde{P}_2]$ | 1      | $\underline{n}_I = n_I \times \frac{C}{C + \alpha_I V_2}$ | $\underline{n}_P = n_P \times \frac{\beta}{\beta + \alpha_P \frac{V_2}{P}}$ |

Table 2: How to transpose the basic scenario's results to scenario NF.

The following notation will be used:  $V_2 := \text{Var}[\tilde{P}_2]$ , an exogenous parameter.

The optimal position of the storer becomes:

$$\underline{x}^* = \frac{1}{C + \alpha_I V_2} \max\{E[\tilde{P}_2] - P_1, 0\}. \quad (48)$$

The storer holds inventory if the expected price is higher than the current spot price.

An easy computation gives the optimal decision of the processor, namely:

$$\underline{y}^* = \frac{1}{\beta P + \alpha_P V_2} \max\{P - E[\tilde{P}_2], 0\}. \quad (49)$$

Unless the forward price of the final product is higher than the expected spot price of the commodity, the processor has no activity.

**Theorem 3** (Existence conditions). *With the notation  $\xi_1 := \mu_1 - \omega_1$ ,  $\xi_2 := E[\tilde{\mu}_2 - \tilde{\omega}_2]$  and  $V_2 := \text{Var}[\tilde{\mu}_2 - \tilde{\omega}_2]$ , existence conditions on  $\xi_1$ ,  $\xi_2$  and  $V_2$  are stricter in scenario NF than in the basic case.*

*Region 1 in NF is included in basic region 1. Region 2 in NF gains on basic region 1 and is cut on its left border. Region 3 doesn't change. Region 4 in NF gains on basic region 1 and is cut on its bottom border.*

See Figure 6.

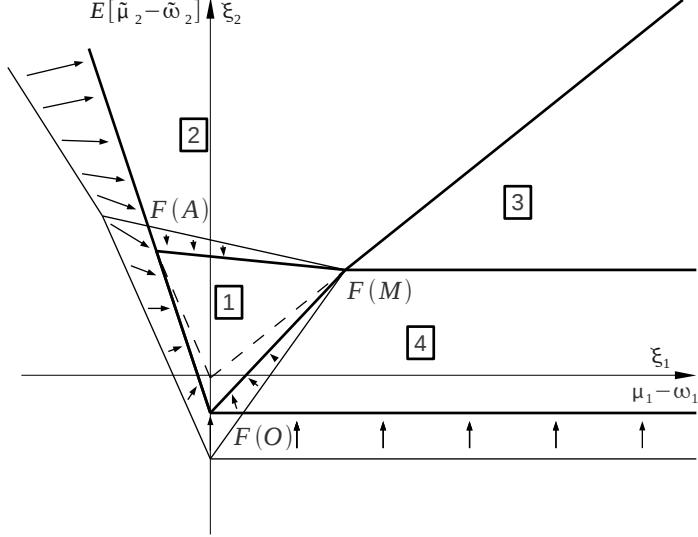


Figure 6: Changes on regions when the futures market is closed.

*Proof.* To prove this theorem, we begin by taking the expected form of equation (7):

$$E[\tilde{P}_2] = \frac{1}{m} (E[\tilde{\mu}_2 - \tilde{\omega}_2] - N_I y^* + N_P y^*).$$

We get the equation:

$$mE[\tilde{P}_2] + N_I x^* - N_P Y^* = E[\tilde{\mu}_2 - \tilde{\omega}_2]. \quad (50)$$

Hence  $\text{Var}[\tilde{P}_2] = V_2$ , a constant. Consider the mapping  $F_{NF} : \mathbb{R}_+^2 \rightarrow \mathbb{R}^2$  defined by:

$$F_{NF}(P_1, E[\tilde{P}_2]) = \begin{pmatrix} mP_1 - \frac{n_I C}{C + \alpha_I V_2} \max\{E[\tilde{P}_2] - P_1, 0\} \\ mE[\tilde{P}_2] + \frac{n_I C}{C + \alpha_I V_2} \max\{E[\tilde{P}_2] - P_1, 0\} - \frac{n_P P}{P\beta + \alpha_P V_2} \max\{P - E[\tilde{P}_2], 0\} \end{pmatrix}.$$

Formally, the analysis is identical to the one done in the basic case. We see now the validity of Table 2. The images  $F_{NF}(O)$ ,  $F_{NF}(A)$ , and  $F_{NF}(M)$  are easily computed:

$$\begin{aligned} F_{NF}(O) &= (0, -\underline{n}_P P), \\ F_{NF}(A) &= P(-\underline{n}_I, m + \underline{n}_I), \\ F_{NF}(M) &= mP(1, 1). \end{aligned}$$

From this, one can find the images of all four regions. In other words, the image of  $O$  moves towards the origin. The image of  $M$  doesn't move. The image of  $A$  goes to the south-east (this is directly visible in the expressions above), in a way that is characterized in detail to prove the theorem. The properties concerning  $O$  and  $M$  are obvious. Concerning the image of  $A$ , we need to check two facts: (1)  $OF_{NF}(A)$  is steeper (in absolute value) than  $OF(A)$ ; (2)  $MF_{NF}(A)$  is flatter (in absolute value) than  $MF(A)$ . (1) We compare  $\rho + m/n_I$  with  $1 + m/\underline{n}_I$ . The former can be rearranged as

$$1 + m \frac{V_2}{\frac{N_P}{\alpha_P} + \frac{N_I}{\alpha_I} + \frac{N_S}{\alpha_S}} + \frac{m}{n_I} = 1 + m \frac{C + \frac{V_2}{\frac{1}{\alpha_I} + \frac{1}{N_I} \left( \frac{N_P}{\alpha_P} + \frac{N_S}{\alpha_S} \right)}}{n_I C}$$

$$= 1 + m \frac{C + \frac{\alpha_I V_2}{1 + \frac{\alpha_I}{N_I} \left( \frac{N_P}{\alpha_P} + \frac{N_S}{\alpha_S} \right)}}{n_I C} = 1 + \frac{m}{n_I} \frac{C + \frac{\alpha_I V_2}{1 + \frac{\alpha_I}{N_I} \left( \frac{N_P}{\alpha_P} + \frac{N_S}{\alpha_S} \right)}}{C + \alpha_I V_2}$$

which is clearly smaller than the latter. (2) We have to compare the two vectors:

$$\begin{aligned} F(M) - F(A) &= P(m + n_I, -\rho n_I), \\ F_{NF}(M) - F_{NF}(A) &= P(m + \underline{n}_I, -\underline{n}_I) \end{aligned}$$

Clearly, the ratios of coordinates are such that the latter is flatter than the former because  $\rho > 1$  and  $\alpha_I V_2 > 0$ .  $\square$

**Prices and volatility.** The absence of a futures market impacts quantities as well as price levels and volatilities. For instance equations (6) and (7) suggest that lower values of  $\underline{X}^*$  and  $\underline{Y}^*$  lead to lower levels of the spot price at time 1, and also, possibly, at time 2.

Prices and quantities in the NF scenario are the following (they can be retrieved directly, or using Table 2 on the expressions in Appendix A):

$$P_1 = \frac{(m + (\underline{n}_I + \underline{n}_P)) \frac{\xi_1}{m} + \underline{n}_I \frac{\xi_2}{m} + \underline{n}_I \underline{n}_P m^{-1} P}{m + (\underline{n}_I + \underline{n}_P) + \underline{n}_I + \underline{n}_I \underline{n}_P m^{-1}}, \quad (51)$$

$$E[\tilde{P}_2] = \frac{\underline{n}_I \frac{\xi_1}{m} + (m + \underline{n}_I) \frac{\xi_2}{m} + (m + \underline{n}_I) \underline{n}_P m^{-1} P}{\underline{n}_I + (m + \underline{n}_I) + (m + \underline{n}_I) \underline{n}_P m^{-1}}, \quad (52)$$

$$\tilde{P}_2 = \frac{\tilde{\xi}_2}{m} + \frac{\underline{n}_I \frac{\xi_1}{m} - ((m + \underline{n}_I) \underline{n}_P m^{-1} + \underline{n}_I) \frac{\xi_2}{m} + (m + \underline{n}_I) \underline{n}_P m^{-1} P}{\underline{n}_I + (m + \underline{n}_I) + (m + \underline{n}_I) \underline{n}_P m^{-1}}, \quad (53)$$

$$X^* = \frac{-(m + \underline{n}_P) \frac{\xi_1}{m} + m \frac{\xi_2}{m} + \underline{n}_P P}{\underline{n}_I + (m + \underline{n}_I) + (m + \underline{n}_I) \underline{n}_P m^{-1}}, \quad (54)$$

$$Y^* = \frac{-\underline{n}_I \frac{\xi_1}{m} - (m + \underline{n}_I) \frac{\xi_2}{m} + (m + 2\underline{n}_I) P}{\underline{n}_I + (m + \underline{n}_I) + (m + \underline{n}_I) \underline{n}_P m^{-1}}. \quad (55)$$

Remark that  $\xi_1$  should be seen as random, notwithstanding the fact that we supposed in this model that markets are open only after its value has been revealed. In this view, the empirical variability of the prices depend directly on the variability of  $\xi_1$ . Let's assume first that  $\xi_1$  and  $\tilde{\xi}_2$  are independent. Later on, we can reason intuitively on positive or negative correlations.

We thus compare in the two scenarios the variability of  $P_1$  attributable to  $\xi_1$ :

$$\text{Var} \left[ \frac{\underline{n}_I \frac{\xi_1}{m}}{\underline{n}_I(\rho + 1) + m + n_P \rho + \underline{n}_I \underline{n}_P \rho m^{-1}} \right]$$

with

$$\text{Var} \left[ \frac{\underline{n}_I \frac{\xi_1}{m}}{2\underline{n}_I + m + \underline{n}_P + \underline{n}_I \underline{n}_P m^{-1}} \right].$$

As we have  $\rho > 1$  and  $n_P > \underline{n}_P$ , the volatility of  $P_1$  is reduced by the presence of the futures market, at least if  $n_I$  and  $\underline{n}_I$  are not too different. In fact, the reduction of the

volatility by futures markets is not guaranteed. Assume for example that  $\alpha_I$  is large (or any other reason making global risk aversion substantial). Then,  $\underline{n}_I$  is very small, clearly making the volatility in the NF scenario smaller.

This case illustrates that financial market may “destabilize” markets, though the term is inappropriate since it only refers to a statistical property. Of course a higher price volatility doesn’t mean a lower welfare, quite the contrary: more volatility means that prices are more effective/informative signals. The impact of markets on prices volatilities is often a naïve aspect of welfare analysis. We shall go further in section 6.

To understand the higher volatility, remark that the storers could react quite differently to different  $\xi_1$  under different market organizations. If the absence of futures frightens the storers, so that they hardly store, their impact on the market in the second period would be negligible. On the contrary, if speculators were not very risk averse, they would accentuate the dependency of actual inventories on the current  $\xi_1$ . These more reactive actions will transport, in a sense, the volatility of the first period to the second one. This explanation is close to the analysis in [Newbery \(1987\)](#).