Speculation in commodity derivatives markets: A simple equilibrium model*

Ivar Ekeland† Delphine Lautier‡ Bertrand Villeneuve§

December 9, 2013
Still preliminary

Abstract

We propose a simple and yet comprehensive equilibrium model of the interaction between the physical and the derivative markets of a commodity. To represent all basic economic functions, we take three types of agents: industrial processors, inventory holders and speculators. Only the two first of them operate in the physical market. All of them, however, may initiate a position in the paper market, for hedging and/or speculation purposes. First, we give the necessary and sufficient conditions on the fundamentals of this economy for a rational expectations equilibrium to exist and we show that it is unique. Second, we propose a generalized framework for the analysis of price relationships: the model exhibits a surprising variety of behaviors at equilibrium which connects the normal backwardation theory and the storage theory. Third, the model addresses the regulatory issues of speculators' presence in the market and their influence on prices.

JEL Codes: D40; D81; D84; G13; Q00.

1 Introduction

In the field of commodity derivative markets, some questions are as old as the markets themselves, and they remain open today. Speculation is a good example: in his famous article about speculation and economic activity, Kaldor (1939) wrote: "Does speculation exert a price-stabilising influence, or the opposite? The most likely answer is that it is neither, or rather that it is both simultaneously."

*The authors acknowledge conversations with Bruno Biais, Eugenio Bobenrieth, Alexander Guembel, Roger Guesnerie, Larry Karp, Jean-Charles Rochet, and Sophie Moinas, remarks from Gabrielle Demange and Sven Rady, and comments from audiences at Paris-Dauphine (FIME Lab and Chair Finance and Sustainable Development), Zurich (ETH), Montreal (IAES), Toulouse (TSE), Santiago (CMM), Lyon (AFFI), Bonn (Mathematical Institute), Toronto (Fields Institute). This article is based upon work supported by the Chair Finance and Sustainable Development and the FIME Lab.

†CEREMADE, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75016 Paris.
Email: ekeland@ceremade.dauphine.fr.

‡DRM-Finance, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75016 Paris.
Email: delphine.lautier@dauphine.fr.

§LEDa, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75016 Paris.
Email: bertrand.villeneuve@dauphine.fr.

More than 70 years later, in June 2011, the report of the G20 (FAO et al. (2011)) states: "The debate on whether speculation stabilizes or destabilizes prices resumes with renewed interest and urgency during high price episodes. [...] More research is needed to clarify these questions and in so doing to assist regulators in their reflections about whether regulatory responses are needed and the nature and scale of those responses." Our simple (perhaps the simplest possible) model of commodity trading provides insights into this question. It also proposes a way to understand how these markets function and how the futures and spot prices are formed. Finally, it illustrates the interest of a derivative market in terms of the welfare of the agents.

In this model, the financial market interacts with the physical market. There are two periods, a single commodity, a numéraire and two markets: the spot market at times $t = 1$ and $t = 2$, and the futures market, where contracts are traded at $t = 1$ and settled at $t = 2$. The spot market is physical (there is a non negativity constraint on inventories), while the futures market is financial (shorting is allowed). There are three types of traders: inventory holders and industrial processors of the commodity, both of which operate on the two markets, and speculators who operate on the futures market only. All of them are utility maximizers and have mean-variance utility (this choice is discussed in the presentation of the model). There is also a price sensitive background demand (or supply) attributed to spot traders, which helps clear the spot market. 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. All decisions are taken at $t = 1$ conditionally on expectations about $t = 2$.

Our main contributions are three: existence and uniqueness of the equilibrium, extended comparative statics, and regulatory implications. They are the consequences of the tractability of the model.

Despite nonlinear equilibrium equations, 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. Moreover, it provides a unified framework for the theory of price relations in commodity futures markets, whereas in the literature this analysis is usually split into two strands: the storage theory and the normal backwardation theory (also named the hedging pressure theory after De Roon et al. (2000)). The former focuses on the cost of storage of the underlying asset, the latter on the risk premium. Although they are complementary, to the best of our knowledge these two strands have remained apart up to now.

We characterize the four possible equilibrium regimes. While each of these four regimes is simple to relate to concrete facts, we believe that our model is the first comprehensive analysis to give explicit conditions on the fundamentals of the economy determining which one will actually prevail in equilibrium. We also 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, as is done in the storage theory, we can explain why there is a contango (in such a case, 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) on the futures market, and how this could change.

Towards this analysis, we give insights into the question of the informational content of the futures price and the price discovery function of futures markets.

As done in the normal backwardation theory, 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. For example, the futures price can be predicted to be lower (resp. higher) than the expected spot price if a synthetic index (denoted by γ) says that there are relatively more (resp. less) storers compared to processors and that they are relatively more (resp. less) risk averse. The precise thresholds depend on the number of speculators and their risk aversion. So the model depicts the way futures markets are used to reallocate risk between operators, the price to pay for such a transfer, and thus provides insights into the main economic function of derivative markets: hedging.[1]

Our model allows for new types of comparative statics. For example, we show that when the number of speculators increases, say because access to the futures market is relaxed, the volatility of spot prices at date 2 goes up. This effect sounds undesirable. 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.

Beyond these descriptive predictions, we use our model to perform a welfare analysis and to draw regulatory implications. This question, again, is as old as derivatives markets. Newbery (2008) summarizes well the usual yet dual appreciation of the impact of derivatives markets on welfare. The author makes a difference between what he calls the "layman" and "the body of informed opinion." He explains that to the first, "the association of speculative activity with volatile markets is often taken as proof that speculators are the cause of the instability," whereas to the second, "volatility creates a demand for hedging or insurance." Our model also exhibits a dual conclusion about welfare, but it is differently stated. First, the model allows for a clear separation between the utility of speculation and that of hedging. Then, the analysis of the impact of an increasing number of speculators shows that, storers and processors, as far as their hedging activities are concerned, have opposite views on the desirability of speculators. They are useless when the positions of storers and processors match exactly; but when one type of agents has needs higher than what the other type can supply, then the former wants more (the latter wants less) speculators because this reduces his costs of hedging. To the best of our knowledge, such an effect has never been clarified before.

Short literature review. Of course, the questions we have raised have been investigated before. Contrary to what is done in this paper, the literature on commodity prices however separates the question of the links between the spot and the futures prices and that of the bias in the futures price.

[1] It is worth noticing that our model operates even without any risk-aversion at all: if we assume that all operators (or even a single one) are risk-neutral, then our model is still valid and gives the four regimes described earlier.

The latter has been investigated first by Keynes (1930) through the theory of normal backwardation whereas the former is usually associated to the theory of storage, initiated by Kaldor (1940), Working (1949) and Brennan (1958). The same separation 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. (2013). Anderson and Danthine (1983a) is an important source of issues and modeling ideas.

Compared with this work, ours is simpler (the producers are not directly modeled) and completely specified. This gives us the possibility to obtain explicit formulas for the equilibrium prices and to investigate further economics issues, like welfare for example. The models developed by Hirshleifer (1988) and Hirshleifer (1989) are also inspired by Anderson and Danthine (1983a). In these papers, Hirshleifer analyzes a point which is interesting for our model but that we leave aside: the coexistence of futures and forward markets. Hirshleifer (1989) also asks whether or not vertical integration and futures trading can be substitute means of diversifying risk.

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. Anderson and Danthine (1983b) is the "inter-temporal" extension of Anderson and Danthine (1983a): they allow the futures position to be revised once within the cash market holding period. To obtain results while keeping tractable equations, the authors however must simplify their model so that only one category of hedger remains in the new 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 inter-temporal analysis. It is related to the literature on equilibrium models which focuses on the *current* spot price and the role of inventories in the behavior of commodity prices, as in Deaton and Laroque (1992), and in Chambers and Bailey (1996). In these models, however, there is no futures market: markets are complete and there is in fact a single type of representative agent. Risk allocation being optimal, these models are not fit for the political economy of regulatory changes.

Beyond the question of the risk premium, equilibrium models have also been used in order to examine the possible destabilizing effect of the presence of a futures market and to analyze welfare issues. This is the case of Guesnerie and Rochet (1993), Newbery (1987), and Baker and Routledge (2012). As the model proposed by Guesnerie and Rochet (1993) is devoted to the analysis of mental ("eductive") coordination strategies, it is more stripped down than ours. As in Newbery (1987), our explicit formulas for equilibrium prices allows for interesting comparisons depending on the presence or absence of a futures market. Finally, contrary to Baker and Routledge (2012), we are not primarily interested in Pareto optimal risk allocations: we focus instead on comparative market performance as measured by utilities *per head*.

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) [2]. 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. The tests undertaken by Acharya et al. (2013) could be used as in fruitful source of inspiration for further developments. As far the analysis of the risk premium is concerned, the empirical tests performed by Hamilton and Wu (2012) and Szymanowska et al. (forth.), as well as the simulations proposed by Bessembinder and Lemmon (2002) are other possible directions.

2 The model

This is a two-period model. There is one commodity, a numéraire, and two markets: the spot market at times $t = 1$ and $t = 2$, and a futures market, in which contracts are traded at $t = 1$ and settled at $t = 2$. 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, short positions are not allowed. In other words, the futures market is financial, while the spot markets are physical.

There are three types of traders.

- *Processors (P)*, or industrial users, 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.

- *Storers (I for inventory)*, 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 $t = 1$ and at $t = 2$. They also operate on the futures market.

- *Speculators (S)*, or money managers, 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

[2] For a recent and exhaustive study on this question, see for example Symeonidis et al. (2012)

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 superscript $\sim$ indicates a random variable.

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

- For $t = 1$, the commodity is in total supply ω_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 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 carried by the storers 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 $\tilde{P}_2 - F$.

There are N_P processors, N_S speculators, 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 $(P_1, F, \tilde{P}_2)$ 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 ω_1 and $\tilde{\omega}_2$ which reach the spot markets at times $t = 1$ and $t = 2$ are exogenous to the model. Traders know ω_1 and μ_1, and share the same priors about $\tilde{\omega}_2$ and $\tilde{\mu}_2$.

- 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 number μ_1 (demand when $P_1 = 0$) is the level at which the economy saturates: to induce spot traders to demand quantities larger than μ_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 types. 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 Optimal positions and market clearing

3.1 Profit maximization

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

$$E[\tilde{\pi}] - \frac{1}{2}\alpha_i \mathrm{Var}[\tilde{\pi}] \tag{1}$$

where α_i is the risk aversion parameter of a type i individual.

Beside their mathematical tractability, there are good economic reasons for using mean-variance utilities. They are not of von Neumann-Morgenstern type, i.e. formula (1) cannot be put in the form $Eu(\tilde{X})$ for some function u. However, Mean-variance utilities capture well the behavior of firms operating under risk constraints. The capital asset-pricing model (CAPM) in finance, for instance, consists in maximizing $E[\tilde{R}]$ under the constraint $\mathrm{Var}[\tilde{R}] \leq \nu$, where $\tilde{R}$ is the return on the portfolio, which is equivalent to maximizing $E[\tilde{R}] - \lambda \mathrm{Var}[\tilde{R}]$, where λ is the Lagrange multiplier. In financial markets, as in commodities markets, agents are mostly firms, not individuals, and they have risk constraints imposed on them from inside (managers controlling traders) and from outside (regulators controlling the firm). This is what formula (1) captures. For the sake of simplicity, we have kept the variance as a measure of risk, but we expect that our results could be extended to more sophisticated ones (coherent risk measures), at the cost of mathematical complications.

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

$$\tilde{\pi}_S(f_S) = f_S\,(\tilde{P}_2 - F),$$

and the optimal position is:

$$f_S^\star = \frac{\mathrm{E}[\tilde{P}_2] - F}{\alpha_S \mathrm{Var}[\tilde{P}_2]}. \tag{2}$$

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 α_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.:

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

The optimal position on the physical market is:

$$x^\star = \frac{1}{C}\max\{F - P_1, 0\}. \tag{3}$$

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:

$$f_I^\star = \frac{\mathrm{E}[\tilde{P}_2] - F}{\alpha_I \mathrm{Var}[\tilde{P}_2]} - x^\star. \tag{4}$$

This position can be decomposed into two elements. First, a negative position $-x^\star$, 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 $(y - \frac{\beta}{2}y^2)\,Z$, where Z 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.:

$$\tilde{\pi}_P(y, f_P) = \left(y - \frac{\beta}{2}y^2\right) Z - y\tilde{P}_2 + f_P(\tilde{P}_2 - F).$$

An easy computation then gives his optimal decisions, namely:

$$y^\star = \frac{1}{\beta Z} \max\{Z - F, 0\}, \tag{5}$$

$$f_P^\star = \frac{\mathrm{E}[\tilde{P}_2] - F}{\alpha_P \mathrm{Var}[\tilde{P}_2]} + y^\star. \tag{6}$$

The futures market is also used by the processor to plan his production, all the more so if the price of his input F is below that of his output Z. The position on the futures market, again, can be decomposed into two elements. First, a positive position $y^\star$, 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.

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 $\frac{N_I}{C} \max\{F - P_1, 0\}$, total inventories would be $(\sum_i 1/C_i) \max\{F - P_1, 0\}$. 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 $X^\star$ defined by:

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

$$X^\star := \max\{F - P_1, 0\}.$$

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

$$n_P := \begin{cases} \frac{N_P}{\beta Z} & \text{if processors are identical,} \\ \frac{1}{Z}\sum_i \frac{1}{\beta_i} & \text{otherwise,} \end{cases}$$
$$Y^\star := \max\{Z - F, 0\}.$$

3.2 Market clearing

The spot market at time 1. On the supply side we have the harvest ω_1. On the other side we have the inventory $n_I X^\star$ bought by the storers, and the demand of the spot traders. Market clearing requires:

$$\omega_1 = n_I X^\star + \mu_1 - m P_1,$$

hence:

$$P_1 = \frac{1}{m}\left(\mu_1 - \omega_1 + n_I X^\star\right). \tag{7}$$

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

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

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

$$\tilde{P}_2 = \frac{1}{m}\left(\tilde{\mu}_2 - \tilde{\omega}_2 - n_I X^\star + n_P Y^\star\right). \tag{8}$$

The futures market. Market clearing requires:

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

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

$$\mathrm{E}[\tilde{P}_2] - F = \frac{\mathrm{Var}[\tilde{P}_2]}{\frac{N_P}{\alpha_P} + \frac{N_I}{\alpha_I} + \frac{N_S}{\alpha_S}}\,\mathrm{HP}, \tag{9}$$

where Hedging Pressure HP is defined as

$$\mathrm{HP} := n_I X^\star - n_P Y^\star. \tag{10}$$

Remark that if different agents of the same type K ($K = P, I, S$)) had different risk aversions α_{Kj} (for $j = 1, \ldots, N_K$), then we would see $\sum_j 1/\alpha_{Kj}$ instead of N_K/α_K in Equation (9). 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. See our synthetic index γ below.

Equation (9) gives a formal expression for the bias in the futures price, which confirms and refines the findings of Anderson and Danthine (1983a). It shows indeed that the bias depends primarily on fundamental economic structures (storage and production costs embedded in $X^\star$ and $Y^\star$, and the number of operators), secondarily on subjective parameters (agents' risk aversions), and thirdly on the volatility of the underlying asset. Note also that the sign of the bias depends only on the sign of HP, which of course is endogenous. As the risk aversion of the operators only influences the speculative part of the futures position, it does not impact the sign of the bias, at least in this partial equilibrium equation. Finally, when HP $= 0$, there is no bias in the futures price, and the risk transfer function of markets is entirely undertaken between hedgers, because their positions on the futures market are opposite and matching exactly. Thus the absence of bias is not exclusively the consequence of risk neutrality but may have other structural causes.

4 Existence and uniqueness of the equilibrium

The equations characterizing the equilibrium are the optimal choices on the physical market (Equations (3) and (5)), the clearing of the spot market at dates 1 and 2 (Equations (7) and (8)), as well as the clearing of the futures market (9):

$$\begin{cases} X^\star &= \max\{F - P_1, 0\} & (3) \\ Y^\star &= \max\{Z - F, 0\} & (5) \\ P_1 &= \frac{1}{m}(\mu_1 - \omega_1 + n_I X^\star) & (7) \\ \tilde{P}_2 &= \frac{1}{m}(\tilde{\mu}_2 - \tilde{\omega}_2 - n_I X^\star + n_P Y^\star) & (8) \\ F &= \mathrm{E}[\tilde{P}_2] - \frac{\mathrm{Var}[\tilde{P}_2]}{\frac{N_P}{\alpha_P} + \frac{N_I}{\alpha_I} + \frac{N_S}{\alpha_S}}(-n_I X^\star - n_P Y^\star) & (9) \end{cases}$$

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

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

where m is the elasticity of demand; γ is a useful synthesis of several parameters that simplifies the calculations and the economic analysis.

From (8), we can derive useful moments:

$$\mathrm{E}[\tilde{P}_2] = \frac{1}{m}(\xi_2 - n_I X^\star + n_P Y^\star), \tag{8E}$$

$$\mathrm{Var}[\tilde{P}_2] = \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}. \tag{8V}$$

We assume $\mathrm{Var}[\tilde{\xi}_2] > 0$, so there is uncertainty on future availability of the commodity. It is the only source of uncertainty in the model. Likewise, we assume (for the time being) that α_P, α_I and α_S all are non-zero numbers, a restriction that is readily lifted.

4.1 Definitions

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

Technically speaking, $(X^\star, Y^\star, P_1, F, \tilde{P}_2)$ is an equilibrium if Equations (3), (5), (7), (8), and (9) are satisfied, with $X^\star \geq 0$, $Y^\star \geq 0$, $P_1 \geq 0$, $F \geq 0$ 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 $(X^\star, Y^\star, P_1, F, \tilde{P}_2)$ 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

$$(X^\star, Y^\star, P_1, F, \tilde{P}_2) \in \mathbb{R}_+^4 \times L^0(\Omega, \mathcal{A}, P)$$

such that Equations (3), (5), (7), (8) and (9) 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* (ξ_1, ξ_2) *verifies:*

$$\xi_2 \geq -n_P \gamma Z \qquad \text{if } \xi_1 \geq 0, \tag{11}$$

$$\xi_2 \geq -n_P \gamma Z - \frac{m + (n_I + n_P)\gamma}{n_I}\xi_1 \qquad \text{if } -n_I Z \leq \xi_1 \leq 0, \tag{12}$$

$$\xi_2 \geq -\frac{m + n_I \gamma}{n_I}\xi_1 \qquad \text{if } \xi_1 \leq -n_I Z, \tag{13}$$

and then it is unique.

Proof. To prove this theorem, we begin by substituting Equation (8E) in Equation (9). We get:

$$mF - \gamma(n_P Y^\star - n_I X^\star) = \xi_2. \tag{14}$$

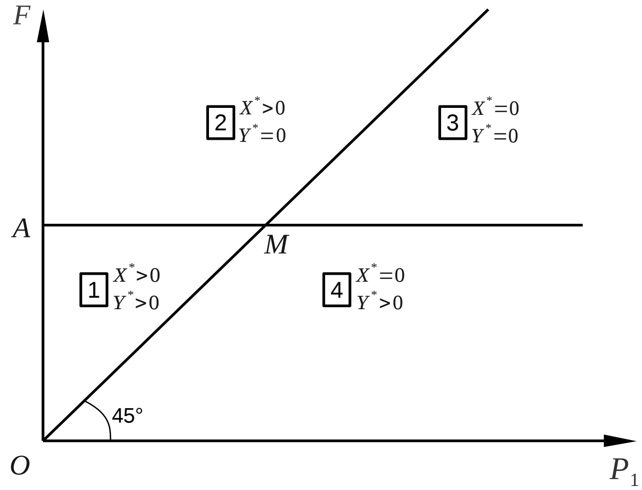

Figure 1: Physical and financial decisions in space (P_1, F): the 4 regions.

We now have two equations, (7) and (14) for P_1 and F. Replacing $X^\star$ and $Y^\star$ by their values, given by (3) and (5), we get a system of two nonlinear equations in two variables:

$$mP_1 - n_I \max\{F - P_1, 0\} = \xi_1, \tag{15}$$

$$mF + \gamma \left(n_I \max\{F - P_1, 0\} - n_P \max\{Z - F, 0\}\right) = \xi_2. \tag{16}$$

Remark that if we can solve this system with $P_1 > 0$ and $F > 0$, we get $\tilde{P}_2$ from (8). So the problem is reduced to solving (16) and (15). Consider the mapping $F : \mathbb{R}^2_+ \to \mathbb{R}^2$ defined by:

$$\varphi(P_1, F) = \begin{pmatrix} mP_1 - n_I \max\{F - P_1, 0\} \\ mF + \gamma \left(n_I \max\{F - P_1, 0\} - n_P \max\{Z - F, 0\}\right) \end{pmatrix}.$$

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

- Region 1, where $F > P_1$ and $F < Z$. In this region, both $X^\star$ and $Y^\star$ are positive.
- Region 2, where $F > P_1$ and $F > Z$. In this region, $X^\star > 0$ and $Y^\star = 0$.
- Region 3, where $F < P_1$ and $F > Z$. In this region, $X^\star = 0$ and $Y^\star = 0$.
- Region 4, where $F < P_1$ and $F < Z$. In this region, $X^\star = 0$ and $Y^\star > 0$

Moreover, in the regions where $X^\star > 0$, we have $X^\star = F - P_1$ and in the regions where $Y^\star > 0$, we have $Y^\star = Z - F$. 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 $(0, Z)$, and by M the point (Z, Z) (so, for instance, Region 1 is the triangle OAM). In Region 1, we have:

$$\varphi(P_1, F) = \begin{pmatrix} mP_1 - n_I(F - P_1) \\ mF + \gamma(n_I(F - P_1) - n_P(Z - F)) \end{pmatrix}.$$

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

$$\begin{aligned} \varphi(O) &= (0, -\gamma n_P Z), \\ \varphi(A) &= Z\,(-n_I, m + \gamma n_I), \\ \varphi(M) &= m\,Z\,(1, 1). \end{aligned}$$

From this, one can find the images of all four regions (see Figure 2). The image of Region 1 is the triangle $\varphi(O)\varphi(A)\varphi(M)$.

The image of Region 2 is bounded by the segment $\varphi(A)\varphi(M)$ and by two infinite half-lines, one of which is the image of $\{P_1 = 0, F \geq Z\}$, the other being the image of $\{P_1 = F, F \geq Z\}$. In Region 2, we have:

$$\varphi(P_1, F) = \begin{pmatrix} mP_1 - n_I(F - P_1) \\ mF + \gamma n_I(F - P_1) \end{pmatrix}.$$

The first half-line emanates from $\varphi(A)$ and is carried by the vector $(-n_I,\, m + \gamma n_I)$. The second half-lines emanates from $\varphi(M)$ 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 $\varphi(O)\varphi(M)$ and by two infinite half-lines, one of which is the image of $\{F = 0\}$, the other being the image of $\{P_1 \geq Z, F = Z\}$. In Region 4, we have:

$$\varphi(P_1, F) = \begin{pmatrix} mP_1 \\ mF - \gamma n_P\,(Z - F) \end{pmatrix},$$

so the first half-line emanates from $\varphi(O)$ and is horizontal, with vertical coordinate $-\gamma n_P Z$, and the second emanates from $\varphi(M)$ 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 (16) and (15) has a unique solution. It can be rewritten as:

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

and it has a unique solution if and only if the right-hand side belongs to the image of F, which

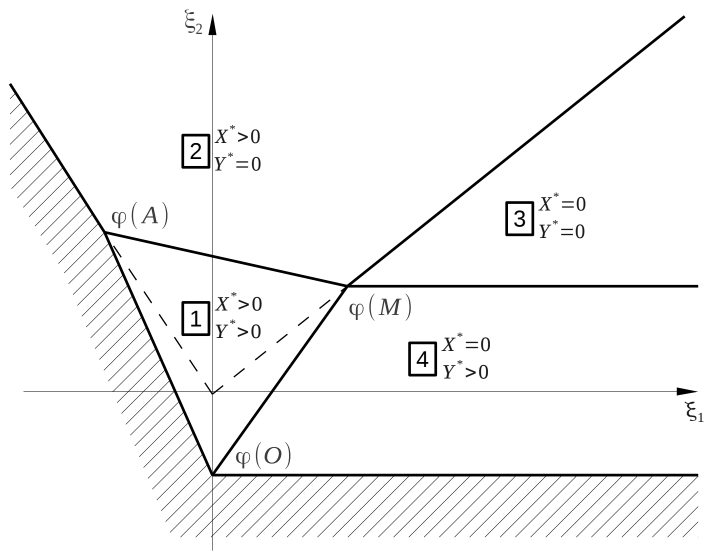

Figure 2: Physical and financial decisions in space (ξ_1, ξ_2): the 4 regions.

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. □

4.3 Equilibrium

The analysis above only considers the existence of a quasi-equilibrium. What happens when the sign of $\tilde{P}_2$ is also considered? Clearly, an equilibrium fails to exists for any (ξ_1, ξ_2) provided some realizations of $\tilde{\xi}_2$ are sufficiently low: states of extreme abundance are inconsistant with positive prices. Theorem 2 gives the exact restrictions on the distribution of $\tilde{\xi}_2$ for an equilibrium to exist.

Theorem 2. *If (ξ_1, ξ_2) supports a unique quasi-equilibrium under the terms of Theorem 1, it supports a unique equilibrium if and only if $\tilde{\xi}_2$ satisfies an additional condition, namely:*

$$\inf\{\tilde{\xi}_2\} \geq m \frac{m n_I \left(\frac{\xi_2}{m} - \frac{\xi_1}{m}\right) - n_P \left(m + n_I\right) \left(Z - \frac{\xi_2}{m}\right)}{m(m + n_I + (n_I + n_P)\gamma) + n_I n_P \gamma} \qquad \textit{in Region 1,}$$

$$\inf\{\tilde{\xi}_2\} \geq \frac{m n_I \left(\frac{\xi_2}{m} - \frac{\xi_1}{m}\right)}{m + n_I(1 + \gamma)} \qquad \textit{in Region 2,}$$

$$\inf\{\tilde{\xi}_2\} \geq 0 \qquad \textit{in Region 3,}$$

$$\inf\{\tilde{\xi}_2\} \geq -\frac{m n_P (Z - \frac{\xi_2}{m})}{m + n_P \gamma} \qquad \textit{in Region 4.}$$

Proof. In Region 1, given Equation (23) in Appendix A, $\tilde{P}_2 \geq 0$ is equivalent to

$$0 \leq \frac{\tilde{\xi}_2}{m} + \frac{mn_I\frac{\xi_1}{m} - ((m+n_I)n_P + mn_I)\frac{\xi_2}{m} + (m+n_I)n_P Z}{mn_I\gamma + m(m+n_I) + (m+n_I)n_P\gamma},$$

which gives the expression of the Theorem after rearrangement.

Remark that the condition for Region 1 is general in the following sense. One can take $n_P = 0$ to get the condition for Region 2; $n_I = 0$, to get the condition for Region 4; $n_I = n_P = 0$ to get the condition for Region 3. This simple shortcut works for other analytical results. □

To be complete, we adopt another perspective and answer another question: for a given (ξ_1, ξ_2), is there a distribution of $\tilde{\xi}_2$ such that an equilibrium exist?

Theorem 3. *If (ξ_1, ξ_2) supports a unique quasi-equilibrium under the terms of Theorem 1, there is a distribution of $\tilde{\xi}_2$ supporting an equilibrium if and only if*

$$\begin{aligned}
&\frac{\xi_2}{m} \geq -\frac{n_I\xi_1 + (m+n_I)n_P Z}{m(m+(\gamma-1)n_P) + n_I(m\gamma + (\gamma-1)n_P)} && \textit{in Region 1,}\\
&\frac{\xi_2}{m} \geq -\frac{n_I}{m+n_I\gamma}\frac{\xi_1}{m} && \textit{in Region 2,}\\
&\frac{\xi_2}{m} \geq 0 && \textit{in Region 3,}\\
&\frac{\xi_2}{m} \geq -\frac{n_P Z}{m+(\gamma-1)n_P} && \textit{in Region 4.}
\end{aligned}$$

Proof. Starting from Theorem 2, and remarking that the limit case allowing us to draw the frontier is when $\inf\{\tilde{\xi}_2\} = \xi_2$, we find the conditions above after rearrangement. □

These additional constraints also have two characteristic points, which are denoted by convention

$$\begin{aligned}
\varphi'(O) &= \left(0, -\frac{mn_P}{m+n_P(\gamma-1)}Z\right);\\
\varphi'(A) &= \left(-n_I Z, \frac{n_I^2}{m+n_I\gamma}Z\right).
\end{aligned}$$

Remark that $\varphi'(O)$ is above $\varphi(O)$ (both with the same negative abscissa), and that $\varphi'(A)$ is below $\varphi(A)$ (both with the same positive abscissa). The crossing point of the two sets of constraints is in Region 1, with coordinates:

$$\left(-\frac{n_I n_P(\gamma-1)Z}{m+(n_I+n_P)(\gamma-1)}, -\frac{mn_P Z}{m+(n_I+n_P)(\gamma-1)}\right).$$

The whole set of constraints is pictured in Figure 3.

When all constraints are considered, it appear that for low ξ_1, the constraint $P_1 \geq$ matters (excessive abundance at $t=1$ must be avoided), whereas the constraints for an equilibrium matter for for a high ξ_1, the constraint $\tilde{P}_2 \geq 0$ matters (excessive abundance at $t=2$ should be avoided).

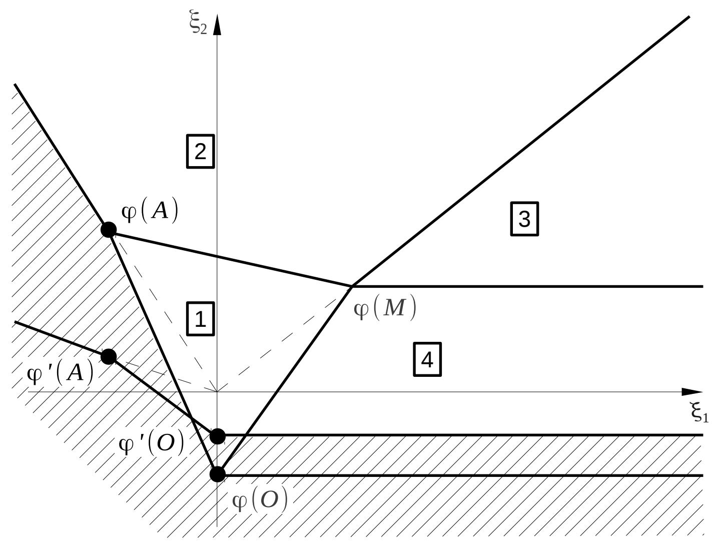

Figure 3: Conditions for the existence of an equilibrium in space (ξ_1, ξ_2): the 4 regions.

5 Equilibrium analysis

In this section we analyze the equilibrium in two steps. First, we examine the four regions depicted in Figure 1. They correspond to very different types of interactions between the physical and the financial markets. Second, we turn to Figure 2 on which we read directly the impact of "initial net scarcity" (ξ_1) and "expected net scarcity" (ξ_2).

5.1 Prices, physical and financial positions

A first general comment on Figure 1 is that in Regions 1 and 2 where $X^\star > 0$, the futures market is in contango: $F > P_1$. Inventories are positive and they can be used for inter-temporal arbitrages. In Regions 3 and 4, there is no inventory ($X^\star = 0$) and the market is in backwardation: $F < P_1$. These configurations are fully consistent with the theory of storage.

The other meaningful comparison concerns F and $E[\tilde{P}_2]$. From Equation (9), we know that hedging pressure HP gives the sign and magnitude of $E[\tilde{P}_2] - F$, i.e. the way risk is transferred between the operators on the futures market. The analysis of the four possible regions, with a focus on Region 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. $F < E[\tilde{P}_2]$. More interestingly, we show why the reverse inequality is also plausible, as mentioned by several empirical studies.[3]

The equation HP $= 0$ cuts Region 1 into two parts, 1U and 1L. It passes through M as can be

[3]For extensive analyses of the bias in a large number of commodity markets, see for example Fama and French (1987), Kat and Oomen (2007) and Gorton et al. (2013).

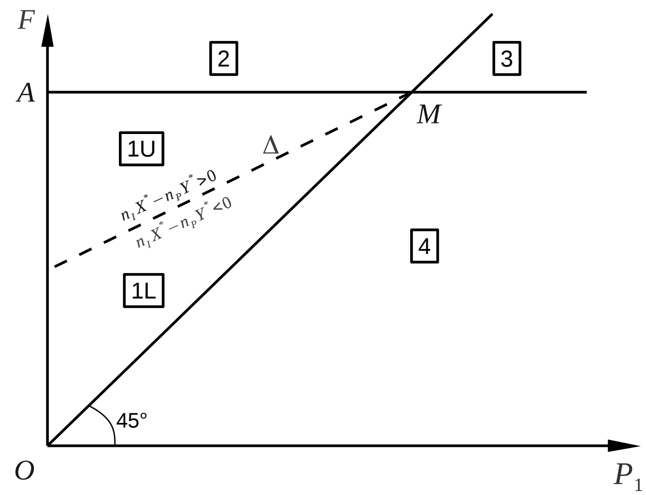

Figure 4: Physical and financial decisions in space (P_1, F) (zoom on Region 1).

seen in Figure 4. This frontier can be rewritten as:

$$\Delta : \quad n_I(F - P_1) - n_P(Z - F) = 0. \tag{17}$$

- Along the line Δ, 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 and they insure each other.

- Above Δ, HP > 0 and $F < E[\tilde{P}_2]$. This concerns the upper part of Region 1 (Subregion 1U) and Region 2.

  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. This backwardation on the expected basis corresponds to the situation depicted by Keynes (1930) as the normal backwardation theory.

- Below Δ, HP < 0 and $F > E[\tilde{P}_2]$. This concerns the lower part of Region 1 (Subregion 1L) and Region 4.

  The net hedging position is long and the speculators must be short, which requires that expected spot price be lower than the futures price.

Table 1 summarizes for each region the relationships between the prices and the physical and financial positions. Attentive scrutiny of the table shows very contrasted regimes.

For example, in Region 2, we have simultaneously a contango on the current basis and a backwardation on the expected basis (or a positive bias). In short, $P_1 < F < E[\tilde{P}_2]$. In Region 3, in the

$\boxed{2}$	$P_1 < F$	$F < E[\tilde{P}_2]$	$F > Z$
	$X^\star > 0$	$f_S > 0$	$Y^\star = 0$
$\boxed{1U}$	$P_1 < F$	$F < E[\tilde{P}_2]$	$F < Z$
	$X^\star > 0$	$f_S > 0$	$Y^\star > 0$
$\boxed{\Delta}$	$P_1 < F$	$F = E[\tilde{P}_2]$	$F < Z$
	$X^\star > 0$	$f_S = 0$	$Y^\star > 0$
$\boxed{1L}$	$P_1 < F$	$F > E[\tilde{P}_2]$	$F < Z$
	$X^\star > 0$	$f_S < 0$	$Y^\star > 0$
$\boxed{4}$	$P_1 > F$	$F > E[\tilde{P}_2]$	$F < Z$
	$X^\star = 0$	$f_S < 0$	$Y^\star > 0$
$\boxed{3}$	$P_1 > F$	$F = E[\tilde{P}_2]$	$F > Z$
	$X^\star = 0$	$f_S = 0$	$Y^\star = 0$

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

absence of hedging of any sort, the futures market is dormant, and this is no bias on the expected basis. Region 4 is the opposite of Region 2: the market is in backwardation and, as $X^\star = 0$, the net hedging position is long, the net speculative position is short and the bias is negative. In short, $P_1 > F > E[\tilde{P}_2]$.

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 ω_1 fails short of the demand of spot traders, and $\xi_2 = \mathrm{E}[\tilde{\mu}_2 - \tilde{\omega}_2]$ is the (expected) extent to which future production will fall short of the demand of spot traders.

Assume that no markets are open before ξ_1 is realized and assume that ξ_1 brings no news about ξ_2. We can take ξ_2 as fixed, and see what happens on equilibrium variables, depending on ξ_1. To fix ideas suppose that we expect moderate scarcity at date 2 ($\xi_2 = \overline{\xi}_2$ in Figure 5, which is supposed to be common knowledge). In the case of a low ξ_1 (abundance in period 1), we are in Subregion 1U. If ξ_1 is bigger, we are in Subregion 1L, and if ξ_1 is even bigger, the equilibrium is in Region 4.

The interpretation is straightforward. If period 1 experiences abundance (Subregion 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 processors' needs. Thus, there is a positive bias in the futures price and speculators have a long position. For a less marked abundance (Subregion 1L), storage is more limited. The storers' hedging needs diminish while the processors' increase. So the net hedging position is long, the bias in the futures price becomes negative and the speculators have a short position. If the commodity is even scarcer (Region 4), there is no storage, only the

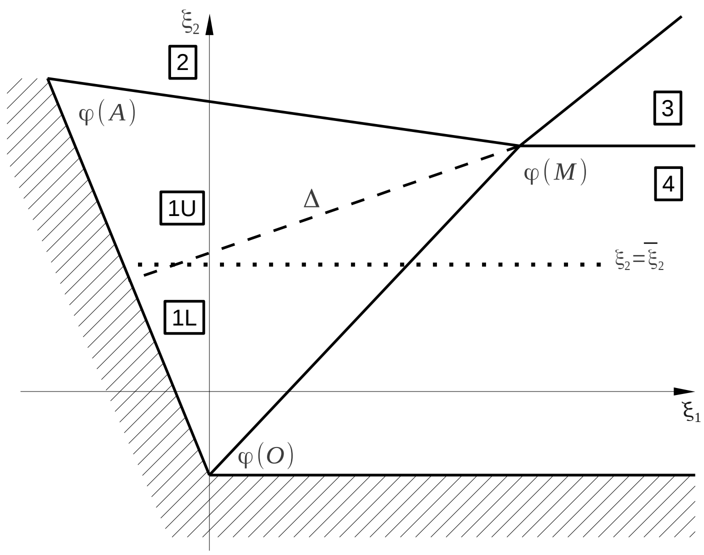

Figure 5: Physical and financial decisions in space (ξ_1, ξ_2) (zoom on Region 1).

processors are active and they hedge their positions.

This example illustrates quite simply why, when there is a contango on the current basis, we can have either expected backwardation or expected contango.

6 The impact of speculation

The impact of speculation can be studied in two ways.

First, the difference between having and not having speculators. This is the approach taken in particular by Newbery (1987). We propose results in this vein in Appendix D. (Summary of existing and new results to be written.)

Second, one can analyze the effect of "increasing speculation." One may think either of a relaxed access to the futures markets, or of a sudden rise in risk appetite. We can translate these changes as an increase in the number N_S, or as a decrease of the risk aversion α_S, or even via the decrease of the risk aversion of any of the other actors (α_I or α_P). The key observation is that all these possible causes impact the synthetic index γ in the same way: it decreases. Indeed, N_S, α_P, α_I and α_S and $\mathrm{Var}[\tilde{\xi}_2]$ appear only through the single parameter:

$$\gamma = 1 + \frac{1}{m} \frac{\mathrm{Var}[\tilde{\xi}_2]}{\frac{N_P}{\alpha_P} + \frac{N_I}{\alpha_I} + \frac{N_S}{\alpha_S}}.$$

This suggests us a simple strategy to perform the comparative statics of speculation.

In the following, we mean by "increased speculation" any of the above mentioned causes decreasing γ. This expression is used for convenience, but the reader should keep in mind that this mean several slightly different things.

Subsection 6.1 shows the impact on prices and quantities; Subsection 6.2 prolongs with detailed welfare analysis and the political economy of speculation.

6.1 Speculators' impact on prices and quantities

This subsection must be seen as a big proposition, the proofs being in Appendix C.1.

Increased speculation has *equilibrium* effects, so that the notion of causality must be used with care. We propose however a sequence of concomitant theoretical facts that are easy to understand and memorize.

The analysis assumes that ξ_1 is known when markets open. In order to perform variance analysis, we consider prices too as random variables at a previous stage, say date 0. Thus now both $\tilde{\xi}_1$ and $\tilde{\xi}_2$ are random, with $\xi_i = \mathrm{E}[\tilde{\xi}_i]$. Moreover, we assume that they are independent. We focus in particular on $\mathrm{Var}_0[\cdot]$ instead of $\mathrm{Var}_1[\cdot|\ \xi_1]$, as was done implicitly up to now. The subscript gives the date at which the statistics are calculated, except if there is no ambiguity.

The analysis of increased speculation is studied as a decrease of γ. A topical case is an increase of N_S but other causes via the above mentioned parameters are also worth considering.

References below are to the columns of Table 2. Remark that in the table the absolute value $|\mathrm{E}[\tilde{P}_2] - \tilde{F}|$ gives the (equilibrium) cost of risk coverage to physical agents. This cost is the starting point of our economic analysis.

We study first Regions 2 and 1U, where $\mathrm{E}[\tilde{P}_2] > \tilde{F}$, and where the physical agents are sellers in aggregate in period 2 (they prefer the sure—as of date 1—$\tilde{F}$ to the random $\tilde{P}_2$).

Let's see first prices and quantities in level.

- Increasing speculation increases in fact capacity to absorb risk. In the competitive setting that we have, this means that risk coverage becomes cheaper: the expected margin $\mathrm{E}[\tilde{P}_2] - \tilde{F} > 0$ decreases. (See column $|\mathrm{E}[\tilde{P}_2] - \tilde{F}|$.)

- This renders risk management cheaper for storers: they increase their inventories whatever the shock in period 1. (See column $\tilde{X}^\star$.)

- To processors, risk coverage was a double win: risk reduction and increased expected profit. The subsidy, or rent, they were receiving is diminished by the increased speculation; thus they reduce their takes. (See column $\tilde{Y}^\star$.)

- Increased inventories means an increased demand in period 1, thus a price increase. (See columns $\tilde{P}_1$.) Quite logically, the effect is a lower price in period 2, due to the extra units drawn from inventories. (See column $\tilde{P}_2$.)

Let's turn now to the variances.

- The decrease of risk coverage cost enables storers to be more reactive to first-period prices, so that overall their opportunistic purchases attenuate even more production or demand shocks on prices: the covariance of inventories and price is negative and it increases in absolute value with speculation. This explains the lower variance of $\tilde{P}_1$. (See column $\mathrm{Var}[\tilde{P}_1]$.)

- The consequence of the previous effect is that there is more variance of the quantity of the commodity delivered in period 2. This adds noise to the current shocks and thus the variance of $\tilde{P}_2$ increases. (See column $\mathrm{Var}[\tilde{P}_2]$.)

Concerning $\tilde{F}$, the effects are driven by the convergence toward $\tilde{P}_2$: they get closer as speculation increases. (See columns $\tilde{F}$ and $\tilde{P}_2$.) Convergence also means that their variances have the same sense of variation with respect to speculation. (See columns $\mathrm{Var}[\tilde{F}]$ and $\mathrm{Var}[\tilde{P}_2]$.)

Whenever $\mathrm{E}[\tilde{P}_2] < \tilde{F}$, the effects are similar but reverse. Processors are most in need of speculators, and they increase they position as speculation increases. Storers in contrast lose part of their rent they have from being contrarians.

	$\lvert\mathrm{E}[\tilde{P}_2] - \tilde{F}\rvert$	$\tilde{F}$	$\tilde{X}^\star$	$\tilde{Y}^\star$	$\tilde{P}_1$	$\tilde{P}_2$	$\mathrm{Var}[\tilde{F}]$	$\mathrm{Var}[\tilde{P}_1]$	$\mathrm{Var}[\tilde{P}_2]$	
$\boxed{2}$	$\searrow$	$\nearrow$	$\nearrow$	0	$\nearrow$	$\searrow$	$\searrow$	$\searrow$	$\nearrow$	$\mathrm{E}[\tilde{P}_2] - \tilde{F} > 0$
$\boxed{1U}$	$\searrow$	$\nearrow$	$\nearrow$	$\searrow$	$\nearrow$	$\searrow$	$\searrow$	$\searrow$	$\nearrow$	
$\boxed{1L}$	$\searrow$	$\searrow$	$\searrow$	$\nearrow$	$\searrow$	$\nearrow$	$\searrow$	$\searrow$	$\nearrow$	$\tilde{F} - \mathrm{E}[\tilde{P}_2] > 0$
$\boxed{4}$	$\searrow$	$\searrow$	0	$\nearrow$	$\longleftrightarrow$	$\nearrow$	$\longleftrightarrow$	$\longleftrightarrow$	$\longleftrightarrow$	
$\boxed{3}$	$\longleftrightarrow$	$\longleftrightarrow$	$\longleftrightarrow$	$\longleftrightarrow$	$\longleftrightarrow$	$\longleftrightarrow$	$\longleftrightarrow$	$\longleftrightarrow$	$\longleftrightarrow$	$\tilde{F} = \mathrm{E}[\tilde{P}_2]$

Table 2: Impact of speculators on prices and quantities

Speculation, prices and quantities in summary. Table 2 shows that Regions 2 and 4 can be viewed as mere subcases of Subregions 1U and 1L. Remark for example that $\tilde{P}_1$ decreases in Subregion 1L whereas it is constant in Region 4. This is due to the fact that the storers are active in Subregion 1L but not in Region 4, and underlines how important the stocks are for the functioning of a commodity market. Inventories indeed appear as the transmission channel for shocks in the space (between the paper and the physical markets) and in the time (between dates 1 and 2). For this is through inventories that a shock appearing in the paper market (i.e. the rise in N_S) impacts the level and variances of the physical quantities and the prices. This result is close to the analysis in Newbery (1987).

As far as the level of the different variables is concerned, our model shows that the impact of an increase in N_S depends, in the end, on which side of the hedging demand dominates. The physical quantities, for example, increase for the operators benefiting from lower hedging costs whereas they decrease for the others. This amplifies the difference in the positions of the operators and consequently their market impact.

The analysis of the variances is less straightforward. The most simple effect is the impact on $\mathrm{Var}[\tilde{\mathrm{F}}]$, which always diminishes under the pressure of a more intense speculative activity (provided that there are stocks in the economy). As regards to the spot prices, a raise in N_S has a stabilizing effect at time 1 and a destabilizing one at time 2. The latter result however might be modified in a three-period model, where the quantities at time 2 would be influenced by the futures price of a contract expiring at time 3. It could also be changed if the price of the output, Z, could be adjusted as an answer to a shock. Up to now, indeed, there is nothing in the model that could absorb a shock at time 2. This version of the model illustrates the fact that financial markets may "destabilize" the underlying 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 will go further on this point in the next subsection.

Note finally that Appendix D completes this analysis with comparisons based on another scenario where the futures market is closed.

6.2 Speculators' impact on utilities

In this section, we express the equilibrium indirect utilities of the various types of agents, and we compute their sensitivities with respect to the parameters, in particular the number of speculators. We proceed in two steps. First, we compute the indirect utilities as functions of equilibrium prices P_1 and F. Second, we compute the elasticities of P_1 and F to deduce the elasticities of the indirect utilities.

We restrict ourselves to the richer case, *i.e.* Region 1, where all agents are active. For the sake of simplicity, we return to an analysis where ξ_1 is known when markets open. Recall that then we have $F < Z$ and $P_1 < F$.

The speculators' indirect utility is given by:

$$U_S = f_S^\star(\mathrm{E}[\tilde{P}_2] - F) - \frac{1}{2}\alpha_S f_S^{\star 2}\mathrm{Var}[\tilde{P}_2],$$

where we have to substitute the value of $f_S^\star$ and and $\mathrm{Var}[\tilde{P}_2]$, which leads to:

$$U_S = \frac{\left(\mathrm{E}[\tilde{P}_2] - F\right)^2}{2\alpha_S \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}}. \tag{18}$$

The storers' indirect utility is given by:

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

where we substitute the values of $f_I^\star$, $x^\star$ and $\mathrm{Var}[\tilde{P}_2]$:

$$U_I = \frac{\left(\mathrm{E}[\tilde{P}_2] - F\right)^2}{2\alpha_I \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}} + \frac{X^{\star 2}}{2C}. \tag{19}$$

Simarly, the processors' indirect utility is:

$$U_P = \frac{\left(\mathrm{E}[\tilde{P}_2] - F\right)^2}{2\alpha_P \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}} + \frac{Y^{\star 2}}{2\beta Z}. \tag{20}$$

For all agents, we see a clear separation between the two components of the indirect utilities. The speculative component is associated with the level of the expected basis. The hedging component changes with the category of agent considered. For the storers, it is positively related to the current basis $F - P_1$, and for the processors, it rises with the margin on the processing activity $Z - F$.

We can use directly Table 2 to produce Table 3, in which there are relatively few ambiguities left.

	U_S	U_I	U_P	
$\boxed{2}$	$\searrow$	See App.	$\searrow$	
$\boxed{1\text{U}}$	$\searrow$	See App.	$\searrow$	$\mathrm{E}[\tilde{P}_2] - \tilde{F} > 0$
$\boxed{1\text{U near }\Delta}$	$\searrow$	$\nearrow$	$\searrow$	
$\boxed{1\text{L near }\Delta}$	$\searrow$	$\searrow$	$\nearrow$	
$\boxed{1\text{L}}$	$\searrow$	$\searrow$	See App.	$\tilde{F} - \mathrm{E}[\tilde{P}_2] > 0$
$\boxed{4}$	$\searrow$	$\searrow$	See App.	
$\boxed{3}$	$\longleftrightarrow$	$\longleftrightarrow$	$\longleftrightarrow$	$\tilde{F} = \mathrm{E}[\tilde{P}_2]$

Table 3: Impact of speculators on welfare. Effects of 1U and 1L are expressed close to line Δ. See App. = See Appendix for complete results.

To conclude, all agents are speculators in their ways. Quite intuitively, the speculative component is all the more important as the futures market is biased, whatever the sign of the bias. All agents lose on this part of their utilities from having more speculators. Consequently, the pure speculator always loses from more competition.

As far as storers and processors are concerned, remark that the second (hedging-related) terms go in strictly opposite directions in Region 1, and weakly opposite directions in Regions 2 and 4. Agents lose (gain) on the hedging terms from having more speculators when they have the smaller (bigger) physical position; namely, storers in Regions 4 and 1L, processors in Regions 2 and 1U.

All terms considered, Table 3 gives unambiguous cases and those where terms draw in opposite directions.

A case is worthy of mention. Around line Δ separating Subregions 1U and 1L, speculation is small since physical positions of storers and processors almost match.[4] In that case, agents with the smaller physical position see their privileged position (they play the best part in the physical

[4] In the neighborhood of the frontier Δ cutting Region 1 into two parts, the speculation term is of second order with respect to the hedging term.

market) eroded by an increased number of speculators. Hence the difference between 1U and 1L around Δ.

In terms of political economy (in the sense that economic interests may determine political positions), we can simplify the message as follows: the interests of storers and processors are systematically opposed. Thus, we cannot assign a pro or con position to either of these groups, the prediction is that in a given position, they will take opposite positions.

7 Conclusion

Our model, although extremely simple (perhaps the simplest possible), shows the interaction between spot markets and a futures market, and exhibits a surprising variety of behaviors. In equilibrium, there may be a contango or a backwardation, the futures price may be higher or lower than the expected spot price, inventory holders may or may not hold inventory, industrial processors may or may not sell forward, adding speculators may increase or decrease the hedging benefits of inventory holders and of industrial processors. All depends, in a way we determine, on market fundamentals and the realization of shocks in the physical market. This rich variety of behaviors can be found in commodities markets as they go, and we have not found in the literature another model which encompasses them all.

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. For this reason, and also because we want to take into account possible differences in the investment horizons of the operators, developing an inter temporal approach is the next step. It would be interesting to see how the conclusions of the two-period model fare in a multi-period or even in a infinite-horizon models.

References

Acharya, V. V., Lochstoer, L. A., and Ramadorai, T. 2013. Limits to Arbitrage and Hedging: Evidence from Commodity Markets. *Journal of Financial Economics* 109:441–465.

Anderson, R. W. and Danthine, J.-P. 1983a. Hedger Diversity in Futures Markets. *Economic Journal* 93:370–389.

Anderson, R. W. and Danthine, J.-P. 1983b. The Time Pattern of Hedging and the Volatility of Futures Prices. *Review of Economic Studies* 50:249–266.

Baker, S. D. and Routledge, B. R. 2012. The Price of Oil Risk. *Working Paper, Carnegie Mellon University, October* .

Bessembinder, H. and Lemmon, M. L. 2002. Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets. *Journal of Finance* 57:1347–1382.

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.

De Roon, F. A., Nijman, T. E., and Veld, C. 2000. Hedging Pressure Effects in Futures Markets. *Journal of Finance* LV:1437–1456.

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.

FAO, IFAD, IMF, UNCTAD, WFP, the World Bank, the WTO, IFPRI, and the UN HLTF 2011. Price Volatility in Food and Agricultural Markets: Policy Responses. Technical report, G20.

Gorton, G. B., Hayashi, F., and Rouwenhorst, K. G. 2013. The Fundamentals of Commodity Futures Returns. *Review of Finance* 17:35–105.

Guesnerie, R. and Rochet, J.-C. 1993. (De)stabilizing Speculation on Futures Markets: An Alternative View Point. *European Economic Review* 37:1046–1063.

Hamilton, J. D. and Wu, J. 2012. Risk Premia in Crude Oil Futures Prices. *Working Paper, University of California, San Diego* .

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. 1939. Speculation and Economic Stability. *Review of Economic Studies* 1:1–27.

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. 2007. What Every Investor Should Know About Commodities, Part I. *Journal of Investment Management* 5.

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.

Newbery, D. M. 2008. Futures markets, hedging and speculation. *In* S. N. Durlauf and E. Blume, Lawrence (eds.), The New Palgrave Dictionary of Economics. Palgrave Macmillan, London.

Routledge, B. R., Seppi, D. J., and Spatt, C. S. 2000. Equilibrium Forward Curves for Commodities. *Journal of Finance* 55:1297–1338.

Symeonidis, L., Prokopczuk, M., Brooks, C., and Lazar, E. 2012. Futures Basis, Inventory and Commodity Price Volatility: An Empirical Analysis. *Economic Modelling* 29:2651–2663.

Szymanowska, M., De Roon, F., and Nijman, T. An Anatomy of Commodity Futures Risk Premia. *Journal of Finance* Forthcoming.

Working, H. 1949. The Theory of the Price of Storage. *American Economic Review* 31:1254–1262.

A Prices and quantities: explicit expressions

Note that $\xi_1 := \mu_1 - \omega_1$, $\tilde{\xi}_2 := \tilde{\mu}_2 - \tilde{\omega}_2$, $\xi_2 := \mathrm{E}[\tilde{\mu}_2 - \tilde{\omega}_2]$, $n_I := N_I/C$ and $n_P := \frac{N_P}{\beta Z}$, and

$$\gamma = 1 + \frac{1}{m}\frac{\mathrm{Var}[\tilde{\xi}_2]}{\frac{N_P}{\alpha_P} + \frac{N_I}{\alpha_I} + \frac{N_S}{\alpha_S}}.$$

The regime is determined by (ξ_1, ξ_2), and the final expressions of equilibrium prices are as follows.

Region 1

$$P_1 = \frac{m(m + (n_I + n_P)\gamma)\frac{\xi_1}{m} + mn_I\frac{\xi_2}{m} + n_I n_P \gamma Z}{m(m + (n_I + n_P)\gamma) + mn_I + n_I n_P \gamma}, \tag{21}$$

$$F = \frac{mn_I\gamma\frac{\xi_1}{m} + m(m + n_I)\frac{\xi_2}{m} + (m + n_I)n_P\gamma Z}{mn_I\gamma + m(m + n_I) + (m + n_I)n_P\gamma}, \tag{22}$$

$$\tilde{P}_2 = \frac{\tilde{\xi}_2}{m} + \frac{mn_I\frac{\xi_1}{m} - ((m + n_I)n_P + mn_I)\frac{\xi_2}{m} + (m + n_I)n_P Z}{mn_I\gamma + m(m + n_I) + (m + n_I)n_P\gamma}, \tag{23}$$

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

Note that

$$\mathrm{E}[\tilde{P}_2] - F = (\gamma - 1)\frac{mn_I\left(\frac{\xi_2}{m} - \frac{\xi_1}{m}\right) - n_P\,(m + n_I)\left(Z - \frac{\xi_2}{m}\right)}{m(m + n_I + (n_I + n_P)\gamma) + n_I n_P \gamma}. \tag{24}$$

Quantities:

$$X^\star = \frac{m(m+n_P\gamma)\left(\frac{\xi_2}{m}-\frac{\xi_1}{m}\right)+mn_P\gamma\left(Z-\frac{\xi_2}{m}\right)}{m(m+n_I+(n_I+n_P)\gamma)+n_In_P\gamma}, \tag{25}$$

$$Y^\star = \frac{mn_I\gamma\left(\frac{\xi_2}{m}-\frac{\xi_1}{m}\right)+m(m+n_I(1+\gamma))\left(Z-\frac{\xi_2}{m}\right)}{m(m+n_I+(n_I+n_P)\gamma)+n_In_P\gamma}. \tag{26}$$

Note that starting from Region 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 Region 2 can be directly retrieved by posing $n_P = 0$ in Equations (21)-(26).

Region 2

$$P_1 = \frac{(m+n_I\gamma)\frac{\xi_1}{m}+n_I\frac{\xi_2}{m}}{m+n_I(1+\gamma)}\ ;\ F = \frac{n_I\gamma\frac{\xi_1}{m}+(m+n_I)\frac{\xi_2}{m}}{m+n_I(1+\gamma)}\ ;\ \tilde{P}_2 = \frac{\tilde{\xi}_2}{m}+\frac{n_I\left(\frac{\xi_1}{m}-\frac{\xi_2}{m}\right)}{m+n_I(1+\gamma)}\ ;$$
$$X^\star = \frac{m\left(\frac{\xi_2}{m}-\frac{\xi_1}{m}\right)}{m+n_I(1+\gamma)}\ ;\ Y^\star = 0.$$

Remark that

$$\mathrm{E}[\tilde{P}_2]-\tilde{F} = (\gamma-1)\frac{n_I\left(\frac{\xi_2}{m}-\frac{\tilde{\xi}_1}{m}\right)}{m+n_I(1+\gamma)} > 0. \tag{27}$$

Region 3

$$P_1 = \frac{\xi_1}{m}\ ;\ F = \frac{\xi_2}{m}\ ;\ \tilde{P}_2 = \frac{\tilde{\xi}_2}{m}\ ;\ X^\star = 0\ ;\ Y^\star = 0\ ;\ \mathrm{E}[\tilde{P}_2]-\tilde{F} = 0.$$

Region 4

$$P_1 = \frac{\xi_1}{m}\ ;\ F = \frac{m\frac{\xi_2}{m}+n_P\gamma Z}{m+n_P\gamma}\ ;\ \tilde{P}_2 = \frac{\tilde{\xi}_2}{m}+\frac{n_P\left(Z-\frac{\xi_2}{m}\right)}{m+n_P\gamma}\ ;\ X^\star = 0\ ;\ Y^\star = \frac{m\left(Z-\frac{\xi_2}{m}\right)}{m+n_P\gamma}.$$

Remark that

$$\tilde{F}-\mathrm{E}[\tilde{P}_2] = (\gamma-1)\frac{n_P\left(Z-\frac{\xi_2}{m}\right)}{m+n_P\gamma} > 0. \tag{28}$$

B Comparative statics on the existence of the equilibrium

This appendix depicts what happens with the existence of an equilibrium when γ decreases. This synthetic parameter is very interesting as $\mathrm{Var}[\tilde{\xi}_2]$, α_P, α_I, α_S and N_S appear only through γ. Of particular interest is the case $\gamma = 1$, which happens when one of the following conditions is verified:

- $\mathrm{Var}[\tilde{\xi}_2] = 0$: the future is deterministic and known.

- One category of agents (or more) is risk neutral. Remark that even in this case, the four market configurations remain.

- At least one sector is extremely competitive: $N_i = +\infty$ for some i.

See Figure 3 for a starting point. Figure 6 illustrates what happens as γ increases. Point $\varphi(M)$ remains fixed, $\varphi(O)$, $\varphi'(O)$, $\varphi(A)$ and $\varphi'(A)$ move vertically, as easy calculations show. Note that the most relevant points as identified in Subsection 4.3, namely$\varphi'(O)$ and $\varphi(A)$, move vertically upwards. This means that existence conditions are restrained in Regions 2 and 4.

Effect on Region 1 seems ambiguous. To be explored further (easy). While the size of Region 1 is clearly enlarged for $\xi_1 \geq 0$, but ambiguous for $\xi_1 \leq 0$.

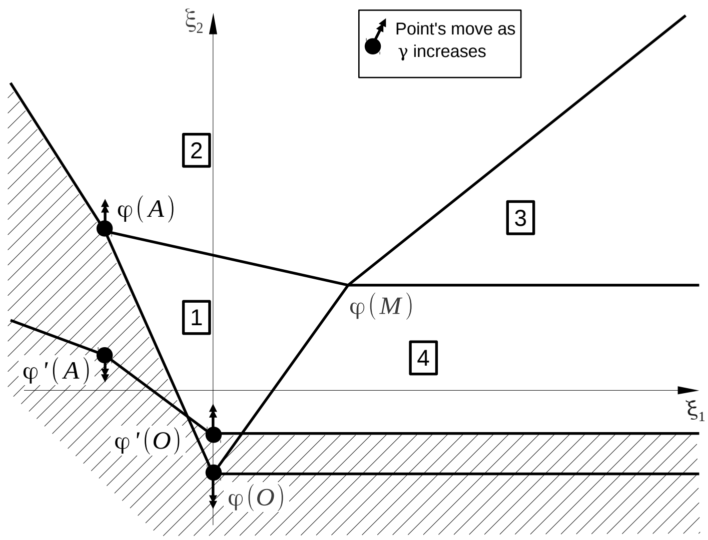

Figure 6: Existence conditions of the equilibrium: comparative statics on γ.

The comparative on γ is now clear: the more frictious the markets, the tighter the existence conditions in Region 2 and Region 4, where only one type of actors actually has a physical position. In Region 1 where physical positions of storers and processors compensate (more or less) each other, the conclusion is not clear cut.

C Prices' and utilities' sensitivities in Section 6

C.1 Speculators' impact on prices and quantities

To perform the comparative statics, we focus first on Regions 2 and 4, in order to examine simple mechanisms, then discuss Region 1, in which the previous effects are mixed in interesting ways. The equilibrium prices are drawn from Appendix A.

Region 2. The facts that $\tilde{P}_1$ and $\tilde{F}$ are weighted averages of $\tilde{\xi}_1/m$ and ξ_2/m, and that $\frac{\xi_2}{m} \geq \frac{\tilde{\xi}_1}{m}$ in Region 2, determine immediately the variations of the 3 prices given in Table 2.

Moreover

$$\mathrm{Var}[\tilde{P}_1] = \left(\frac{m + n_I\gamma}{m + n_I(1+\gamma)}\right)^2 \frac{\mathrm{Var}[\tilde{\xi}_1]}{m^2}, \qquad \mathrm{Var}[\tilde{F}] = \left(\frac{n_I\gamma}{m + n_I(1+\gamma)}\right)^2 \frac{\mathrm{Var}[\tilde{\xi}_1]}{m^2},$$
$$\mathrm{Var}[\tilde{P}_2] = \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2} + \left(\frac{n_I}{m + n_I(1+\gamma)}\right)^2 \frac{\mathrm{Var}[\tilde{\xi}_1]}{m^2}, \qquad \mathrm{Var}[\tilde{X}^\star] = \left(\frac{-1}{m + n_I(1+\gamma)}\right)^2 \mathrm{Var}[\tilde{\xi}_1],$$

with obvious comparative statics.

Remark that we can conclude directly from (27) that $\mathrm{E}[\tilde{P}_2] - \tilde{F}$ decreases as speculation increases.

Region 4. Again, $\tilde{P}_1$ and $\tilde{F}$ are weighted averages of $\tilde{\xi}_1/m$ and ξ_2/m. This, in addition to the fact that $\frac{\xi_2}{m} \leq Z$ in Region 4, determine the comparative statics on the 3 prices.

Remark that we can conclude directly from (28) that $\tilde{F} - \mathrm{E}[\tilde{P}_2]$ decreases as speculation increases.

Region 1. The two Subregions 1U and 1L are separated by the line Δ, already encountered, defined by $n_I\tilde{X}^\star - n_P\tilde{Y}^\star = 0 \Leftrightarrow \mathrm{E}[\tilde{P}_2] - F = 0$ (see Equation 24).

Taking into account the fact that prices and quantities have the form $\frac{A+B\gamma}{C+D\gamma}$, with positive numerators and denominators, and that such expressions are increasing with respect to γ if $BC - DA \geq 0$, it is easy (but tedious) to show that 1U and 1L are the relevant subregions, the former resembling Region 2 and the latter Region 4. This is true for the levels of $\tilde{P}_1$, $\tilde{F}$, $\tilde{P}_2$, $\tilde{X}^\star$, and $\tilde{Y}^\star$, whose changes are summarized in Table 2.

Note that

$$\mathrm{Var}[P_1] = \left(\frac{m(m + (n_I + n_P)\gamma)}{m(m + n_I + (n_I + n_P)\gamma) + n_I n_P\gamma}\right)^2 \frac{\mathrm{Var}[\tilde{\xi}_1]}{m^2},$$
$$\mathrm{Var}[\tilde{P}_2] = \left(\frac{m n_I}{m(m + n_I + (n_I + n_P)\gamma) + n_I n_P\gamma}\right)^2 \frac{\mathrm{Var}[\xi_1]}{m^2} + \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2},$$
$$\mathrm{Var}[\tilde{F}] = \left(\frac{m n_I\gamma}{m(m + n_I + (n_I + n_P)\gamma) + n_I n_P\gamma}\right)^2 \frac{\mathrm{Var}[\tilde{\xi}_1]}{m^2}.$$

The position with regard to Δ is not relevant for the variances, which are monotonic in the same way whatever subcase is concerned. See Table 2.

C.2 Utilities

We will now particularize formulas (18), (19) and (20) to the case when the markets are in equilibrium. In that case, $\tilde{P}_2$ becomes a function of (P_1, F), and the formulas become (after replacing the

n_i by their values in terms of the N_i):

$$U_S = \frac{\mathrm{Var}[\tilde{\xi}_2]}{2m^2\alpha_S\left(\sum \frac{N_i}{\alpha_i}\right)^2}\left(\frac{N_I}{C}(F-P_1)-\frac{N_P}{\beta Z}(Z-F)\right)^2; \tag{29}$$

$$U_I = \frac{\mathrm{Var}[\tilde{\xi}_2]}{2m^2\alpha_I\left(\sum \frac{N_i}{\alpha_i}\right)^2}\left(\frac{N_I}{C}(F-P_1)-\frac{N_P}{\beta Z}(Z-F)\right)^2 + \frac{(F-P_1)^2}{2C}; \tag{30}$$

$$U_P = \frac{\mathrm{Var}[\tilde{\xi}_2]}{2m^2\alpha_P\left(\sum \frac{N_i}{\alpha_i}\right)^2}\left(\frac{N_I}{C}(F-P_1)-\frac{N_P}{\beta Z}(Z-F)\right)^2 + \frac{(Z-F)^2}{2\beta Z}. \tag{31}$$

Formulas (29), (30) and (31) give us the indirect utilities of the agents at equilibrium in terms of the equilibrium prices P_1 and F. These can in turn be expressed in terms of the fundamentals of the economy, namely ξ_1 and $\tilde{\xi}_2$ (see Appendix A): substituting formulas (21), (22) and (23), we get new expressions, which can be differentiated to give the sensitivities of the indirect utilities with respect to the parameters in the model.

To investigate whether an increase in the number of speculators increases or decreases the welfare of speculators, of inventory holders, and of industry processors, we found simpler to work directly with formulas (29), (30) and (31) and to take the sensitivities of P_1 and F with respect to the varying parameter N_S. The calculations are in Appendix C. In contrast, the complete substitution of equilibrium values seems unworkable. Such analysis has never been done before.

Sensitivity of U_S. Differentiating formula (29) yields:

$$\begin{aligned}
\frac{dU_S}{dN_S} =& \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2\alpha_S\left(\sum \frac{N_i}{\alpha_i}\right)^2}\left(n_I(F-P_1)-n_P(Z-F)\right)\left(m+n_P\left(1+\frac{m}{n_I}\right)\right)\frac{dP_1}{dN_S} \\
&- \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2\alpha_S^2\left(\sum \frac{N_i}{\alpha_i}\right)^3}\left(n_I(F-P_1)-n_P(Z-F)\right)^2 \\
=& -\frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2\alpha_S^2\left(\sum \frac{N_i}{\alpha_i}\right)^3}\left(1-\frac{\mathrm{Var}[\tilde{\xi}_2]}{m\sum\frac{N_i}{\alpha_i}}\frac{\left(m+n_P\left(1+\frac{m}{n_I}\right)\right)}{\left(\frac{m}{n_I}+1\right)(m+\gamma n_P)+\gamma m}\right) \\
&\times\left(n_I(F-P_1)-n_P(Z-F)\right)^2. \\
=& -\frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2\alpha_S^2\left(\sum \frac{N_i}{\alpha_i}\right)^3}\,\frac{m^2+n_In_P+m(2n_I+n_P)}{m^2+mn_I+\gamma(n_In_P+m(n_I+n_P))} \\
&\times\left(n_I(F-P_1)-n_P(Z-F)\right)^2.
\end{aligned} \tag{32}$$

The sign of $\frac{dU_S}{dN_S}$ is constant in Region 1: it is negative. Adding speculators decreases the remuneration associated to risk bearing.

Sensitivity of U_I. Differentiating formula (30) yields:

$$
\begin{aligned}
\frac{dU_I}{dN_S} &= -\frac{\operatorname{Var}[\tilde{\xi}_2]}{m^2\alpha_S\alpha_I\left(\sum\frac{N_i}{\alpha_i}\right)^3}\,\frac{m^2+n_In_P+m(2n_I+n_P)}{m^2+mn_I+\gamma(n_In_P+m(n_I+n_P))}\\
&\qquad\times\left(n_I(F-P_1)-n_P(Z-F)\right)^2+\frac{F-P_1}{C}\left(\frac{dF}{dN_S}-\frac{dP_1}{dN_S}\right)\\
&= -\frac{\operatorname{Var}[\tilde{\xi}_2]}{m^2\alpha_S\alpha_I\left(\sum\frac{N_i}{\alpha_i}\right)^3}\,\frac{m^2+n_In_P+m(2n_I+n_P)}{m^2+mn_I+\gamma(n_In_P+m(n_I+n_P))}\\
&\qquad\times\left(n_I(F-P_1)-n_P(Z-F)\right)^2\\
&\quad+\frac{F-P_1}{C}\frac{1}{n_I\alpha_S}\frac{\operatorname{Var}[\tilde{\xi}_2]}{\left(\sum\frac{N_i}{\alpha_i}\right)^2}\frac{n_I(F-P_1)-n_P(Z-F)}{\left(\frac{m}{n_I}+1\right)(m+\gamma n_P)+\gamma m}.
\end{aligned}
\tag{33}
$$

As mentioned before, the utility due to speculative activities decreases when N_S increases. As far as the utility of hedging is concerned, the effect depends on the sign of: $n_I(F-P_1)-n_P(Z-F)$. Remind that this line separates Region 1 into two subcases. In Subregion 1U, the utility of hedging increases for the storers, because they need more hedging than processors. The opposite conclusion arises in Subregion 1L.

As far as the total utility is concerned, we will not pursue the calculations further, noting simply that $n_I(F-P_1)-n_P(Z-F)$ factors, so that the result is of the form:

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

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

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

Both Δ and D go through the point M where $P_1=F=Z$. 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 Δ, 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 Δ, and the sign changes across Δ. 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 (31) yields:

$$\begin{aligned}\frac{dU_P}{dN_S} =& -\frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2\alpha_S\alpha_P\left(\sum \frac{N_i}{\alpha_i}\right)^3}\frac{m^2+n_In_P+m(2n_I+n_P)}{m^2+mn_I+\gamma(n_In_P+m(n_I+n_P))} \\ & \times (n_I(F-P_1)-n_P(Z-F))^2+\frac{F-Z}{\beta Z}\frac{dF}{dN_S} \\ =& -\frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2\alpha_S\alpha_P\left(\sum \frac{N_i}{\alpha_i}\right)^3}\frac{m^2+n_In_P+m(2n_I+n_P)}{m^2+mn_I+\gamma(n_In_P+m(n_I+n_P))} \\ & \times (n_I(F-P_1)-n_P(Z-F))^2 \\ & +\frac{F-Z}{\beta Z}\left(\frac{m}{n_I}+1\right)\frac{1}{m\alpha_S}\frac{\mathrm{Var}[\tilde{\xi}_2]}{\left(\sum \frac{N_i}{\alpha_i}\right)^2}\frac{n_I(F-P_1)-n_P(Z-F)}{\left(\frac{m}{n_I}+1\right)(m+\gamma n_P)+\gamma m}.\end{aligned} \tag{34}$$

Again, the utility due to speculation decreases and that linked with hedging depends on the sign of: $n_I(F-P_1)-n_P(Z-F)$. In Subregion 1U, the utility of hedging decreases for the processors, and it increases in Subregion 1L.

We will not pursue the calculations further, noting simply that $n_I(F-P_1)-n_P(Z-F)$ factors again, so that:

$$\frac{dU_P}{dN_S} = A^\star(n_I(F-P_1)-n_P(Z-F))(K_1^\star(F-P_1)+K_2^\star(Z-F))$$

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

D Comparison with the no-futures scenario (NF)

This appendix is devoted to the case where there is no futures market (scenario NF): speculators are inactive and there remains three kinds of operators: storers, processors, and spot traders. The optimal position of the storer becomes:

$$\underline{x}^\star = \frac{1}{C+\alpha_I\frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}}\max\{\mathrm{E}[\tilde{P}_2]-P_1,0\}. \tag{35}$$

The storer holds inventory if the expected price is higher than the current spot price. The processor's activity depends on the fact that the forward price of the output is higher than the expected spot price of the commodity:

$$\underline{y}^\star = \frac{1}{\beta Z+\alpha_P\frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}}\max\{Z-\mathrm{E}[\tilde{P}_2],0\}. \tag{36}$$

When there is no futures market, uncertainty on the future spot price determines the decisions undertaken in the physical market. In this scenario, the latter necessarily have a speculative aspect.

Theorem 4 (Existence conditions)**.** *Existence conditions on ξ_1 and ξ_2 are stricter in scenario NF than in the basic case.*

Proof. To prove this theorem, we begin by taking Equation (8) depicting the expected equilibrium at date 2:

$$\mathrm{E}[\tilde{P}_2] = \frac{1}{m}\left(\xi_2 - N_I \underline{x}^\star + N_P \underline{y}^\star\right).$$

Hence $\mathrm{Var}[\tilde{P}_2]$ is a constant $\frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}$. Consider the mapping $\varphi_{NF} : \mathbb{R}^2_+ \to \mathbb{R}^2$ defined by:

$$\varphi_{NF}(P_1, \mathrm{E}[\tilde{P}_2]) = \begin{pmatrix} mP_1 - \frac{n_I C}{C+\alpha_I \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}} \max\{\mathrm{E}[\tilde{P}_2] - P_1, 0\} \\ m\mathrm{E}[\tilde{P}_2] + \frac{n_I C}{C+\alpha_I \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}} \max\{\mathrm{E}[\tilde{P}_2] - P_1, 0\} - \frac{n_P \beta Z}{\beta Z+\alpha_P \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}} \max\{Z - \mathrm{E}[\tilde{P}_2], 0\} \end{pmatrix}.$$

Formally, the analysis is identical to the one done in the basic case. We reuse previous calculations by applying on variables and parameters the transposition given in Table 4.

Basic scenario:	F	γ	n_I	n_P
	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
NF scenario:	$\mathrm{E}[\tilde{P}_2]$	1	$\overbrace{\underline{n}_I = n_I \times \frac{C}{C+\alpha_I \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}}}$	$\overbrace{\underline{n}_P = n_P \times \frac{\beta Z}{\beta Z+\alpha_P \frac{\mathrm{Var}[\tilde{\xi}_2]}{m^2}}}$

Table 4: Variables and parameters transposition.

We can now turn to existence conditions, the proof is in three steps. We use the complete conditions of existence from Subsection 4.3.

1. Remark that $\varphi'_{NF}(O)$ is above $\varphi'(O)$, where we use the obvious convention that $\varphi'_{NF}(O)$ is the equivalent in the NF scenario of $\varphi'(O)$ in the basic model. We will use similar conventions in the following. Calculations are a bit tedious, but unambiguous (checked with Mathematica).

2. Note that the slope of the frontier of region 2 is steeper (in absolute value) in the case NF than in the basic model. This also restricts existence possibilities in case NF. Another tedious but unambiguous calculation proved with Mathematica.

3. $\varphi_{NF}(A)$ has a higher abscissa and a lower ordinate than $\varphi(A)$, leading also to more restrictive conditions in scenario NF. This calculation is immediate.

These three properties are given in Figure 7. □

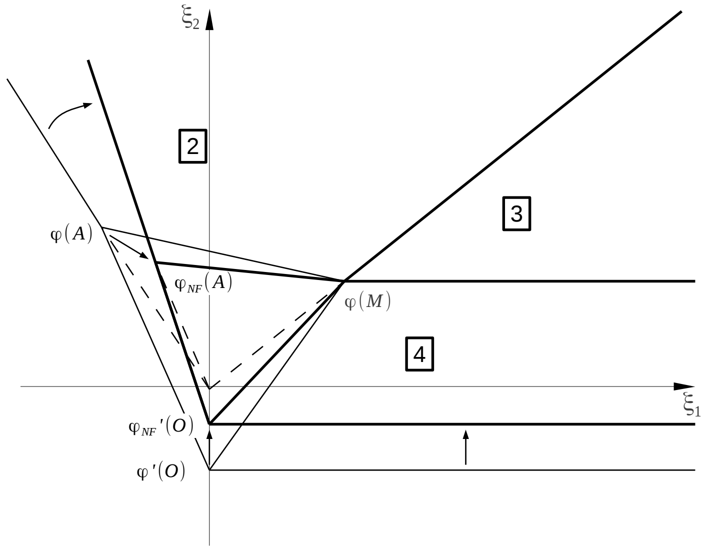

Figure 7: Existence conditions: comparison between and without futures market.

See Figure 7.

The four regions in the NF scenario are included in those of the basic scenario. Region 1 diminishes. Region 2 gains on the basic Region 1 and it is cut on its left border. Region 3 doesn't change. Region 4 gains on the basic Region 1 and it is cut on its bottom border.

Prices and volatility. The absence of a futures market also impacts price levels and volatilities. For instance Equations (7) and (8) suggest that lower values for inventories and production lead to lower levels of the spot price at date 1, and also, possibly, at date 2.

In order to analyze the variances, we will consider ξ_1 as random, as we did in Section 6. Other things equal, having futures or not can change the region in wich the equilibrium is. Yet, for simplicity we compare variances region by region.

Prices in the NF scenario are the following (they can be retrieved directly, or with Table 4 and the equations of Appendix A):

$$\underline{\tilde{P}}_1 = \frac{m(m+\underline{n}_I+\underline{n}_P)\frac{\tilde{\xi}_1}{m} + m\underline{n}_I\frac{\xi_2}{m} + \underline{n}_I\underline{n}_P Z}{m(m+\underline{n}_I+\underline{n}_P) + m\underline{n}_I + \underline{n}_I\underline{n}_P}, \tag{37}$$

$$\mathrm{E}[\underline{\tilde{P}}_2] = \frac{m\underline{n}_I\frac{\tilde{\xi}_1}{m} + m(m+\underline{n}_I)\frac{\xi_2}{m} + (m+\underline{n}_I)\underline{n}_P Z}{m\underline{n}_I + m(m+\underline{n}_I) + (m+\underline{n}_I)\underline{n}_P}, \tag{38}$$

$$\underline{\tilde{P}}_2 = \frac{\tilde{\xi}_2}{m} + \frac{m\underline{n}_I\frac{\tilde{\xi}_1}{m} - ((m+\underline{n}_I)\underline{n}_P + m\underline{n}_I)\frac{\xi_2}{m} + (m+\underline{n}_I)\underline{n}_P Z}{m\underline{n}_I + m(m+\underline{n}_I) + (m+\underline{n}_I)\underline{n}_P}. \tag{39}$$

Let us compare the variance of $\tilde{P}_1$ in Region 1 in the two scenarios, i.e.:

$$\mathrm{Var}[\tilde{P}_1] = \left(\frac{m(m+(n_I+n_P)\gamma)}{m(m+n_I+(n_I+n_P)\gamma)+n_I n_P \gamma}\right)^2 \frac{\mathrm{Var}[\tilde{\xi_1}]}{m^2}.$$

with :

$$\mathrm{Var}[\underline{\tilde{P}}_1] = \left(\frac{m(m+\underline{n}_I+\underline{n}_P)}{m(m+2\underline{n}_I+\underline{n}_P)+\underline{n}_I \underline{n}_P}\right)^2 \frac{\mathrm{Var}[\tilde{\xi_1}]}{m^2}.$$

After tedious calculations, it appears that the latter is unambiguously bigger than the former.[5] Markets are stabilizing the price in period 1, as we saw in Subsection 6.1, because purchases are countercyclical. In the absence of a futures market, the effect is attenuated.

Concerning period 2, we have to compare:

$$\mathrm{Var}[\tilde{P}_2] = \left(\frac{m n_I}{m(m+n_I+(n_I+n_P)\gamma)+n_I n_P \gamma}\right)^2 \frac{\mathrm{Var}[\tilde{\xi_1}]}{m^2} + \frac{\mathrm{Var}[\tilde{\xi_2}]}{m^2}.$$

with

$$\mathrm{Var}[\underline{\tilde{P}}_2] = \left(\frac{m \underline{n}_I}{m(m+2\underline{n}_I+\underline{n}_P)+\underline{n}_I \underline{n}_P}\right)^2 \frac{\mathrm{Var}[\tilde{\xi_1}]}{m^2} + \frac{\mathrm{Var}[\tilde{\xi_2}]}{m^2}.$$

In region 4, they are identical: due to the absence of storage, the absence of futures leaves the two periods independent statistically. Yet, quantities are higher if there are futures.

In region 2, the variance is bigger in the base scenario. This is the effect underlined in Newbery (1987): the facilitation of storage transports shocks from the first period to the second one.

The comparison is ambiguous in region 1, and our attemps to factorize the difference has not produced particularly interesting conditions. One reason is that passing from one scenario to the other is a qualitative step that doesn't have smooth effects on mathematical expression.

[5] We analyzed the numerator after reduction of the difference to the same denominator. All terms have the same sign. Calculations have been verified with a formal calculator. A copy of the file is available upon request.