1. Dynamic Hedging

Summary: *This chapter introduces the theoretical framework for the analysis of the execution of dynamic hedging. A discussion of the issues related to the application of financial theory to the microstructure of dynamic hedging is provided.. Among these issues is the "continuous time problem", the "delta paradox". This chapter also presents results related to the simplification of the risk neutral argument.*

1.1. Introduction

The bad news is that neoclassical economics cannot easily handle the activity of dynamic hedging in an economy in which there are market frictions, asymmetric information, and where the adjustments need to be made in discrete time, in other words the world we truly live in. The good news is that we may not have to. Indeed the very same results can be obtained once one uses some of the accretions to the neoclassical economics of the post Arrow-Debreu era, like industrial organization, the theory of the firm. This chapter will thus focus on the bad news, with the good news to come later chapters.

Recall from the general introduction the problem of dynamic hedging, which we describe as the attempt to analyze dynamic hedging using tools that are not made for it. In other words putting a little bit of microstructure, as has been done, leads to strange results and paradoxes. Completing the analysis miraculously puts it back on its feet.

The problem of dynamic hedging is that it needs to be done in real time, in a real marketplace, between real economic agents, each with his own set of information, some of which are employed in firms. Most of the analysis of continuous time finance creates problems when applied with microstructure constraints. Ignoring microstructure is possible in some cases of the Arrow-Debreu two-period model, but is no longer possible when one addresses the wrinkles of dynamic hedging. The simplifications of the neoclassical paradigm, discussed in chapter 6, while convenient to prove the existence of equilibrium, causes some non-trivial distortions that need to be remedied. This chapter will include most of the foundations on which the rest of the document will be based, by establishing the notion of market structure. Market structure is characterized with the following notions:

- Discrete time: there is an economic lower bound on time, and not a constant one at that at it varies according to economies[1]. Time not being continuous there is an uncertainty attending the execution period.
- Discrete prices: prices are not continuous. There is an economic lower bound on minimum price variations, the minimum "tick" increment.

[1] See Geman and Ané (1996, 1997). This point is further discussed in the general appendix

- Information is never complete. We need to set the framework for the discussion of asymmetric information that will be discussed in chapter 3.
- There exist pronounced asymmetries in the way operators need to execute their hedges; a long option operator's hedges do not correspond to the reverse of those of a short options counterpart.
- Owing to the absence of continuous prices (and time) the operational hedge ratios can no longer be derived using continuous time methods.
- Transaction costs exist. But they are complexity (a discussion of this issue will be left to the next chapter and the general appendix).

Market structure impacts the risk neutral argument, as the existence of a non-satiated marginal arbitrageur between the cash and forward prices can set some characteristics for the distribution.

We note that options are not the only instrument that require dynamic hedging. All financial instruments with convex payoff[2], which we call nonlinear, need dynamic replication. Such expansion of the space of the security aims at including within the framework of microstructure analysis non-option instruments that are non-linear in their price, such as bonds with extreme convexity, or nonlinear baskets, tools that need to be dynamically hedged but do not present strictly speaking a strike price.

---

[2] By convex payoff is meant a payoff that is convex with respect to a linear state variable.

This introductory chapter is organized as follows. We start with a broader definition of a derivative security. We present the microstructure backdrop to the notion of dynamic hedging, with a discussion of the informational distinctions between varieties of markets. We discuss the effect of the discreteness of time of portfolio revision policies, a point that will dominate the text and justify it with a presentation of the "continuous time problem". We present a simplification of the risk-neutral argument[3] by linking it to Keynes' insight and Working's notion of convenience yields. Finally the problem of convergence of the higher moments, the "delta paradox" is presented. The appendix includes some formal definitions – rephrased .

1.2. Nonlinear Derivative Security

Definition 1-1 A nonlinear derivative security is a security the value of which, $V(S,t_0)$, $(S,t_0) \in R^{n+} \times R^{+}$, $V \in R^{+}$ *, depends on the underlying asset price S (also called "primary asset") and time and satisfies either of the following valuation rules:*

i) The diagonal elements of the Hessian $D = \partial^2 V / \partial S_i^{\,2}$ *includes, for some S, at least a non 0 component.*

ii) $V(S,t_0)$ *depends, for some S, on at least one element of* Σ *the matrix with elements* $Cov(S_i,S_j)$*.*

---

[3] The risk neutral argument is defined in this document as the **economic rationale** that allows us to use the risk-neutral distribution, or, in more mathematical terms, the equivalent martingale measure, in pricing financial instruments.

iii) There exists an ξ *and an* $\alpha \in R^+$ *for which* $|V(S,t_0) - V(S,t_0+\xi)| > \alpha$..

By value is meant here arbitrage value.

The requirement i) (the "gamma" condition) corresponds to the casual definition that any derivative security presents a positive or negative "gamma", a second derivative, with respect to at least one of the underlying instruments.

The requirement ii) (the "vega" condition) means that a derivative security will have a valuation that depends on the volatility of at least one of the underlying assets.

The final requirement iii) implies that a derivative security will have time value and that such time value will change over time. iii) makes a distinction between the value of a linear security that is time dependent (like a simple forward contract) and that of a time dependent nonlinear one (a convex bond) in that a time dependent nonlinear one depends on t_0, the initial time, whereas the valuation of a linear one does not depends on t_0[4], only t (expiration).

Thus: $i \Leftrightarrow ii \Leftrightarrow iii$ a simple application of Itϒ's lemma[5].

---

[4] A forward contract for example, is linear but the price in a risk neutral environment is not expected to change at any t_0 owing to the martingale property.

[5] It suffices to see that the forward equation requires a derivative with respect to time.

1.3. Dynamic Hedging

Definition *1-2:* *Dynamic hedging corresponds to any discrete time self financing strategy pair* countable *sequence* $(Q_{ti}, B_{ti})_{i=0}^{n}$ *,*$(R^n \times R)$ *where* Q_{ti} *is the quantity of units (or shares) of the primitive asset S held at time* t_i, $t_0 \le t_i \le t_n$ *and* B_{ti} *are the cash balances held in a default-free interest bearing money market account that satisfies all of the following:*

i) Q_{ti} *is* $\mathbf{F}_{ti}$ *adapted,* $\mathbf{F}_{ti}$ *being the filtration*[6] $\mathbf{F}_{tj} \subset \mathbf{F}_{ti}$, $t_j < t_i$ *, all information related to* S_{ti}.

ii) $\mathrm{E}_{tj,\, j<i}(V(S_t,t_i) - B_{ti} + Q_{ti}\, S_{ti})_{i=0}^{n} = Order(\gamma)$ *for all* t_i *and* t_j.

iii) $\mathrm{E}(U(V(S_t,t) - B_{ti} + Q_{ti}\, S_{ti})))_{i=0}^{n} \ge \mathrm{E}(U(V(S_t,t) - B`_{ti} + Q`_{ti}\, S_{ti})))_{i=0}^{n}$*, with* $(B`_{ti},Q`_{ti})_{ti=0}^{t}$ *any different sequence from* $(B_{ti},Q_{ti})_{ti=0}^{t}$.

iv) $n < \infty$

Here x is said to be of order(γ) if limit x/γ bounded when $x \to 0$.

The notion "of order(γ)" is an artifact introduced here to accommodate the use of a neoclassical marginalist approach without being stopped by some of the fundamental problems – as a large share of option theory (like transaction costs analysis) was designed outside of the economics of supply and demand. Intuitively $O(\gamma)$ means 0 in a purely competitive economy. But in some cases where, either absence of risk-neutrality, or other changes in the basic assumptions require the introduction of risk premium may allow us to put a price on γ. The fact that

---

[6] We assume that the underlying process remains continuous (it is monitored continuously) while the revision of the portfolio constitution is discretely operated. This point is discussed in the appendix.

operators need to make a profit but that the competition of the marketplace should push profits asymptotically to 0 can thus be represented by such symbol. In other places in the text where the Walras "*ni benefices ni pertes*" is invoked the notion of "of order gamma" should be understood[7].

Condition i) states that the quantity of hedge in the underlying securities held is not anticipating with respect to S_{ti}. Intuitively it means that no portfolio constitution will depend on the immediate price: There will be a time lag between prices triggering the portfolio revisions and the actual obtained price for the revision.

Condition ii) states that the expectation from the dynamic hedging strategy is a martingale ("of order γ") along all possible trajectories of the path of S_{ti} (according to the well established fact that any self financing strategy involving an underlying martingale is a martingale -- see the modern rephrasing in Harrison and Pliska, 1981,or our discussion in the appendix to the chapter). Also note that the

---

[7] This allows us to escape the argument that the valuation process can be described as an incentive-submartingale, or otherwise the dynamic hedger would not attempt any replication of the process. It also helps in dealing with the issue of transaction costs. In other words there is a minimum requirement of positive return from the hedges in order to jolt the trader into action, similar to that of the expected profit that motivates holder of any portfolio. Such positive return will be small and vanish asymptotically (under the competition in the marketplace) to the point of conventional asset pricing levels. Also note that "of order gamma" provides a hint of a flexible analytical tool to solve for an aspect of economic equilibrium highlighted by the Grossman-Stiglitz paradox.

self financing condition eliminates the interest rates from the equation so the expectation is limited to the profits over the transaction costs.

Condition iii) states that the strategy sequence followed by the dynamic hedger needs to be the optimal in utility for the operator involved. The risk-neutral agent will be a special case of such utility function. This condition will be `seful later on once transaction costs will be introduced. It means, roughly, that, should the operator involved have a Von Neuman-Morgenstern utility of wealth function U(x), member of a class of monotonically positive first derivative U'(x)>0 and U''(x) < 0 for all x, then the operator would have a revision of hedges as small as possible to lower the variance of the hedged package. We do get here into details of the utility function; we just mention that the literature treats it as time separable (i.e. $U(W_t+W_{s>t}) = U(W_t)+U(W_s)$).

Note that we will make the case, in several different ways, with the study of transaction costs, on why the special case of total risk neutrality should prevail, including in economies with transaction costs.

Condition iv) states that a strategy is a discrete undertaking, as will be discussed next section

Definition *1-3 A positive gamma dynamic hedging strategy over an interval* $[S-\alpha, S+\beta]$ *between periods* t *and* $t+\Delta t$ *is a sequence where* $sign(Q_t-Q_{t-\Delta t}) = sign(S_t-S_{t-\Delta t})$ *and a negative gamma strategy corresponds to a strategy where* $sign(Q_t-Q_{t-\Delta t}) = sign(S_{t-\Delta t}-S_t)$.

The gamma is thus locally defined between two points: A portfolio of options can change in sign below $S-\alpha$ or above $S+\beta$. The distinction between positive and negative gamma is not trivial, owing to the informational content of the trades executed, as will be shown.

Take a portfolio period t, assuming hedging with continuous time derivatives (by hewing to the Black-Scholes hedging policy): The portfolio contains a short European option (on one asset) long risk-free of the asset in order to be delta neutral:

(1-1 $$P = \sum q_i V_{it} + Q_t S_t + B_t$$

with $Q_t=\sum q_i \partial V_{it}/\partial S_t$, where V_i and q_i are the n derivatives securities on S and their quantities (positive or negative).

If $\sum q_i \partial^2 V_{it}/\partial S_t^2 > 0$ (respectively <0) instantaneously then $\partial Q_t/\partial S_t>0$ (respectively <0).

In discrete time, with $\Delta S< \alpha$ and $\Delta S<\beta$, if $\sum q_i(V_i(S+\Delta S,t)+V_i(S-\Delta S,t)-2V_i(S,t))/\Delta S^2 >0$ then $Q_t(S+\Delta S,t)- Q_t(S,t)>0$.

There is no mathematical difference between a positive or negative second derivative with respect to the primary asset, owing to the perfect asymmetry. A positive $\partial^2 V/\partial S^2$ has the properties of -1 $\partial^2 V/\partial S^2$. However there are market microstructure reasons, as will be presented throughout this volume, where the expectation of a positive gamma dynamically hedged portfolio is not the equivalent of that of an equivalent negative gamma one under the same rebalancing policy. There needs to be a different rebalancing policy for the negative gamma; even then the nature of the transaction costs are markedly different. These are expounded in chapter 2.

Chapter 2, in addition, discusses the limitations of the discrete time revisions, since operators can now go bankrupt owing to the lack of support of the distribution. We prove that there is always a positive probability of exceeding any capital when the operator has a delta hedge.

1.3.1. Example of Non-Option Nonlinear Instruments

Among the members of the class of instruments considered that are not options *stricto-sensu* but require dynamic hedging can be rapidly mentioned a broad class of convex instruments:

I. Low coupon long dated bonds. Assume a discrete time framework. Take B(r,T,C) the bond maturing period T , paying a coupon C where $r_t = \#^t r_s$ ds. We have the convexity $\partial^2 B/\partial r^2$ increasing with T and decreasing with C.

II. Contracts where the financing is extremely correlated with the price of the Future such as the Eurodollar contracts such as the Chicago Mercantile Exchange Eurodollar contract and the Paris MATIF (*Marché ✌ Terme des Instruments Financiers*) PIBOR (Paris Interbank Borrowed Rate). This will be discussed further down.

III. Baskets with a geometric feature in its computation such as the US Dollar Index listed on the Financial Exchange FINEX in New York.

IV. A largely neglected class of assets is the “quanto-defined” contracts (in which the payoff is not in the native currency of the contract), such as the Chicago Mercantile Exchange listed Japanese NIKEI Future where

> the payoff is in U.S. currency. Such Future contract presents convexity owing to the correlation between the performance of the contract and the U.S. Dollar/Japanese Yen parity. In short, while a Japanese Yen denominated NIKEI contract is linear, a US dollars denominated one is nonlinear and requires dynamic hedging. See Duffie(1992), for the intuition of the issue.

It is easy to explain such convexity with the following: take at initial time t_0, the final condition

$$V(S,T) = S_T$$

where T is the expiration date. More simply, the security just described is a plain forward, assumed to be linear. There appears to be no Ito term there yet. However should there be an intermediate payoff such that, having an accounting period *i/T*, the variation margin is paid in cash disbursement, some complexity would arise. Assume $\Delta(t_i)$ the changes in the value of the portfolio during period (t_i,t_{i-1}).

(1-2 $$\Delta(t_i)= (V(S,t_i)-V(S, t_{i-1}))$$

If the variation is to be paid at period t_i, then the operator would have to borrow at the forward rate between periods t_i and T, here $r(t_i,T)$. This financing is necessary to make $V(S,T)$ and S_T comparable in present value. In expectation, we will have to discount the variation using forward cash flow method for the accounting period between t_{i-1} and t_i . Seen from period T, the value of the variation becomes

(1-3 $$E_t\,[exp[-r(t_i,T)(T-t_i)]\ \Delta(t_i)]$$

where E_t is the expectation operator at time *t* (under the risk-neutral probability measure) .

Therefore we are delivering at period T, in expectation, as seen from period t_0, the expected value of a stream of future variations

(1-4 $$E_{t0}[\sum exp[-r(t_i,T)(T-t_i)]\ \Delta(t_i)]$$

However we need to discount to the present using the term rate *r(T)*. (1-4 becomes

(1-5 $$V(S,T)|_{t=t0}= V[S,t_0]+ exp[r(T)]\ E_{to}\ [\sum exp[-r(t_i,T)(T-t_i)]\ \Delta(t_i)]$$

which will be different from S_T when any of the interest rate forwards is stochastic.

Result : *When the variances of the forward discount rate $r(t_iT)$ and the underlying security S_T are strictly positive and the correlation between the two is lower than 1,*

$$V(S,T)|_{t=t0} \neq S_T$$

We can easily prove it by examining the properties of the expectation operator. See Taleb(1997) or Duffie (1996) for further discussion of what is commonly called the "quanto" problem[8].

To conclude, a forward contract is linear but the price in a risk neutral environment is not expected to change at any t_0 owing to the martingale property. Therefore

$$F(S, t_0))= F(S,t_0+\Delta t)$$

---

[8] See the exposition of the issue in Burghardt et al (1993).

while a nonlinear instrument will satisfy

$$E[V(S,t_0)]=E[V(S,t_0+\Delta t)]$$

(where the expectation operator is taken under the risk neutral probability measure), but

$$V(S,t_0)\neq V(S,t_0+\Delta t).$$

1.4. Principle of Finite Order Flow and Lower Bound of Variations

Remark: *A sequence of dynamic strategies* $(S_{ti})_{i=0}^{n}$ *can be only implemented in discrete time. The elements of the sequence need to be finite or countably infinite. Thus limit analysis can only be used while re-scaling to take into account such framework. There will thus be a lower bound on* Δt *that is the smallest possible increment feasible in the economy, smaller than which operators cannot transact.*

The notion of physical implementation of a hedging policy is in contradiction with that of continuous time limit analysis: no policy can be formulated in infinitesimal time steps[9]. Thus a policy becomes a discriminating issue between time steps thus requiring countable activities.

---

[9] From an operational standpoint, there cannot be a processing back-office capable of recording, clearing and wiring over a time interval some amount of funds an infinite number of times. Analyzing the process in terms of production capabilities, it is not feasible to manufacture infinite

Another lower bound needs to be imposed on the observed price change. This principle is more casually called in the business world the "law of the indivisible tick", or minimum price increment. Markets have a minimum price increment over which they are legally allowed to move and varies between 1/32 for units of bonds, 1/8 for stocks, .01 for foreign exchange transactions, for SP500 futures and so on. Likewise there is an official "tape" limiting the number of transactions over some time interval. Rule 101 of The Chicago Mercantile Exchange in Chicago prevents the prearranged ratio, or an agreement to trade at two prices that can potentially allow the operators to agree on an average. Thus they disallow the bargaining that involves an α such that the pre-agreed price becomes $P= \alpha P_1 +(1-\alpha) P_2$ in the same transaction[10], where $\alpha \in (0,1)$ and P_1 and P_2 distinct prices.

Most exchanges operate as in succession of rapid auctions where the local equilibrium for the supply and demand is established (or supposed to be established) as a condition for the price clearing: Operators shout orders while awaiting for a better fill prior to transacting. Preventing the establishment of such local equilibrium is considered, on listed exchanges, a blatant violation of the rules. By most exchange rules (as discussed later) and unofficial (but seriously enforced

---

amounts of discrete goods within a finite interval. For a presentation of the practical issues in clearing a trade, (see the *Chicago Mercantile Exchange Rules and Regulations Document*).

[10] This does not mean that an operator cannot obtain a final price that is $\alpha P_1 + (1-\alpha) P_2$. It simply means that it cannot be legally pre-agreed upon.

as well as self policed pit etiquette) operators need to wait before filling a customer order a "reasonable time" in order to give the best bid or the best offer a chance to materialize.

Harris (1991) argues that a non-zero tick simplifies trader's information set, shrinks the bargaining time, and allows for more efficiency in the transaction. Angel(1997) studies the effect of the tick size on the activity in the marketplace, and makes references to the "cognitive value of rounded numbers". The general appendix discusses the various econometric methods of inferring from a fixed tick size as well as the stochastic properties of the fixed tick process. Clearly the notion of minimum price increment is not trivial, as will be discussed. This minimum tick issue can be further formalized as follow (another use of Zeno's paradox).

Take two sole players a and b in the marketplace, with a quantity *q* exogenous (and no substitutability between the goods). Assume a non-informational[11] successive bargaining competitive equilibrium between the two players. Non-informational in this context means that the counter-price in the bargaining process does not lead of any change in the demand function of the counter party.

Proposition 1-1: The assumption of a fixed time interval Δt requires a set of finite possible prices.

---

[11] The absence of information can be explained as follows. Assume time in discrete periods indexed by n. If p_{An}, the price offered by A leads to p_{Bn+1} , a counteroffer by B, then p_{An+2} would not be affected by a conditional valuation of the security. Such condition, enforcing path independence of beliefs, is used for purposes of comparability with the information-free neoclassical equilibrium.

A corollary would be that discreteness of time means discreteness of prices. To prove this, take their bargaining as leading to a successive t☉tonnement-style fixed point such that

$$f(p^*)=p^*$$

Assume that the initial starting prices are $p_a > p^*$ and $p_b < p^*$. Price clear upon convergence when

$$p^* = p_a = p_b$$

Assume δt the minimum time between offer and counteroffer. Unless there is a finite set of possible prices between p_a and p_b , P being the vector of dimension $(n,1)$,

$$P^T=[p_{1,}\, p_2,\ldots,\, p_n]$$

the time to clear the trade will be infinite, as it would take an infinite number of iterations. The total time being finite, $T= n\ \delta t$, then since δt is finite, n become an integer and there need to be a finite set of prices.

The minimum time increment as well as the distribution of time between price changes is examined in the empirical study in chapter 8.

There is a rich econometric literature for the estimation of the continuous processes from discretely sampled data. We discuss, again in chapter8, the following question: Is the data generating process a fully discrete one, or is it a continuous one, but with discrete realizations? Among the earlier works, Lo (1988) examines the inverse problem: he views the diffusion as the underlying data generating process and views the price realizations as samples from it. He characterizes the DGP as an Ito process of the form:

(1-6 $$\mu(S,t)\, dt + \sigma(S,t)\, dZ$$

and makes some maximum likelihood inference , as the realizations over a time period Δt of the vector $S=[S_0,S_1,\ldots S_t]$, not necessarily equally spaced in time, are discrete. On later studies on how to extract the parameters of the distribution we mention Hansen and Scheinkman (1995), and Duffie and Glynn (1996). Hansen and Scheinkman (1995), particularly, examines the properties of the theoretical moments of the Dynkin operator $\frac{\partial}{\partial t} E_0(S)$ (the time derivative of the expectation of the Ito process).

While these studies concern the econometric inference, with some data generating process in mind, this thesis approaches the problem from the standpoint of an option replicating agent: the economic implications of the data generating process are the final object of this study, rather than the parameters. The idea of whether it is a diffusion or a true discrete time process, while it matters, matters less than the properties of the scaling of the data and the effect the properties can have on the economic decision making. Also note that Appendix A presents a review of the literature on the discreteness of price increments.

At a more philosophical level, there is the discreteness of human actions and decisions: Continuous time should not be abused and mistaken for a goal for financial markets but as a simple tool of mathematical analysis. Finally the activity can be countably infinite if projected over an infinite time horizon (such as the valuation of a nonlinear perpetual physical investment such a gold mine[12]) .

Boyle and Emanuel (1980) provides a seminal presentation of the issue of discrete time revision of a portfolio containing the three Black-Scholes components (options, delta hedge and risk-free cash balance)[13]. Indeed while price changes can be examined at a continuous time limit, activities of market participants cannot be considered at such limit if there is the smallest amount of transaction costs.

1.5. Presentation of the Continuous Time Problem

Hedging policies in continuous time invariably lead to the pathological case of infinite transaction costs. The principle of finite order flow makes such case a non sequitur. It can be easily shown that any positive transaction cost, proportional or fixed, cause the volatility of the underlying security to be either infinite or degenerate if it is measured in continuous time.

Assume, to simplify (next page footnote), ΔS a *standard* Brownian motion (mean 0 and unit variance) and

---

[12] A “real option” like a gold mine can be valued as a *dynamic call option* the strike of which corresponds to the costs of extraction from the ground. Taking into account dynamic strategies one can assume a sequence of infinitely countable time steps. For a discussion of real options see Dixit and Pyndik (1994).

[13] Boyle and Emanuel(1980) show that the tracking error of a portfolio of options revised at discrete intervals is a Chi-squared error (in proportion to the square of the ΔS). The mistracking error will therefore be a function of the magnitude of the second derivative of the option price.

$$\sum_{i=0}^{n} \Delta S^2 = n\ \Delta t$$

its pre-transaction costs (non-central) observed variance over a period $(t\text{-}t_0) = n\ \Delta t$. Assume k arbitrarily fixed transaction costs (they are assumed here to be non proportional, without a loss of generality).

- First consider the case of negative gamma. Operator needs to place orders of the same sign as the market move (to buy the rallies and sell the dips). Rehedging the position with transaction costs increases the mean square movement ΔS^2 , by buying the asset when ΔS is positive and selling it when ΔS is negative. Paying k an infinite number of times leads to non-boundedness and non satisfaction of the regularity condition. We can use symmetry, since we assumed arithmetic Brownian motion as a simplification, and consider that $\Delta S>0$ takes place $n/2$ times[14] and that $\Delta S<0$ takes place $n/2$ times. We would then look for the limit of Δt approaching 0, and n approaching infinity of

  $$½ \sum((\Delta S+k)^2 + (\Delta S\text{-}k)^2)$$

  which equals $\sum \Delta S^2 + \sum k^2$ .Since

  $$\sum \Delta S^2 + \sum k^2 = n\ \Delta t + n\ k^2$$

  At the limit, the adjusted variance becomes infinite. The square integrability condition, that

---

[14] P(ΔS>0) is slightly lower than ½ in a logarithmic return framework and ½ in a standard Brownian. More specifically, P(ΔS)>- ½ σ^2 = ½ when S lognormally distributed. The distinction loses in significance if the operator reacts to Log[$S_{t+\Delta t}/S_t$] increments.

$$\int_0^t \sigma_s^2 ds < \infty$$

where σ_s is the adjusted volatility (see Leland, 1985, and the general appendix), no longer holds.

- **Positive Gamma**: Operator needs to sell the rallies and buy the dip. But being rational, he will only sell the rally if the total ΔS-k remains positive if $\Delta S > 0$ and negative if $\Delta S < 0$. Therefore he will act according to *Max* $[|\Delta S\text{-}k|,0]>0$. The average positive ΔS

$$E[\Delta S>0] = \sqrt{(2/\pi)}\ \Delta t$$

and the average negative ΔS

$$E[\Delta S>0] = -\sqrt{(2/\pi)}\ \Delta t$$

$\sum$ Max $[|\Delta S\text{-}k|,0]$ can be approximated with

$$\tfrac{1}{2}\ \text{Max}\ [\ (1/\sqrt{(2/\pi)}\ \text{-}n\ k)^2),0]\ +\tfrac{1}{2}\ \text{Min}\ [\ (1/\sqrt{(2/\pi)}+n\ k)^2),0]$$

the limit of which tends to 0 when n tends to infinity. Therefore, taking

$$\int_0^t \sigma_s^2 ds = 0$$

when dS is dampened by buying and selling against the market direction.

Note that should the operator behave irrationally, the value of the option would then turn negative (see Avellaneda and Paras, 1994).

Another, more formal way to analyze this is through the local time theorems[15]. A Brownian motion augmented by a transaction cost will spend an infinite amount of swings between two price levels.

---

[15] See Karatzas and Shreve (1994).

1.6. The Economic Dimension of the Continuous Time Problem

The economic dimension of the existence of time adjustments in an economy is larger than expected: It applies to practically all aspects of neoclassical economics. Canterbery (1995) dubs it *the case of the missing auctioneer*, as it was known to be after the critique by the post-Keynes Keynesians Leijonhufvud and Clover of the process of price adjustments, in support of Keynes' approach . We can also see a connection with the bounded rationality issue, of which Keynes was a precursor, as he insisted in the aftermath of his *General Theory* on uncertainty of knowledge and foresight as the cause for chronic unemployment of resources. Putting time into an economy creates some more than slightly annoying process of adjustment, which creates markets.

1.6.1. Endogeneity of the Sample Path

At an financial level, the economics toolkit in the analysis of dynamic hedging does not accommodate the continuous-time Arrow-Debreu problem, which will be phrased as follows. In the simple Arrow-Debreu model, an equilibrium is reached in a Walrassian tatonnement as agents simultaneously determine their optimal quantities[16]. By simultaneous here is meant that, in the absence of time, it is possible to reach the optimum without the actual process of bargaining. Introducing continuous time into the equation leads to the issue of the endogeneity

---

[16] As Merton(1992) points out, in one of that author's usual potent footnotes, "which state [of nature] is unaffected by the actions of economic agents, either individually or collectively".

of the sample path. The Arrow-Debreu framework can only accommodate *exogenous* dynamic hedgers since the dynamic hedging activity will impact the path. Grossman(1988) proved that the introduction of information into the equation causes options to no longer become redundant securities. Later, the seminal paper of Grossman and Zhou (1996) proved that dynamic hedging caused mean reversion in asset returns, raises the volatility in some Arrow-Debreu states, and causes volatility to be correlated with volume. His tools applied to portfolio insurance, which is the analog of the replication in Definition 1-2.

1.7. The Fundamental Problem of Option Execution

The mere distinction between positive and negative gamma execution, coupled with the assumption of discrete time leads to the following

Proposition 1-1 A positive gamma dynamic hedger can use either market and limit orders. A negative gamma dynamic hedger can only use limit orders.

A limit order is defined as an order to buy or sell a security at a predetermined price, above the market if the order is a sale and below the current market price if the order is a buy. The limit order is characterised best by the fact that the operator knows the price (though not the time) at which he will transact. A market order is a price independent order to buy or sell a given quantity that is triggered by the market price of the security reaching some price on the screen. The operator does not know the price at which he will transact.

The proof of the proposition is simple: the positive gamma hedger has the choice of putting limit orders since

$$\text{sign}(Q_t\text{-}Q_{t\text{-}\Delta t}) = \text{sign}(S_t\text{-}S_{t-\Delta t})$$

and, with a limit sell order, assuming the sell order is at predetermined level $S_\tau*$

$$Q_t\text{-}Q_{t\text{-}\Delta t}>0 \text{ when } S_\tau*\text{-}S_{t-\Delta t}>0 \text{ since } S_\tau*>S_{t\text{-}\Delta t}$$

where τ is the stopping time to reach $S*$. As to the negative gamma hedger, there will be a lag between $S_\tau*$ and the obtained price $S**$. The reason is that since

$$\text{sign}(Q_t\text{-}Q_{t\text{-}\Delta t}) = \text{-sign}(S_t\text{-}S_{t-\Delta t})$$

the sell order will be below the market, $S_\tau**<S_{t\text{-}\Delta t}$. Owing to the discreteness of time, and given that the orders are put in place during one period and traded the next period, and denoting s the time during which the actual order is executed, we have

$$S_s \neq S_\tau$$

since

$$s>\tau$$

This problem permeates the microstructure of dynamic hedging since time is no longer continuous and there is a lower bound on Δt, as will be discussed.

A negative gamma execution is not anticipating with respect to the price at which the execution is done, while a positive gamma one can be anticipating.

1.8. Back to Basics: Risk-Neutrality and Absence of Convenience Yield

It is a generally forgotten (or unnoticed) fact that the intuition of the risk neutral argument[17] and the linear pricing rule already existed in Keynes (1923). We note that the linear pricing rule culminated with the Black-Scholes-Merton option pricing formula, as the payoff of an option became the limit of a portfolio of dynamic hedges, with presents no variance, rather than that of a single security. The definition below and the lemma allow for a heuristic (and simpler) restatement of the risk-neutral argument. We argue that should there be no possibility for an operator to conduct an operation of the sort Keynes-Working cash-and carry (on both sides, the buying and the selling, as further defined) then the pricing kernel one will apply needs to have its expectation equal to be that of the risk-neutral distribution[18]. We will then use it to prove the convergence of the arbitrage value,

---

[17] The risk-neutral argument can be easily defined as follows: a financial security that can be hedged in a manner to eliminate its sensitivity to market direction, and in a way to present no variance, should not include the market price of risk. The Black-Scholes-Merton formula's breakthrough is based on their argument that μ, the expected return of the stock, can be replaced by m, the difference between its carry and the risk-free rate (assuming both known and constant). More modern approaches (more notable the French School of mathematical finance, whose representative paper on the subject is Geman, El Karoui, Rochet, 1995) change the probability measure rather than the return itself, by use of the Girsanov theorem. See also the next footnote.

[18] A similar approach was presented by Ross (1976) in his formulation of the Associated Martingale Process, AMP. Also see Dybvig and Ross (1987) for a full presentation of the issue.

but not its derivatives (in the next section). The argument's potency becomes more obvious once one sees that had there been a forward market for stocks in 1973, at the time of the Black and Scholes (1973) paper, then the point of risk neutrality would have been obvious owing to the put-call parity rules.

Definition 1-4: *Local fungibility conditions for quantity* q *between period* t *and period* τ *are:*

i) *That the security can be moved through time via a frictionless static cash-and-carry operation.*

ii) *That there exists no regulatory constraints for frictionless borrowing and lending of a security.*

iii) *That there exists a marginal arbitrage operator non-satiated for the quantity q.*

This is a back-to basics extension of the Keynes intuition of the *cash-and-carry* arbitrage[19]. Take S_t, a security offering a yield d. Take a continuously compounded risk-free rate r in the economy, fixed for period starting at t and ending at $\tau > t$. Assume that there exists a *marginal* operator who can borrow at the risk-free rate r for period τ and buy the asset S_t for investment earning d. Then the valuation of the security for delivery at period τ needs to be

[19] See DeRosa (1996), and Grabbe (1996). It is remarkable that the most potent argument in the history of financial theory appeared in the Sunday edition of the Manchester Guardian, later restated in Keynes' *Tract on Monetary Reform* (1923). See Riley and Hirshleifer (1992) for a discussion of the issue.

(1-7 $$S_\tau = S_t\, e^{\,(-r+d)(\tau-t)}$$

Beyond Keynes (1923), Working (1949) links future prices to contemporaneous spot prices by the per-period cost to the marginal trader of holding the underlying asset in inventory. A "convenience yield" can arise otherwise – where the owner of a Future or a forwards can in effect earn extra yield for providing the facility to the speculator or the hedger. In a more modern framework, it is the marginal absence of convenience yield that is defined here as risk neutrality[20].

Define a kernel $g(S_\tau)$ as a Arrow-Debreu price, the value of a security paying 1 unit should the state S_τ take place and 0 otherwise, and $dG(S_\tau)$ a the derivative of a security paying 1 unit if the realizations are higher than S_τ (hence functionally the respective equivalents of a density and a cumulative distribution). Assume a two period economy with no trading in between. Define $g^*(S_\tau)$ as another distribution (different "beliefs"). We call (for obvious reasons) $g(S_\tau)$ the *risk-neutral* pricing kernel for S period τ. We can thus set the conditions for the Rieman-Stieljes integral :

(1-8 $$\int_0^\infty dG(S_\tau) = \int_0^\infty g(S_\tau)\, dS_\tau = 1$$

***Remark**: In a two period model, an option on a security S that meets all the local fungibility conditions between period t and τ can only be valued using*

[20] Note that convenience yields can arise from structural shortage, or from physical limitations on the warehousing of securities.

a pricing kernel that presents the first moment of the associated risk-neutral kernel.

Take a Call on S expiring period τ, C(K,τ), with all the conventional definitions of a call except for the fact that there is no continuous trading.

The expected first moment for S_τ under g becomes:

(1-9 $$\mathrm{E}(S_\tau) \equiv \int_0^\infty S_\tau \, g(S_\tau) dS_\tau = S_t \exp(r-d)(\tau - t)$$

The value under distribution g* is:

(1-10 $$\mathrm{C}(S_\tau, K, g^*) = e^{r(\tau-t)} \int_0^\infty \max(S_\tau - K, 0) \, g^*(S_\tau) dS_\tau$$

therefore:

(1-11 $$\mathrm{C}(S_\tau, K, g^*) = e^{r(\tau-t)} \int_K^\infty (S_\tau - K) \, g^*(S_\tau) dS_\tau$$

By a similar argument, the put value is:

(1-12 $$\mathrm{P}(S_\tau, K, g^*) = e^{r(\tau-t)} \int_0^\infty \max(K - S_\tau, 0) \, g^*(S_\tau) dS_\tau \, ,$$

therefore:

(1-13 $$\mathrm{P}(S_\tau, K, g^*) = e^{r(\tau-t)} \int_0^K (K - S_\tau) \, g^*(S_\tau) dS_\tau$$

Given the standard option algebra that

(1-14 $$C(S,K,\tau,g^*) - P(S,K,\tau,g^*) = \exp[r(\tau-t)] \, E[S_\tau, ,g^*]$$

if $S_\tau = E(S_\tau, g)$, i.e. if S_τ is priced under *g*, then puts and calls need to be priced under a kernel that has the same expectation as *g*. Another way to prove it is by seeing that: if $C(S,K,\tau,g^*) > C(S,K,\tau,g)$, then $P(S,K,\tau,g^*) > P(S,K,\tau,g)$, which is incompatible with the notion of higher expectation than that through the risk-neutral kernel.

The risk-neutral expectation of a security's value, however, only applies for the first moment. We will see that it does not necessarily apply to the first mathematical derivative with respect to the strike price, the delta, ironically, with the discussion below[21].

1.9. The Convergence Problem in Continuous Time and the Delta Paradox[22].

Using continuous-time pricing methods when market price changes have a lower bound can be annoying. We can add to that the fact that operators may not have the motivation to revise their portfolios at the lower bound, owing to transaction costs (a definition of transaction costs can include simple time investment). We will discuss below the problem of the convergence of the derivatives once one takes into account the smallest amount of discreteness.

---

[21] We note that the same results would hold under discrete time. Take $\pi(S\tau_i)$ the equivalent discrete "risk-neutral" Arrow-Debreu price (see elaboration in the general appendix). We satisfy $\sum_{i=1}^{n} \pi_{\tau_i} = 1$ across all states of nature. We also satisfy $\sum_{i=1}^{n} \pi_{\tau i} S_i = S_t \exp(r-d)(\tau - t)$.

[22] This topic has been discussed by the author in Taleb(1997) and debated in various presentations. Eric Benhamou wrote an unpublished master's thesis on the subject, Benhamou (1997).

Take a position ∏= α_1C(K_1)+ α_2 C(K_2)+α_3 S, K_1>S exp(r-d)(τ-t)>K_2, where C is priced on any state price density that satisfies standard smoothness conditions. Select α_1>0 and α_2<0 such that:

(1-15
$$\alpha_1 \left.\frac{\partial^2 C(K_1)}{\partial S^2}\right|_{S=S_0} = -\alpha_2 \left.\frac{\partial^2 C(K_2)}{\partial S^2}\right|_{S=S_0}$$

which satisfies "gamma neutrality". and select α_3 such that:

$$\alpha_1 \left.\frac{\partial C(K_1)}{\partial S}\right|_{S=S_0} + \alpha_2 \left.\frac{\partial C(K_2)}{\partial S}\right|_{S=S_0} + \alpha_3 = 0$$

which satisfies "delta neutrality". It is easy to see that if such portfolio satisfies the conditions above, that

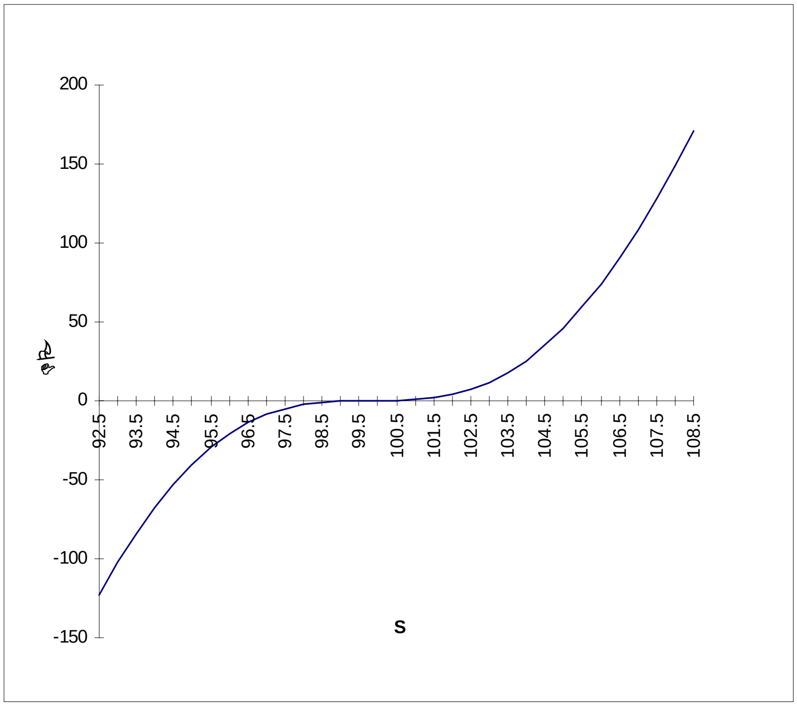
(1-16
$$\left.\frac{\partial \Pi}{\partial S}\right|_{S=S_0} = 0 \text{ and } \left.\frac{\partial^2 \Pi}{\partial S^2}\right|_{S=S_0} = 0$$

and

$$\left.\frac{\partial \Pi}{\partial S}\right|_{S\neq S_0} > 0\,,\ \left.\frac{\partial^2 \Pi}{\partial S^2}\right|_{S>S_0} > 0 \text{ and } \left.\frac{\partial^2 \Pi}{\partial S^2}\right|_{S<S_0} < 0$$

In other words, the portfolio "neutrality" was entirely illusory, even in continuous time. The set S=S_0 has a zero Lebesgue measure, P[S=S_0] = 0, which completes the paradox, as illustrated in Figure 1. As the Figure shows, the delta neutrality is only satisfied when S=100.

Figure 1: The "Delta' Neutral Trade. The vertical axis plots the change in value of the package Π (in 000), where $\alpha_1 = \alpha_2 = \$1000,000$, $K_1 = 105$, $K_2 = 95.5$, $\sigma = .157$, expiration 30 days.



Why is this a problem ? Because under discrete time distribution π such a problem does not occur. Pricing options under π obtains

$$\frac{\Delta\Pi}{\Delta S} \equiv \frac{\Pi(S+\Delta S) - \Pi(S-\Delta S)}{2\Delta S} > 0\,,$$

owing to the fact that the valuation under continuous kernel is the same as the valuation under discrete distribution (but not the opposite since we can have a continuous distribution that can satisfy all requirements for a discrete distribution, without the opposite being true).

Remark 1-2 Π(S) using a continuous pricing kernel will always approximate the valuation Π(S) using a discrete kernel, but such equivalence will not apply to the derivatives $\frac{\partial \Pi}{\partial S}$ *and* $\frac{\Delta \Pi}{\Delta S}$

As to the second derivative of the option with respect to the asset price, the gamma,

$$\frac{\Delta^2 \Pi}{\Delta S^2} \equiv \frac{\Pi(S+\Delta S)+\Pi(S-\Delta S)-2\Pi(S)}{\Delta S^2} = 0$$

which shows that, when prices are discrete, using a continuous distribution provides adequate results for valuation, but not hedging. This concludes the weak convergence problem.

Economic application: Since we have established that dt (continuous time) is not possible to attain, what Δt to chose? Should it be the minimum possible economic Δt? How small does this smallest Δt need to be? It is indeed difficult to study the issue without falling into a utility framework. It will be explained in later chapters that using an inventory framework effectively reduces the Δt to a small holding period for a class of operators (MEDH) that are providers of the service of dynamic hedging.

1.10. Microstructure Focus on Bourses and Equity Instruments

The long exclusion in the traditional research in market microstructure of non-equity markets has affected our understanding of the dynamics of the trading of most non-linear financial instruments. The conventional market microstructure literature, started with Garman (1976, see Biais, Foucault and Hillion ,1997 or Hamon 1995 for a review), is generally limited to equity products, a remnant of the finance literature of the 60s[23]. It is to note that such share represents less than 3% of total volume of contracts traded, which weakens the economic significance of such research. The other issue is that by limiting the analysis to the type of "niche" market making, much of the nonlinear instruments was missed. Perhaps the first major microstructure paper to examine the trading outside of the narrowly defined dealer market was Grossman (1992), as discussed in the next section.

Since the sixties, particularly since 1973, there has been an overwhelming growth in financial Futures and the so-called "upstairs" or over-the-counter trading. The numbers are quite eloquent: Foreign exchange transactions worldwide is estimated to clear $1,000 billion per day : In New York alone the volume is estimated at $400 billion per day on average whereas the New York Stock exchange transactions total less than $25 billion. The total volume of worldwide traded contracts in the Bonds and Eurodollars exceeds $900 billion daily: the bond Futures themselves clear 80 billion and the activity in the underlying swaps and cash equals it[24]. As will be described, these products are mildly nonlinear as there is an It♈ term that puts them in the category of dynamic hedging.

---

[23] There has been a recent trend to break away from the conventional products, as witnessed by recent papers in the *Journal of Finance* focusing on more fixed-income type securities. See recent work by Peiers (1997) that examines the microstructure of foreign exchange transactions.

[24] Sources: Wall Street Journal volume pages (hence WSJ), CME CBOE and CBT data. It is worthy to note that the October 15 issue of the WSJ announcing the Nobel price granted to Robert

Professional-to-professional trading represents generally a larger volume of transactions than those limited to retail or individual investors. Again: while we need to study microstructure to properly formulate and understand option theory, it may also be necessary to study option theory to understand microstructure –a field that should lead to interesting applications in the future. Grossman (1988) discusses what is commonly described as the “positive feedback” effect in the casual literature. The existence of an option in the marketplace will lead to a behavior in the underlying security that is susceptible of changing the conditional variance of the asset.

1.11. Distinction Between Fragmented and Centralized Markets –A Stiglerian Examination

The text will make no clear distinction between listed and “upstairs” trading operations. As we will later introduce the notion of the most efficient dynamic hedger MEDH, we will see that for him the distinction (principally informational) is not relevant to the analysis. Grossman (1992) introduces in the formal literature the distinction “upstairs” and “downstairs” markets. We use here a functional, not physical (i.e. depending on the location of the participants), definition of the

---

C. Merton and Myron Scholes described the pricing formula as having been instrumental on the spread and development of “stock options”, ignoring the volume results in their own pages that would show the preponderance of currency and fixed-income option activity.

difference. The approach here will be Stiglerian, in the sense that the classification will be based on the notion of costs of search.

1.11.1.Downstairs Trading

Definition 1-5 Downstairs, or listed trading, corresponds to operations whose consummation is information officially made available to the general public upon the agreement.

Downstairs trading is typically done in an "open outcry" system, as in a Futures market or a stock exchange. There are regulations forcing the highest bid and lowest offer to be represented at all times. The Stiglerian costs of search will be insignificant in such a market as the trader will incur a cost to discover the highest bid or the lowest offer. We will argue that with the emergence of electronic trading and the creation of upstairs brokerage makes all trading become downstairs in nature.

1.11.2.Upstairs Trading

Definition 1-6 Upstairs trading correspond to operations whose consummation is the sole immediate knowledge of the two participants in the transaction.

Upstairs operators conduct transactions that are not officially recorded in a central clearing house. It has the effect of scrambling public information and violating the law of one price. Two trades can be consummated simultaneously between two sets of two different participants at two markedly different prices without allowing for arbitrage. In addition there is a cost of search, both in time and direct costs, to discover a price as the dealer needs to sample between a series of participants. Furthermore, this can induce price dispersion. Stigler(1960) distinguishes between embedded quality differential:

> "Price dispersion is a manifestation -and indeed it is the measure- of ignorance in the market(...) There is never absolute homogeneity in the commodity if we include the terms of sale within the context of the commodity."

The impact on dynamic hedging is far from trivial. It means simply that transaction costs for option replication in a full upstairs market include a cost of search heretofore ignored in the literature, simply the cost of "shopping", as will be discussed later. When the trader buys at S(1+k/2) and sells at S(1-k/2) as stated in the transaction costs literature (starting with Leland, 1985), the k/2 does not represent the full transaction costs. For a recent discussion of the informational difference between the two types (assuming the distinction across these lines is clearly marked, which we contest further down), see Biais (1993).

1.11.3.The Mixture: Do pure Upstairs Markets Exist?

It can be argued that the existence of intermediaries and agents between counterparts reduces the existence of pure upstairs operations –even if the agents do not partake of all trades. The fact that there exists a cost of search for all counterparts creates an opportunity for an upstairs broker to step in and derive a profit for offering a search service for a small nominal fee. An upstairs broker is defined here as an agent who "searches" for the prices in a continuous way, either by continuous solicitation of quotes or by acting as an information repository for future trades. The existence of an upstairs broker between dealers creates de facto a centralized market. The existence of broker “boxes” obligating brokers to shout the last traded price would prevent violations of the law of one price. Another consideration is the correlation between instruments. We will attempt to provide below a brief intuition of the issue with a simplified model.

1.11.3.1.A Simplified Model

We start with the general case. Assume a (n,1) vector of securities for period t, $\mathbf{X_t}$ indexed by X_{it} , i={1,…,n}. Denote the corresponding returns over period Δt by

$$r_{it} \equiv \log\left(\frac{X_{it}}{X_{i(t+\Delta t)}}\right)$$

Further assume that the logarithmic returns are (multivariate) normally distributed with mean (n,1) $\boldsymbol{\mu}$ and variance (n,n) Σ. Use a vector (n,1) of weights **w**, w_i taking the value 0 or 1 (therefore non normalized). The value 1 corresponds to the event of the trading of the corresponding component of the vector X in the interval]t, t+Δt[and 0 for no trading. Create a new covariance matrix Σ**’** of rank m such that

$$m = \sum w_i$$

by eliminating the rows and columns of the returns of the assets that did not trade. Take the lower triangular Cholesky decomposition (m,m) matrix **C** such that

$$\mathbf{C}\ \mathbf{C}^{\mathbf{T}} = \Sigma'$$

Take the (1,m) vector **L** as the m^{th} row of **C**, with components l_z, z={1,…,m}.

Next we infer the variance, of non-traded security X_j, assumed from the conditional period return $r_i| \mathbf{w_t} \, \mathbf{X_t}$, i.e. the basket that traded. The conditional volatility of X_j will be the scalar

$$\sigma_j \, \mathbf{\Sigma' \, L}$$

To illustrate, what follows examines the special case of a security whose price is inferred from another, more liquid one, that, in addition, is representative of the vector. Assume that X_i is traded in a public forum, either a liquid box market (with trades announced to the dealers by brokers upon consummation) or on a listed exchange. Traders will be able to get an idea of the state of the vector **X** conditional on the last trade X_i and estimate the conditional expected value of the other instruments. In the event of X_i being highly correlated to X_l a liquid security, the variance of the conditional expectation of $V(X_i|X_l)$ will be sufficiently low and will only affect the share of randomness:

$$\sigma_i \sqrt{1-\rho_{i,l}^2}$$

where σ_i^2 is the unconditional variance of the returns of X_i and $\rho_{i,l}$ the correlation between i and l. Thus the fact that many “fragmented” markets trade a security that presents a strong correlation with another one traded in a centralized market may weaken the differential information for our purpose.

There are also reputational issues weakening the fragmentation of markets. A customer’s trade may go unnoticed by any but the two parties but there are enough checks in place, particularly when institutional customers are obligated to verify the quotes with more than one institution. A customer can verify the closeness to the price at which he transacted with that prevailing in the broker market at similar time and possibly inflict a reputational damage to the bank in the event of a blatant violation. Moreover there are occasional leaks as traders do not have necessarily the legal obligation to keep the information private. The mechanism of broker verification has been successful (see Taleb, 1997) with the recording of trigger prices for barrier options.

A clear example would be swaps and bonds: The fact that there is a bond Future that presents a strong correlation to these instruments makes the analysis of "fragmented markets" inappropriate. Most markets will thus be somewhat centralized.

Appendix: Some Further Definitions

Financial Contracts in the Arrow-Debreu Framework

This section is a brief review of the economic thought attending the theory of financial contingent claims contract. A financial contingent claims transaction is defined in the Arrow and Debreu equilibrium as a contract between two parties leading to an exchange of contingent securities the payoff of which depends on the realization of some state of nature (see appendix for further refinements of the issue) for period T. The economic principles underlying such transaction in the neoclassical model is (see Arrow, 1953,1964 and Debreu, 1954, 1959) that, assuming a static model in which the adjustments do not take place in real time, and, conditional upon a given price,

1) The trade takes place if each of the two parties' welfare is increased as a direct result from the outcome of the transaction,

2) The parties in an unhindered market will keep transacting until some allocation optimality is reached (no gain from trade) , i.e. the joint welfare reaches an optimum.

Two types of agents will thus be facing each other: buyers and sellers of contingent claims. The buyer of contingent claims is eliminating uncertainty as can be seen. The seller of contingent claims needs to eliminate his own uncertainty: in the continuous-time world the Black and Scholes (1973) argument takes care of the problem since he will eliminate his uncertainty through dynamic revision (with variance asymptotically nil), *provided there are no transaction costs.* Such provision has been theoretically difficult to handle (aside from the problem of transforming the static path independent Arrow-Debreu model into a continuous one, as discussed before, without endogenizing the activity of the dynamic hedger).

Internal transaction costs here will correspond to the *friction* in the dynamic replication of Arrow-Debreu states by a specialized (and risk-neutral as will see) agent, only conducting self financing "admissible" strategies (a requirement that can be easily circumvented to allow, as will be described, a small variance provided that the integral of the cash debit equals that of the credit[25]) . If the

[25] Clearly the self financing requirement is a technical requirement to prevent the infinite expectations of an exponential strategy of a Saint-Petersburg type. We can easily obtain the same

security is static, then the costs will be the initial hedge (there is no contingency as the security spans the entire distribution). If the security is defined as nonlinear (i.e. including an Ito term, as discussed in chapter 1), then the costs will correspond to the discounted value of:

Final Payoff + Expected Cost of Portfolio Revision [26]

To allow for the package to be a martingale the price of the security from the standpoint of a bidding agent needs to satisfy the above equation in a convincing way, as is discussed in chapter 1. While the notion of ***marginal break-even*** is granted in economic analysis (the intellectual parent to our notion of martingale), that of risk neutrality, which we will equally retain, has more disturbing features. How can discuss risk-neutrality when accepting friction and infrequent hedging, given the accepted market price of risk ? The operator, between two revisions, as pointed out by Boyle and Emanuel (1980), Gilster(1990) and explained in the general appendix, in addition to the more provocative Gilster (1996), incurs a large share of variance. His systematic risk is not eliminated, thus he may require compensation. We described elsewhere the redundancy problem in the presence of market friction; chapter two proffered a possible solution.

---

results by putting a less restrictive constraint on the portfolio: simply by showing the existence of a reachable limit.

[26] $E^Q f(S_T) - \sum_{i=0}^{n} \mathbf{1}_{ti} Q_{ti}\ k$ where T is the expiration time, $f(S_T)$ the expiration value of the security,

$\mathbf{1}_{ti}$ an indicator function taking value of 0 if there is no revision and 1 if there is a revision..

In this thesis, we will use a marginal analysis of transaction costs, the bedrock of neoclassical economics. Hence as we discuss option replication the same principle of marginal profits will dominate this discussion. There are no reasons to dwell on the argument in favor of risk-neutrality –see Kyle(1989) for a discussion of the assumptions of the Kyle(1985) model where both risk neutrality and zero marginal profits are assumed.

Self Financing Strategies

The self financing requirement was born with the Saint Petersburg paradox, solved by Daniel Bernouilli in 1738. It will be discussed in chapter 7 with the problems of pure probability measurement. A coin is tossed *n* times until the first head appears; 2^n ducats are then paid out. The paradox lies in that the mathematical expectation of gain is infinite although common sense said (even then) that a sum to play the game needs to be finite. In modern terms, it was then discovered that an unbounded martingale has infinite expectations, unless there is a discounting of it that is concave (such as a concave Von-Neuman Morgenstern) or unless there is a cutoff in time (the game cannot be played perpetually). In the rational expectations literature the discount factor took care of the problem (the transversality condition): so long as a future stream of cash flows is discounted, if can be infinite and still have a finite expectation. This prompted Ramsey, in his debate with Keynes, to interpret the mathematical expectation as meaning other than simple discounting operation (an argument that would be later on fully formalized by Savage) –thus setting the ground for today's credence that probability is indissociable from utility.

The importance of self financing trading strategy was stressed in financial economics by Black and Scholes 1973 and Merton 1973 who constrained their portfolio to have no variance at the limit of dynamic revisions. In a more formal way, a strategy θ_t is self financing if and only if (assuming no dividends):

$$\theta_t \,.\, X_t = \theta_0 \,.\, X_0 + \int_0^t \theta_s dX_s$$

where X is a stochastic process in $\Re^N$, under condition that the variance of every component of X , $V[X^i_t]$, is finite for all i and all t. Some technical conditions are necessary (such as square integrability on the stochastic integral). Clearly the condition is that the account at period t does not include any cash infusion and the market value corresponds to initial value plus marked to market changes.

Next we define an arbitrage. A strategy is an arbitrage[27] if:

$$\theta_0.X_0 < 0 \text{ and } \theta_T.X_T \geq 0$$

or

$$\theta_0.X_0 \leq 0 \text{ and } \theta_T.X_T > 0 \,.$$

We can see the connection between arbitrage and the linear pricing rule. Casually, absence of arbitrage is connected to the existence of a linear pricing rule (see

[27] See Dybvig and Ross (1987), Ingersoll (1986), Duffie(1996), Dana and Jeanblanc-Piqué (1994), for reviews.

Dybvig and Ross, 1987)– in $\Re^n$ there cannot be a combination of securities outside of the positive orthant where all securities are inside of it.

Change of Probability Measure and Discrete time Processes

We define equivalent martingale measures as follows: the measure Q defined on a probability space (Ω, F), with F the filtration, is said to be equivalent to P provided, in any event, we have

i) Q(A) >0 if and only if P(A)>0 for all A $\in F$, and
ii) The Radon-Nikodym derivative dQ/dP has a finite variance.

We recall that recent improvement in the fundamental theorem of asset pricing allow to equate absence of arbitrage opportunities with the existence of an equivalent probability measure under which the discounted price process is a martingale. We have the following two implications:

i) Existence of equivalent martingale measure implies absence of arbitrage.
ii) Absence of arbitrage and change of numeraire implies existence of equivalent martingale measures.

The more precise interpretation in terms of change of numéraire was given in Geman, El Karoui, Rochet (1995).

Note that these results were formalized in discrete time. A piece of research by Elliott and Madan (1997) applies the continuous time results of the martingale properties to the discrete time case. Their result simplifies asset pricing under dynamic replication as it allows the bypassing of continuous time portfolio revision to create the risk-neutral bridge between continuous processes and state price density. In other words, there is no need for the Black-Scholes continuous revision to allow for the risk-neutral dynamic replication of an asset price: there exists a probability measure under which such *discrete* process is a martingale.

Taking the continuously compounded return defined by

$$w_{jt} \equiv \ln(1+s_{jt})$$

They prove that the joint density one period ahead $\psi(w_{1,t},\ldots,w_{Nt}|F_{t-1}\}$ under P is equivalent to $\psi'(w_{1,t},\ldots,w_{Nt}|F_{t-1}\}$ under Q. We note the relevance of the work as it does treat the problem from the standpoint of the change of probability measure to accommodate a change in the discrete time pricing deflator. In other words it proves that, in discrete time, there is a risky asset that can be used (just as in continuous time) to deflate risky returns and establish a martingale.

Appendix : Dynamic Hedging Variance Reduction: an Arrow-Debreu State Price Approach

This section shows how the dynamic hedging of a position reduces the risk. Truly this approach is not altogether new to finance: An exactly similar problem was presented in Stigler(1961) pioneering work on information, where uncertainty is examined in terms of reduced information. In the Stiglerian framework, agents know that perfect information leads them to perfect knowledge of price (i.e. minimum variance of result), but it is only a limit. Likewise here the dynamic hedging agent can attain perfect dynamic replication with no variance, as by the Black and Scholes (1973) argument, as a limit of continuous time hedging. Number of searches in the Stiglerian approach and revisions in the Black-Scholes approach play identical roles. We will see in the next chapter how, furthermore, the number of dynamic hedges can result in an increase of the knowledge of the *true* distribution.

The approach followed below can provide more insight than the differential equation as it allows for the calculation of the variance using time dependent parameters. It shows that the dynamic hedging over a period, as by the functional central limit theorem, is multiplied by a factor of $1/\sqrt{n}$, n being the number of rebalancing actions during the activity.

Take a European option expiring period T, examined at period 0. We use $n(X_T)$ as the continuous-time analog to the Arrow-Debreu state price, defined as today's expected value of a security with the following payoff

$$p=1 \text{ if } S_T = X_T$$

$$p=0 \text{ elsewhere}$$

We use $N(X_T)$ as the security with the following payoff

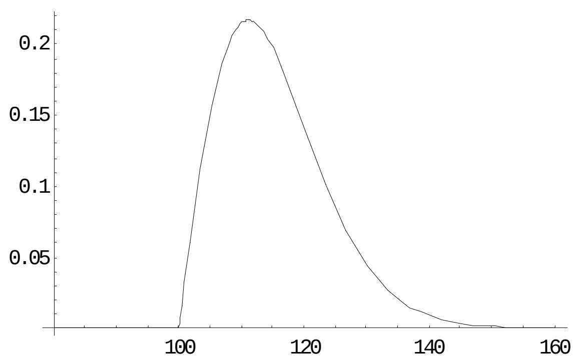
$$p=1 \text{ if } S_T < X_T$$

$$p=0 \text{ elsewhere.}$$

The final unhedged risk neutral payoff for a long option position is $E_0[\Phi(S_t)]$, where $\Phi(S_t) = Max(S_t-K,0)$. We will use the stretched term distribution to denote this pricing kernel.

$C(S_0)$ as the market value of the option expiring period t at period 0, which should correspond to $E_0[\Phi(S_t)]$.

Figure 2: A naked option's price distribution, showing on the vertical axis the density of the payoff multiplied by the payoff, assuming t= 1 year and σ=.157



The variance from the naked package, which we call P^0 becomes:

$$V[P^0] \equiv \int_0^\infty (\Phi(S_t) - C(S_0))^2 n(S_t) dS_t = \int_0^\infty \Phi(S_t)^2 n(S_t) dS_t - C(S_0)^2 .$$

Adding a delta hedge to the operation at period 0, which is defined as the derivative of the option value with respect of the expectation under the pricing kernel, i.e.

$$\frac{\partial C(S_t)}{\partial S_t} \equiv \left.\frac{\partial C}{\partial S}\right|_{S=S_t}$$

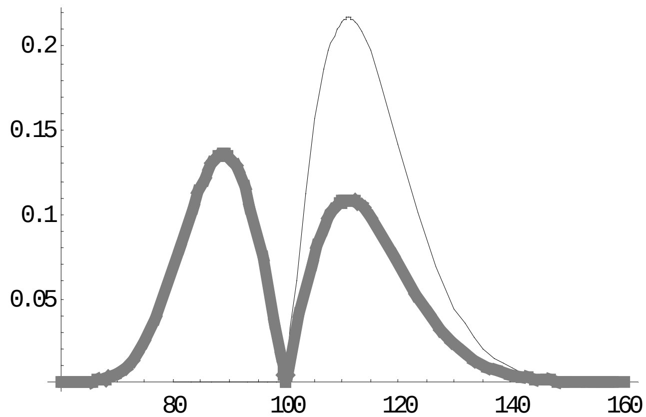
as the delta period t associated with price S and calling the package P^1, the variance becomes

$$V[P^1] \equiv \int_0^\infty \left(\Phi(S_t) - C(S_t) - \frac{\partial C(S_0)}{\partial S_0}\left(S_t - S_0\right)\right)^2 n(S_t) dS_t$$

expanding it

$$V[P^1]=V[P^0]-2\frac{\partial C(S_0)}{\partial S_0}\int \Phi(S_t)(S_t-S_0)n(S_t)dS_t+S\sqrt{t}\left(\frac{\partial C(S_0)}{\partial S_0}\right)^2$$

Figure 3: Comparing a "naked" call to a Delta Hedged One by showing on the vertical axis the density of the payoff multiplied by the payoff of both structures, assuming t= 1 year and σ=.157



Adding some trades in between at period t/2 changes the profile. We now face more than one leg: the distribution between 0 and T of $\Phi(S_t)$ and the distribution of the P/L from gamma hedging, i.e.

$$\frac{\partial C(S_0)}{\partial S_0}\left(S_{t/2}-S_0\right)+\frac{\partial C(S_{t/2})}{\partial S_{t/2}}\left(S_t-S_{t/2}\right)$$

Next we define a conditional Arrow-Debreu security, one that depends on more than one occurrence: $n[S_t|S_{t/2}]$ becomes the conditional value of an Arrow-Debreu state S_t, conditional upon the occurrence of state $S_{t/2}$.

As to the distribution of $S_{t/2}$ conditional on S_t, it becomes simple to examine it in a Brownian bridge framework: Conditional on S = S_t, the density of $S_{t/2}$ can be inferred as follows. Note that $n(S_t)$ is a shortcut here for $n(S_t|S_0)$. We need $n(S_{t/2}|S_0 \cup S_t)$, which we write as shortcut $n(S_{t/2}|S_0)$.

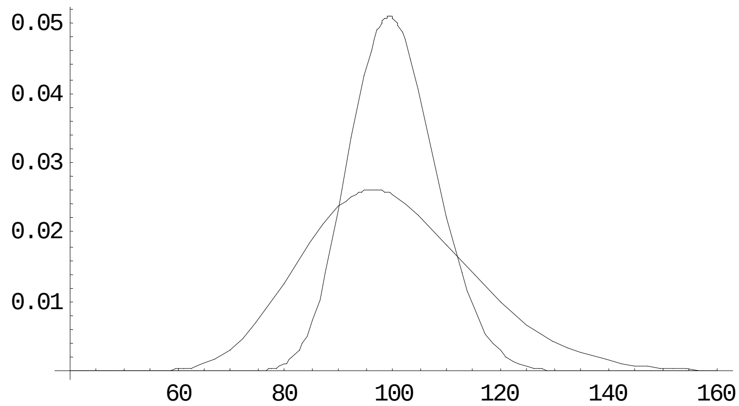
(1-17 $$n(S_{t/2}|S_t) = \frac{n(S_t|S_{t/2})n(S_{t/2})}{n(S_t)}$$

which can be computed as

$$n\left(S_{\frac{t}{2}}|S_t\right) = \frac{\sqrt{\frac{2}{\pi}}}{\sigma S_{t/2}\sqrt{t}} \exp\left(\frac{\left(\frac{1}{2}\sigma^2 t + \log\left(\frac{S_t}{S_0}\right)\right)^2}{2\sigma^2 t} - \frac{\left(\frac{1}{4}\sigma^2 t + \log\left(\frac{S_t}{S_{t/2}}\right)\right)^2}{\sigma^2 t} - \frac{\left(\frac{1}{4}\sigma^2 t + \log\left(\frac{S_{t/2}}{S_0}\right)\right)^2}{\sigma^2 t} \right)$$

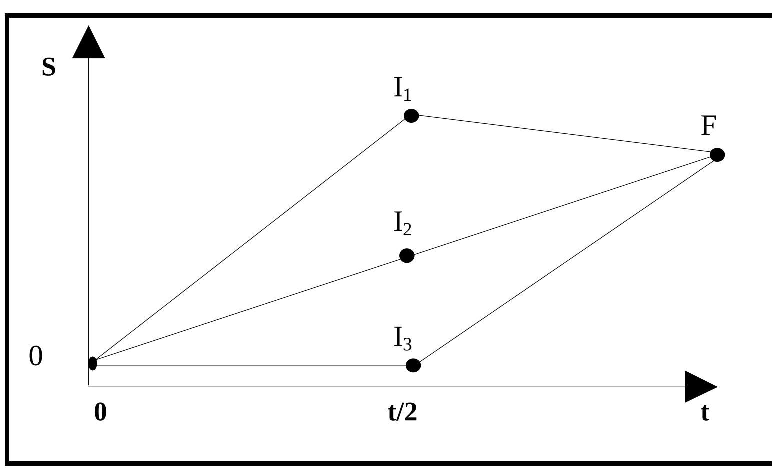
It is obvious that the conditional distribution of $S_{t/2}$, conditional on S_t (itself conditional on S_0) has a lower variance than both that of S_t (as well as that of $S_{t/2}$). Such a fact will be the reason for the variance lowering gamma hedging at period *t/2*. Graphically this shows in the following example, where a one year 15.7 distribution is plotted against the conditional six month where S_t=100 and S_0=100.

Figure 4: Unconditional 1 year and Conditional 6 month state price density. The graph illustrates the decline in conditional variance from one single dynamic revision.



Furthermore it can be easily shown that the mean of $S_{t/2}$ will be located somewhere between S_0 and S_t .

Figure 5: Possible Values for $S_{t/2}$. Revising the delta at I_1 will cause a loss, while I_2 will cause a mild profit and I_3 will cause a larger profit for the option short hedger, conditional on the asset price (started at O) terminating at F.



The intuition shown by Figure 5 is that rehedging on some path will bring a loss, like the lower node, while rehedging on others, like the upper one, will bring a profit. Overall, however, we will see that the variance is reduced.

The variance of the portfolio with just 2 rehedges becomes:

(1-18

$E[\Phi(S)- C(S_t)+(S_{t/2}-S_0)\ \partial C(S_0)/\partial S_0)+(S_t-S_{t/2})\ \partial C(S_{t/2})/\partial S_{t/2})]^2$

And with n hedges:

$$E[\Phi(S)- C(S_t)+\sum (S_i-S_{i-1})\ \partial C(S_i)/\partial S_i]^2$$

(1-18 can be written as:

$$V(0,t)=\int_0^\infty\int_0^\infty\left(\phi(S_t)-C(S_0)+(S_{t/2}-S_0)\frac{\partial C(S_0)}{\partial S_0}+(S_t-S_{t/2})\frac{\partial C(S_{t/2})}{\partial S_{t/2}}\right)^2 n(S_t)n(S_{t/2}|S_t)dS_t dS_{t/2}$$

as to V(0,t/2), it can be calculated as:

$$V(0,t/2)=\int_0^\infty\int_0^\infty\left(C(S_{t/2})-C(S_0)+(S_{t/2}-S_0)\frac{\partial C(S_0)}{\partial S_0}\right)^2 n(S_{t/2})dS_{t/2}$$

Combining with (1-18 shows that, thanks to the assumption of constant and time-independent parameters, we have

(1-19 $$V(0,t)=2\ V(0,t/2)=2\ V(t/2,t)$$

It can be shown that, given the property that $E[\Gamma_t]$ is invariant with t, the variance of a portfolio that is rebalanced twice will be twice that of a portfolio of a period half the length. By induction, we can show a result usually obtained with the functional central limit that:

$$\sigma_{0,t}=1/n \sqrt{\sum \sigma_{0,t/n}}$$

We will present a different approach in the appendix, where a differential equation will be used and will be tied-in with the Boyle and Emanuel (1980) and Leland (1985) approach.