Sentiment Lost: the Effect of Projecting the Empirical Pricing Kernel onto a Smaller Filtration Set

Carlo Sala[*] | Giovanni Barone-Adesi[†]

Swiss Finance Institute and University of Lugano | Swiss Finance Institute and University of Lugano

October 29, 2015

[*]Carlo Sala, Institute of Finance, University of Lugano, Via G. Buffi 13, CH-6900 Lugano, Switzerland. E-mail: carlo.sala@usi.ch.

[†]Giovanni Barone-Adesi, Institute of Finance, University of Lugano, Via G. Buffi 13, CH-6900 Lugano, Switzerland. E-mail: giovanni.baroneadesi@usi.ch.

Sentiment Lost: the Effect of Projecting the Empirical Pricing Kernel onto a Smaller Filtration Set

Abstract

Supported by empirical examples, this paper provides a theoretical analysis to show what is the impact of an improper calibration of the physical measure on the estimation of the empirical pricing kernel. While extracting the risk-neutral measure from option data provides a naturally forward looking measure, extracting the real world probability from a stream of historical returns is only partially informative, thus suboptimal with respect to investors' future beliefs. In virtue of this disalignment, most of papers present in literature are then affected by the a homogeneity bias.

From a probabilistic viewpoints, the missing beliefs are totally unaccessible stopping times on the coarser filtration set. As a consequence, an absolutely continuous local or strict local martingale, once projected on it, becomes continuous with jumps.

As a result of a non fully informative physical measure, the proposed empirical pricing kernel is no longer a true martingale, as required by the classical theory, but a strict local martingale with consequences on the probabilistic nature of the relative risk-neutral measure.

Finally we show how the implied options' moments help in reducing the degree of inaccessibility and shorten the distance between what is theoretically required and empirically accessible.

Keywords: Physical Measure, Radon Nikodym, Empirical Pricing Kernel, Filtration set, Local Martingale, Strict Local Martingale.

1 Introduction

After presenting the econometrical issues associated to the estimation of a fully time varying physical measure, this paper analyzes, from a probabilistic viewpoint, what is the impact of using a suboptimal information set in estimating the real world measure and its effects on the empirical pricing kernel (henceforth: EPK). Although this is not an empirical paper, to increase the quality of the study, the theoretical analysis is supported by empirical results extracted from Sala (2015)[37] and Sala and Barone-Adesi (2015)[38]. Two are the main innovations of the paper. First it extends the recent findings of Protter (2015)[36] to the analysis of the pricing kernel (henceforth: PK). Second, it proposes, for the first time, a probabilistic investigation on the improper use of the information for the estimation of the real world measure that composes the EPK.

Stocks and options market data are the usual inputs used to estimate the real[1] and risk-neutral measures of the EPK. However, while stock returns are, by structure, fully backward looking hence only partially informative, option surfaces are naturally forward looking. Due to this disalignment, it turns out that most of papers present in literature that study the EPK are affected by a non-homogeneity bias: they compare a fully conditional forward looking risk-neutral measure extracted from option surfaces with a partially conditional real world measure extracted from a time series of historical returns.

Following Sala (2015)[37] and Sala and Barone-Adesi (2015)[38] we complete the physical measure using part of the information extracted from the option moments. Once the option moments are estimated empirically through an asymmetric GJR-GARCH-FHS model, they are blended together thorough a fully non-parametric Dirichlet process. The revised obtained measure, now fully conditional, is the new physical measure[2]. In this paper we don't aim to extend the proposed model empirically, instead we analyze the theoretical framework that underpins and justify the use of a fraction of risk-neutral measure at denominator.

2 Intuition behind the theoretical problem

Investors' subjective beliefs are forward looking. An investor decides if and how to trade depending, among the others, on her personal beliefs. Since the investment horizon spans from the present into

[1]The real world measure has many different names in literature. Among the others it is also known as the physical measure, the subjective measure, the personal measure and others. Throughout the paper we refer to is as either the real or the physical measure.

[2]More details in Sala (2015)[37] and Sala and Barone-Adesi (2015)[38].

the future it is, by definition, uncertain. This uncertainty represent the degree of riskiness of an investment. Therefore, the valuation of any risky investment, has to take into account the forward nature of the subjective beliefs.

Counterintuitively with respect to their nature, investors beliefs are estimated with backward looking information (i.e.: Aït Sahalia and Lo (1998)[1], Jackwerth (2000)[26], Brown and Jackwerth (2001)[8], Rosenberg and Engle (2002)[14], Barone-Adesi et al.(2008)[5], Yatchew and Härdle (2006)[40]). Using these data, an important fraction of the investor's risk and preferences are lost. As a consequence, a discrepancy between what is empirically obtainable and what is theoretically required by the neo-classical asset pricing literature arises. The larger the forward looking information bias, the larger is the subsequent mispricing.

From the point of view of the information filtration, this can be translated into a shrunk information set:

$$\mathcal{H}_t \subset \mathcal{F}_t \qquad \forall\, t > 0 \tag{1}$$

where the two information sets are increasing[3] in time and contain all available and potentially usable information:

$$\mathcal{H}_t = \{x_{-\infty}, \ldots, x_{t-1}, x_{t-\Delta_t}\} \qquad \text{and} \qquad \mathcal{F}_t = \{x_{-\infty}, \ldots, x_{t-1}, x_t\} \tag{2}$$

where Δ_t represents the fraction of missing forward looking information that involves any risky decision to undertake from today, t, with respect to a future time, $t+\tau$. From here on we consider $\mathcal{F}_t$ as the theoretical (or complete) information set and $\mathcal{H}_t$ as the suboptimal information set.

Future stochastic events are usually modelled by means of conditional expectations[4]:

$$\mathbb{E}^P(X|\mathcal{F}_t) = \int_{\mathcal{F}} x dP \qquad \text{for each} \quad F \in \mathcal{F} \tag{5}$$

[3]Two assumptions underpin this statement: the first is that information is time-varying, the second is that decision makers keep memory of the entire past data.

[4]Let us assume that at time t we want to forecast the tomorrow's value of a random variable X_{t+1} given the set of available information I_t. This is an optimization problem; more precisely we pick the best predictor among all possible predictors by choosing the one that minimizes the expected quadratic prediction error:

$$\mathbb{E}^P[(x_{t+1} - \widehat{x}_{t+1|t})^2|I_t] \tag{3}$$

Given the problem, the best (minimum mean squared error: MMSE) predictor is the conditional expectation given the information set:

$$\hat{x}_{t+1|t} \equiv \mathbb{E}^P(x_{t+1}|I_t) \tag{4}$$

where the conditionality is with respect to what is known at time t: the information set.

Let us apply the generic case to a finance related problem: the today prediction of the tomorrow price P_{t+1} (a random variable) given, I_t, the information available at time t:

$$P_{t,t+1} = \mathbb{E}^P(P_{t,t+1}|I_t) = \int_{-\infty}^{+\infty} P_{t,t+1} f(P_{t,t+1}|I_t) dP_{t,t+1} \qquad \text{for} \quad t < t+1 \tag{6}$$

where the above integral represents the weighted sum of the averaged possible future values under the physical probability. Given the time horizon of the prediction, here $t, t+1$, the sum of all possible future values of $P_{t,t+1}$ are wighted by their probabilities $f(P_{t,t+1}|I_t)dP_{t,t+1}$ and then summed[5]. From (6) emerges clearly the importance of a right use of the information set. The average is in fact computed using the I_t conditional probabilities. All information available thus enter into the $P_{t,t+1}$ forecast. The poorer the information set, the worse the forecast.

From the FTAP, under some technical assumptions, the same expectation can be estimated using different probability measures i.e.: the physical $\mathbb{E}^P$ and the risk-neutral $\mathbb{E}^Q$. Although always valid theoretically, the equivalence of the EMM may be (and is often) violated empirically. An improper conditioning of the information is usually at the heart of the violations. Being the risk-neutral measure an unbiased estimator of future investor's beliefs[6], an a.s. inequality arises if the information that compose the physical measure is suboptimal as a consequence of possible missestimations. For example, if we compute the today price, $G_{t,T}$, of a generic contingent claim that expires at time T through the fundamental equation of the asset pricing under the two different

[5]The extremes of the integral may also be defined i.e.: put and call options are bounded by the strike price either above or below.

[6]This might not be true in presence of i.e.: options mispricing and/or wrongly estimated implied volatility.

measures:

$$
\begin{aligned}
G_{t,T} &= e^{-r_t\cdot\tau}\mathbb{E}^{Q}[\varphi_{t,T}(S_T)|\mathcal{F}_t] && (7)\\
&= e^{-r_t\cdot\tau}\int_0^\infty M_{t,T}(S_T)\cdot\varphi_{t,T}(S_T)p_{t,T}(S_T|S_t)dS_T && (8)\\
&= e^{-r_t\cdot\tau}\mathbb{E}^{P}[M_{t,T}(S_T)\cdot\varphi_{t,T}(S_T)|\mathcal{F}_t] && (9)\\
&= e^{-r_t\cdot\tau}\mathbb{E}^{P}[M_{t,T}(S_T)\cdot\varphi_{t,T}(S_T)] && (10)\\
&\neq e^{-r_t\cdot\tau}\mathbb{E}^{\widehat{P}}[\widehat{M}_{t,T}(S_T)\cdot\varphi_{t,T}(S_T)] && (11)\\
&= e^{-r_t\cdot\tau}\mathbb{E}^{\widehat{P}}[M_{t,T}(S_T)\cdot\varphi_{t,T}(S_T)|\mathcal{H}_t] && (12)\\
&= e^{-r_t\cdot\tau}\int_0^\infty \widehat{M}_{t,T}(S_T)\cdot\varphi_{t,T}(S_T)p_{t-\Delta,T}(S_T|S_{t-\Delta})dS_T && (13)\\
&= e^{-r_t\cdot\tau}\mathbb{E}^{\widehat{Q}}[\varphi_{t,T}(S_T)|\mathcal{H}_t] && (14)\\
&= \widehat{G}_{t,T} && (15)
\end{aligned}
$$

where, for each $t \in T$, we used the convention:

$$\widehat{x} = \mathbb{E}^{\widehat{P}}(x|\mathcal{H}_t) \tag{16}$$

to represents the conditional expectation under the suboptimal information set and where $\tau = T - t$ represents the time-to-maturity while $\varphi_{t,T}(S_T)$ the terminal payoff of the product given the value of the underlying S_T. The inequality in (11) is due to the missing forward looking information (relative to T) of the today (time t) risk physical density:

$$p_{t,T}(S_T|S_t) \neq p_{t-\Delta,T}(S_T|S_{t-\Delta}). \tag{17}$$

Although the missing information is referred to the future, the impact is today. As a consequence, while the left hand side is fully conditional to all todays values, the right hand side is not. Moreover, is clearly visible how an improper estimation propagates also on the risk-neutral pricing through the PK. The difference in (17) is the focus of the paper.

It follows naturally that a suboptimal information set has also an impact on the relative probability measures:

$$\mathcal{F} \Rightarrow P \qquad \text{and} \qquad \mathcal{H} \Rightarrow \widehat{P} \tag{18}$$

$$\mathcal{F} \Rightarrow Q \qquad \text{and} \qquad \mathcal{H} \Rightarrow \widehat{Q} \tag{19}$$

where P (resp. $\widehat{P}$) represents the physical probability measure for a the optimal $\mathcal{F}$ (resp. sub-optimal $\mathcal{H}$) information set. The same logic applies for the risk-neutral measures. The difference between $G_{t,T}$ and $\widehat{G}_{t,T}$ represents the distance between an empirically biased asset pricing with respect to the theoretical one. This distance represents the sub-optimality of the information set.

In a nutshell: conditional expectations reflect the change in unconditional probabilities given some auxiliary information; the definition and use of these information is then of fundamental importance to correctly evaluate this change. If the filtration used is missing of relevant information, hence is suboptimal, the projection onto the smaller set leads to an inequality:

$$\mathbb{E}^P(X|\mathcal{F}_t) \neq \mathbb{E}^{\widehat{P}}(X|\mathcal{H}_t) = \int_{\mathcal{H}} xdP \qquad \text{for each} \quad H \in \mathcal{H} \tag{20}$$

Conditional expectations are of great importance in finance. The definition of a martingale, upon which are based the FTAP is nothing but a conditional expectation.

2.1 Econometric problems behind the estimation of the real world measure

The PK is defined as the discounted ratio of the risk-neutral over the risk physical measure. It follows by definition that these three stochastic processes are among them tightly related. As advantage, the knowledge of two of them implies automatically and uniquely the knowledge of the third one. As a consequence, as a disadvantage, a misestimation of one of the two also loads on the third one. The literature is abundant of different methodologies, both parametric and non parametric[7] which aim to estimate the two measures and the EPK, either jointly or separately.
While theoretically the EPK is a decreasing function of aggregate resources, many empirical papers found violations in different areas of the functionals (Aït Sahalia and Lo (1998)[1], Jackwerth (2000)[26], Brown and Jackwerth (2001)[8], Rosenberg and Engle (2002)[14], Yatchew and Härdle[40], Ziegler (2007)[41]). The persistency and robustness of these violations put interest to

[7] The latter have to be preferred. From Fama (1965)[15] and Mandelbrot (1966)[32] it is well-known that none of these quantities are neither normal nor have a closed parametric form.

the investigation of the so called pricing kernel puzzle. Since then, researchers have taken great interest in proposing different econometric techniques to estimate the EPK and its measures trying to answer to the puzzle. From the literature it emerges that, among the others, the most problematic econometric task to perform is to assure full conditionality to the time varying estimation of the physical density.

Empirically, it turns out that to propose a day-to-day estimation methodology of the measures that compose the EPK is as much important as econometrically non trivial. The importance of a correct estimation of these measures lie in the wide use of the EPK for many daily operations (i.e.: asset pricing and risk management). The econometric issues, as also partially commented in Bliss and Panigiritzoglou (2004)[7], lie in the nature of the underlying. Using historical data many estimation methodologies put unreal and theoretically not required stationary assumption on the estimation of the measures or on the PK itself. Therefore, it is not by chance that many works only propose monthly or yearly estimations remaining silent for the daily ones. It turns out that, from this viewpoint, the main reason behind a biased estimation is not the technique used for the estimation but rather the data used as input.

Needless to say, models' inputs are of key importance to determine the type and the quality of the final outputs. Market option prices provide a forward looking measure; in fact, by contract, the owner of an option has the right, but not the obligation, to exercise it at expiration (or before it, if it is an American or Bermudian option). This feature is reflected into the option value, which is a non-decreasing function of volatility. Therefore, the market prices of options, through the volatilities, embed important forward looking information of the future distribution of prices of the underlying asset. Also the higher moments of the distribution embed important information. This is particularly true into the tails of the distributions: where most of sentiment lies.

The same richness cannot be achieved by stock and index prices. By their contractual nature these assets are options-free and Markovian, hence only backward looking.

Estimating the EPK and extracting the risk aversion from stock prices is a well-known piece of theory in the literature. Despite their unambiguous superiority in estimation, it is only from the beginning of the millennium that scholars have begun using options data for estimations (Chernov and Ghysels (2000)[9]).

The superiority in estimation of options with respect to stocks (and also futures) is manifold.

First, stock prices have discounting as well as time-horizon problems. By contract definition, stocks do not expire but live infinitely. They are defined over an indefinite time horizon; therefore the discounting process becomes non trivial. As a consequence, additional assumptions (which often times are unreal, i.e.: on the characteristics of future dividends), are needed to determine the discounted cash flow. On top of that, the obtained final outcome is not very much informative since the discounted final cash flow is just a single value; this means that no inference about variations in preferences over different time horizons is possible.

On the contrary, option contracts have, by definition, a bounded life that is defined by a fixed time-to-maturity, T, and is known from the inception of the contract [8]. Moreover, for each time t, we have the so called option surface: a broad spectrum of time-to-maturities, T_i and strikes K_j[9] that covers different states of the world. These characteristics allow for inference on preferences over specific horizons and simultaneously over different horizons and strikes.

These features make options qualitatively superior also to future and forward [10] contracts which, by their nature, do not share with stock prices the discounting but only the time-horizons problem. In fact, even though these contracts have bounded expirations, they have just one single expiration day, thus providing only a single statistic for each expiry date/observation date pair. Therefore, having a single data, a direct density estimation is not possible without further assumptions [11].

Last but probably most important, extracting densities from the option surfaces does not share the previously mentioned stationarity problems. In fact, while we can directly estimate a time varying risk-neutral density from a cross section of single options observations with no need to bound the structure of the data we sample from, the same is not true from the time series of stock prices. For the physical density, in fact, to obtain decent results from just using a time series of stock returns, we need to put a priori unrealistic bounds on the time structure of the data. While the degree of stationarity can be of different length, in no case can be justified economically

Given the above characteristics, it is natural that just using historical stock data it is not possible to properly capture the investors future beliefs. As (a non)alternative, some papers propose to

[8] Only perpetuity options differ from this characteristic, but are an exceptional case, more theoretical than really used in practice. American and Bermudian options instead, have the flexibility to be exercised at any time or at fixed time before T. Although non trivial, a reliable discounting is still possible.

[9] Both indices are finite, i.e. $i = 1, \dots, T$ and $j = 1, \dots, N$.

[10] For our purposes we can ignore the margin requirements required by the marking-to-market provision and consider forward and futures as equal.

[11] As noted by Jackwerth and Rubinstein (2006)[27], starting from at least 8 option prices we have enough information to determine the general shape of the implied distribution. Although possible, to have too a few data in estimation may lead to possibly dangerous small sample bias.

increase the degree of conditionality by increasing the rolling window of the estimations. Jackwerth (2000)[26] tries to "enrich" the informative content of the physical measure by working on a longer time series. He proposes to use ten instead of the classical two to five years of data for estimation. As expectable, the obtained results are qualitatively almost unchanged. Due to the intrinsic backward nature of the dataset, to increase the length of the rolling window used in estimation is not a solution. At some extremes it might also lead to misleading results. In fact, by using more and more data, from a statistical perspective we gain from the point of view of the conditionality but we reduce the informative content of the last data which, also in the case in which it would be informative of the future, it would be negligible and almost totally prevailed by the high amount of past data used.
The difficulty in estimating the objective measure is that it directly depends on the evolution of the underlying process, which time series is not fully informative. The estimations are further complicated by the high amount of possible different model specifications and by basic data-problems (i.e.: data scarcity). For these reasons, some authors fully avoid the density estimation and propose a ratio of estimated risk-neutral measure over an unknown physical measure (Golubev et al. (2008))[18].

While the estimation of the real world densities presents several difficulties, it is by now an empirical and theoretical fact that the most important information embedded in financial instruments is the state price density (SPD), or the Arrow-Debreu state prices.
The time-state preference model of Arrow (1964)[2] and Debreu (1959)[3], which proposes the now named Arrow-Debreu security, models a very basic financial instrument (called pure or primitive security) that pays one unit of numeraire (like a currency or a commodity) on one specific state of nature and zero elsewhere. Passing from discrete to continuous states, Arrow-Debreu securities are defined by the so called state price density (SPD). Under the continuous framework the security pays one unit of numeraire x if the state falls between x and $x + dx$ and zero elsewhere. As a consequence of their high informative content, the Arrow-Debreu securities become one of the key element to work with and understand the general economic equilibrium under uncertainty and to determine the price of any contingent claims. For these reasons the estimation of such SPDs has been a very important topic of research within the financial economics community. Once that a complete set of option prices for a specific time-to-maturity is available [12], there are many para-

[12] By now we have a huge amount of data available. This richness is an important element for the finance research community and it was not available since the end of the millennium. Things changed right after and now there is a very large amount of simultaneously traded liquid options that differ for time-to-maturity and strikes.

metric and non-parametric methods to recover the risk-neutral measure [13].

To conclude: to obtain a time-varying estimation of the real world density it is of key importance, more than to pick the right extraction model, to use the right source of information. Being historical stock returns only backward looking, hence non fully informative, we need to somehow complete the measure by using other sources of information, i.e.: the investors' future sentiment extracted form the implied moments of the option surface.

2.2 PK as the unconditional Radon Nikodym derivative

Pricing in a risk-neutral world has the extraordinary advantage of using a unique probabilistic measure which is neutral and so applicable to all investors. Unfortunately, if the representative investor's risk attitude is not neutral, the obtained values may provide misleading results. As a link among the two measures, there is the PK which is the collector of all beliefs, errors and premiums of the investor: in a nutshell it embeds all the relevant information required to convert the risk-neutral measure into a real world measure and viceversa. Mathematically, given its role, it is convenient to express the PK as a discounted Radon Nikodym derivative. It turns out that the PK is nothing but a discounted kernel: a time adapted operator which, working as a transition function of a stochastic process, allows us to move from a neutral to a subjective world. In this section we first define the unconditional Radon Nikodym derivative, then we move to a time dependent version.

Proposition 2.1. *Defined on a measurable space* $(\Omega, \mathcal{F}, \mathbb{F})$*, the unconditional version of the Radon Nikodym derivative requires that:*

1. *P is a σ-finite measure on $(\Omega, \mathcal{F}, \mathbb{F})$ and is atomless;*
2. *Q is another σ-finite measure on $(\Omega, \mathcal{F}, \mathbb{F})$ and is absolutely continuous with respect to P on $\mathbb{F} : Q \ll P$[14];*
3. *the Radon-Nikodym derivative of Q with respect to P, is a nonnegative Borel measurable function M defined on the extended real line ($M : \mathbb{R}^+ \to \mathbb{R}^+$) and satisfies $P\{x \in \mathbb{R}^+ : M(x) = y\} = 0$ for all $y \in \mathbb{R}^+$.*

If assumptions 1 and 2 are satisfied, then there exists an a.s. (with respect to measure Q)

[13] See Bahara (1997)[4] and Jackwerth (1999)[25] for a complete review of the topic.
[14] More on the absolute continuity of the measures in Appendix (A).

random variable M which follows directly from the application of Appendix (A) and must obey the last point of the above proposition.

It follows that the unconditional version of the Radon Nikodym derivative, is defined as:

$$M = \frac{dQ}{dP} \tag{21}$$

and satisfies the following:

- $Q(A) = \mathbb{E}^P[M \cdot \mathbb{1}_A], \qquad \forall\, A \in \mathcal{F}$
- $\mathbb{E}^P[M] = 1,$
- $M > 0, \qquad$ P - a.s.
- $M \in \mathcal{F}$

The obtained value, is a kernel, an operator. Once discounted, equation (21) leads to the PK. Mathematically, equation (21) is also called the likelihood ratio between Q and P on $\mathcal{F}$ or the change of measure.

The technical requirements of proposition (2.1) have all a natural economic interpretation and the extension to their time dependent counterparts is, at least theoretically, immediate. The importance of the atomlessnes[15] of the underlying probability space is analyzed also by Hens and Reichlin (2011)[22]. They show how, in a complete market, having an atomic probability space, not al bets are possible. As a consequence the whole spectrum of payoffs is not achievable. This leads to a flex on the PK, thus to a violation[16].

The absolute continuity requirement can be translated in economic term as absence of arbitrage. If the assumption about the equality of the two measures with respect to the negligible events were violated, there would exist a set $A \in \mathcal{B}(\mathbb{R}^+)$ with $P(A) = 0$ and $Q(A) > 0$. Although the inequality would apply only to rare events, a violation of the absolute continuity of the events it is all that is necessary for having an arbitrage in the market. In such a scenario, a contingent claim with payoff equal to the indicator function for this set would have an a.s. strictly positive price together with

[15] Roughly speaking, atomlessnes is linked with continuity and atomicity with discreteness of the probability space. An atomless space put zero mass to a given unit, i.e.: the probability of any particular value is equal to zero. Formally: given a measurable space and two measurable sets: $A \in \Omega$ and $B \subset A$, a set $A \in \Omega$ is called an atom if:

$$P(A) > 0 \qquad \text{and} \qquad P(B) = 0 \qquad \text{where} \qquad P(B) < P(A) \tag{22}$$

Is atomless when:

$$P(A) > P(B) > 0 \qquad \text{where} \qquad P(B) < P(A) \tag{23}$$

this lead to a continuity of values for the non-atomic measure.

[16] They perform a test to check for the risk attitude of the investors and test for possible risk seeking behaviours. In a complete market, if the underlying probability space is not atomless, the shape of the PK passes from concave to convex in the central area.

an a.s. zero payoff. Shorting such a product would lead to an a.s. profit with no risk.
Finally, the third requirement ensures no flatness for the PK, hence a unique ordered ranking[17].

3 From a totally continuous to a mixed process: the Radon Nikodym derivative approach

The projection of a martingale onto a smaller filtration is still a martingale, the same holds for super and quasimartingales (Stricker (1977)[39]). Things change for local martingales (Follmer and Protter (2010) [17]) and strict local martingales (Protter (2015)[36]). This is especially true when the process is no longer adapted (hence not measurable) to the smaller filtration.
Due to their widespread use, both local and strict local martingales are of key importance in mathematical finance[18]. Following a paper of Johnson and Helms (1963)[30] in which they refined the concept of "class D" supermartingales, and extending the Doob-Meyer decomposition (1954 -1963)[12][33] the notion of local martingale has been created by Itô and Watanabe (1965)[24]. The concept has then been used extensively in finance through the different theories involving the stochastic integration.
Strict local martingales have been introduced in finance by Herrison and Kreps, Harrison and Pliska (1978-1981)[20][21] and Delbaen and Schachermayer (1994 and 1998)[10][11] on their works on the absence of arbitrage and the no free lunch with vanishing risk (NFLVR) condition. More recently, the concept has been extended to models of financial bubbles and to test the validity of stochastic volatility models.
While the literature on the expansions of filtrations is, since the seminal work of Itô (1978), a pretty rich field of the literature[19], there is not much work related to the shrinkage of filtrations. It is only recently that the subject has gained more attention. In mathematical finance, most of attention on the subject has been put on the existence of financial bubbles[20] and in credit risk (see, among the others: Jarrow at al. (2007)[28]). The literature is still silent on the behaviour of the PK to filtration shrinkage.

[17] See Beare (2011)[6] for more details.
[18] Strict local martingales play an important role in different field of the recent finance literature, among the others: for the Benchmark approach of Platen and Heat (2006)[34], for the analysis of relative arbitrage of Fernzhold and Karatzas (2010)[16] and for the understanding of the financial bubbles of Heston at al.(2007)[23] and Jarrow et al. (2010)[29].
[19] For a review on the subject: see Chapter 6 of Protter (2001)[35].
[20] From a theoretical viewpoint, the existence of financial bubbles is due to the different behaviours of martingales and local martingales to possible filtrations' shrinkage.

Following the terminology of Elworthy, Li and Yor (1998)[13] a strict local martingale is locally a martingale but not a true martingale while a local martingale is a martingale equipped with a sequence of stopping times.

Loosely speaking, a strict local martingale remains such on a smaller filtration if and only if there is a sequence of reduced (projected) stopping times that is common to both filtrations. In case this is not possible, we pass from a totally continuous to a mixed process made of a continuous part and a set of totally unaccessible stopping times. These inaccessibility represents the information lost due to the smaller filtration. This is exactly what happens when one neglects forward looking information from the filtration set of the physical distribution. Section (4) shows empirically this fact. To turn back to a valid strict local martingale, one needs of an absolutely continuous compensator.

As anticipated above, the PK is nothing but a change of measure. The passage from the physical to the risk-neutral measure is mainly a change of drift which passes from being subjective (usually represented with μ) to be an objective and common value (represented with r). In our case, assuming the risk-neutral measure as unbiased, it follows that what one observes under a smaller filtration set is "less" then a subjective drift, what is missing are the subjective beliefs that affect the expected returns of the investor.

Having at denominator of the PK a "non-updated" unconditional physical distribution not only we may have puzzling results, but we also end up with a non continuous process with jumps at the extremes of the function[21]. Completing the measure with forward looking information, not only we solve for the non-monotonicity, but we also reduce the number of jumps, above all in the left tail of the distribution, the one more sensitive to change of investors sentiment[22].

3.1 Projecting the Radon Nikodym derivative onto a smaller filtration set

In this section, by means of the Radon Nikodym derivative, we apply the general theory of Protter (2015)[36] to the financial PK. As a first step, we move toward the conditional version of the Radon Nikodym derivative. If we restrict our attention to the conditional version of the above results we

[21]It is a fact that the PK is moving and, by the Hansen and Jagannathan bounds (1991)[19], it needs of a high enough variance to be valid. Anyhow, this is not to be confused with irrationally jumping values which lead to economically invalid results.

[22]The right tails of the distribution often suffer of low liquidity due to the low trading of call options. Therefore, explosive values in the right tails is not due to a lack of forward looking information into the information set but is a lack of the entire information set. The same happens for the left tails, but with a lower frequency.

have a natural link with the first FTAP.

From (21) we recall[23], without proving, that:

$$Q_t(A) = \mathbb{E}^P[M_t \cdot \mathbb{1}_A], \qquad \forall A \in \mathcal{F}_t \tag{24}$$
$$= \int_A M_t dP_t \tag{25}$$

then, certainly $Q_t \ll P_t$ for all $t \in T$. The Radon Nikodym theorem goes on the opposite direction:

Theorem 3.1. *If P_t and Q_t are two σ-finite measures on $(\Omega, \mathcal{F}, P, \mathbb{F})$ and $(\Omega, \mathcal{F}, Q, \mathbb{F})$ such that $Q_t \ll P_t$ on $\mathcal{F}_t$ for all $t \in T$, then there exists a sequence of strictly positive M_t:*

$$M_t = \mathbb{E}^P(M|\mathcal{F}_t) = \frac{dQ}{dP}\bigg|_{\mathcal{F}_t} > 0 \qquad \forall t > 0 \tag{26}$$

such that $Q_t(A) = \int_A M_t dP_t$ for all $A \in \mathcal{F}_t$.

From theorem (3.1), it follows that:

Proposition 3.1. *If $(M_t)_{t>0}$ is an $\mathcal{F}_t$-adapted sequence of strictly positive martingales on a complete and filtered probability space, with i.e. $\Omega = \mathbb{R}^T$ and $\mathbb{E}^P[M_t] = 1$, then $(M_t dP_t)_{t>0}$ is a consistent family of probability measures on $(\mathbb{R}_t)_{t \in T}$.*

From Harrison and Kreps (1979) [20], the new measure satisfies the necessary conditions to be an equivalent martingale measure (EMM), namely the probabilistic equivalence of the measures, the existence of a non-negative value from the Radon-Nikodym derivative and the martingale property of the price process under the change of measure. This leads to the first FTAP which states that a market has no free lunches with vanishing risk (NFLVR) if and only if there exists an EMM. For a proof, see [10] [11].

As anticipated above, it is a well-known result that some categories of semimartingales remain such, once projected onto a coarser and adapted filtration. Before we move to the main theorem, we analyze the properties of these processes.

Theorem 3.2. *Let M_t be a martingale/supermartingale/quasimartingale on $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, then the*

[23] As a main difference with respect to section (21) here we ripropose the same theories but with time dependency.

optional projections[24] *onto a smaller filtration* $\mathbb{H} = (\mathcal{H}_t)_{t\in T}$ *are still a martingale/supermartingale/quasi-martingale for the filtration* $\mathbb{H}$.

Following [17] and exploiting the high interconnection of the three classes of martingales, we prove the above theorems all together.

Proof: Martingale If $s \leq t$:

$$\mathbb{E}^{\widehat{P}}(\widetilde{M_t}|\mathcal{H}_s) = \mathbb{E}^{\widehat{P}}(\mathbb{E}^{\widehat{P}}(M_t|\mathcal{H}_t)|\mathcal{H}_s) = \mathbb{E}^{\widehat{P}}(M_t|\mathcal{H}_s) \tag{28}$$

$$= \mathbb{E}^{\widehat{P}}(\mathbb{E}^{P}(M_t|\mathcal{F}_s)|\mathcal{H}_s) = \mathbb{E}^{\widehat{P}}(M_s|\mathcal{H}_s) \tag{29}$$

$$= \widetilde{M_s} \tag{30}$$

where the a.s. equality of $\mathbb{E}^{\widehat{P}}(M_t|\mathcal{H}_t) = \widetilde{M}$ is a direct consequence of the uniform integrability of the "class D" processes $(M_t)_{0\leq t\leq n}$ where n is arbitrarily extended to ∞.

Supermartingale: The validity of the proof is based upon the non-negativity of M_t. From the FTAP, a necessary condition for the existence of the PK is its strict non-negativity. If follows that:

$$\mathbb{E}^{\widehat{P}}(\widetilde{M_t}|\mathcal{H}_s) = \mathbb{E}^{\widehat{P}}(\mathbb{E}^{\widehat{P}}(M_t|\mathcal{H}_t)|\mathcal{H}_s) = \mathbb{E}^{\widehat{P}}(M_t|\mathcal{H}_s) \tag{31}$$

$$= \mathbb{E}^{\widehat{P}}(\mathbb{E}^{P}(M_t|\mathcal{F}_s)|\mathcal{H}_s) \leq \mathbb{E}^{\widehat{P}}(M_s|\mathcal{H}_s) \tag{32}$$

$$= \widetilde{M_s} \tag{33}$$

Quasimartingle: A quasimartingale is just a difference of two supermartingales:

$$A_t = B_t - C_t \tag{34}$$

where B_t and C_t are positive and right continuous supermartingales and $A(t)$ is a quasimartingale iff there exists a constant D_t such that:

$$\sup \sum_{1\leq i\leq n} \mathbb{E}(|X(t_i) - \mathbb{E}(X(t_{i+1}|\mathcal{F}(t_i))|) \leq D_t \tag{35}$$

[24] Any nonanticipating and right continuous (cadlag) process is an optional process. If $X = (X_t)_{t\geq 0}$ is a bounded and measurable stochastic process, then the optional projection $\widetilde{X}$ of X is the a.s. unique (and bounded) process such that for any stopping time τ:

$$\mathbb{E}^{P}(X_\tau \mathbb{1}_{\tau<\infty}) = \mathbb{E}^{\widehat{P}}(\widetilde{X}_\tau \mathbb{1}_{\tau<\infty}) \tag{27}$$

.

where the supremum is taken over all finite sets of the partition $t_1 < t_2 \cdots < t_n < t_{n+1}$[25].

As a consequence of the previous proof the result follows. $\square$

Martingales, supermartingales and quasimartingales are examples of semimartingales. Theorem (3.2) extends the Stricker's theorem (1977)[39] for non-adapted semimartingales.

3.2 Filtration shrinkage and information inaccessibility for strict local martingales

As anticipated, things change when we deal with local and strict local martingales. In fact, optional projections of local and strict local martingales need not be local and strict local martingales anymore. Projected onto a smaller filtration set the finite variation part may acquire a drift which could be singular (with respect to $d\langle\widetilde{M}, \widetilde{M}\rangle_t$). As a consequence the stability of the process is no longer guaranteed. Following the recent paper of Protter (2015)[36] we focus the attention on these processes and we extend his results for the PK and EPK.

The NFLVR theorem states that under an equivalent martingale measure (EMM) positive prices are local martingales. A filtration shrinkage may pose stability problems for the cited theorem. As a consequence, if an investor has a suboptimal information set it may be affected by mispricing and "relative" arbitrages.

Using the concept of total inaccessibility, we firstly define when a filtration set is significantly smaller. A totally unaccessible stopping time is the flipping side of a a predictable stopping time:

Definition 3.1. *A map $\tau : \Omega \to \mathbb{R}^+$ is a predictable stopping time if there exists an announcing sequence of stopping times τ_n where a.s. and for each n: $\tau_{n+1} \geq \tau_n$, $\tau > \tau_n$ and $\lim_{n\to\infty} \tau_n = \tau$. We define a totally unaccessible stopping time U if for any predictable stopping time τ: $P(\tau = U) = 0$.*

Examples of totally inaccessible stopping times include the jump times of Poisson processes.

Theorem 3.3. *A sub-filtration is significantly (and sufficiently) poorer with respect to its larger counterpart if it has less stopping times and/or if some of them are present but totally unaccessible.*

[25] Dealing with quasimartingales the partition is usually over $[0, \infty]$ such that: $0 = t_1 < t_2 \cdots < t_n < t_{n+1} = \infty$. The inclusion of ∞ in the index set makes it homeomorphic to $[0, t]$ for $0 < t \leq \infty$.

Keeping the interconnection between theory and empirics alive, an empirical technical detail may help to clarify. Given our framework, the smaller filtration set is not as such because it contains less data, but because these data (stock returns) are qualitatively less informative than the ones forming the larger set (options data).

The effect of not capturing the forward looking information of the investor as required by the theory is that $\mathbb{H}$ has a significantly and/or qualitatively lower amount of information. With respect to the above definition it means that the smaller filtration set loses some or many $\mathbb{F}$ stopping times and some of the remaining become totally unaccessible. Most of them lie into the tails of the distribution[26], where sentiment matters the most. Theorem (3.3) clarifies and better charachterizes the value of Δ defined in (2). A smaller filtration set is then defined as sub-optimal for pricing purposes.

Theorem 3.4. *If M_t is a sequence of absolutely continuous strict local martingales defined on $(\Omega, \mathcal{F}, \mathbb{F})$ for each $t \in T$, its projection onto a significantly and sufficiently poorer smaller filtration $\mathbb{H}$ became a mixed process with random jumps:*

$$M_t = \mathbb{E}^P(M|\mathcal{F}_t) = \frac{dQ}{dP}\bigg|_{\mathcal{F}_t} \neq \frac{d\widetilde{Q}}{d\widetilde{P}}\bigg|_{\mathcal{H}_t} = \mathbb{E}^{\widehat{P}}(M|\mathcal{H}_t) = \widetilde{M_t} \qquad \forall\, t > 0 \tag{36}$$

where $\widetilde{Q}$ and $\widetilde{P}$ are the projection of P and Q onto $\mathcal{H}$.

For a local martingale, to be still such under a smaller filtration, it needs a reducing sequence of stopping times that is common for both filtrations. The extension for strict local martingales is mainly a technical detail. The main message is the same: if the two processes do not stop at common points and if these points are not qualitatively equals they cannot share the same properties. Comparing theorem (3.2) and (3.4) it is immediate to notice that the reason of the instability of the projected processes lies on the reproducibility of the set of stopping times. Since under a sufficiently poorer information set $M_t \neq \widetilde{M_t}$ and since we assume the risk-neutral measure to be unbiased[27] $\widetilde{Q}_t = Q_t$, it follows naturally that the reason of the inequality is to be found into the projected subjective measure $\widetilde{P}_t \neq P_t$.

To summarize in economic terms, the missing forward looking information not captured by a biased physical distribution leads to sufficiently poorer filtration set, thus to missing stopping times and

[26] Although the larger effects are usually manifested into the tails, the inability of capturing meaningful information (Δ) appears also into the central area of the PK. The term pricing kernel puzzle has been originated to explain the non concavity of the functional into the central area of the graph.

[27] We are implicitly assuming, the possibly non true condition of absence of market mispricings.

to jumps at the extremes of the PK function.

4 Empirical Application

To conclude the analysis, in this section we show empirically what just defined theoretically. Figure (1), taken from Sala (2015)[37], shows and compare two fully and partially conditional EPKs estimated from a random day. The left column shows the estimated densities, the right one the relative EPKs obtained taking the ratio of these same densities. The figure is also divided horizontally to shows the entire spectrum of time-to-maturity. As a only difference among the two methodologies the physical measure of the partially conditional EPK is estimated fitting a GJR GARCH model onto 3500 daily stock returns while the real world measure of the fully conditional EPK is made time dependent by blending the same physical measure with the risk-neutral one[28]. Both estimations are proposed with Gaussian and empirical (FHS) innovations. For the densities, the dotted lines represent the "classical" risk physical and neutral measures while the continuous ones the fully conditional physical measure. For the EPKs the continuous (dotted) lines represent the fully (partially) conditional EPKs.

As anticipated theoretically, when the amount of data is high enough so that a meaningful statistical inference is possible and statistically accurate, the fully conditional functionals are well behaved while the partially conditional ones explode. As expectable, the degree of missing stopping times and inaccessibility gets higher as we move to shorter times-to-maturity. This is a classic result and is related to degree of moneyness that gets smaller for short-living options. On the contrary, where the liquidity of the surface gets higher along the entire horizontal support, the information extracted from the options' implied moments makes the stopping times accessible.

To make the picture clearer and as a confirmation of the previous findings, the top panel of figure (2) shows the same events for another random day but from a closer look.

The figure only represents a single day single time-to-maturity. To facilitate the analysis, the horizontal and vertical lines shows respectively the risk-neutral area and the area of no return. As a robustness, and to show the validity of the results from another prospective the bottom panel of figure (2) shows the same estimations but for a day with a higher liquidity of DOTM and OTM

[28]The risk-neutral measure is extracted by minimizing the distance between real and simulated options market data.

call options.

Here it is clearly visible how the inaccessibility of physical measure is translated on the right side of the figure. While the partially conditional EPKs have exploding values thus producing a U shaped functional, the information extracted from the options produce an EPK that is in line with the neoclassical theory.

5 Conclusion

We have shown how a suboptimal filtration affects the measures and the EPK. As defined in proposition (3.1), the martingale property of the Radon-Nikodym derivative process is crucial for a correct pricing under the risk-neutral approach. It is in fact a necessary condition that makes sure that the risk-neutral measure has a total mass of one, allowing it to be a proper probability measure.

By this property it follows that if and only if the PK is a positive[29] martingale process, there exists a valid risk-neutral measure: the previously defined EMM. From the definition of the strict local martingale, it follows that if the PK is only locally a martingale but not a true martingale, it is not guarantee that the related risk-neutral measure Q is a valid probability measure. Due to the central role of the PK and the measures that compose it, these different relations pose challenges in many areas of financial economics. As a demonstration, at the heart of the existence of financial bubbles and illusory arbitrages, there is the above cited relationship. In fact, it is when two investors are differently informed that one ends up valuing an investment as a martingale and one as a strict local martingale. The disalignment may lead to illusory mispricing, hence to non existing arbitrages and to bubbles.

Whereas what we showed empirically only partially explains this phenomenon, it surely gives an important intuition on this direction. What we present is indeed an empirical estimation that, although it substantially improves the degree of conditionality and richness of the information set, it is not immune to errors. Being estimated with a numerically intensive method and using noisy real data the model can only produce approximated final results. It is a classical result in numerical mathematics that algorithmic and data errors lie at the hearth of any numerical model that tries to describe a real world phenomenon. Moreover, specifically for the estimation of the EPK there are different small sample bias (Leisen (2015)[31]) which may affect both the parametric

[29]Positivity is only a requirement to be a valid PK under the FTAP.

and non-parametric estimation methodologies used in estimation. Therefore, we do not claim that our proposed fully conditional EPK is, from a theoretical point of view, a true martingale. Surely, results show that is a martingale for a much larger fraction[30] than the proposed unconditional EPK. What we propose, then, is a less biased EPK.

As a main message, from the evaluation it follows that a better informed physical measure leads to an informatively less biased EPK.

Due to the tight interconnection among the measures and the PK, a proper estimation of the physical measure is of fundamental importance if one would extract the risk-neutral measure from the estimated PK[31].

[30]This is true both horizontally, domain of the gross returns, and vertically, domain of the EPK

[31]The risk-neutral measure is biased if derived from a biased physical measure. In our case, the risk-neutral density is estimated using a cross section of market option data, hence independently form the risk physical measure obtained from a time series of stock returns. As a consequence, unless the market data are mispriced, its estimate is unbiased.

6 Bibliography

[1] Y. Äit-Sahalia and A. W. Lo. Non-parametric estimation of state-price densities implicit in financial asset prices. *The Journal of Finance*, 53(2):499–457, Apr 1998.

[2] J.K Arrow. The role of securities in the optimal allocation of risk-bearing. *The Review of Economic Studies*, 31(2):91–96, Apr 1964.

[3] J.K Arrow and G. Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, 22(3):265–290, Jul 1954.

[4] B. Bahara. Implied risk-neutral probability density functions from options prices: Theory and application. *Bank of England ISSN 1368-5562*, 1997.

[5] G. Barone-Adesi, R. Engle, and L. Mancini. A garch option pricing model in incomplete markets. *Review of Financial Studies*, 21(1223-1258), 2008.

[6] B.K. Beare. Measure preserving derivatives and the pricing kernel puzzle. *Journal of Mathematical Economics*, 47:689–697, 2011.

[7] R.R. Bliss and N. Panigirtzoglou. Option-implied risk aversion estimates. *Journal of Finance*, 59:407–446, 2004.

[8] D.P. Brown and J. Jackwerth. The pricing kernel puzzle: Reconciling index option data and economic theory. *Working Paper*, 2004.

[9] M. Chernov and E. Ghysels. A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of Financial Economics*, 56:407–458, 2000.

[10] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathemathische Annalen*, 300:463–520, 1994.

[11] F. Delbaen and W. Schachermayer. Fundamental theorem of asset pricing for unbounded stochastic processes. *Mathemathische Annalen*, 334:215–250, 1998.

[12] J.L. Doob. *Stochastic Processes.* Wiley, 1953.

[13] K.D. Elworthy, X.-M.Li, and M. Yor. The importance of strictly local martingales: Applications to radial ornstein-uhlenbeck processes. *Probability Theory and Related Fields*, 115:325–355, 1999.

[14] R. Engle and V. Rosenberg. Empirical pricing kernels. *Review of Financial Studies*, 64(341-372), 2002.

[15] E. Fama. The behavior of stock-market prices. *The Journal of Business*, 38(1):34–105, Jan 1965.

[16] D. Fernzhold and I. Karatzas. On optimal arbitrage. *Annals of Applied Probability*, 20:1179–1204, 2010.

[17] H. Follmer and P. Protter. Local martingales and filtration shrinkage. *ESAIM Probability and Statistics*, 14:825–838, 2011.

[18] Y. Goulbev, W. Härdle, and R. Timonfeev. Testing monotonicity of pricing kernels. *SFB 649 Working paper*, 2008.

[19] L. Hansen and R. Jagannathan. Implications of security market data for models of dynamic economies. *Journal of Political Economy*, 99(2), 1991.

[20] J. M. Harrison and D.M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

[21] J.M Harrison and S.R. Pliska. Martingales and sochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:215–260, 1981.

[22] T. Hens and C. Reichlin. Three solutions to the pricing kernel puzzle. *Working paper, University of Zurich*, 2011.

[23] S. Heston, M. Lowenstein, and G. Willard. Options and bubbles. *Review of Financial Studies*, 20:359–390, 2007.

[24] K. Itô and S. Watanabe. Transformation of markov processes by multiplicative functionals. *Annals de l'institut Fourier*, 15:15–30, 1965.

[25] J. Jackwerth. Option-implied risk-neutral distributions and implied binomial trees: A literature review. *Journal of Derivatives*, 7(2):66–82, Winter 1999.

[26] J. Jackwerth. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies*, 13:433–451, 2000.

[27] J. Jackwerth and M. Rubinstein. Recovering probability distributions from option pirces. *The Journal of Finance*, 51(5):1611–1631, Dec 1996.

[28] R. Jarrow, P. Protter, and A.D. Sezer. Information reduction via level crossings in a credit risk model. *Finance and Stochastics*, 11(2):195–212, April 2007.

[29] R. Jarrow, P. Protter, and K. Shimbo. Asset price bubbles in incomplete markets. *Mathematical Finance*, 20:145–185, 2010.

[30] G. Johnson and L. Helms. Class "d" supermartingales. *Bulletin of the American Mathematical Society*, 69(1):59–62, 1963.

[31] D. Leisen. Small sample bias in pricing kernel estimations. *Available at SSRN: http://ssrn.com/abstract=2409661*, Jan 2015.

[32] B. Mandelbrot. Forecast of future prices, unbiased market, and "martingale" models. *Journal of Business*, 39(1):242–255, 1966.

[33] P-A. Meyer. Decomposition of supermartingales: the uniqueness theorem. *Illinois Journal of Mathematics*, 7(1):1–17, 1963.

[34] E. Platen and D. Heath. *A Benchmark Approach to Quantitative Finance.* MR22653260. Springer Finance, Berlin, 2006.

[35] P. Protter. *Stochastic Integration and Differential Equations.* Version 2.1. Springer Verlag, Heidelberg, second edition edition, 2005.

[36] P. Protter. Strict local martingales with jumps. *Stochastic Processes and their Applications*, 125(4):1352–1367, April 2015.

[37] C. Sala. Does the pricing kernel anomaly reflect forward looking beliefs? *Working Paper*, 2015.

[38] C. Sala and G. Barone-Adesi. The pricing kernel density: the case of the information that did not bark. *Working Paper*, 2015.

[39] C. Stricker. Quasimartingales, martingales locales, et filtrations naturalles. *Wahrscheinlichkeitstheorie verw. Gebiete*, 39:55–63, 1977.

[40] A. Yatchev and W. Härdle. Non-parametric state price density estimation using costrained least squares and the bootstrap. *Journal of Econometrics*, 133(2):579–599, 2006.

[41] A. Ziegler. Why does implied risk aversion smile? *Review of Financial Studies*, 20(3):859–904, 2007.

A Absolute Continuity and the Radon Nikodym Derivative

If ς and φ are two measures on a σ-algebra $\mathcal{B}$ of subsets of X, then φ is absolutely continuous with respect to ς if $\varphi(A) = 0$ for any $A \in \mathcal{B}$ such that $\varsigma(A) = 0$. In symbol: $\varphi \ll \varsigma$.

A stronger generalisation of the above is: if the measure φ is a.s. finite, i.e. $\varphi(X) < \infty$, the property $\varphi \ll \varsigma$ is equivalent to the following: for any $\alpha > 0$ there is a $\gamma > 0$ such that $\varphi(A) < \alpha$ for every A with $\varsigma(A) < \gamma$ (this follows from the Radon-Nykodin theorem, see below, and the absolute continuity of the integral).

The generic definition can be extended to signed measures and to vector-valued measures φ. The vector-valued extension can be generalized to vector-valued ς's: in such a case the absolute continuity of φ with respect to ς amounts to the requirement that $\varphi(A) = 0$ for any $A \in \mathcal{B}$ such that $|\varsigma|(A) = 0$, where $|\varsigma|$ is the total variation of ς.

Assuming that ς is σ-finite, the Radon-Nikodym theorem characterizes the absolute continuity of φ with respect to ς with the existence of a nonnegative Borel function $f \in L^1(\varphi)$ such that $\varphi = f\varsigma$, i.e.:

$$\varphi(A) = \int_A f d\varsigma \qquad \text{for every} A \in \mathcal{B}$$

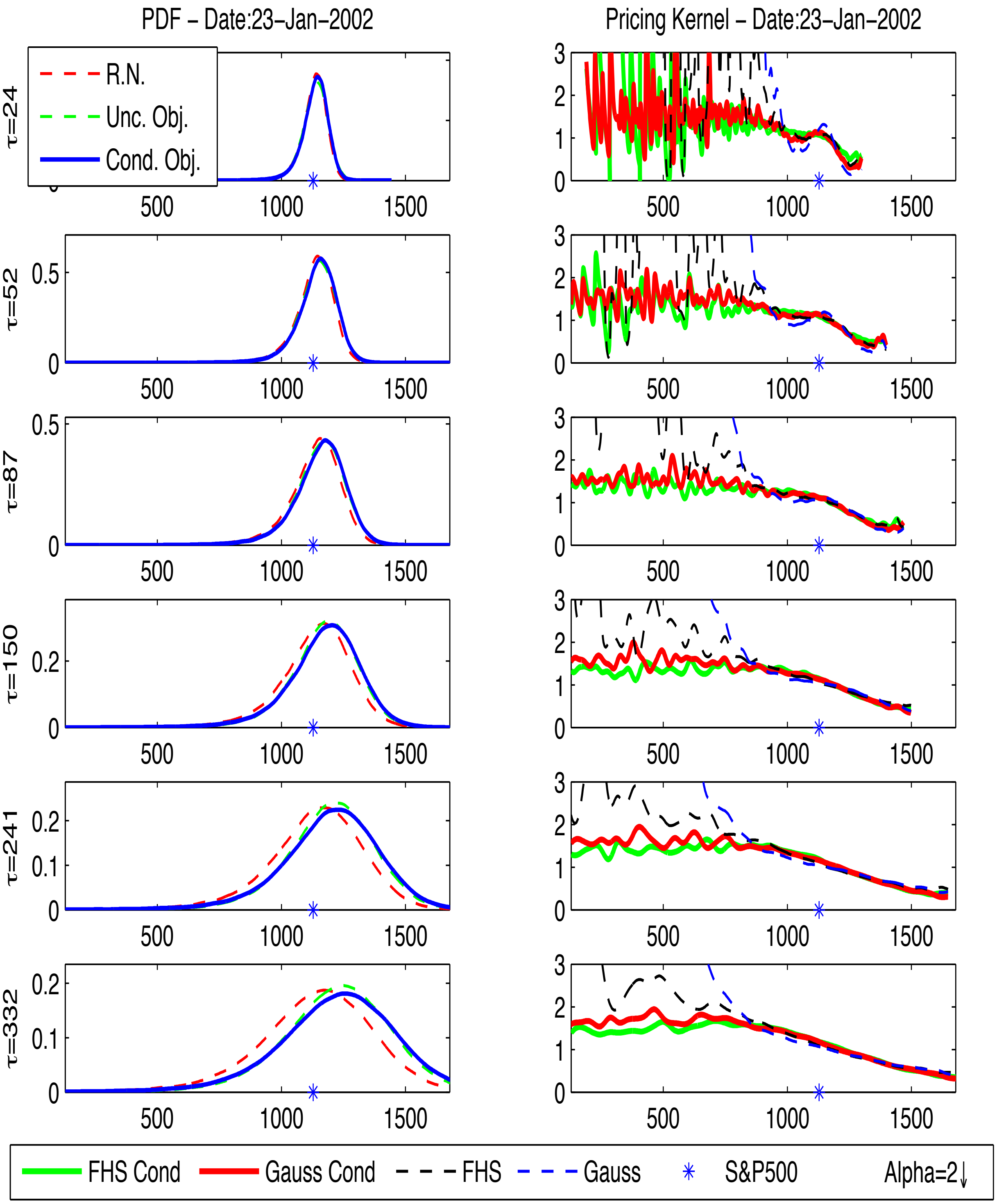
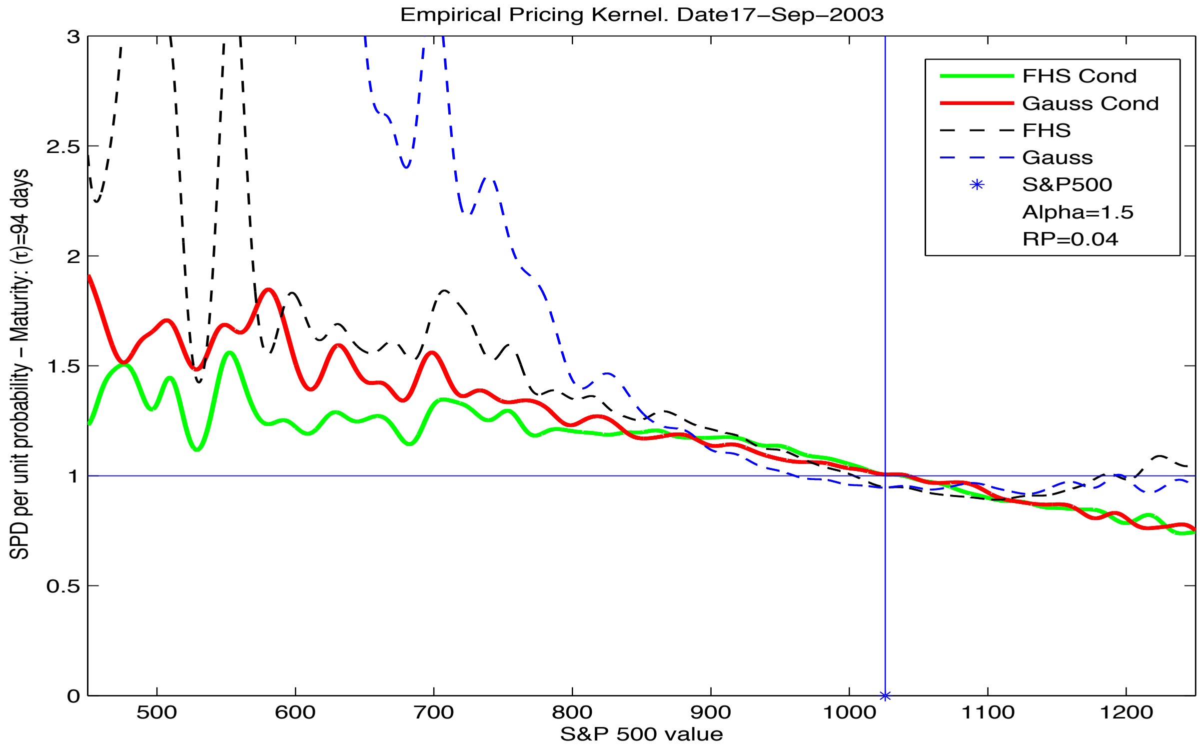


Figure 1: Single day ($t = 4$) all range of times-to-maturity probability density functions (PDFs) (left column). Partially ($M_{t-\Delta,t+\tau}$) and fully ($M^{\dagger}_{t,t+\tau}$) conditional empirical pricing kernels EPKs (right column) estimated with $\alpha^* = 2$ and decreasing, 50.000 simulations and risk premium (R.P.)= 8%.



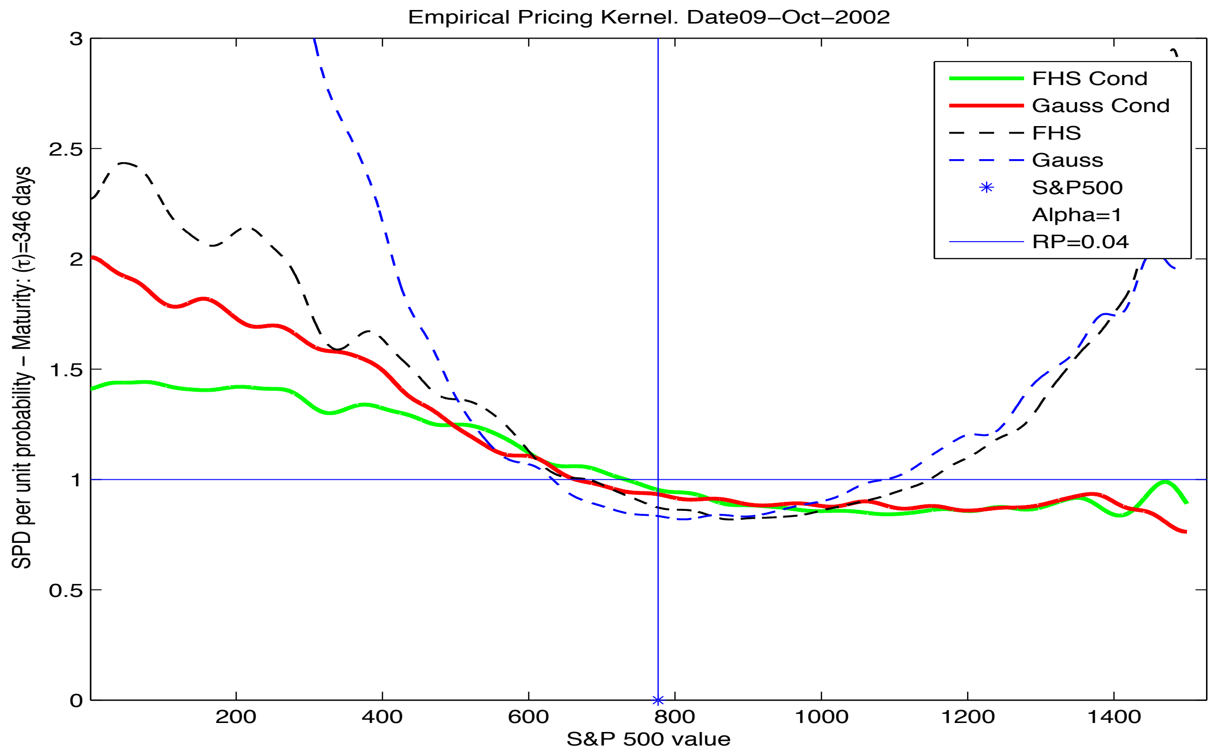


Figure 2: Top: Single day ($t = 90$) single time-to-maturity ($\tau = 94$) conditional ($M^{\dagger}_{t,t+\tau}$) and partially conditional ($M_{t-\Delta,t+\tau}$) EPKs with $\alpha^* = 1.5$, 50.000 simulations and R.P.=4%.
Bottom: Single day ($t = 41$) and single time-to-maturity (τ= 346) Conditional EPKs ($M^{\dagger}_{t,t+\tau}$) and Partially Conditional EPKs ($M_{t-\Delta,t+\tau}$) with $\alpha^* = 1$, 50.000 simulations and R.P.=4%