Puzzle solver

Christian, Mueller-Kademann

Zurich University of Applied Sciences

15 October 2009

Puzzle solver

Christian Müller*

Zurich University of Applied Sciences
School of Management and Law
CH-8401 Winterthur, Switzerland
Tel.: +41 58 934 68 87
Fax: +41 58 935 71 10
Email: much@zhaw.ch

First version: August 2005

This version: October 15, 2009

Abstract

This paper presents a model for asset markets with a subjectively rational solution for the price of the traded asset. Traders cannot act objectively rational and an increase in the number of traders does not enlarge the information set neccessary for determining the "true" price. Consequentely, many well-known "puzzles" vanish as there is no objective truth to which data could live up. An empirical test is conducted which demonstrates the relevance of the argument across time, space, and markets.

JEL classification: F31, F47, C53

Keywords: rational expectations, uncertainty, Tobin tax, financial crisis

*I do thank Mike Müller, Stefan Issler, Daniel Spescha, conference participants of the 10th ICMAF conference 2006 in Rethymno, the 4th INFINITY conference Dublin 2006, the 5th Halle macroeconometric workshop 2006, the Sydney QMF 2007, the annual EEA conference Milan 2008, the 40th money, macro and finance conference London 2008, and the CEA annual meeting 2009 for many helpful comments. All mistakes are mine.

1 Introduction

Recurring economic crises painfully demonstrate that researchers have not yet found satisfactory answers to fundamental questions like what determines prices and quantities on markets. This is especially true for financial markets such as foreign exchange markets or asset markets in general. In fact, hardly any other branch in the literature knows so many puzzles. These puzzles generally arise because observations are far away from what the economic community would accept as reasonable considering the theoretical explanations.

The lack of sound economic explanations of, for example stock or foreign exchange prices, even has the intriguing effect of economists loosing ground in the very centre of economic activity that is at the trading floors. A look at the literature of the past decades might likewise be helpful. A great many articles, especially in the fields of empirical macro finance prove but one thing: the dire retreat of mainstream economics from being able to explain prices at asset markets, let alone quantities.

In this paper I argue that a basic economic concept which constitutes the nucleus of most modern approaches is at the heart of the problem. This nucleus is the concept of rational expectations.

During the past couple of decades the economics literature has been significantly shaped by the notion of rationality. Hardly any peer-reviewed research output will make it to the press unless it lives up to the standards of rationality such as providing a rational solution to whatever problem is posed (see Conlisk, 1996). While proving quite useful in theoretical discussions empirical evidence for rationality is rather rare. Quite to the contrary, especially for those markets where stakes are at the highest and where rationality should yield the largest pay-offs, economists struggle to provide convincing evidence for rational expectations. For example, foreign exchange models are famous for their various failures better known as puzzles such as the difficulties to predict spot rates by forward rates (Wang and Jones, 2003; Salvatore, 2005) and the hassles in beating the naive random walk hypothesis in forecasting spot rates (Taylor, 1995; Obstfeld and Rogoff, 2000; Cheung, Chinn and Pascual, 2005). Likewise, certain "volatility" puzzles relate to the inexplicable behaviour of the second moments of the exchange rate. Very similar problems with economic models arise when stock prices are under consideration (Shleifer and Summers, 1990). Keywords such as irrational exuberance, irrational bubbles, noise trading and so forth all describe but one thing, the impossibility to match theoretical models with data.

In this paper I argue that the common concept of rationality is part of the problem. In the various economic contexts rationality has two components. The first is the ability of individuals to process all available and relevant information. This can be considered as the positive aspect of rationality. However, without the second, the first component appears rather hollow. This second part constitutes the benchmark to which the information processing is compared to. This benchmark usually is an objective functional relation between some variables (Nerlove, 1983; Pesaran, 1987, p.11) which according to the first component individuals know, or are able to discover, or can learn about, or behave accordingly due to some potentially unknown mechanism. Consequently, instead of modelling the behaviour of all subjects it suffices to characterise the behaviour of a representative agent or (heterogeneous) groups of agents. Objectivity hence means the independence of the functional relation from the individuals' (subjective) minds, actions etc. as well as the existence of the representative agent, or groups of agents, as well as the existence of the functional relation.

So far the bulk of criticism on the rationality paradigm and hence the suggestions for overcoming its empirical problems has rested with the first component. In his impressive survey of the literature on bounded rationality Conlisk (1996) already lists four major reasons why individuals can hardly be expected to shoulder the task of really fully exploiting all available information. He quotes not only the many economic papers that demonstrate the failures of the rationality hypothesis but also discusses some contributions to the psychology literature. Similarly, Tirole (2002) puts forth four reasons as to why we may observe deviations from rational behaviour. More recent contributions extend this list towards learning and sentiments (see e.g. Grauwe and Kaltwasser, 2007; Bacchetta and van Wincoop, 2005; Sims, 2005).

In this paper I will focus instead on the second component which is so far dominated by the concept of objectivity. Interestingly, the objectivity presumption seems so "natural" that it is hardly scrutinised. However, Nerlove (1983) and Pesaran (1987) have already pointed out the difficulties of applying the (traditional) rationality paradigm empirically when objectivity is not assumed. I will expand on that criticism a little further by first sketching a market solution that establishes equilibrium without requiring this solution to be objective. The second part of the paper shows some data and estimates that support the model.

2 The subjective asset market pricing model

The model will show the possibility of a world in which the stochastic future states of the world are not objective. It is therefore important to understand that this modelling attempt is just that: a means of demonstrating that many features of asset markets can be captured in a model that does not assume objective probability distributions for events in the future. As will be argued below, in some aspects the model even captures more characteristics compared to simple standard alternatives.

Let us assume an asset market that is characterised by infinite liquidity from an individual's point of view. Infinite liquidity might be justified by noting that a single investor is always small compared to the total or by supposing that credit markets work perfectly in the sense that a convincing investment idea will always meet sufficient means of financing.

Second, let us further assume a very large asset market such that there are always enough assets available for selling or buying. For example, foreign exchange and stocks of large multinational companies would fall into this category. As we are only considering professional trading, each investor may switch her role between buyer and seller of the asset at any point in time.

2.1 The median investment

2.1.1 The investor's problem

The investor at the asset market is assumed to act as an inter–temporal arbitrageuse. She buys (sells) if she thinks the asset price in the future to be higher (lower) than today's: $s_t < s_{t+1}$ ($s_t > s_{t+1}$), where s denotes the price and t time. Accordingly, the sum invested, x_t is either positive or negative.

Thus, the investor's sole objective is to make profit which probably characterises today's financial markets pretty well. Putting this approach in context one might remember the seven reasons for trading foreign exchange given by Friedman (1953). Only one of them (the seventh) was speculation. All other motives Friedman considered are related to some "real" economic activity such as raising the means for cross-border goods trade.

Friedman goes on explaining the price mechanisms for all reasons except the last one. Let us therefore look at number seven.

Unfortunately, at time t, only s_t is known while s_{t+1} is not. The investor can, however, attach to each possible future price a probability giving rise to a probability distribution function $M_t(s_{t+1} \mid \mathcal{I}_t)$ with $\mathcal{I}_t$ representing all the information available at t. M is a distribution function about all possible values for s_{t+1} and its shape depends on $\mathcal{I}_t$. Of course, $M_t(s_{t+1} \mid \mathcal{I}_t)$ is again an element of $\mathcal{I}_t$ and hence, $M_t(s_{t+1} \mid \mathcal{I}_t)$ is in general not identified. This phenomenon is known as the "infinite regress of expectations" problem (Pesaran, 1987; Conlisk, 1996).

Let me suggest a pragmatic approach to cope with this problem. Instead of providing a solution let us assume that the recursive expectation formation process converges, or it stops for other reasons such as information processing capacity constraints. It may be interesting to note that the latter reason can be seen as an analogy to both rational learning and imperfect knowledge, two recent suggestions for getting a grip on the forward premium puzzles in the foreign exchange market, for example. In particular, if it is assumed that the individual trades on the basis of some non-convergent expectation, then its decisions can be looked at as being based on limited information which may come about through imperfect knowledge or learning.

The investment decision is made about the amount of x_t to invest. Obviously, there are three opportunities for the investor. She can either buy, sell, or do nothing. In order to progress the last option is not considered, it could, however, analytically be included by noting that inaction may incur opportunity costs. There are thus two possibilities left. If she sells the asset and tomorrow's price is higher than today's then she will have made a profit, otherwise she loses money, and vice versa.

It is possible to further extend the model by a feedback from the amount invested to the perception of the risk involved. Such an extension would not alter the main findings while explaining individual investment plans. For the sake of brevity I continue on a more conventional road side-lining the feedback issue for the time being. I thus use the standard case which has it that given expected prices and the existing portfolio the investor plans to exchange a given amount of asset against a (minimum) price depending on her expectation about future prices. We are now ready to establish market equilibrium by a standard order book mechanism.

2.1.2 The investment rule

The individual rule is to invest as much and as long in the market as there is a difference between a predetermined quantile of $M_t(s_{t+1})$, denoted $m_{\mathcal{I}_t}(s_{t+1})$, and the spot price. Thus, the general investment rule can be given as

1. Asset supply: $x_t^+ = x_t$ if $s_t > m_{\mathcal{I}_t}(s_{t+1})$.
2. Asset demand: $x_t^- = -x_t$ if $s_t \leq m_{\mathcal{I}_t}(s_{t+1})$.

2.2 The median investor

Assume a finite number $J \geq 1$ of *distinct* investors $j = 1, 2, \ldots, J$, who individually form beliefs about the future asset price. They offer and demand the asset according to the aforementioned rule. Demand and supply coincide under the following conditions.

DEFINITION 1 (Market clearing and equilibrium price)**.** *The market clears if*

$$\sum_{i=1}^{J} x_{t,i} = 0.$$

The market is in equilibrium if

$$s_t \leq s_t^+ = min\{m_{\mathcal{I}_t}^{(1)}(s_{t+1}), m_{\mathcal{I}_t}^{(2)}(s_{t+1}), \ldots, m_{\mathcal{I}_t}^{(i)}(s_{t+1}), \ldots\} \, \forall i = 1, 2, \ldots, J^+.$$

and

$$s_t \geq s_t^- = max\{m_{\mathcal{I}_t}^{(1)}(s_{t+1}), m_{\mathcal{I}_t}^{(2)}(s_{t+1}), \ldots, m_{\mathcal{I}_t}^{(J)}(s_{t+1}), \ldots\} \, \forall j = 1, 2, \ldots, J^-.$$

where J^+ *and* J^- *count the suppliers and the sellers of the asset respectively. Any price* $s_t, s_t^- \leq s_t \leq s_t^+$ *is an equilibrium price.*

Ordering all sets $\{m_{\mathcal{I}_t}^{(i)}(s_{t+1}), x_{t,i}\}$, $i = 1, 2, 3 \ldots, J$ from smallest to largest $m_{\mathcal{I}_t}^{(\cdot)}(s_{t+1})$ obtains the market price as $s_t = m_{\mathcal{I}_t}^{(*)}(s_{t+1})$ where $m_{\mathcal{I}_t}^{(*)}(s_{t+1})$ corresponds to the median $x_{t,*}$ of the ordered sequence.

An interesting case for the market solution is $J = 1$. It follows that $x_{t,1} = 0$ and hence no transaction takes place. Nevertheless, the spot price is defined. However, as no transaction takes place, the 'true' spot price cannot be observed, instead the last period's will feature in the statistics.

This situation would arise if *all* agents expect the same future spot price, have the same probability distribution in mind including identical attitudes toward risk and accordingly want to either go short or long. The latter makes sure that the asset is neither supplied nor demanded. It is therefore a matter of taste to call this situation a break down of the market or not. As an example consider international investment banks which trade collateralised debt obligations. Since these papers are in general not traded at exchanges quantitative methods are employed to price them. The more investors rely on similar or even identical pricing models the lower J will be. Therefore, the spread of (similar) quantitative pricing models may have contributed to the 2007 / 2008 financial crises, when the securities' market effectively collapsed.[1]

Borrowing from the public choice literature (Downs, 1957) I suggest to call the market price the median investor result. This is because the spot price is defined by the price that complies with part one of the market clearing condition and hence by the investor who offers or demands the pivotal investment.

2.2.1 A generalisation

So far I have considered some s_{t+1} as if the future spot price was the only concern of the investor. However, one could look at some $s^{\star}_{t+1} := g(s_{t+1})$. The analysis would nevertheless be applicable as long as the investor's (net) gain is given by $s^{\star}_{t+1} - s_t$ and hence an (adjusted) investment rule can be defined.

2.3 Objectivity and subjectivity

Under the standard rational expectation hypothesis (REH), there exists an 'objective' probability distribution that all agents either know (simple version of REH) or are able to discover (rational learning, see Pesaran, 1987). This concept implies that (rational) investors use the *same* functional form for expectation formation, *i.e.* the same probability distribution, the same set of exogenous or state variables and the same set of information, possibly up to an unpredictable individual error. The existence of such an 'objective' probability distribution is crucial and difficult to maintain, or as Branch (2004) has it:

[1] These markets finally died when it became apparent that all models generated too high prices. As a result only one model ($J = 1$) survived. This model made all owners of collateralised debt obligations trying to sell.

> ...even econometricians must approximate the true structure of the economy. Given the inability of econometricians to estimate the economic model perfectly, it is unrealistic to expect agents to have such ability.
>
> Branch (2004, p. 593)

Nevertheless, economic theory maintains the existence of an objective probability distribution as the workhorse for both theoretical and empirical analysis. The results of those exercises are the well-known "puzzles". This becomes evident from terms like "irrational" investors as compared to "rational" ones, "fundamentalists" versus "chartists" versus "perfect markets", or "excess volatility" all of which imply that there exists a higher order truth economists possess. Consequently, anybody who behaves differently from whatever is considered theoretically sound will be called "irrational" or the like. A typical example provide for example Verma and Verma (2007) who regress trader's sentiments on a list of "fundamentals" to the effect that the unexplained part of this regression is called "irrational". Even Branch (2004) maintains the existence of an objective solution although he admits the existence of fully justified heterogeneity within investors. The median approach instead allows for an infinite variety of opinions and suggests that it is this variety which is responsible for the price process. In fact, the price is nothing but the outcome of an agreement between subjects. Hence subjectivity is a constituent element of the market process, not an annoying stain on an otherwise perfect economic model world.

It may be noteworthy that the existence of an objective probability distribution function implies that any subject who would not act on the basis of the objective probability distribution function may be called irrational. On the other hand, if there was no objective distribution function, rational investors would know that and they would call subjects irrational who assume its existence.

In other words, the standard notion of rationality is inextricably linked to the assumption of an objective probability distribution function. Hence, heterogeneity and thus irrationality of (some) agents does only make sense with respect to objective probability distribution functions. For example Sims (2005) notes it might indeed be rational for agents not to be (fully) rational due to costs of information gathering and processing. However, this kind of irrationality could be dealt with by delegating the investment decision to rational individuals and sharing the implicit gains. If, by contrast, this irrationality was genuine in the sense that losses do not matter, delegation would not be an option. Such investors might be central banks trying to target exchange rates or governments buying banking sector stocks to save the industry.

In that perspective, the median approach can be regarded as a general version of the standard rational expectation hypothesis approach. Under standard REH with homogeneous agents we would find $M_t^{(j)}(s_{t+1}) = M_t^{(i)}(s_{t+1}) = M_t(s_{t+1})$, $\forall j \neq i$ and hence $J = 1$. Therefore, the standard rational expectation solution is a special case of the median approach.

I now look at the possibility to obtain an objective probability distribution for s_{t+1} with $J > 1$. Objectivity is obtained when the subject does not play a role, that is J does not affect the market outcome. Such a situation is commonly characterised as the presence of small, negligible investors, or as atomised markets and so on. In the median approach the first distinction has to be made between a probability distribution conditional on $x_t^j = 0$, $\forall j$ and $x_t^j \neq 0$. In the first case, I suppose that there may exist a probability distribution over all $m_{\mathcal{I}}^j(s_{t+1} \mid x_t^j = 0)$, $\forall j$. In this case we would necessarily find $s_t = s_{t+1}$ and spot price determination would not be an issue.

In the alternative situation, given the median model, all information sets, all individual probability distributions and the market solution, s_t can be calculated. I define

$$\mu^{(J-1)} \quad := \quad \frac{1}{J-1}\sum_{i=1}^{J-1} m_{\mathcal{I}}^{i}.$$

The $\mu^{(J)}$ and $s_t^{(J)}$ are defined accordingly. Being the Jth investor the pre-condition for objectivity would therefore be

$$(1) \qquad \mu^{(J)} = \mu^{(J-1)}.$$

Equation (1) implies that the individual medians converge to a fixed number, such that if the pool of investors was growing, the observed quantiles would converge to a stationary number.

The answer to the question whether there is a J for which condition (1) hold is, however, no. Notice first that the market solution requires $s_t^{(J)} - s_t^{(J-1)} = \gamma \neq 0$.[2] Then write

$$\begin{aligned}
(2) \qquad \mu^{(J)} &= \frac{1}{J-1}\sum_{i=1}^{J-1} s_t^{(J-1)} - \frac{1}{J(J-1)}\sum_{i=1}^{J-1} s_t^{(J-1)} + \frac{1}{J}s_t^{(J-1)} + \frac{1}{J}\sum_{i=1}^{J}\gamma \\
(3) \qquad\qquad &= \mu^{(J-1)} + \frac{1}{J}s_t^{(J-1)} - \frac{1}{J}\mu^{(J-1)} + \gamma
\end{aligned}$$

to see that only the two middle terms in (3) disappear for large J, whereas the last remains no matter how large J gets. Therefore, for non-degenerate values of γ a limiting value for $\mu^{(J)}$, $J \to \infty$ does not exist. Hence, s_t is nonstationary in J and an objective distribution probability does not exist. Instead, the distribution always depends on J implying that it is inherently subjective. Nonstationarity w.r.t. time of the first moment and hence every higher moment follows directly.

[2] Strictly speaking, $\gamma \neq 0$ only holds for sure if the new investor's investment exceeds $x_{t,*}$, the median investor's investment. Otherwise, several investors have to enter. The line of argument is neither affected by this special case nor by conditioning γ on J.

Summarises, rationality in the median approach means that the subjective character of the market mechanism should be understood and so should be ones' inability to provide an objective probability distribution for future spot prices. Hence, the hunt for rationality is irrational.

3 Empirical hypotheses and test results

Nonstationarity of the prices with respect to the number of investors really is the key implication of the median model. In the following, I therefore consider the implied relationship between J and the market solution and present empirical evidence.

The test of the median approach builds on (3) which implies that $s_t^{(J)}$ can likewise be written as

$$(4) \quad s_t^{(J)} = \mu^{(1)} + (J-1)\gamma$$

Using standard results it is clear that the variance of $s_t^{(J)}$ increases with J for finite and non-degenerate γ.

In contrast, if it was true that an objective probability distribution for s_{t+k} existed, we should observe a lower variance of the average market price the more investors are active. This is because the larger the number the less important is an individual's inability to perfectly forecast. Hence, the more investors are trading, the more information should be available about the 'true' probabilities. By the law of large numbers the asset price should thus oscillate around its expected (objectively 'true') value with an ever smaller variance as the number of investors grows. A straightforward test would thus investigate the relationship between J and the variance of s_t.

There are two principle ways for doing so. The preferred way certainly is laboratory experiments. As of today no such experiment has been conducted. To my best knowledge starting with the seminal contribution by Smith, Suchanek and Williams (1988) all experiments run so far pose an objective price process in the first place. No wonder therefore, that the underlying process can be discovered and the results seem to support the objectivity view. Unfortunately, I have not been able to run an appropriate experiment. Therefore, I turn to data analysis.

3.1 Test setup

I use the following procedure to test the model. Consider the objective price process in its estimable form with $\tilde{f}(\tilde{\mathcal{I}}_t)$ approximating $m_{\mathcal{I}_t}^{(*)}(s_{t+1})$

$$\begin{aligned} s_t &= f(\mathcal{I}_t) \\ &= \tilde{f}(\tilde{\mathcal{I}}_t) + \varepsilon_t \\ \varepsilon_t &\sim (0, \sigma_t^2), \sigma_t^2 < \infty \end{aligned}$$

and contrast it to its subjective counterpart

$$\begin{aligned} s_{i,t}^* &= f(\mathcal{I}_t) + \varepsilon_{i,t}, i = 1, \ldots, J \\ \frac{1}{J}\sum_{i=1}^{J} \varepsilon_{i,t} &= 0 \\ \bar{\sigma}_{i,t}^2 &\equiv \frac{1}{J}\sum_{i=1}^{J} \varepsilon_{i,t}^2 \end{aligned}$$

where the subscript i indicates the individual investor. Obviously, if we would average over all investors, the mean of $s_{i,t}^*$ should approach s_t, the objective solution.[3] Basically

[3] For the subjective model we cannot provide a probability distribution function but require that the residual average is always zero. This requirement invokes a restriction on $f()$, but this restriction is very mild.

all macro finance models operate under this assumption and try to identify the expected value of s_{t+1} conditioning on the information set $\mathcal{I}_t$ and the behavioural model, $f(\cdot)$. The well-known trouble though is that according estimation

> ... can be carried out only conditional on the behavioural model .... This means that conclusions concerning the expectations process will not be invariant to the choice of the underlying behavioural model. Pesaran (1987, p.22)

The reason for this hassle is threefold. Firstly, ergodicity of the process during the observation period is required, second, knowledge about the conditioning variables (the set $\mathcal{I}_t$) is needed, and thirdly, the form of f must be known. An alternative is to not condition on $\mathcal{I}$ and $f(\cdot)$ but on t. The remaining condition would be ergodicity, but this time with respect to the individuals rather than time. Notice that the variance of s_t always has two sources. The first is time, that is the variation in $\mathcal{I}$ and possibly in $f(\cdot)$ while the second is $\varepsilon_{i,t}$. As outlined before, this latter source is commonly assumed away by the notion of many and individually negligible investors. Moreover, by the very estimation approaches such as using monthly, weekly, quarterly or even annual data, the underlying variation in the number of investors is effectively ignored. The burden of explaining s_t always rests with $\mathcal{I}_t$ (and $f(\cdot)$). Even worse, in macroeconomic analyses the potential information of the disaggregated data is often thrown away by defining equally spaced points in time. The loss in information arises because after aggregation there is seemingly no variation in the amount of information entering s_t between two points in time.

In contrast to the traditional macroeconomic estimation approach I am going to condition on t. That is, I will observe $s^*_{i,t}, i = 1, \ldots, N_t$, or rather some approximation, and check whether or not its mean estimates collapses to s_t the larger N_t which would support the well–established estimation approaches. In other words, the number of trades is used as a proxy for the number of investors. In order to mark the difference between the number of (independent) investors and its approximation, the latter will be referred to as N_t. More formally, observe the sample mean

$$\begin{aligned}
s_t^* &= \frac{1}{N_t}\sum_{i=1}^{N_t} s_{i,t}^* \\
&= \frac{1}{N_t}\sum_{i=1}^{N_t} f_t(\mathcal{I}_t) + \frac{1}{N_t}\sum_{i=1}^{N_t} \varepsilon_{i,t} \\
&= f_t(\mathcal{I}_t) + \frac{1}{N_t}\sum_{i=1}^{N_t} \varepsilon_{i,t} \\
&= s_t + \frac{1}{N_t}\sum_{i=1}^{N_t} \varepsilon_{i,t}
\end{aligned}$$

and since $\frac{1}{N}\sum_{i=1}^{N} \varepsilon_{i,t} = 0$ the key question becomes whether or not $s_t^* \rightarrow s_t$ as N_t increases. Notice that for $\sigma_{i,t}^2 < \infty \; \forall i = 1, \ldots, N_t$ and independent investors we obtain for the variance of s_t^*

$$\begin{aligned}
Var(s_t^*) &= s_t^2 + Var\left(\frac{1}{N_t}\sum_{i=1}^{N_t} \varepsilon_{i,t}\right) \\
&= s_t^2 + \frac{1}{N_t}\bar{\sigma}_{t,N}^2
\end{aligned}$$

where the last term approaches zero as N_t grows. Thus, the standard approach is perfectly justified under the conditions used.

There exists a simple way to assess whether or not s_t^* approaches s_t in the limit. We only need to know whether or not $\bar{\sigma}_t^2$ does indeed obey to an upper limit. In other words if we could establish that this estimate is independent of N_t we could trust in using s_t in macro finance modelling. Of course, there is another possibility, namely that the variance degenerates with an increasing number of investors. If so, the standard approach could still be justified, even though the assumption of independent investors could be questioned.

Therefore, we only need to consider these two cases. Let

$$\hat{\sigma}_t^2 \quad = \quad \frac{1}{N_t} \sum_{i=1}^{N_t} \left(s_{i,t}^* - s_t^* \right)^2$$

and notice that the model of section 2 implies that on average $\hat{\sigma}_t^2(N_{t+j}) > \hat{\sigma}_{t+j}^2(N_t)\ \forall N_{t+j} > N_t$ and $j \neq 0$. By contrast, the size of N_t should not be systematically related to $\hat{\sigma}_t^2$ on average in the standard approach. In essence, the empirical test boils down to checking whether or not there is a link between N_t and $\hat{\sigma}_t^2$, or, more precisely, if there is a positive association between these two.

To perform the test I suggest the following strategy. By t I denote a five minutes time interval at the stock market and $s_{i,t}^*$ is approximated by a single tic, such that the average of $s_{i,t}^*$ over $i = 1, \ldots, N_t$ obtains s_t under the null hypothesis, that is the standard approach.[4] I thus operate under the assumption that during this time span the (public) information set does not change and that each tic represents an independent investor. I thereby avoid the impossibility to define the information set and functional form with certainty. Notice, if we would instead assume that both do indeed change considerably, any aggregation of the information sets to lower frequencies would be questionable right from the start. It may be interesting to notice that this approach very closely mimics the results obtained by Barndorff-Nielsen and Shepard (2002) who prove that the variance of stochastic volatility processes can be consistently estimated even though the exact form of drift and volatility are unknown.

Hence, by calculating the variance of s_t *within* each of these five minutes' bins, the only source of variation must arise from the individual investors, yet not from the change in the information set.

[4] For one of the data sets I use ten minutes intervalls.

Interestingly the setup covers random-walk like price processes and covariance stationary price processes. Assume that the observed price data is a discrete time approximation to time continuous processes. In such models information arrives continuously. If true, there is no possibility to condition on the point in time of the analysis since those points would be infinitely small and therefore unobservable. Each integral over the prices would, however, feature a variance that would increase with its support. Since we fix the support of the integral the expected variance of the resulting prices thus generated will again be finite. In other words, the variance does not depend on the number of observations we observe. It does, however, depend on how much time has passed within the five / ten minutes intervall. As we draw randomly from within the intervalls this dependency does not affect the validity of the null hypothesis (see also the simulation results reported in figure 6 on p. 39). Therefore, no matter whether or not the prices are generated by a random walk or a covariance stationary process there should be no systematic link between the variance and the number of observations under the null hypothesis that the standard approach holds.

3.2 The data

In the first exercise I use stock prices of frequently and internationally traded stocks: Nestlé and Credit Suisse. Nestlé, is a Swiss company which is one of the largest enterprises in Europe. Likewise, Credit Suisse is one of the biggest banks on the continent.

The data at hand covers two distinct periods. The first stretches over January and February 2007 (Nestlé only) which can be considered a quiet and 'normal' market period. The second sample starts on January, 1st and ends on 31 July 2009.

These three data sets comprise more than 84'400 (Nestlé, 2007 sample) and more than

Table 1: Data characteristics

per bin	mean	min	max	variance
Nestlé share prices January and February 2007				
Total number of bins: 1650				
no. tics	49.38	.0	188.0	729.07
variance	.09	.0	2.91	0.04
Nestlé share prices January – July 2009				
Total number of bins: 14675				
no. tics	89.47	2.0	773.0	3327.3
variance	.0007	0.0	.059	1.8e-6
Credit Suisse share prices January – July 2009				
Total number of bins: 14675				
no. tics	88.684	2.0	955.0	4268.9
variance	.003	0.0	.280	348.7e-6

Sources: Swiss stock exchange, own calculations.

1.3 million (Nestlé and Credit Suisse each, 2009 sample) observations which are aggregated into 1650 (Nestlé, 2007 sample) and 14675 (Nestlé and Credit Suisse each, 2009 sample) ten and five minutes bins respectively. Table 1 on page 19 summarizes the data characteristics, while figure 1 provides a plot of the 2007 Nestlé data. In the top panel we see a cross plot of the data while the bottom panel presents a non-parametric density estimate of the number of observations, that is the sizes of the bin. These two plots already do suggest that the variance tends to increase with the number of observed trades. Turning to formal methods this impression is corroborated.

3.3 Empirical evidence across time, ...

Denoting the empirical price variance estimate $\hat{\sigma}_t^2$ by ζ_t the following regressions analysis sheds light on the relationship between variance and number of tics. Because the functional form of this relation is unknown I use a seventh order ($i_{max} = 7$) Taylor approximation of

Table 2: Estimation results: Coefficient estimates and residual standard deviation

i_{max}	α_0	α_1	α_2	α_3	α_4	α_5	$\hat{\sigma}_\epsilon$
Nestlé share prices January and February 2007							
1	-0.059	0.003	n.a.	n.a.	n.a.	n.a.	0.178493
	(-6.45)	(19.0)	n.a.	n.a.	n.a.	n.a.	–
4	0.171	-0.01	0.0002	-1.8e-6	5.5e-9	n.a.	0.172287
	(4.29)	(-3.60)	(3.73)	(-3.29)	(3.41)	n.a.	–
Nestlé share prices January – July 2009							
3	0.054	0.006	6.6e-6	2.1e-8	n.a.	n.a.	1.22
	(1.53)	(7.91)	(1.60)	(3.67)	n.a.	n.a.	–
Credit Suisse share prices January – July 2009							
5	0.0	0.0196	0.0003	-1.96e-6	4.5e-9	-2.8e-12	5.16271
	–	(7.75)	(7.39)	(-9.93)	(12.8)	(-14.2)	–

Coefficient estimates and corresponding t-values in parentheses below.

the true functional relationship between ζ_t and N_t to begin with:

$$(5) \quad \zeta_t = \alpha_0 + \alpha_1 N_t + \alpha_2 N_t^2 + \cdots + \alpha_{i_{max}} N_t^{i_{max}} + \epsilon_t$$

$$\epsilon_t \sim i.i.d.(0, \sigma^2)$$

Although the variables exhibit a time subscript the regressions are essentially cross section regressions. There may be occasions on which there are periods of generally higher or particular low ζ_t around but under the null hypothesis (the standard approach) this should not be related to N_t.

Applying standard model reduction technologies such as general-to-specific F-testing and selection criteria (Akaike, Final Prediction Error, Schwarz) I derive a suitable representation of the data. In most cases the optimal order seems to be four. Next, the first derivative with respect to the number of observations within each bin is calculated and evaluated for the data range. The following table collects the optimally fitting models and in one instance (Nestlé share prices in 2007) also a model variant where a simple linear model is estimated ad hoc.

Yet another, nonparametric estimation of the relationship between bin size and variance is reported in the appendix. All methods deliver the same results qualitatively.

In the case where a simple linear model is estimated a standard t-test can be used for evaluating the validity of the subjective model. The null hypothesis maintains the standard case while the alternative corresponds to the subjective asset pricing model.

(6) $H_0 : \alpha_1 = 0 \quad vs. \quad H_1 : \alpha_1 > 0.$

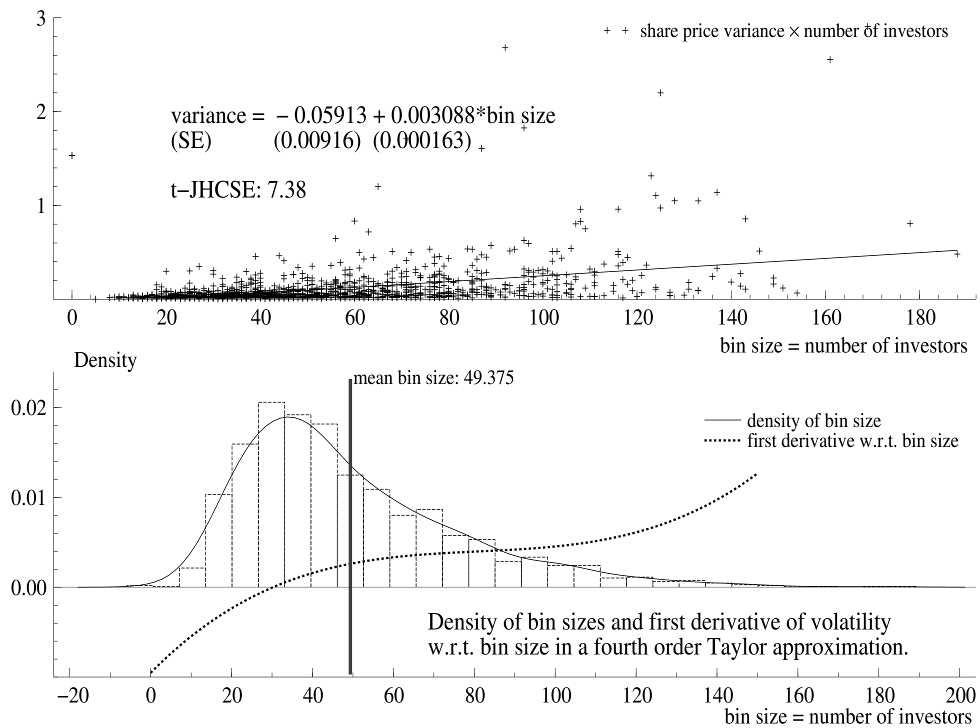


Figure 1: Nestlé 2007: Data plot and graphical estimation results

The estimation results are reported in table 2. They point strongly to a positive relationship between the number of trades and the variance of the price. This is in stark contrast to the usual conviction that more trades would reveal more information about the true price. Instead of increasing the precision with which we measure the price by using more observations it does in fact decrease.

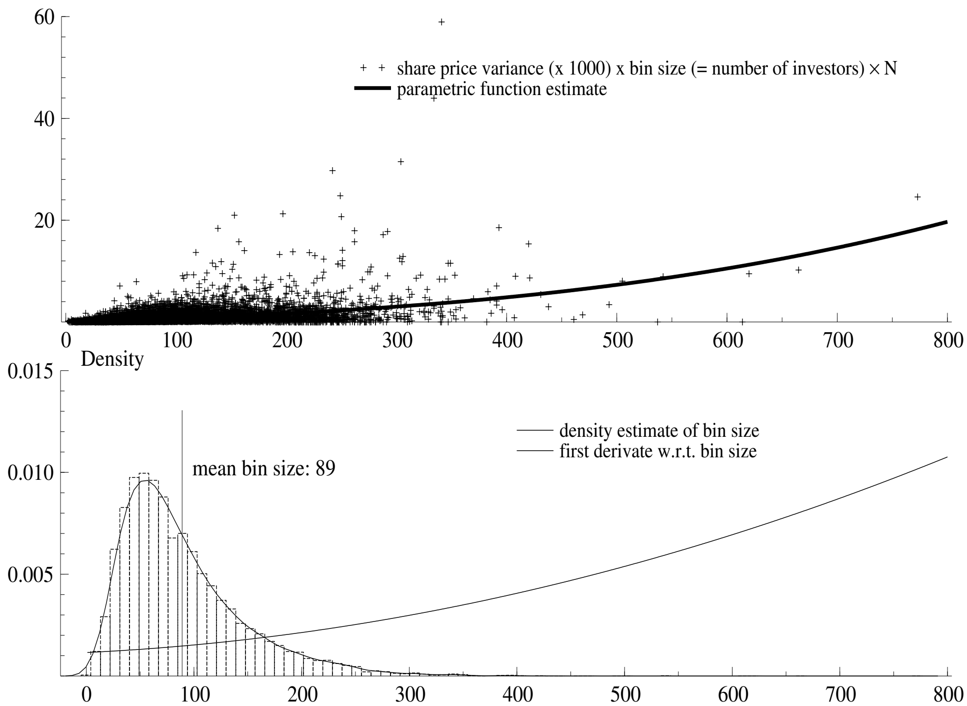


Figure 2: Nestlé 2009: Data plot and graphical estimation results

Because it is not easy to gauge the first derivative with respect to the number of observations from the coefficient estimates, I provide plots of the derivatives. It turns out that the first derivative is positive around the mean. This can be inferred from the lower panel of figures 1 to 3 where the dotted (figure 1), or smooth solid line (figures 1, 2) marks the function of the *first derivative*.[5] All in all there is little doubt that instead of increasing the precision of our price measure the precision decreases when more trades take place for any fixed information set.

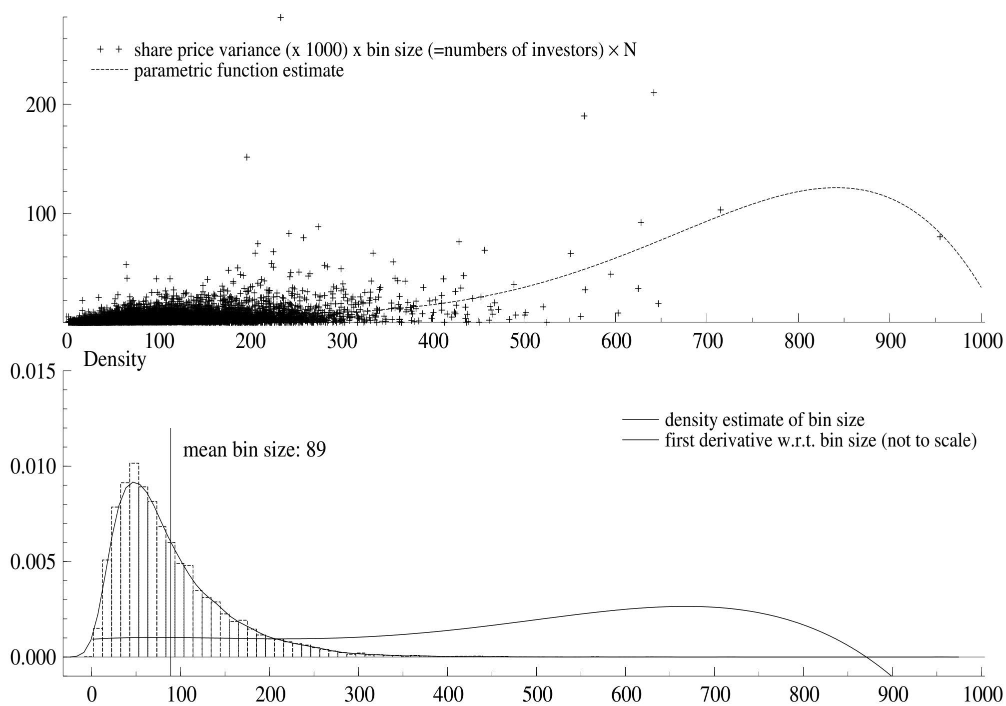


Figure 3: Credit Suisse 2009: Data plot and graphical estimation results

[5] The first derivative is normalised to match the density estimate scale. This adjustment does not affect its position relative to the zero line.

Interestingly, Lyons (2001), and Evans and Lyons (2002) observe similar effects when they report the tremendous increase in the measure of fit of their exchange rate model. The key variable they introduce is order flow data leading to an increase of up to 64% in the measure of regression fit. Moreover, the variables which are in line with economic theory are insignificant on all but one occasion. The result is similar to the present since (cumulated) order flows are under fairly plausible assumptions proportionate to the number of investors. Given the median model no wonder therefore, that Evans and Lyons are able to explain a larger share of the variance.

The empirical results support the view that discouraging market participation by appropriate means would reduce the price volatility. One such method could be the so-called Tobin tax. As yet, it is too early to conclude that a Tobin tax would also be the optimal tool, however.

The evidence presented here, could be challenged on grounds of endogeneity bias. If the number of investors was dependent on the variance of the price process, then the regression coefficients of equation (5) would not be reliable. Therefore, recent papers such as Ané and Ureche-Rangau (2008) investigate the hypothesis that both number of trades (rather: trading volume) and volatility are jointly determined by a latent number of information arrivals. In our context this would imply that the five (ten) minutes time interval was not short enough for keeping the information set constant. In the particular case of Ané and Ureche-Rangau the data is daily price and volume of stocks which certainly justifies modelling information arrivals. However, the general question whether or not trading volume / number of traders is exogenous to the volatility remains.

In support of my regression approach I would like to point to the well-known lunchtime volatility decline. In fact, for every major asset market, be it stock markets, foreign exchange markets, or bond markets intra-day volatility assumes an U-shape (see e.g., Ito, Lyons and Melvin, 1998; Hartmann, Manna and Manzanares, 2001, and the references therein). Thus, following an exogenously determined decline in the number of investors (traders) the volatility decreases justifying the assumption of weak exogeneity of numbers of investors. The same U-shape pattern can be found in my data. For the sake of brevity I do not report the details. They are available on request, however.

In sum, the empirical evidence is more in favour of the model presented in section 2 than in line with the traditional approach.

3.4 ..., space and markets

The previous sections provide evidence for abandoning the standard macro finance approach in favour of an alternative model that maintains individual rational behaviour while emphasising the role of subjective rationality on the macro level.

However, there are at least two possibilities to match the data evidence with the traditional view. One possibility is offered by infinite variance Lévy processes as price generating processes. These processes also feature a higher variance the more data we observe holding the information set constant. As regards the discrimination between the subjective model and Lévy processes there is little one can do except from experiments. Therefore, objective Lévy processes and the subjective asset pricing model probably generate data with very similar basic characteristics.

The second explanation could be that the five / ten minutes time interval is not short enough for actually keeping the information set constant. If so, the increase in the variance as more observations enter the intervall might simply be a reflection of a variation in the information set.

Is this argument sufficient word of comfort for returning to the standard approach? In my opinion it is not. The reason is very simple. While shares like Nestlé's are traded every other second many of those assets which can be considered alternative investments and hence conditioning variables in portfolio models, for example, may be traded far less frequently. As an example consider the Swiss bond market. A safe alternative to the Swiss shares would be Swiss government bonds. It can happen that those bonds are not traded at all within hours. Therefore, the assumption made before finds support that within the five / ten minutes time interval the information set remains constant.

Turning the argument around we would need to carefully synchronise the data of interest and the information set, before we take up the standard approach again. Therefore, an inevitable test of macro finance model would have to look at the high frequency data and make sure that during those time spells where the conditioning variables *do not* change the corresponding number of investors *do not* have explanatory power for the variance of the dependent variable. So far, the standard procedure would be to synchronise observation data by using "suitable" time aggregates such as days, weeks, months, or quarters. I do hazard the guess that the synchronisation exercise, however laborious, would always produce the same result namely nonstationarity with respect to the number of trades.

Luckily, high quality data which permits such synchronisation exercise is becoming more readily available. Very recently Akram, Rime and Sarno (2008) have investigated arbitrage on foreign exchange markets, for example. Their high frequency data set consists of matched spot, forward (forward swap) and deposit interest rate data for the currency pairs British Pound / US Dollar, Euro / US Dollar, Japanese Yen / US Dollar. This data will be used in the follwoing to corroborate the previous findings.

Akram et al.'s (2008) main data source is Reuters which is an advantage for the British

Pound but less so for the Euro and the Yen as Reuters is not the main trading platform in these latter two cases. Moreover, the Japanese Yen is most heavily traded when Reuters does not collect the data. Therefore, we will only look at the British Pound and the Euro pairs.

Even though Akram et al. (2008) collect observations at the highest possible frequency available to them there are occasions on which quotes for the swap, the spot, and the interest rates do not occur simultaneously. Therefore, the variables with the lowest trading activity set the limits. The most important effect on the data sample is a difference in the number of observations despite an exact match of the sample period.

Of course, in order to test the model we need to track the market activity as closely as possible. Whenever there are quotes for, say, the spot rate while there are no changes in the interest rate we lose information. That's why we again restrict our analysis to the largest information sets.

The variable of interest is the arbitrage opportunity defined by the covered interest parity condition given below

$$(7) \quad fx_t = fx_t^e \frac{i_t}{i_t^*} + e_t.$$

Equation (7) has it that the spot exchange rate (denoted fx_t) must equal the forward rate (fx_t^e) up to deposit interest rate (i_t) on domestic assets discounted by the foreign interest rate (i_t^*) of the same maturities as the forward contract. As regards the actual data bid and ask prices are available. Using ask and bid qoutes provides a much more reliable picture of true arbitrage opportunities. Consequently, for each currency pair we obtain two deviation measures.

A nonzero e_t indicates arbitrage opportunities. Akram et al.'s (2008) analysis focusses on the properties of e_t. They show for example that sizeable arbitrage opportunities exist

Table 3: CIP deviation data characteristics Feb – Sep 2004

per bin	mean	min	max	variance
POUND / USD ask 12 months				
Total number of bins: 19727				
no. tics	139.21	1.0	1524.0	7323.25
variance	9.25	0.0	1772.89	519.85
POUND / USD bid 12 months				
Total number of bins: 19727				
no. tics	139.21	1.0	1524.0	7323.25
variance	10.38	0.0	1649.45	531.97
POUND / USD ask 6 months				
Total number of bins: 19711				
no. tics	131.61	1.0	1521.0	7239.00
variance	2.05	0.0	2282.09	358.70
POUND / USD bid 6 months				
Total number of bins: 19711				
no. tics	131.61	1.0	1521.0	7239.00
variance	2.09	0.0	2182.55	337.41
EURO / USD ask 12 months				
Total number of bins: 19735				
no. tics	129.74	1.0	558.0	5955.00
variance	6.20	0.0	9695.46	4858.83
EURO / USD bid 12 months				
Total number of bins: 19735				
no. tics	129.74	1.0	558.0	5955.00
variance	5.14	0.0	9594.91	4728.03
EURO / USD ask 6 months				
Total number of bins: 19713				
no. tics	117.14	1.0	559.0	5781.86
variance	0.91	0.0	33.08	2.67
EURO / USD bid 6 months				
Total number of bins: 19713				
no. tics	117.14	1.0	559.0	5781.86
variance	0.91	0.0	29.80	2.06

Sources: Akram et al. (2008), own calculations.

but these are all very short lived. For the sake of brevity I do not describe the data in detail. All those details are reported in Akram et al. (2008), the data has been downloaded from Dagfinn Rime's website. Rime also kindly provided advise in handling and interpretating the data.

In what follows we will look at derived values for e_t for the two currency pairs Pound / US Dollar, and Euro / US Dollar. For each of these two pairs e_t is calculated for bid and ask spot rates respectively. I investigate forward contracts for twelve and six months because these are the most liquid markets and we therefore most likely obtain a fair picture of the whole market. Taken together, eight data sets are available for analysis.

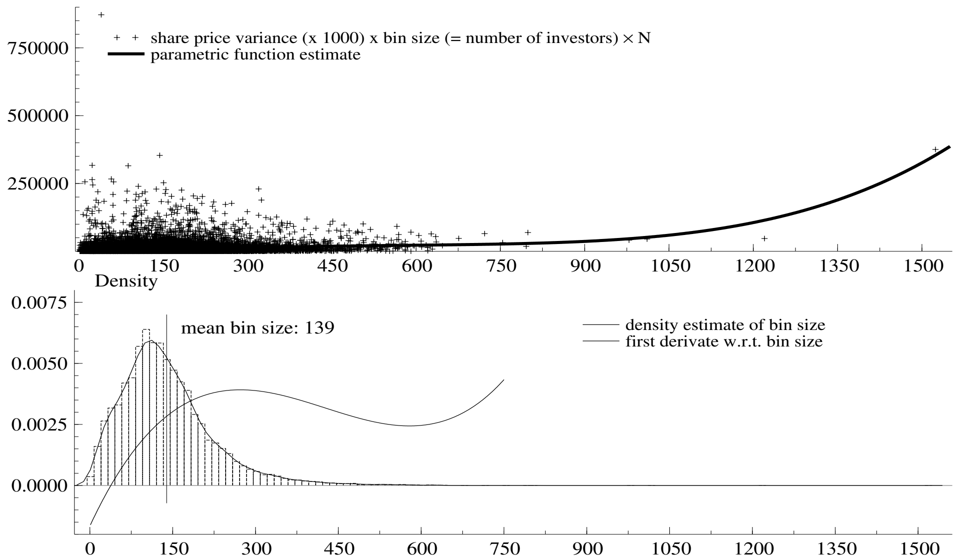
The observation period is February 13 to September 30, 2004, weekdays between 07:00 and 18:00 GMT which provides up to 2.7 million observations per currency pair and quote (bid or ask). This data is again bundled into five minutes bins.

After going through the same steps of analysis as before it turns out that the standard approach can again be rejected in basically all cases. The first derivative of the function describing the relationship between bin size and variance is positive around mean / median, and relying on nonparametric analysis, there is convincing evidence for this derivative to be significantly positive.

Table 4: Estimating CIP deviation variance: Coefficient estimates and residual standard deviation

i_{max}	α_0	α_1	α_2	α_3	α_4	$\hat{\sigma}_\epsilon$
POUND / USD ask 12 months						
4	7485.59	-15.31	0.2237	-0.0004	2.36e-7	18705.7
	(522.3)	(8.94)	0.0456	7.78e-7	3.558e-8	–
POUND / USD bid 12 months						
1	7095.85	20.677	n.a.	n.a.	n.a.	16786.22
	(380.6)	(2.51)	n.a.	n.a.	n.a.	–
POUND / USD ask 6 months						
1	1478.29	3.562	n.a.	n.a.	n.a.	3466.82
	(74.46)	(0.52)	n.a.	n.a.	n.a.	–
POUND / USD bid 6 months						
3	1008.62	12.875	-0.0455	7.0e-5	n.a.	3718.51
	(196.1)	(3.451)	(-0.017)	(2.41e-5)	n.a.	–
EURO / USD ask 12 months						
2	1682.37	6.86	6.05	n.a.	n.a.	6799.21
	(86.96)	(2.320)	(0.012)	n.a.	n.a.	–
EURO / USD bid 12 months						
3	1667.63	38.67	-0.1445	0.0003	n.a.	7269.83
	(188.8)	(3.95)	0.0233	0.839e-5	n.a.	–
EURO / USD ask 6 months						
4	240.727	14.25	-0.0955	0.0003	-2.19e-7	1590.91
	(45.56)	(1.476)	0.0144	5.11e-5	5.765e-8	–
EURO / USD bid 6 months						
4	229.241	13.285	-0.086	0.0002	-2.03e-7	1404.12
	(40.21)	(1.303)	0.0127	4.51e-5	5.088e-8	–

Regressions of variance on number of traders (tics). Coefficient estimates and corresponding t-values in parentheses below.



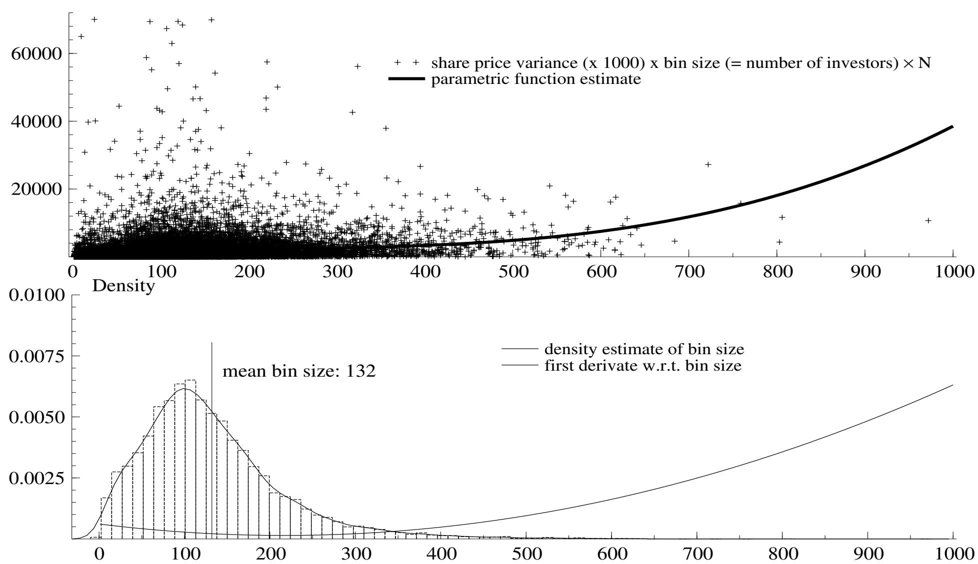


Figure 4: UIP one year ask (top panel) and six months bid (bottom panel) British Pound / US Dollar 2004: Data plot and graphical estimation results

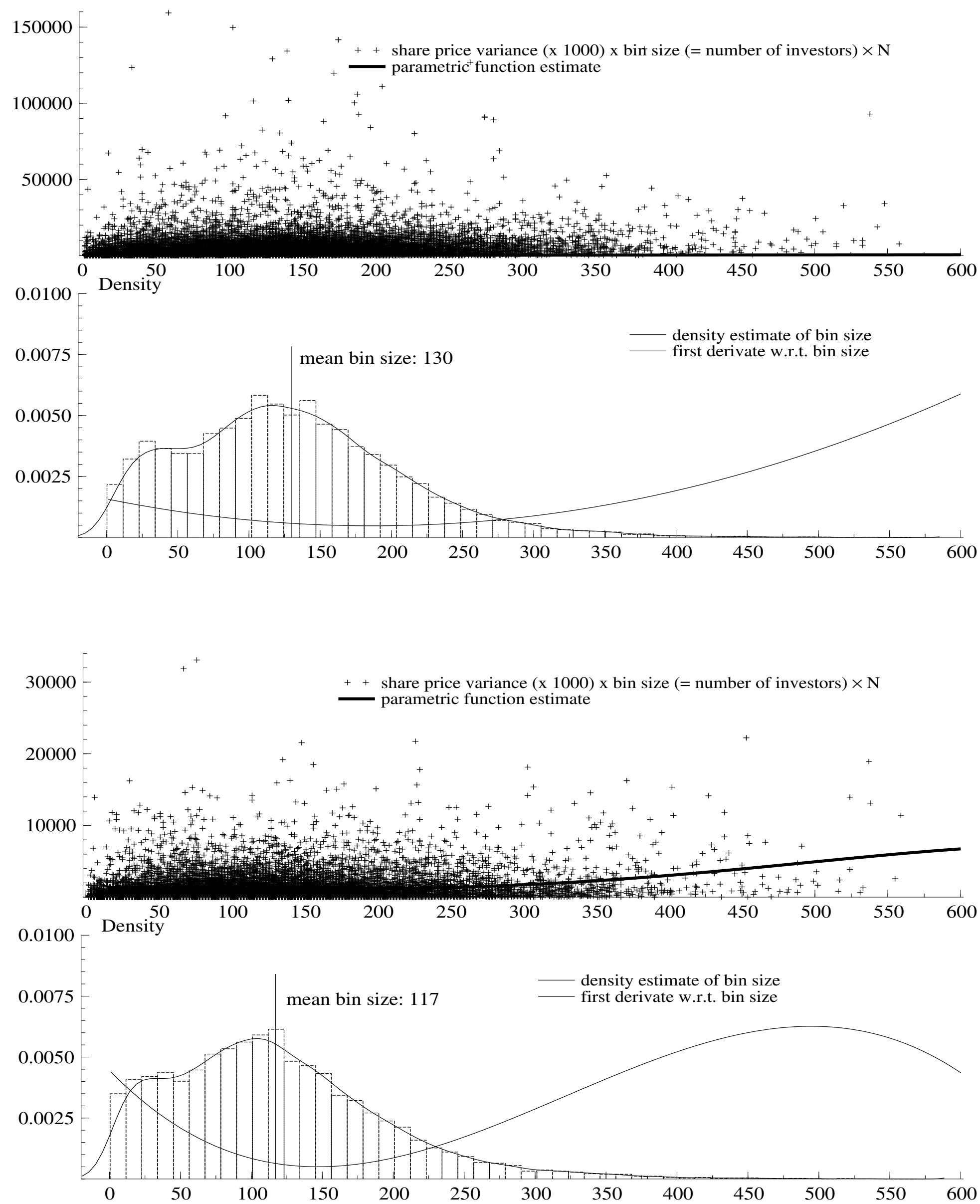


Figure 5: UIP one year bid (top panel) and six months ask (bottom panel) Euro / US Dollar 2004: Data plot and graphical estimation results

3.5 Reconciling evidence from experiments

The empirical investigation showed that the data has properties which one would expect if investors behaved according to the median model. However, the median model poses the absence of an objective price process and this absence is impossible to prove since the non-existing can not be proven to not exist. Therefore, the empirical evidence can be interpreted as an *as if* behaviour. Investors behave *as if* there was no objective price process. This interpretation also holds the key for reconciling experimental evidence of investors' behaviour.

The first stylised fact is the so-called irrational behaviour in artificial asset markets (see inter alia Smith et al., 1988; Cipriani and Guarino, 2005). It has likewise be demonstrated that experienced traders can push the market price towards its fundamental value and hence eradicate irrational prices (see e.g. Dufwenberg, Lindqvist and Moore, 2005; Drehmann, Oechssler and Roider, 2005; Hussam, Porter and Smith, 2008, to name but a few). Notably, all these experiments use a design in which an (implicit) objective price process is induced. For example, the traded asset may yield a return with a given probability each period. Therefore, irrationality in such a situation might be used as an argument against the median model. I prefer a different interpretation, however. The participants in these experiment behave exactly as they would have done in the real world: they trade as if there was no objective price process. By contrast, expert traders are able to discover the induced pricing rule and hence tend to behave rationally. Therefore, these experiments do not lend support to the standard approach. The decisive question is how do experts trade in the absence of an objective price process? Thus, the need for an accordingly set up experiment remains and economists might have a closer look at optimal decision making under the subjective probability approach in general.

Therefore, in the light of more realistic properties of the theoretical model I consider the conclusion of rational "irrationality" of asset markets the more plausible one.

Alternatively, one may regard each tic as a piece of information itself. In the logic of my argument each investor would represent an indispensable piece of information. The standard REH approach would thus have to include each investor in the information set.[6] Then, the standard approach and my model would generate data which would be observationally equivalent.

4 Summary and conclusions

The observation of a price for an asset is no proof for this price to follow a discoverable, objective stochastic process. In this paper I discuss a model for the determination of asset prices which emphasises the role of subjective probability distribution functions for the price process. It generalises models which are based on concepts of objective probability functions. Next to being less restrictive the new model can be regarded as simpler. A test has been suggested with the traditional view as the null hypothesis and the new approach in the alternative. Empirical investigations covering many data sets across time, space, and markets found strong support for rejecting the Null. Allowing for subjectivity in financial markets helps understanding major pricing puzzles. It also lends tentative support to the Tobin tax for reducing asset price volatility. The new model implies an alternative research agenda which focuses on optimal decision making under fundamental uncertainty.

[6] Arguably, with such modification the idea of a representative agent disappears.

References

Akram, Q. F., Rime, D. and Sarno, L. (2008). Arbitrage in the foreign exchange market: Turning on the microscope, *Journal of International Economics* **76**(2): 237–253.

Ané, T. and Ureche-Rangau, L. (2008). Does trading volume really explain stock returns volatility?, *Journal of International Financial Markets, Institutions and Money* **18**(3): 216–235.

Bacchetta, P. and van Wincoop, E. (2005). Rational Inattention: Solution to the Forward Discount Puzzle, *Research Paper 156*, International Center for Financial Asset and Engineering.

Barndorff-Nielsen, O. E. and Shepard, N. (2002). Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Modells, *Journal of the Royal Statistical Society, Series B* **64**: 253 – 280.

Branch, W. A. (2004). The theory of rationally heterogenous expectations: Evidence from survey data, *The Economic Journal* **114**: 592 – 621.

Cheung, Y.-W., Chinn, M. D. and Pascual, A. G. (2005). Empirical exchange rate models of the nineties: Are any fit to survive?, *Journal of International Money and Finance* **19**: 1150 – 1175.

Cipriani, M. and Guarino, A. (2005). Herd Behavior in a Laboratory Financial Market, *American Economic Review* **95**(5): 1427–1443.

Conlisk, J. (1996). Why bounded rationality?, *Journal of Economic Literature* **34**(2): 669 – 700.

Downs, A. (1957). *An Economic Theory of Democracy*, Harper and Row, New York.

Drehmann, M., Oechssler, J. and Roider, A. (2005). Herding and Contrarian Behavior in Financial Markets: An Internet Experiment, *American Economic Review* **95**(5): 1403–1426.

Dufwenberg, M., Lindqvist, T. and Moore, E. (2005). Bubbles and experience: An experiment, *American Economic Review* **95**(5): 1731–1737.

Evans, M. D. D. and Lyons, R. K. (2002). Order flow and exchange rate dynamics, *Journal of Political Economy* **110**(1): 170 – 180.

Friedman, M. (1953). The Case of Flexible Exchange Rated, *in* M. Friedman (ed.), *Essays in Positive Economics*, University of Chicago Press, Chicago, pp. 157–203.

Grauwe, P. D. and Kaltwasser, P. R. (2007). Modeling optimism and pessimism in the foreign exchange market, *CESifo Working Paper Series 1962*, CESifo GmbH.

Hartmann, P., Manna, M. and Manzanares, A. (2001). The microstructure of the euro money market, *Journal of International Money and Finance* **20**(6): 895–948.

Hussam, R. N., Porter, D. and Smith, V. L. (2008). Thar She Blows: Can Bubbles Be Rekindled with Experienced Subjects?, *American Economic Review* **98**(3): 924–937.

Ito, T., Lyons, R. K. and Melvin, M. T. (1998). Is There Private Information in the FX Market? The Tokyo Experiment, *Journal of Finance* **53**(3): 1111–1130.

Lyons, R. K. (2001). Foreign exchange: macro puzzles, micro tools, *Pacific Basin Working Paper Series 01-10*, Federal Reserve Bank of San Francisco.

Nerlove, M. (1983). Expectations, Plans, and Realizations in Theory and Practice, *Econometrica* **51**(2): 1251 – 1280.

Obstfeld, M. and Rogoff, K. (2000). The six major puzzles in international macroeconomics: Is there a common cause?, *Working Paper 7777*, National Bureau of Economic Research.

Pesaran, H. M. (1987). *The Limits to Rational Expectations*, Basil Blackwell, Oxford.

Salvatore, D. (2005). The euro-dollar exchange rate defies prediction, *Journal of Policy Modelling* **27**(2): 455 – 464.

Shleifer, A. and Summers, L. H. (1990). The noise trader approach to finance, *Journal of Economic Perspectives* **4**(2): 19 – 33.

Sims, C. (2005). Rational inattention: A research agenda, *Discussion Paper, Series 1: Economic Studies 34*, Deutsche Bundesbank.

Smith, V. L., Suchanek, G. L. and Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets, *Econometrica* **56**(5): 1119–51.

Taylor, M. P. (1995). The economics of exchange rates, *Journal of Economic Literature* **33**(1): 13 – 47.

Tirole, J. (2002). Rational irrationality: Some economics of self-management, *European Economic Review* **46**(4 – 5): 633 – 655.

Verma, R. and Verma, P. (2007). Noise trading and stock market volatility, *Journal of Multinational Financial Management* **17**(3): 231–243.

Wang, P. and Jones, T. (2003). The Impossibility of Meaningful Efficient Market Parameters in Testing for the Spot–Foreward Relationship in Foreign Exchange Markets, *Economics Letters* **81**: 81 – 87.

A Nonparametric estimation of the bin size – variance relationship

Equation (5) defines a parametric function of the relation between bin size (the approximation of number of traders) and the variance of the asset price within those five / ten minute time bins. The according results lend support to the hypothesis of a positive association between the number of trades and the variance of the asset price. Nevertheless, one may wonder to what extent these results depend on the specific parametric functional forms used. Therefore, I report the outcome of a nonparametric, local quadratic estimation of the relation between bin size and variance.

The estimation is based on the software XploRe which is specifically designed for analysing financial market data by means of non- and semi-parametric functions.[7] In particular, I make use of the procedure "lplocband" of the "smoother" library applying the Epanechnikov kernel. The kernel bandwidth is chosen manually because the automatic procedures always selected the lowest possible bandwidth within the pre-defined range. These lower bands were close to the minimum distance between any two explanatory variable data points. The results do not change qualitatively, however, within a large range of bandwidths.

[7] The software is available free of charge from http://lehre.wiwi.hu-berlin.de/Professuren/quantitativ/ statistik/xplore, the code and the data are available on request from the author.

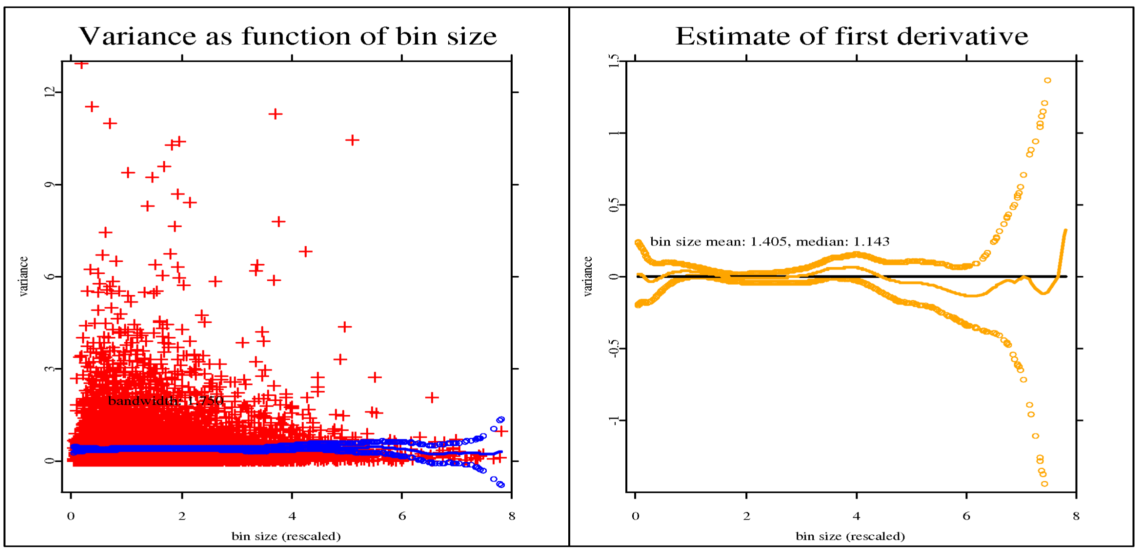


Figure 6: Random walk: Nonparametric estimation of bin size – variance relation and its first derivative

Before turning to the empirical evidence let me reconcile the results which could be expected under the null hypothesis, the standard approach. Figure 6 plots observations that are generated by simulating 14675 random walks of length 955. In the next step, between 2 and 955 data points of these random walks are selected randomly from each of the 14675 data sets. These observations mimic the five minutes bins. Accordingly, the variance of these bins is estimated and set in relation to the number of artificial observations entering the bin. This simulation procedure thus draws on the actual Credit Suisse data and clearly demonstrates that even under the random walk hypothesis for price data the relationship between bin size and variance should be completely stochastic; the first derivative estimate frequently crosses the zero line, and the 95 percent confidence bands safely enclose zero.

By contrast, the empirical relationships do look pretty different. For example figure 7 shows that the estimated first derivative is significantly larger than zero around the mean bin size in the case of the 2007 Nestlé data. Very similar pictures emerge for the other data sets.

In some instances (see figure 11 on p. 46), there are also hints for another phenomenon. In these instances the relationship between variance and bin size seems to be negative. This situation occurs when trading volume is low (small bin sizes) and gives rise to the possibility of dependent observations. For example, when trading activity is low, several consecutive trades may be exercised by the same trader(s).

In order to render the estimation feasible, i.e. avoiding numerical problems, the independent variable was divided by twice the maximum value of the bin size. If that was not sufficient to overcome numerical problems both variables were normalised by their respective empirical standard deviations. This linear transformation cannot affect the relation

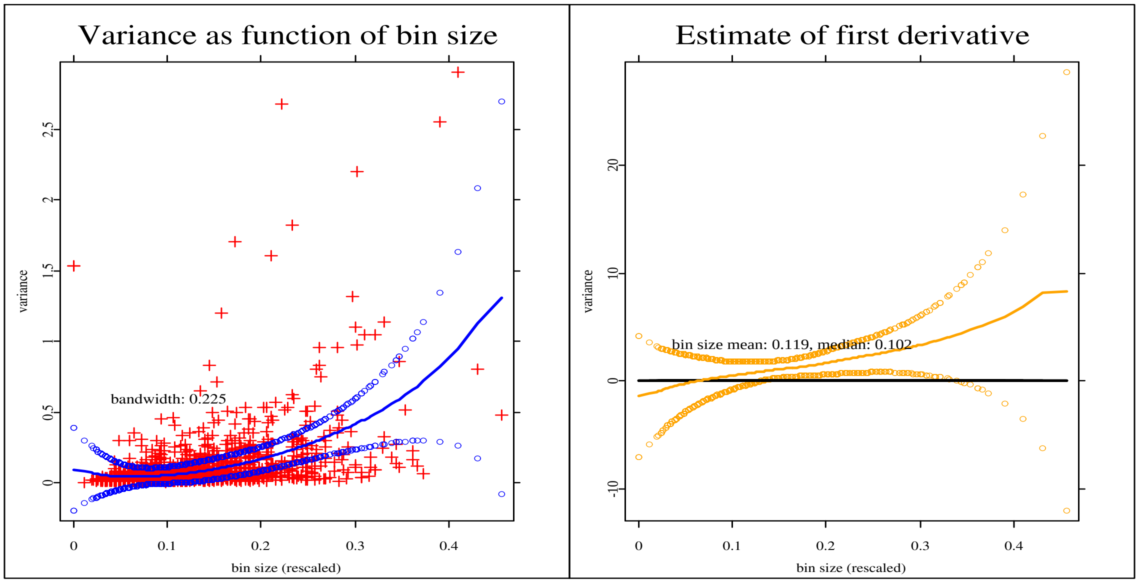
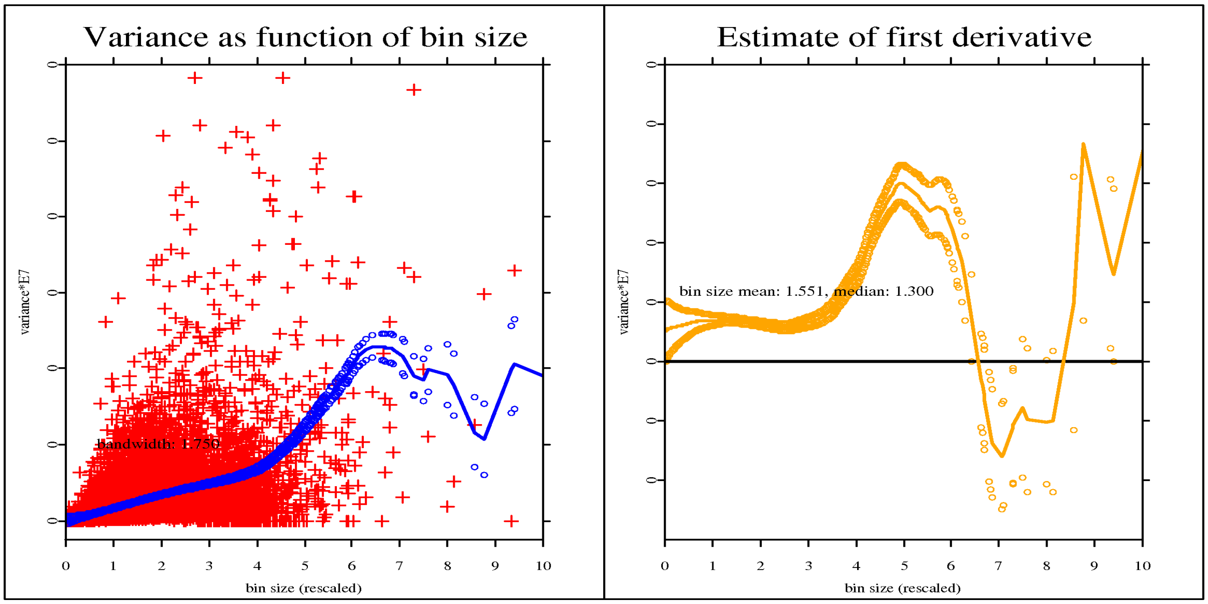


Figure 7: Nestlé tic data 2007: Nonparametric estimation of bin size – variance relation and its first derivative

between the independent and the dependent variable.



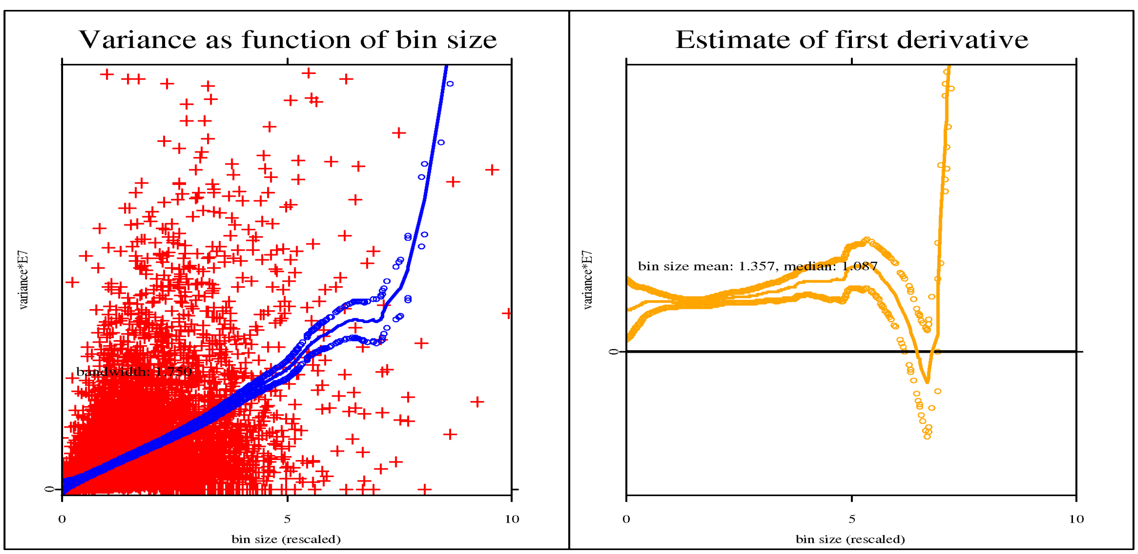


Figure 8: Nestlé tic data 2009 (top panel) and Credit Suisse tic data 2009 (bottom panel): Nonparametric estimation of bin size – variance relation and its first derivative

Pretty much in line with the parametric estimation the variance increases with the bin size. The panel on the left shows an upward trend in the variance for growing bin sizes and the panel to the right confirms that the first derivative of the relationship is significantly larger than zero around the mean bin size and for sizes larger than the mean. Therefore, the hypothesis derived from the median model receives support once more.

B Evidence from foreign exchange markets

The following graphs depict the results for the data compiled by Akram et al. (2008). Here, the data is always standardized such that the empirical variance of dependent and independent variable is one. As before, this linear transformation cannot affect their relationship.

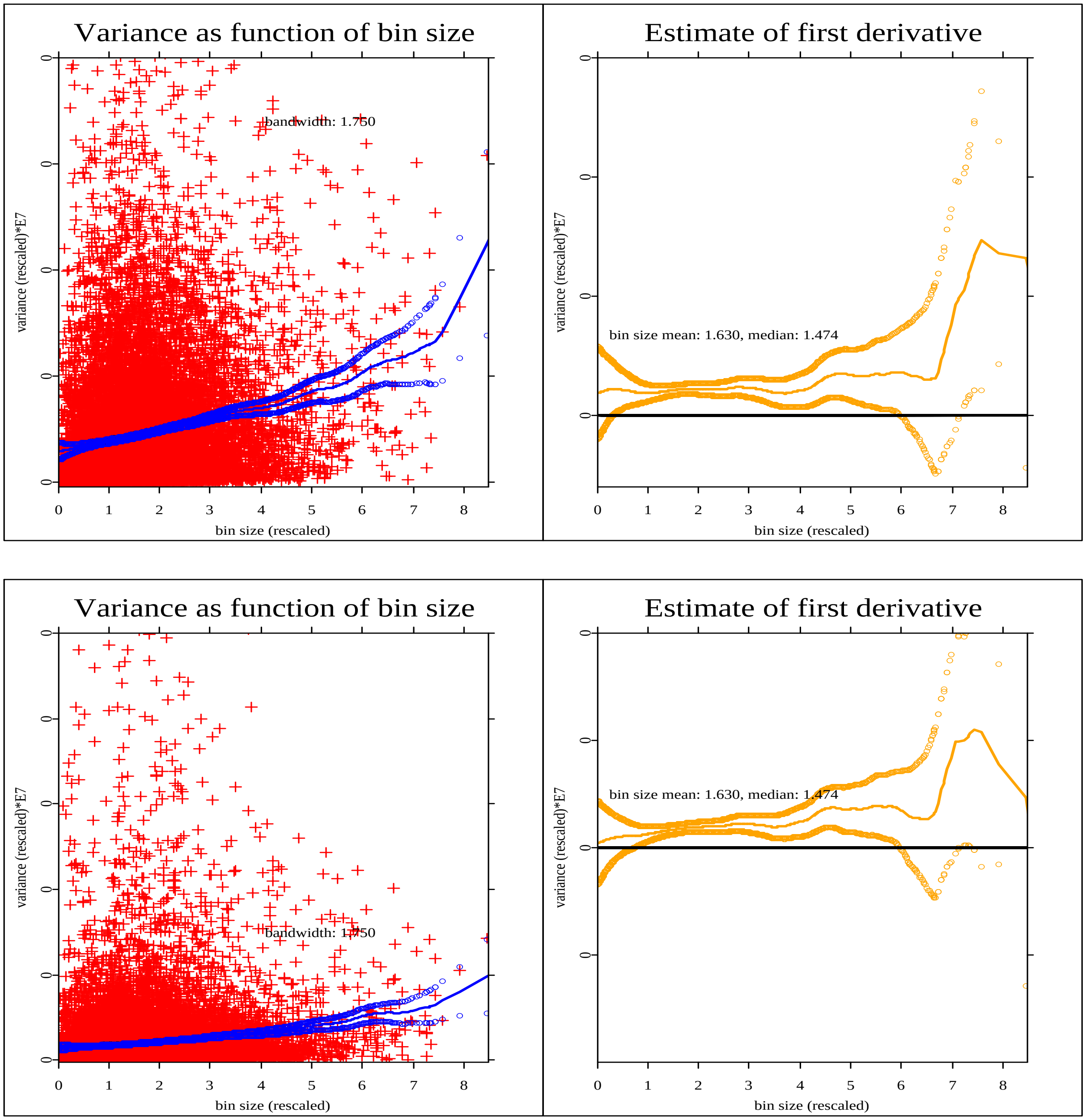


Figure 9: Pound one year ask (top panel) and bid (bottom panel): Nonparametric estimation of bin size – variance relation and its first derivative

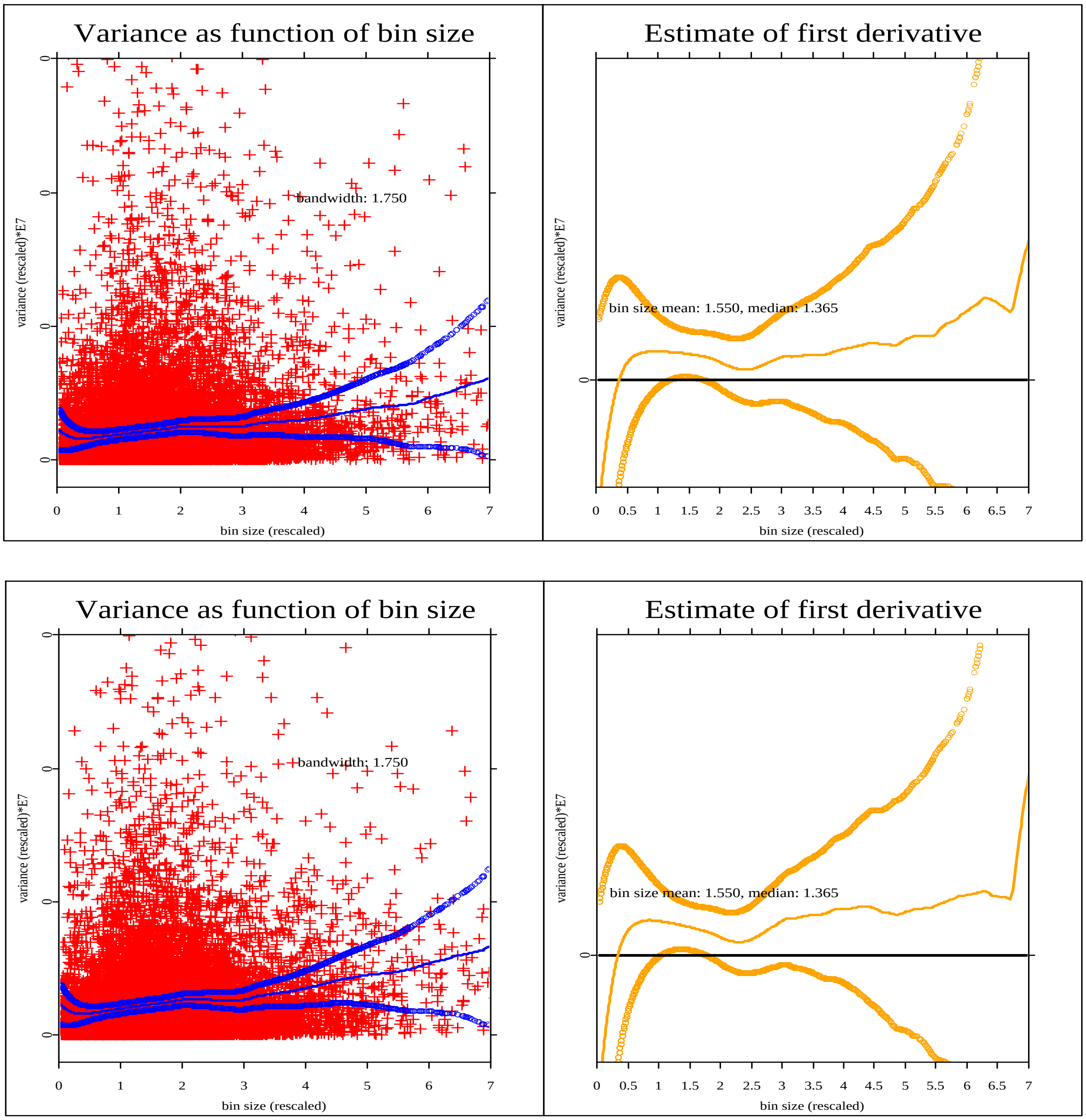


Figure 10: Pound 6 months ask (top panel) and bid (bottom panel): Nonparametric estimation of bin size – variance relation and its first derivative

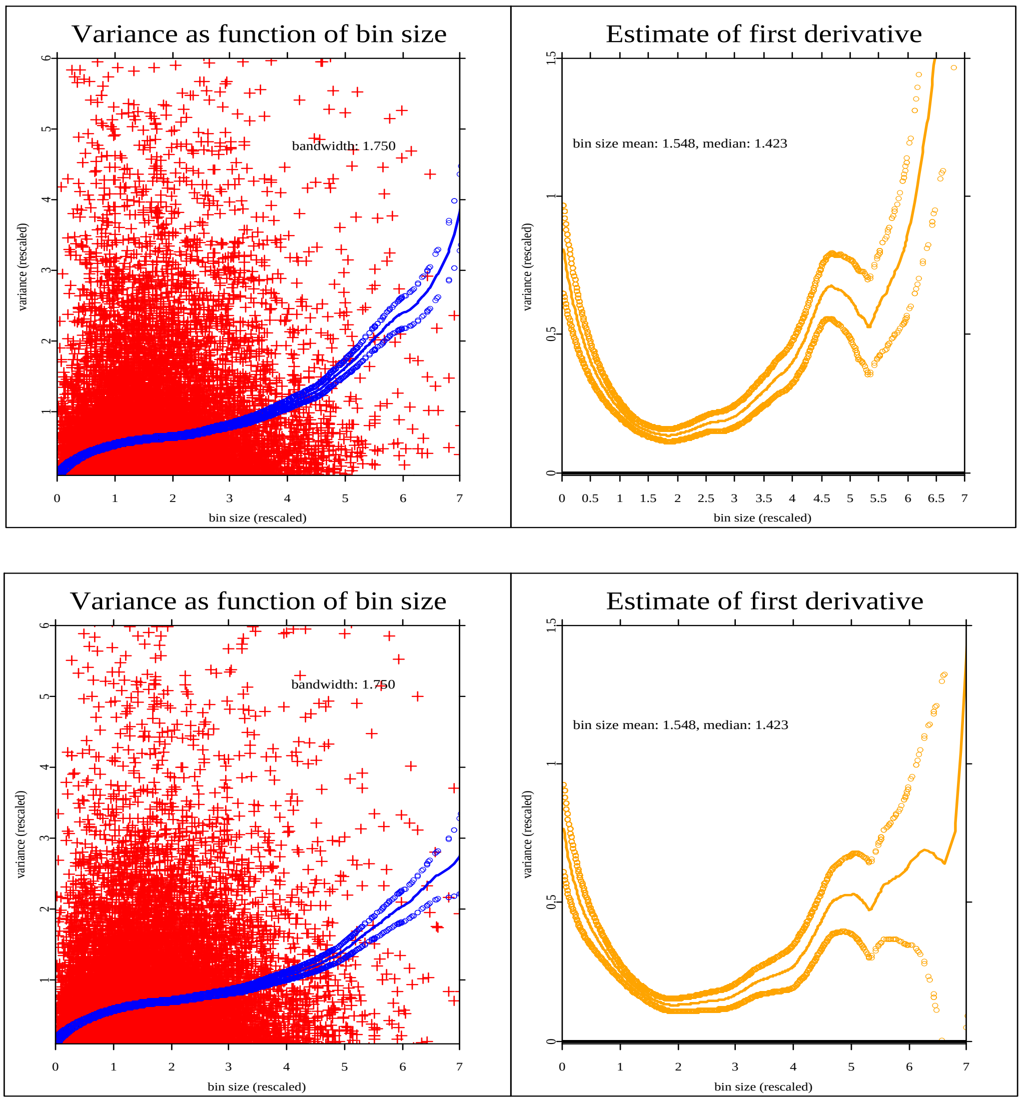


Figure 11: Euro 6 months ask (top panel) and bid (bottom panel): Nonparametric estimation of bin size – variance relation and its first derivative

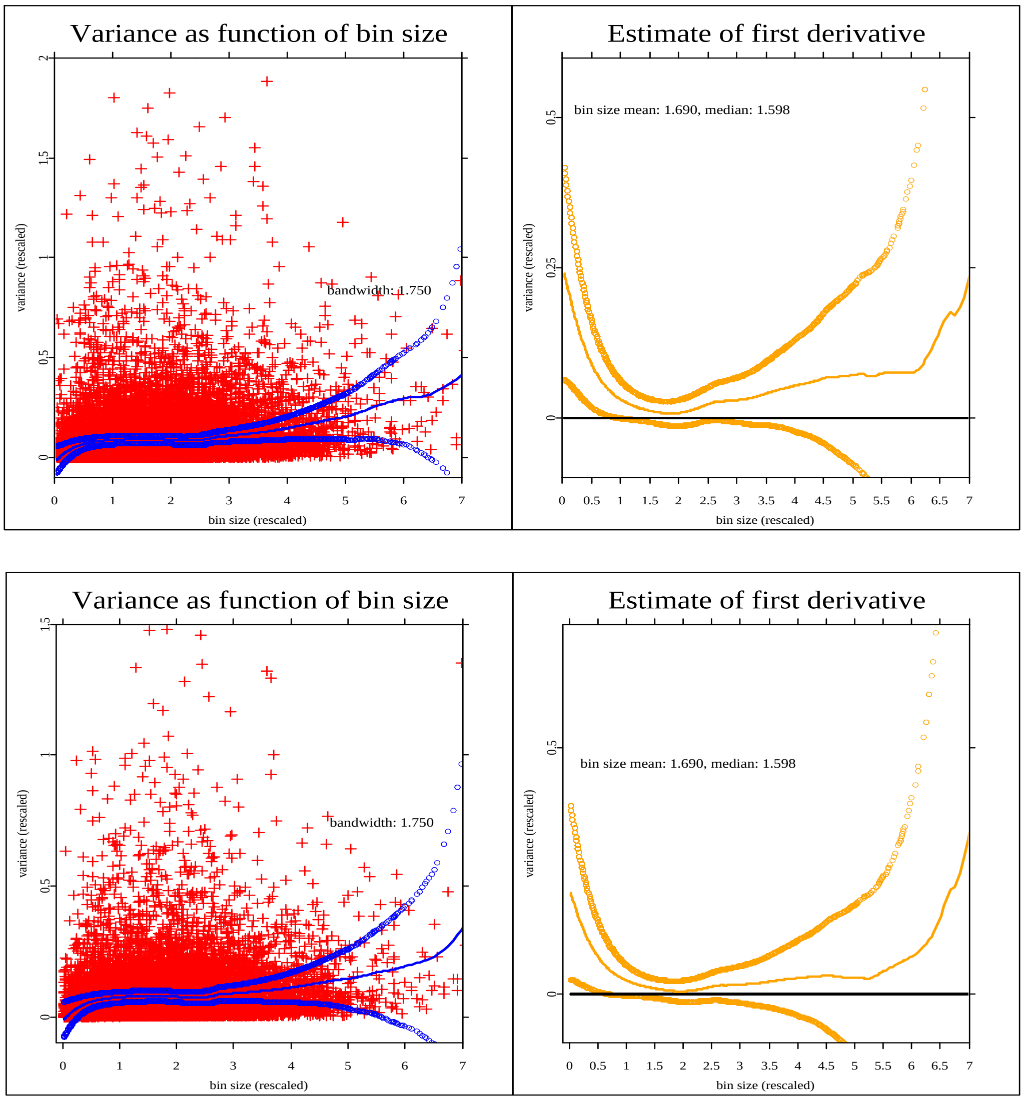


Figure 12: Euro 12 months ask (top panel) and bid (bottom panel): Nonparametric estimation of bin size – variance relation and its first derivative