

LIES, DAMNED LIES, AND STATISTICS?

Examples From Finance & Economics

LIES, DAMNED LIES, AND STATISTICS?

Examples From Finance & Economics

Reliable data analysis is one of the hardest tasks in sciences and social sciences. The need for it is pervasive:

“The best thing about being a statistician is that you get to play in everyone’s backyard”

John Tukey (1915–2000).

LIES, DAMNED LIES, AND STATISTICS?

Examples From Finance & Economics

Reliable data analysis is one of the hardest tasks in sciences and social sciences. The need for it is pervasive:

“The best thing about being a statistician is that you get to play in everyone’s backyard”

John Tukey (1915–2000).

A host of modern statistical tools are constantly being developed to cope with a wealth of new data.

Some methods need to be tailored to cope with different environments. (Not all ailments are cured by the same medicine!)

Today's "backyard" will be financial econometrics and empirical macro-economics. In particular:

1. Statistical distribution theory: how to quantify randomness.

Option-pricing requires the specification of a distribution of likely future prices. An arbitrarily-chosen form will give misleading results.

“You can't fix by analysis what you bungled by design”

Light, Singer, and Willett (1990).

We need new flexible specifications for distributions.

2. The relation between interest rates on different maturities, now and in the future; the “term structure of interest rates”:

(a) Rare uncharacteristic events can generate one-off “outliers”.

They have distortionary effects on traditional methods of estimating relations between these interest rates.

(b) Short-term interest rates do *not* have the “Markovian” dynamics that prevail in the literatures on finance and time series.

We first need new methods to deal with such time-series data.

3. These new dynamics arise from a general-equilibrium economic model.

Because of this link, it turns out that exchange rates, stock market indexes, and all macroeconomic variables are well characterized by this new process.

Implications for trading (momentum, cycles, etc.),
but also for macroeconomic stabilization.

1. Flexible specification of probability distributions

Loosely speaking, a probability density function (or simply “density”) depicts the relative frequency of outcomes for a variable.

1. Flexible specification of probability distributions

Loosely speaking, a probability density function (or simply “density”) depicts the relative frequency of outcomes for a variable.

A flexible parametric form for densities is needed to:

1. cover a wide range of possible descriptions and shapes;
2. achieve efficiency gains over nonparametric estimation, gains that are useful for cases where there is not a lot of data.

This is needed for many statistical applications, well beyond the current context of option pricing.

A “call option” gives you the right (but not the obligation) to buy an asset at a predetermined price until an expiration (or maturity) date.

Its value depends on the distribution of likely future prices.

A “call option” gives you the right (but not the obligation) to buy an asset at a predetermined price until an expiration (or maturity) date.

Its value depends on the distribution of likely future prices.

Let x_t be the price of an asset at time t . Assume, for the moment, that the asset does not pay dividends. Suppose that this asset is underlying a European call option with expiration date T and strike price k .

Then, the intrinsic value of this option at expiration is $\max\{x_T - k, 0\}$.

In an arbitrage-free economy, there exists a risk-neutral density f such that the price of this call option at time t can be written as

$$c_t(k) = e^{-(T-t)r} \mathbf{E}_t(\max\{x_T - k, 0\}) \equiv e^{-(T-t)r} \int_k^\infty (u - k) f(u) du,$$

where r is the continuously-compounded risk-free interest rate and \mathbf{E}_t is the expectation taken at time t .

Differentiating the integral gives

$$\frac{dc_t(k)}{dk} = -e^{-(T-t)r} \int_k^\infty f(u) du \equiv -e^{-(T-t)r} (1 - F(k)),$$

where F is the c.d.f. corresponding to the p.d.f. f . The second derivative is given by

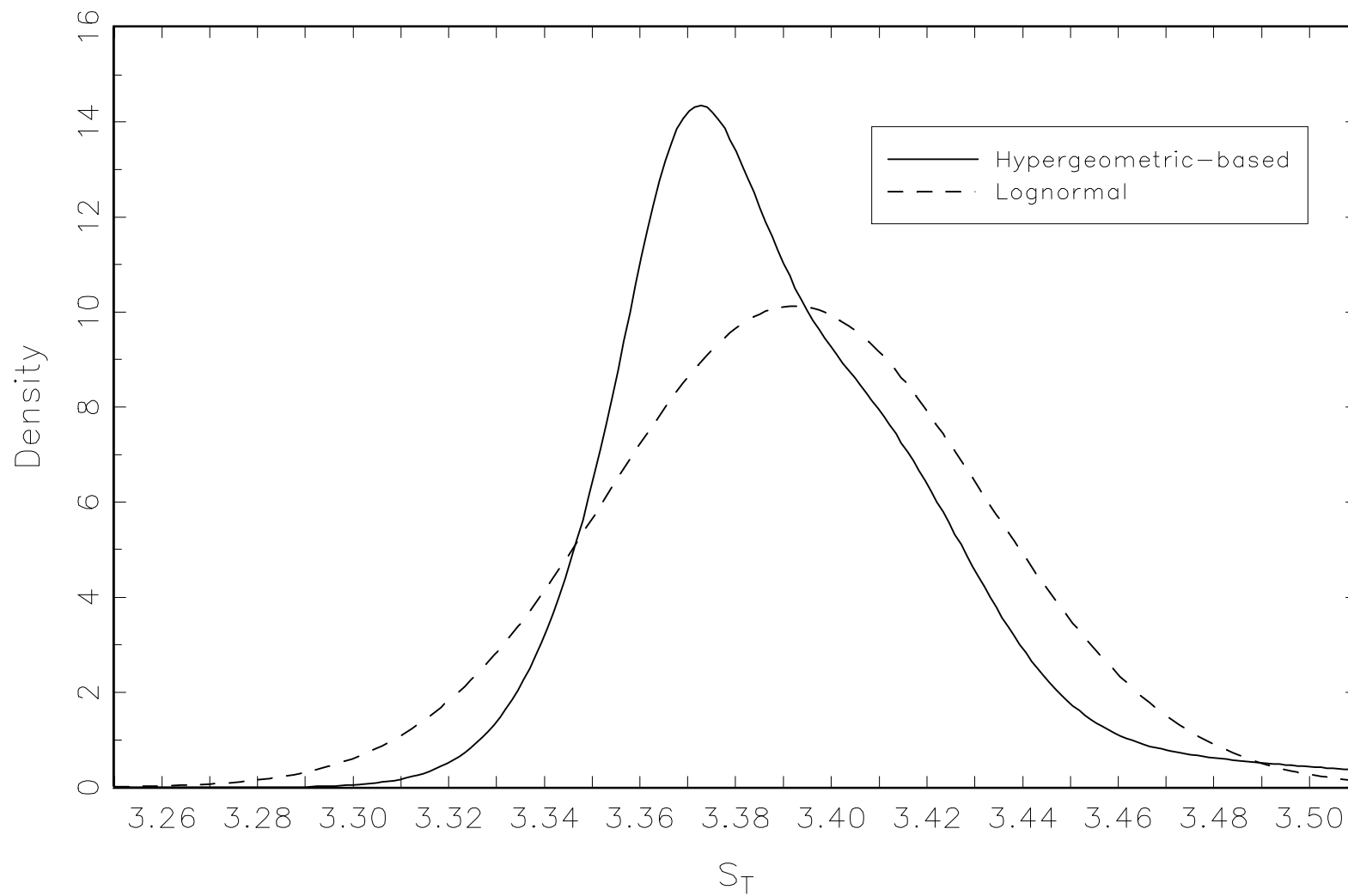
$$\left. \frac{d^2 c_t(k)}{dk^2} \right|_{k=u} = e^{-(T-t)r} f(u),$$

which reveals the required density f . If the asset pays a dividend yield of r^\dagger , then it should be subtracted from r and the RHSs should be multiplied by $e^{(T-t)r^\dagger}$.

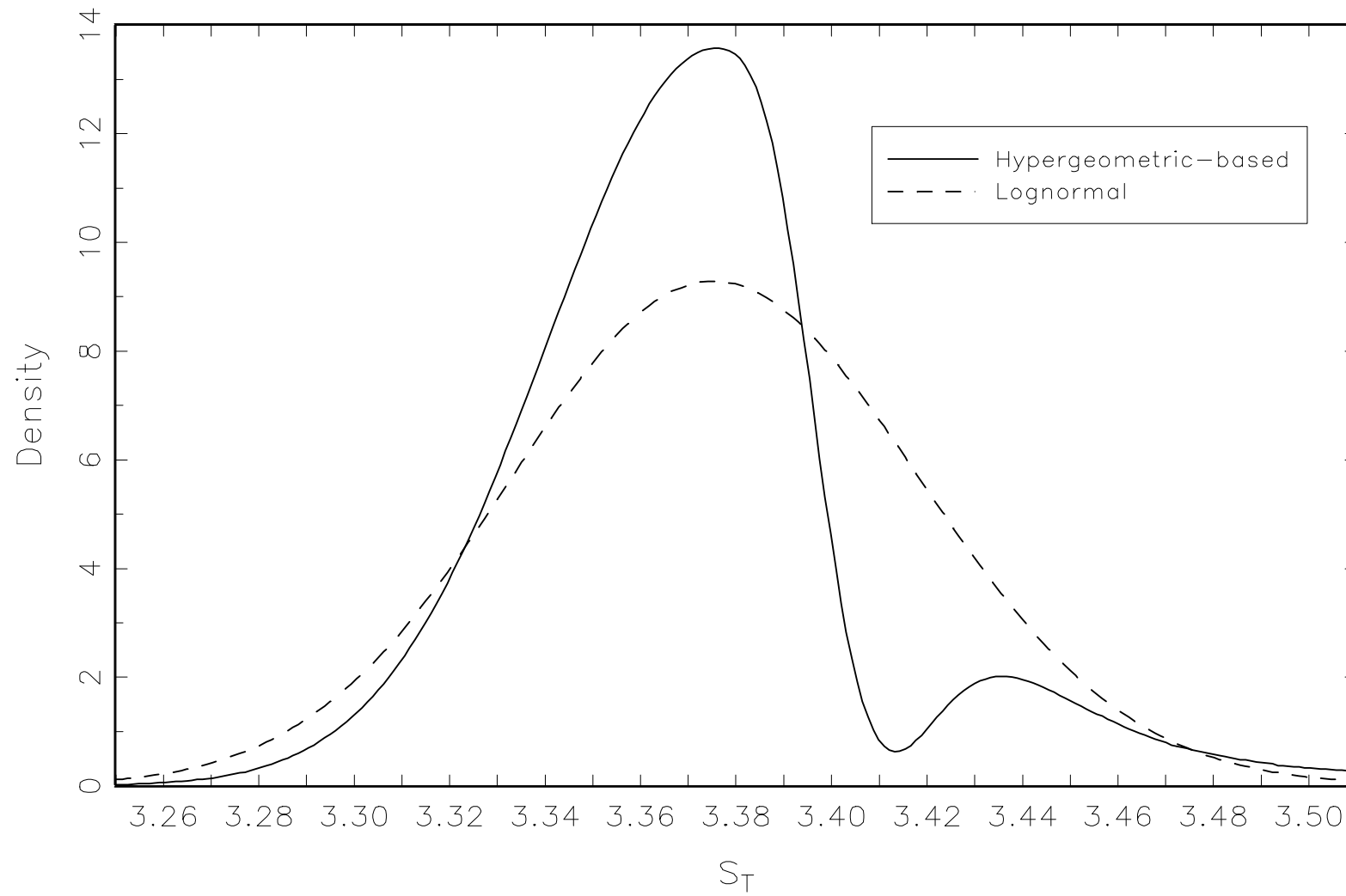
The Black-Scholes (1973, *JPE*) and Garman-Kohlhagen (1983, *JIMF*) models assume that f is log-normal.

Using a construction from special functions in mathematics (“hypergeometric functions”), first introduced in Abadir and Rockinger (2003, *ET*), we get the following densities about future asset prices (exchange rates and S&P500), as implied by current option prices...

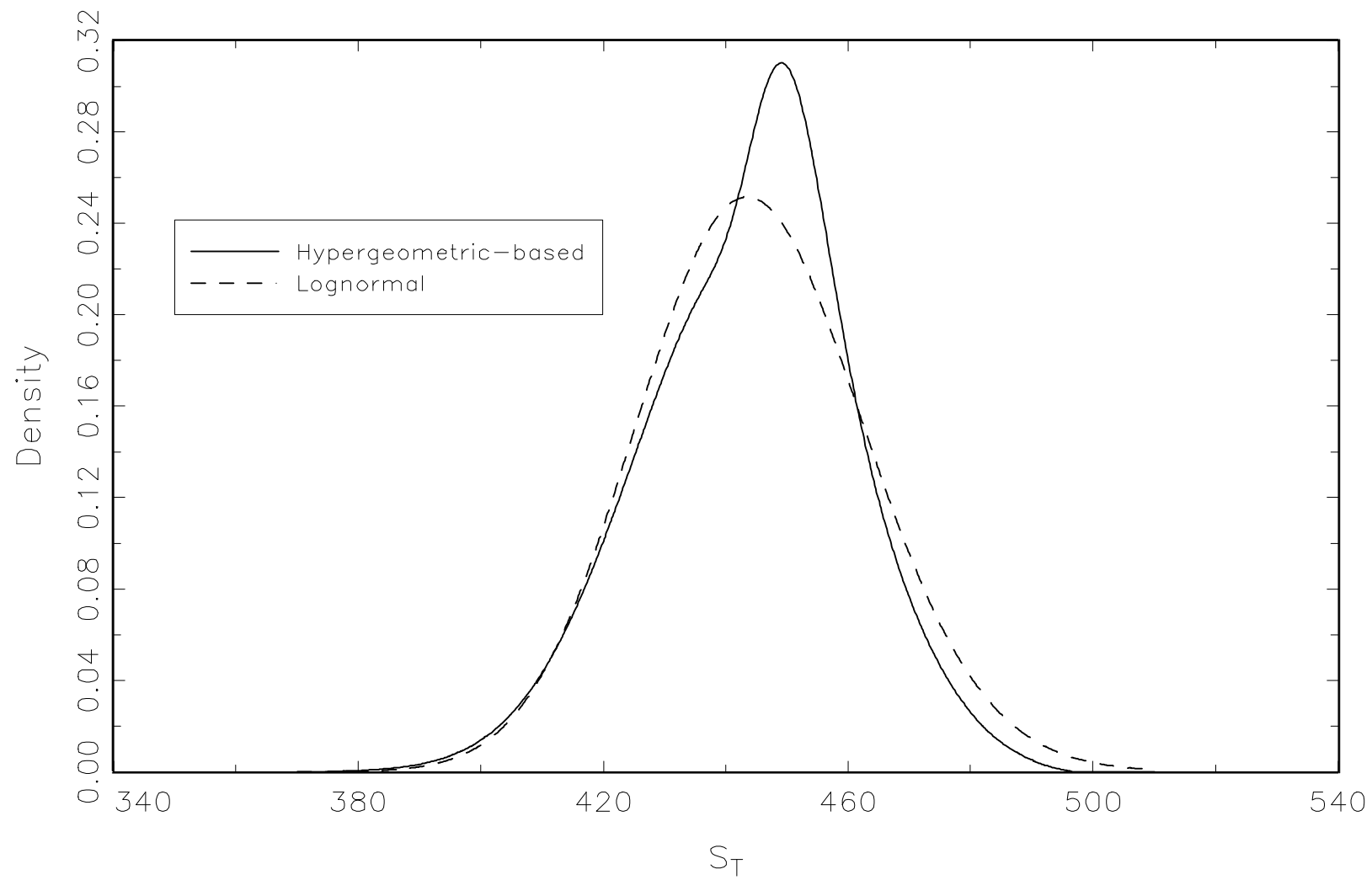
Lognormal and Hypergeometric-based Risk-Neutral-Densities
17 May 96 3-months to maturity



Lognormal and Hypergeometric-based Risk-Neutral-Densities
28 April 97 3-months to maturity

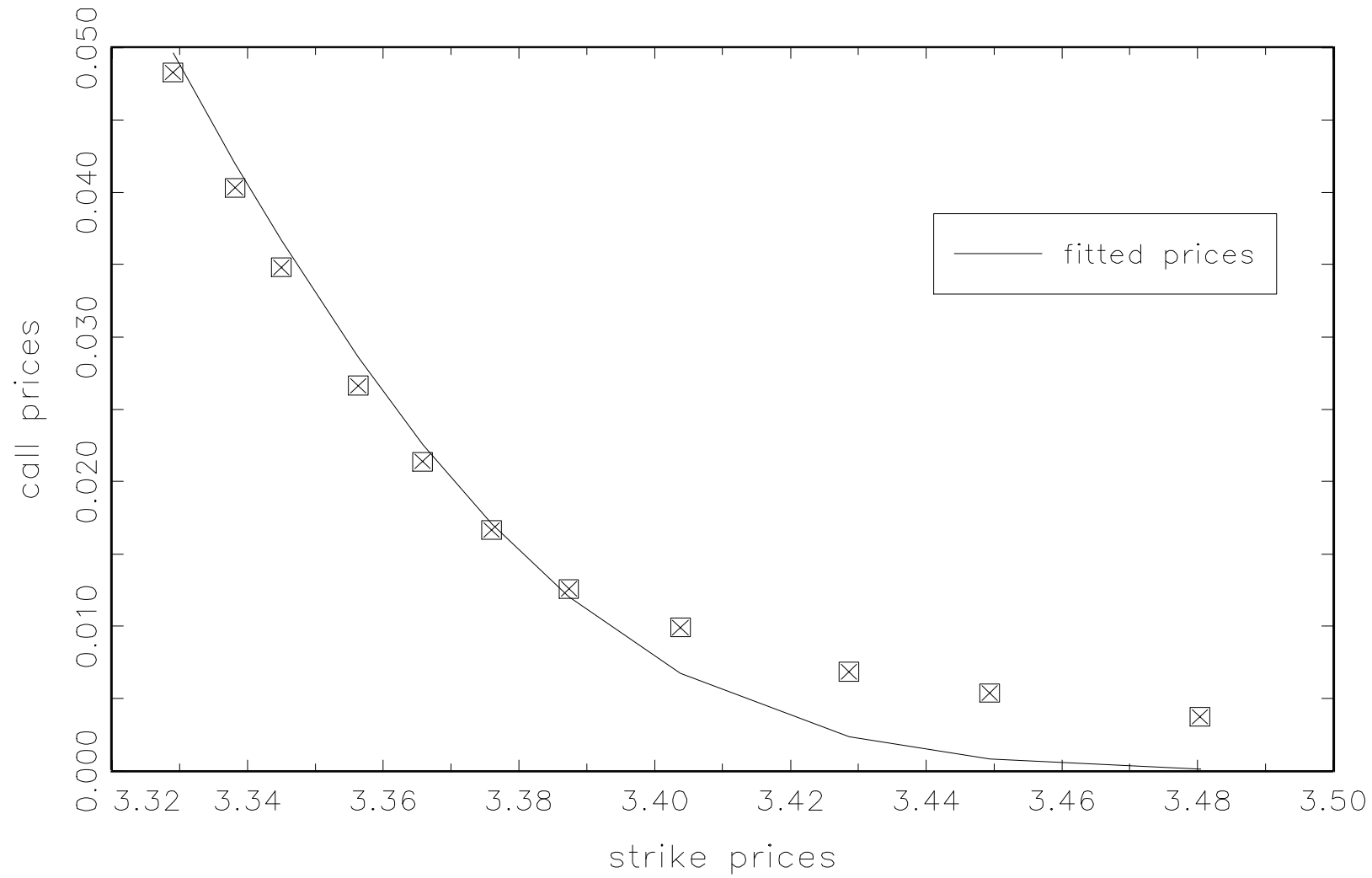


Lognormal and Hypergeometric-based Risk-Neutral-Densities
S&P500 index options, 3 May 1993, 46-days to maturity

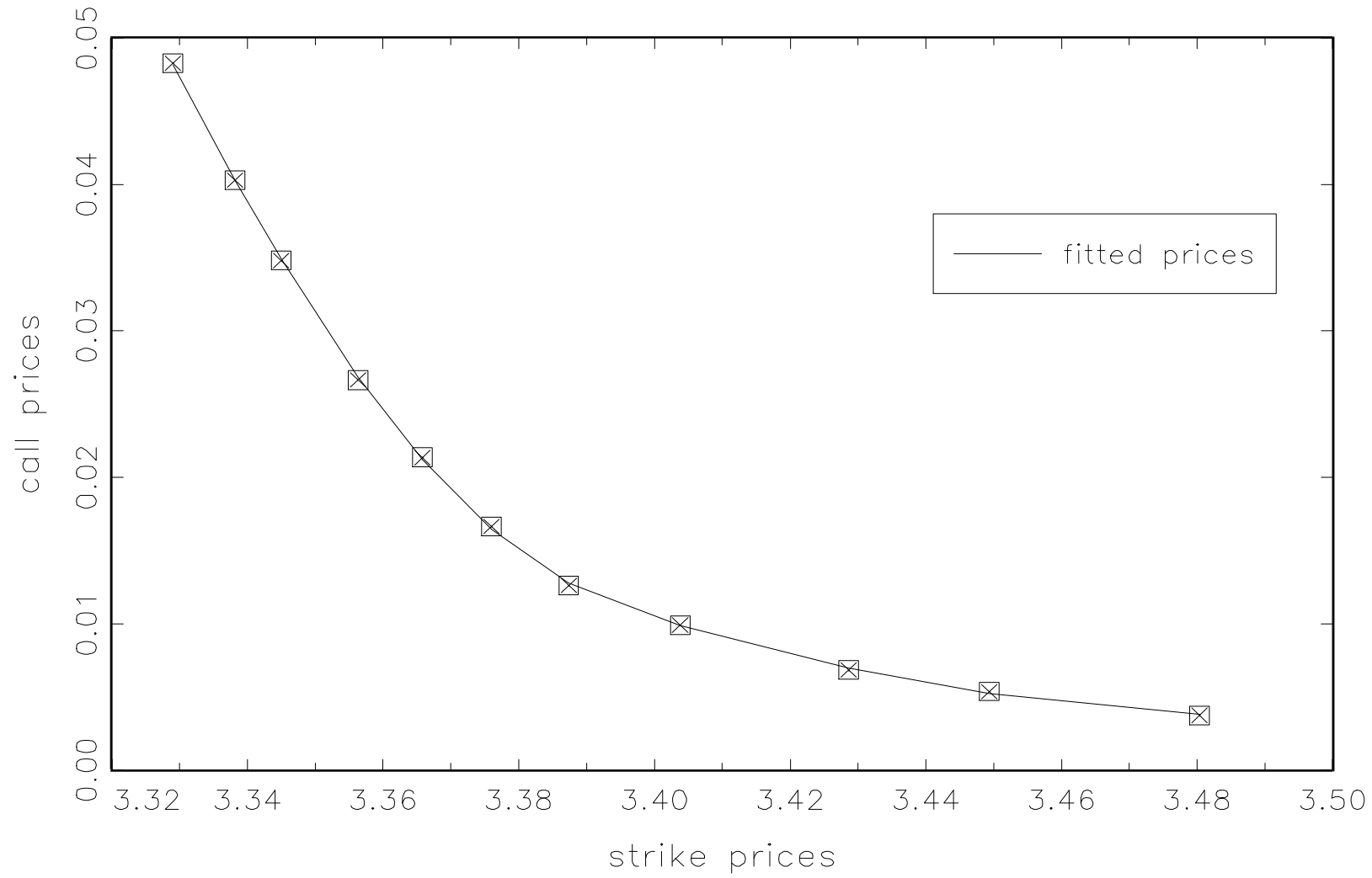


The representative fit for option prices implied by these densities, as opposed to the Garman-Kohlhagen (1983, *JIMF*) or Black-Scholes (1973, *JPE*) benchmark is...

Original and fitted Garman–Kohlhagen call option prices
28 April 97 3-months to maturity



Original data and Hypergeometric-based fit
28 April 97 3-months to maturity



There are other methods too. But they do not fit as well, and they do not detect the:

1. growing polarization in the currency market when an important event happens;
2. negative skew of the density of S&P500 (higher prevalence of negative corrections in shares).

There are other methods too. But they do not fit as well, and they do not detect the:

1. growing polarization in the currency market when an important event happens;
2. negative skew of the density of S&P500 (higher prevalence of negative corrections in shares).

Also, other methods force unjustified features on the estimates (e.g. number of modes, oscillatory tails,...) or have no explicit structure (a drawback when forecasting and/or analyzing tails, i.e. extreme events).

The new method does very well, and uniformly so for:

1. volatile days, as well as for quiet ones;
2. different maturities; and
3. options on very different underlying assets.

The estimated densities' functional forms (i.e. shapes) vary over the dates, maturities, and assets. Therefore, other parametric methods which restrict functional forms will not do uniformly well.

The new method does very well, and uniformly so for:

1. volatile days, as well as for quiet ones;
2. different maturities; and
3. options on very different underlying assets.

The estimated densities' functional forms (i.e. shapes) vary over the dates, maturities, and assets. Therefore, other parametric methods which restrict functional forms will not do uniformly well.

The estimates are stable and robust, as revealed by sensitivity analysis to bid-ask spread, dependence on each data point, etc.

Talking of robustness...

2. The term structure of interest rates

a. The expectation hypothesis (EH)

Simple statement of EH: “an investor who holds a bond for a long time realizes an interest which is an "average" of the oscillating rates of those who speculate during the interim”. Irving Fisher (1896).

2. The term structure of interest rates

a. The expectation hypothesis (EH)

Simple statement of EH: “an investor who holds a bond for a long time realizes an interest which is an "average" of the oscillating rates of those who speculate during the interim”. Irving Fisher (1896).

Time and time again, the EH has been rejected empirically.

Up to today, many papers treat deviations from the EH as model specification errors, and try to fix them.

2. The term structure of interest rates

a. The expectation hypothesis (EH)

Simple statement of EH: “an investor who holds a bond for a long time realizes an interest which is an "average" of the oscillating rates of those who speculate during the interim”. Irving Fisher (1896).

Time and time again, the EH has been rejected empirically.

Up to today, many papers treat deviations from the EH as model specification errors, and try to fix them.

But is this really the case? If so, what causes EH to break down?

2. The term structure of interest rates

a. The expectation hypothesis (EH)

Simple statement of EH: “an investor who holds a bond for a long time realizes an interest which is an "average" of the oscillating rates of those who speculate during the interim”. Irving Fisher (1896).

Time and time again, the EH has been rejected empirically.

Up to today, many papers treat deviations from the EH as model specification errors, and try to fix them.

But is this really the case? If so, what causes EH to break down?

It will turn out that the empirical rejection of the EH is largely due to a handful of outliers, corresponding to a couple of rare events that are mostly politically-driven.

Once these are accounted for, there is a dramatic change in the estimates of term-structure regressions, bringing them much closer to the EH.

Once these are accounted for, there is a dramatic change in the estimates of term-structure regressions, bringing them much closer to the EH.

Let us focus on the following version of EH:

forward rates are unbiased predictors of futures rates.

Define $f(t, u)$ as the instantaneous forward rate that satisfies the pricing

$$P(t, T) = \exp \left(- \int_t^T f(t, u) du \right), \quad T > t,$$

for a zero-coupon bond with face value 1.

Define $f(t, u)$ as the instantaneous forward rate that satisfies the pricing

$$P(t, T) = \exp \left(- \int_t^T f(t, u) du \right), \quad T > t,$$

for a zero-coupon bond with face value 1.

Let $\pi(t, u) := f(t, u) - E_t[r(u)]$, where $r(u)$ is the spot rate at $u > t$ and E_t is the expectation taken at time t .

Define $f(t, u)$ as the instantaneous forward rate that satisfies the pricing

$$P(t, T) = \exp \left(- \int_t^T f(t, u) du \right), \quad T > t,$$

for a zero-coupon bond with face value 1.

Let $\pi(t, u) := f(t, u) - E_t[r(u)]$, where $r(u)$ is the spot rate at $u > t$ and E_t is the expectation taken at time t .

The EH with constant risk premium states that $\pi(t, u) = \bar{\pi}$, hence

$$y(t, T) = \frac{\int_t^T (E_t[r(u)] + \bar{\pi}) du}{T - t} = \bar{\pi} + \frac{\int_t^T E_t[r(u)] du}{T - t},$$

where $y(t, T)$ is the yield at time t of a bond maturing at time T .

Define $f(t, u)$ as the instantaneous forward rate that satisfies the pricing

$$P(t, T) = \exp \left(- \int_t^T f(t, u) du \right), \quad T > t,$$

for a zero-coupon bond with face value 1.

Let $\pi(t, u) := f(t, u) - E_t[r(u)]$, where $r(u)$ is the spot rate at $u > t$ and E_t is the expectation taken at time t .

The EH with constant risk premium states that $\pi(t, u) = \bar{\pi}$, hence

$$y(t, T) = \frac{\int_t^T (E_t[r(u)] + \bar{\pi}) du}{T - t} = \bar{\pi} + \frac{\int_t^T E_t[r(u)] du}{T - t},$$

where $y(t, T)$ is the yield at time t of a bond maturing at time T .

A discretization gives a restriction that has been extensively tested:

$$y(t, t + n) = \bar{\pi} + \frac{\sum_{i=1}^n E_t[y(t + i - 1, t + i)]}{n}, \quad (1)$$

where $n := T - t$ is the time to maturity.

Reproducing (1),

$$y(t, t + n) = \bar{\pi} + \frac{1}{n} \sum_{i=1}^n E_t [y(t + i - 1, t + i)],$$

we see that the n -period rate is the average of a series of one-period expectations of future rates (recall Fisher's quote), plus a constant risk premium $\bar{\pi}$.

Reproducing (1),

$$y(t, t+n) = \bar{\pi} + \frac{1}{n} \sum_{i=1}^n E_t [y(t+i-1, t+i)],$$

we see that the n -period rate is the average of a series of one-period expectations of future rates (recall Fisher's quote), plus a constant risk premium $\bar{\pi}$.

Here, changes in yields can only be explained by changes in the expectation of the future interest rates.

If the risk premium were time-varying, yields would also depend on changes of attitudes towards risk or changes in interest rate volatility.

Interest rates are highly persistent, as we shall see in Section 2.b.

To reduce the possibility of spurious results, most empirical tests of EH focus on interest-rates' differences or spreads, rather than levels.

Interest rates are highly persistent, as we shall see in Section 2.b.

To reduce the possibility of spurious results, most empirical tests of EH focus on interest-rates' differences or spreads, rather than levels.

Two relations can be derived from (1). One of them is

$$E_t[y(t+1, T)] - y(t, T) = \alpha_n + \beta_n \frac{y(t, T) - y(t, t+1)}{n-1},$$

where $\beta_n = 1$ if EH is satisfied. A positive risk premium translates into negative values for α , while Jensen's convexity effect increases α .

The equation relates future changes in long-term interest rates (call this LHS z) to the slope of the term structure (call this RHS fraction x) after correction for a constant term premium.

Least squares (LS) regression has been used to check this hypothesis.

Least squares (LS) regression has been used to check this hypothesis.

But the LS regression of z on x is mathematically equivalent to the calculation of the conditional expectation $z = E(z | x) + \varepsilon$ from the joint density of (x, z) .

Least squares (LS) regression has been used to check this hypothesis.

But the LS regression of z on x is mathematically equivalent to the calculation of the conditional expectation $z = E(z | x) + \varepsilon$ from the joint density of (x, z) .

Conditional expectations are linear for distributions like elliptical ones (including normal and Student t), but nonlinear for most others.

Nonparametric (NP) kernel regression reveals the functional form of the conditional expectation and how far it deviates from linearity and from $\beta_n = 1$.

Least squares (LS) regression has been used to check this hypothesis.

But the LS regression of z on x is mathematically equivalent to the calculation of the conditional expectation $z = E(z | x) + \varepsilon$ from the joint density of (x, z) .

Conditional expectations are linear for distributions like elliptical ones (including normal and Student t), but nonlinear for most others.

Nonparametric (NP) kernel regression reveals the functional form of the conditional expectation and how far it deviates from linearity and from $\beta_n = 1$.

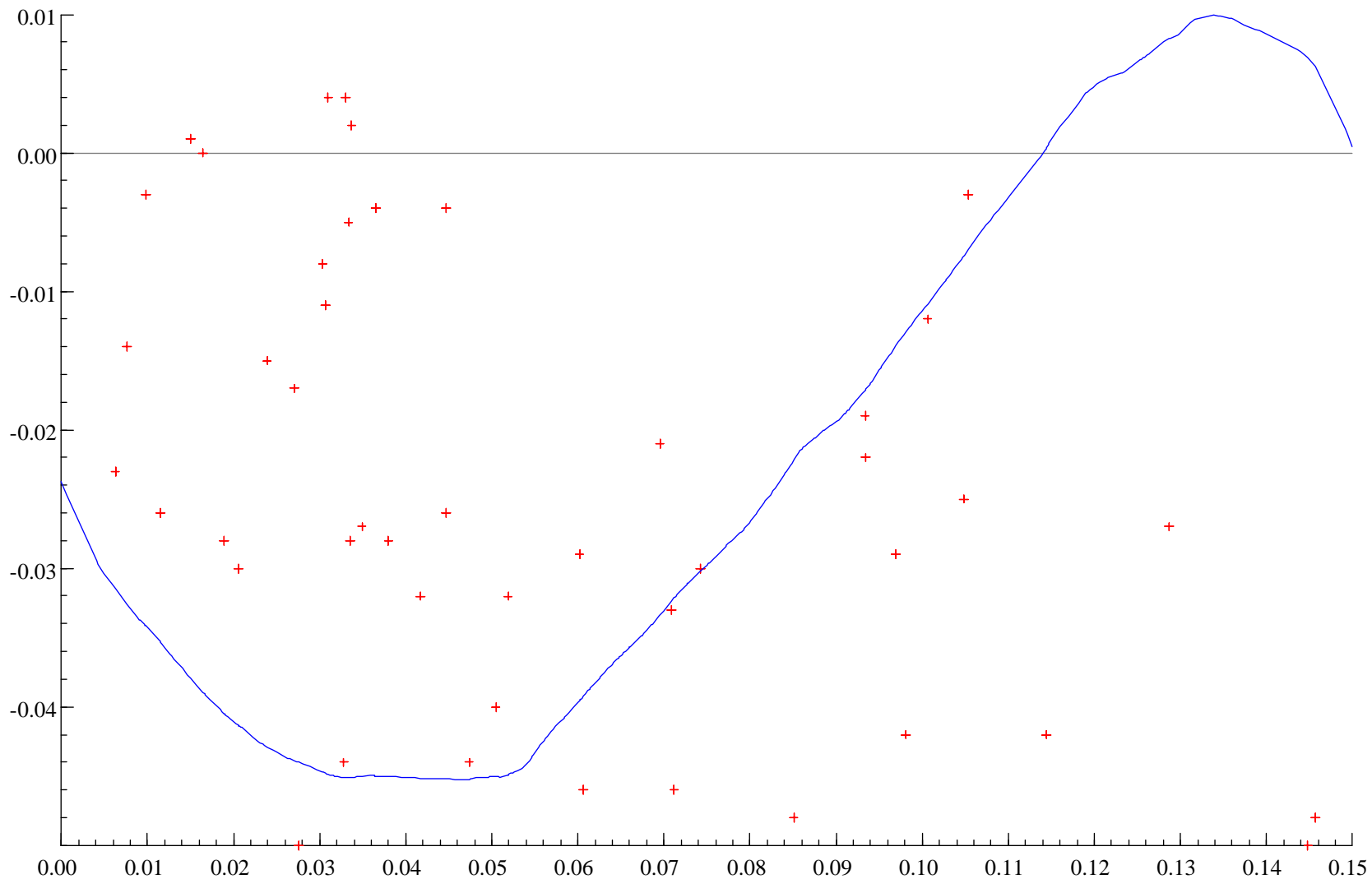
Being a local-LS device, it also reveals the extent to which earlier results in the literature depended on a handful of extreme observations driven by unusual events: in 1981/1982 (the Volcker experiment) and 1987 (the stock market crash).

From the CRSP database, one can get monthly yields on Treasury Bills with 1 to 12 months to maturity from January 1959 to December 2000, and yields on artificial zero-coupon Treasury bonds with 1 to 10 years to maturity.

The results are qualitatively similar across the board, so only one-year estimates are reported.

Two graphs for the same NP regression will follow: one showing all the data points, the other zooming into the same picture to show in more detail the fitted NP regression curve for the bulk of the data.

Actually, it is even better (for EH) if we zoom into the body of the graph and enlarge it...



For the x-axis ranging from 0.05 to 0.12, the NP regression is amazingly linear (this has not been imposed at the outset) with a slope of just over 0.7. Even when taking all the points on the last graph, the “average” slope is little changed.

This is in sharp contrast to the massively negative number obtained by LS regression, namely -1.2 with a standard error of 0.45 leading to the rejection of EH.

We can now see that the puzzling LS result was due to the handful of extreme values (in the first figure) that are driving the “average” slope down, and that the slope for the bulk of the data is statistically indistinguishable from the unit coefficient that is implied by EH.

b. Up & down it goes (dynamics of the short rate)

It is well known that interest rates are persistent.

Further, it has been acknowledged relatively recently that the short-rate process contains nonlinearities. For example, Aït-Sahalia (1996, *RFS*) finds nonparametrically that the process is:

1. very persistent when rates are close to their historical norm;
2. more mean-reverting when rates are far from their norm.

NP provides an exploratory tool, but it does not help with forecasts (NP is a “local” smoothing tool and is too flexible for projections into the future).

We need a parametric model to do so.

NP provides an exploratory tool, but it does not help with forecasts (NP is a “local” smoothing tool and is too flexible for projections into the future).

We need a parametric model to do so.

To add to the troubles, we will see that the process is not Markovian:

1. one cannot summarize the bulk of the dynamics by conditioning on a fixed number of previous values;
2. the representation of the conditioning structure evolves over time if (as is always the case) the process has started at a fixed point in time.

So we cannot use such models, which are currently dominating the field.

NP provides an exploratory tool, but it does not help with forecasts (NP is a “local” smoothing tool and is too flexible for projections into the future).

We need a parametric model to do so.

To add to the troubles, we will see that the process is not Markovian:

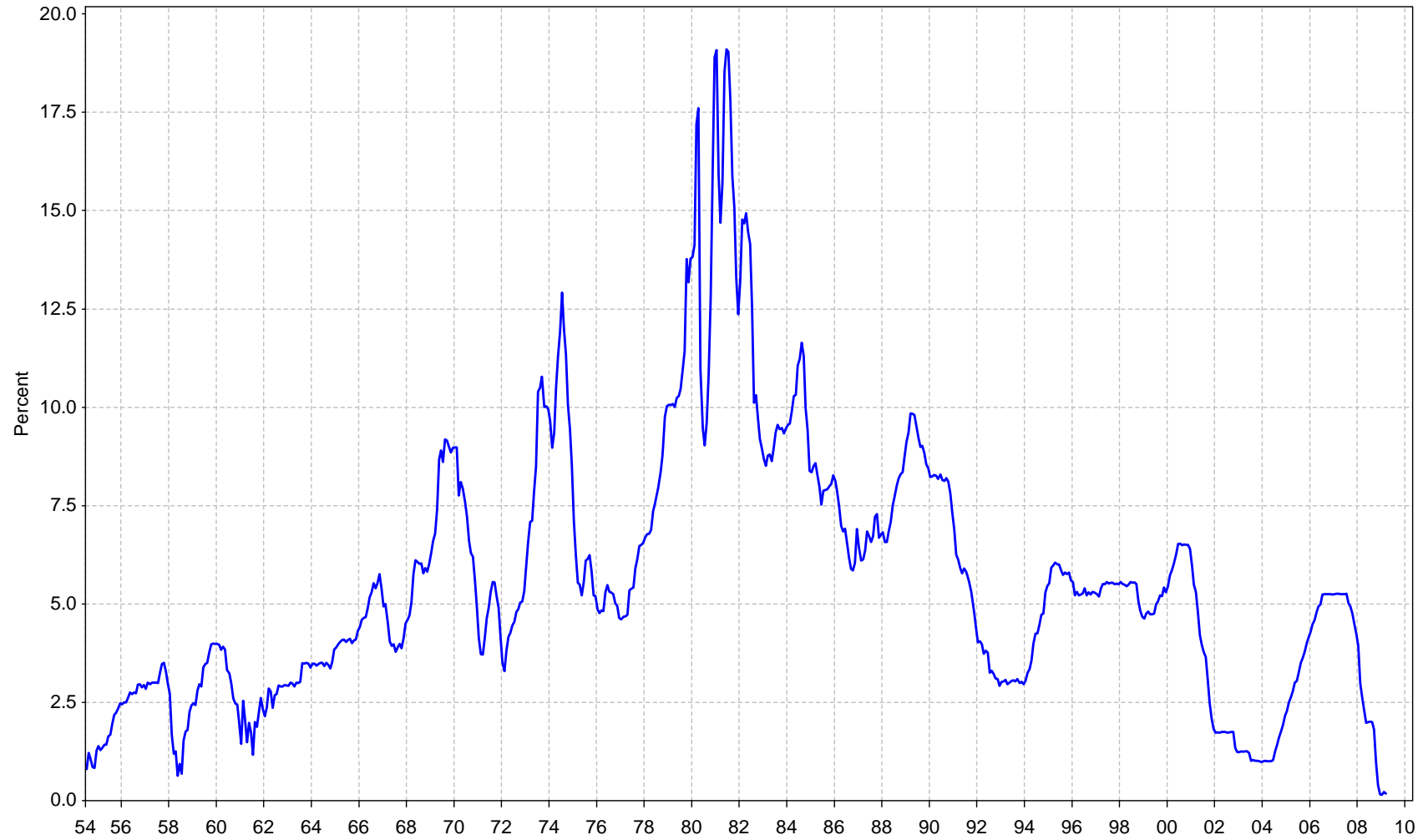
1. one cannot summarize the bulk of the dynamics by conditioning on a fixed number of previous values;
2. the representation of the conditioning structure evolves over time if (as is always the case) the process has started at a fixed point in time.

So we cannot use such models, which are currently dominating the field.

Furthermore, the process does not look like any existing non-Markovian model that we can pull off the shelf: we need new technology!

Here is a picture illustrating what we're up against...

United States, Policy Rates, Fed Funds Effective Rate, Average, USD



Source: Reuters EcoWin

What a horrible animal! All jagged and seemingly patternless.

What a horrible animal! All jagged and seemingly patternless.

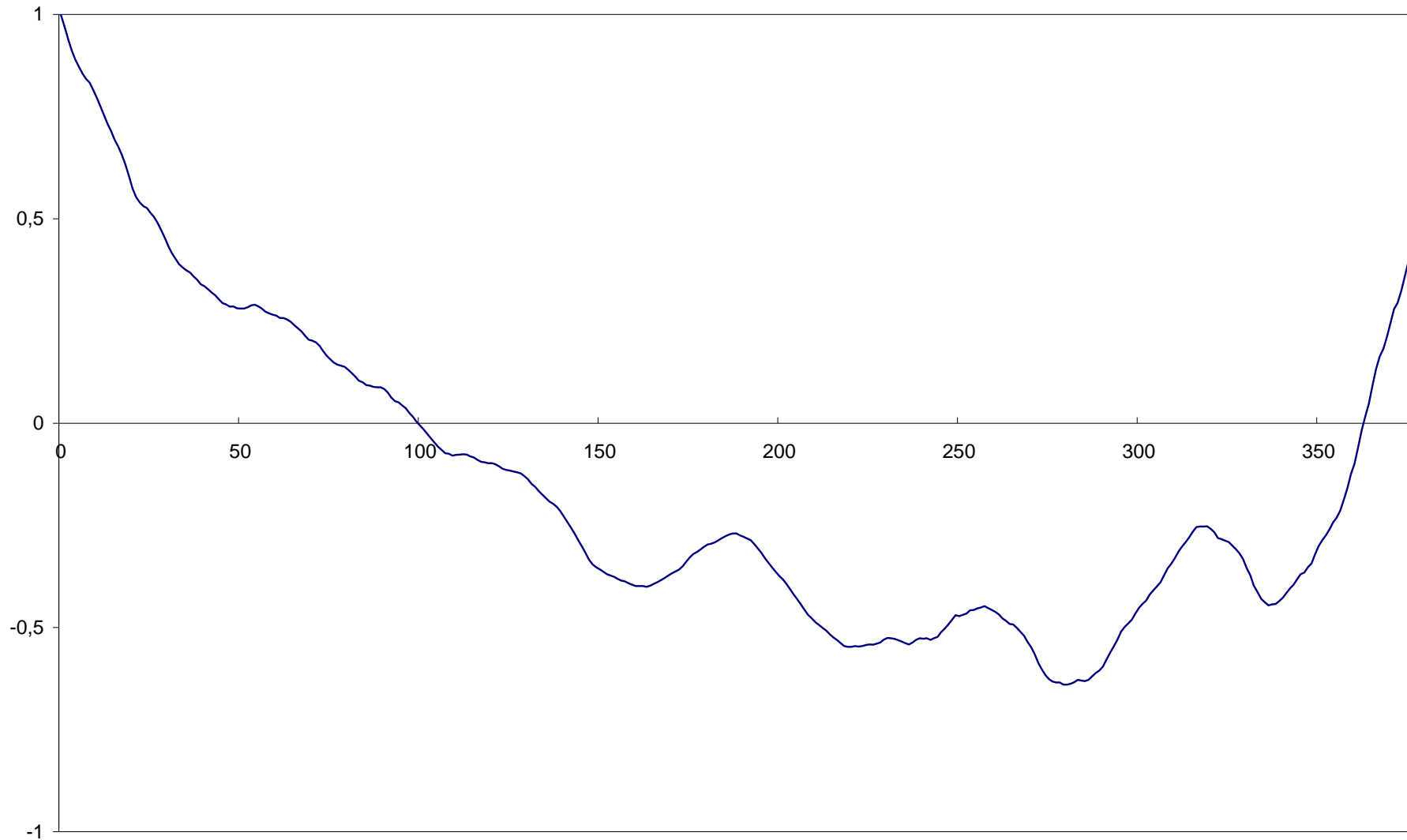
Really?!

What a horrible animal! All jagged and seemingly patternless.

Really?!

Here is how each point correlates with the previous values (the auto-correlation function or ACF)...

... a smoothie (with a wiggly tail)!



Interest rates go in long cycles.

“The evidence in Fama and Bliss (1987) that forward interest rates forecast future spot interest rates for horizons beyond a year repeats in the out-of-sample 1986–2004 period. But the inference that this forecast power is due to mean reversion of the spot rate toward a constant expected value no longer seems valid. Instead, the predictability of the spot rate captured by forward rates seems to be due to mean reversion toward a time-varying expected value that is subject to a sequence of apparently permanent shocks that are on balance positive to mid-1981 and on balance negative thereafter.”

Eugene Fama (2006, *RFS*)

What next?!

It turns out that this interest-rate process has a parsimonious representation in the ACF and frequency domains. The ACF is

$$\rho_\tau := \frac{\text{cov}(z_t, z_{t-\tau})}{\sqrt{\text{var}(z_t) \text{var}(z_{t-\tau})}} \approx \frac{\cos(\omega\tau)}{(1 + b\tau)^c} \quad (b, c > 0).$$

Its Fourier-inversion produces a spectrum $f(\lambda)$ that is proportional to $|\lambda - \omega|^{c-1}$; that is, at frequency ω , there is a singularity when $c \in (0, 1)$.

It turns out that this interest-rate process has a parsimonious representation in the ACF and frequency domains. The ACF is

$$\rho_\tau := \frac{\text{cov}(z_t, z_{t-\tau})}{\sqrt{\text{var}(z_t) \text{var}(z_{t-\tau})}} \approx \frac{\cos(\omega\tau)}{(1 + b\tau)^c} \quad (b, c > 0).$$

Its Fourier-inversion produces a spectrum $f(\lambda)$ that is proportional to $|\lambda - \omega|^{c-1}$; that is, at frequency ω , there is a singularity when $c \in (0, 1)$.

For linear ARIMA(p, d, q) processes having $d \in (0, \frac{1}{2})$, the spectrum has a singularity at the origin that is proportional to $|\lambda|^{-2d}$, giving the correspondence $c = 1 - 2d$ if $\omega = 0$ but *not* otherwise!

It turns out that this interest-rate process has a parsimonious representation in the ACF and frequency domains. The ACF is

$$\rho_\tau := \frac{\text{cov}(z_t, z_{t-\tau})}{\sqrt{\text{var}(z_t) \text{var}(z_{t-\tau})}} \approx \frac{\cos(\omega\tau)}{(1 + b\tau)^c} \quad (b, c > 0).$$

Its Fourier-inversion produces a spectrum $f(\lambda)$ that is proportional to $|\lambda - \omega|^{c-1}$; that is, at frequency ω , there is a singularity when $c \in (0, 1)$.

For linear ARIMA(p, d, q) processes having $d \in (0, \frac{1}{2})$, the spectrum has a singularity at the origin that is proportional to $|\lambda|^{-2d}$, giving the correspondence $c = 1 - 2d$ if $\omega = 0$ but *not* otherwise!

This representation is not arbitrarily chosen.

It arises from a general-equilibrium economic model that allows for heterogeneity of firms, introduced by Abadir and Talmain (2002, *REStud*).

Not surprisingly, most macroeconomic series and financial aggregates share this common structure, to which we now turn.

3. The new evolution

Once nonlinear dynamics are accounted for, gone is the apparent unit root that has dominated econometrics for the past 30+ years!

3. The new evolution

Once nonlinear dynamics are accounted for, gone is the apparent unit root that has dominated econometrics for the past 30+ years!

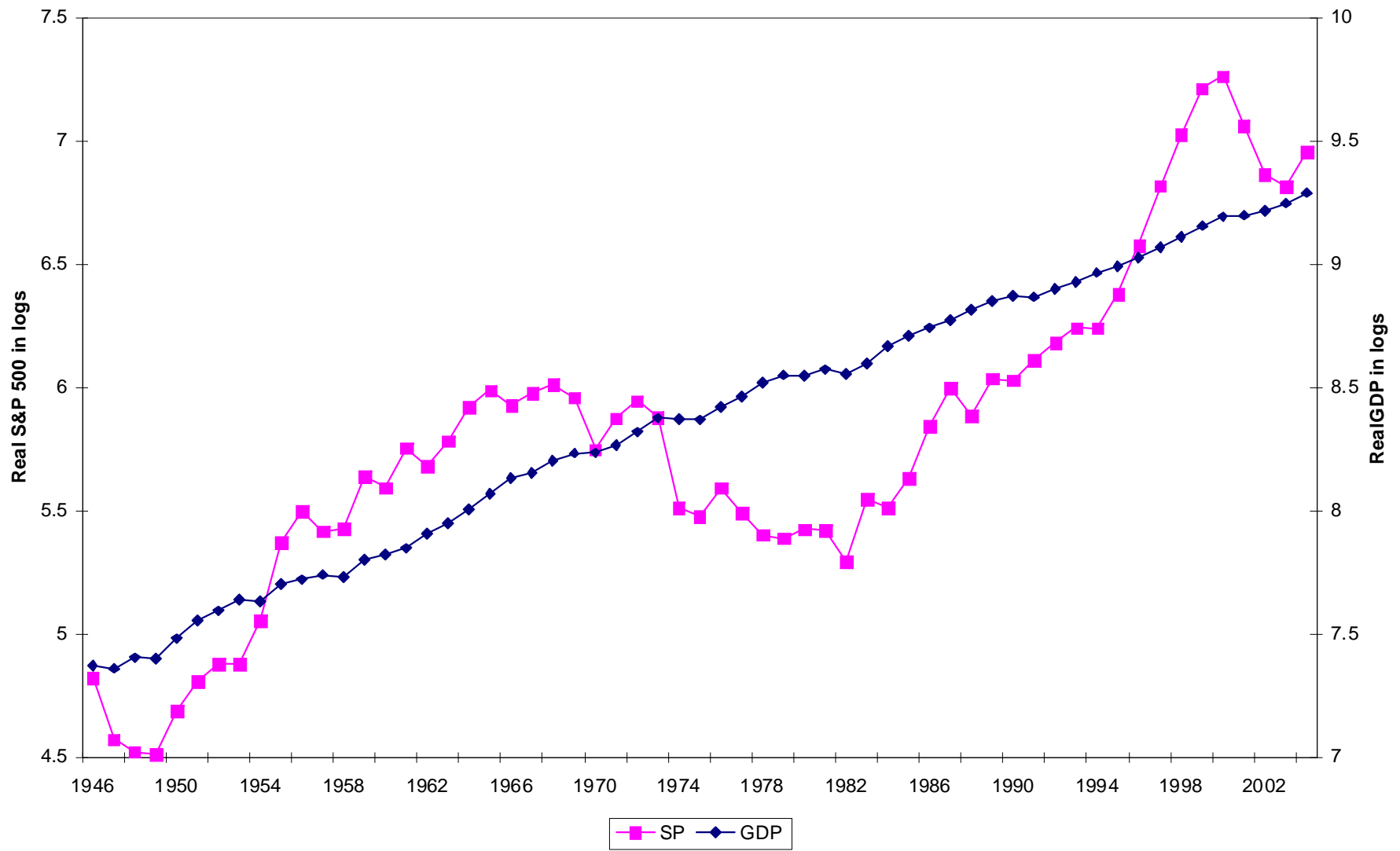
This applies to exchange rates, stock market indexes, and all macroeconomic variables.

3. The new evolution

Once nonlinear dynamics are accounted for, gone is the apparent unit root that has dominated econometrics for the past 30+ years!

This applies to exchange rates, stock market indexes, and all macroeconomic variables.

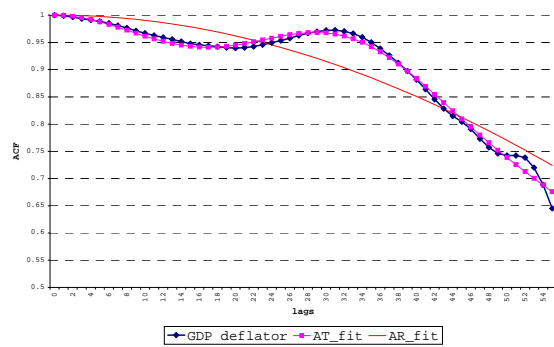
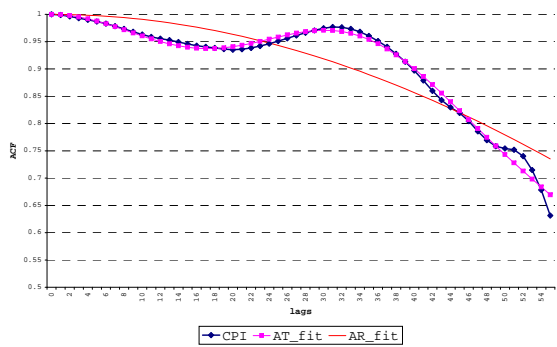
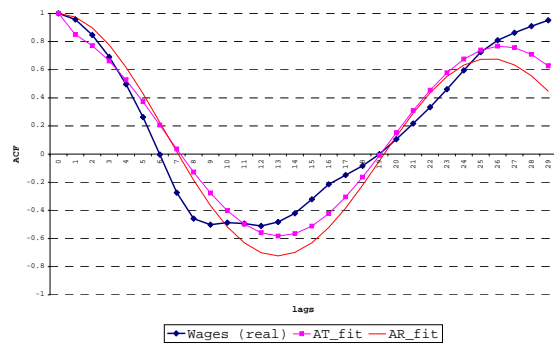
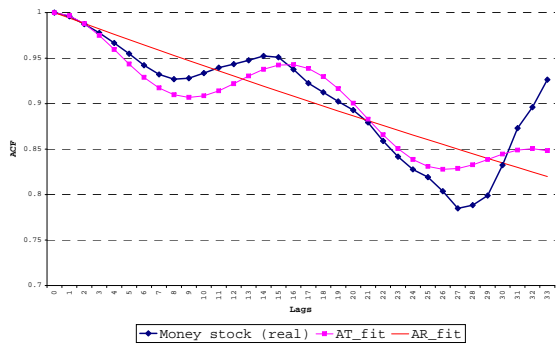
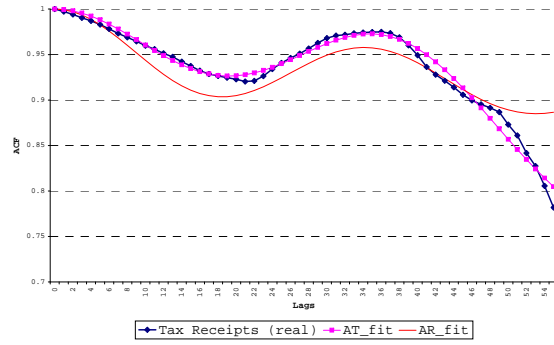
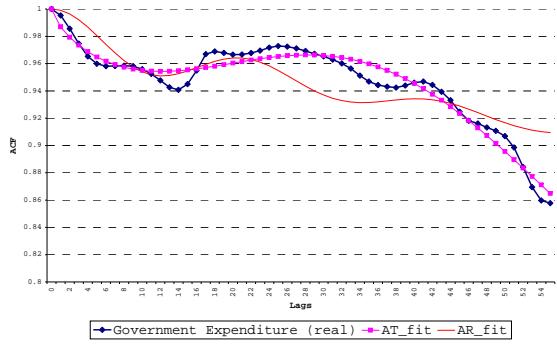
First, we can illustrate this in the time domain in the next graph, where we see that the variables are evolving around a time trend, well within finite variance bounds that don't expand over time (unlike unit-root processes).

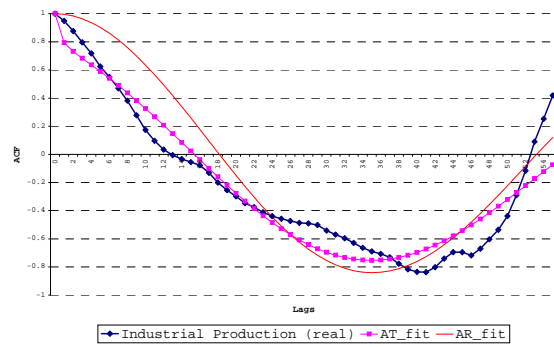
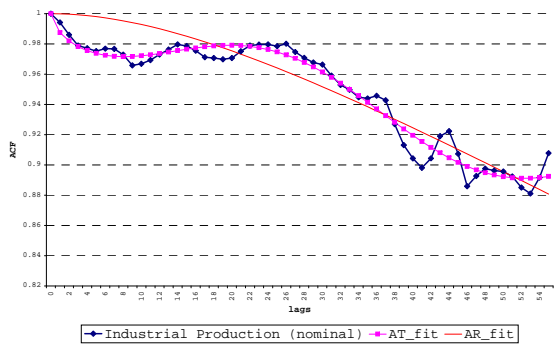
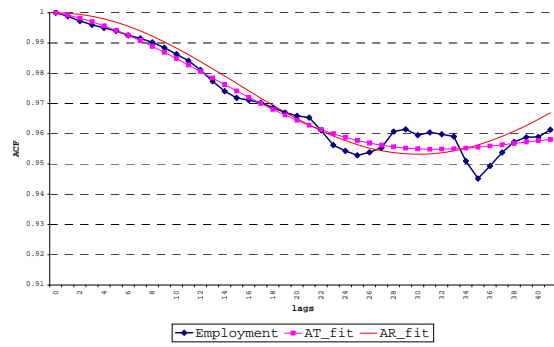
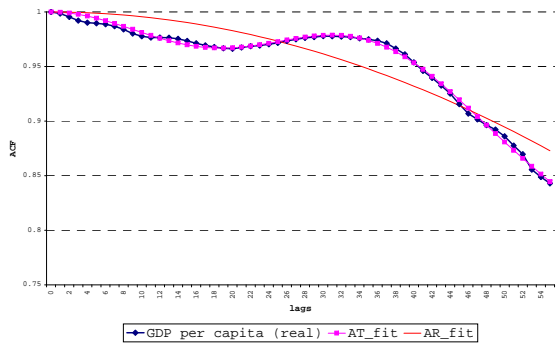
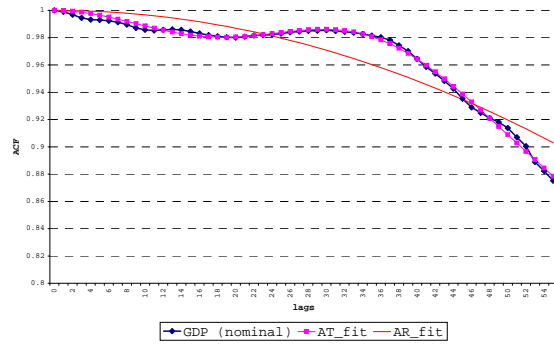
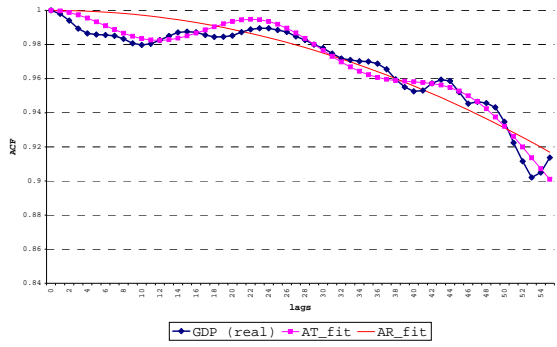


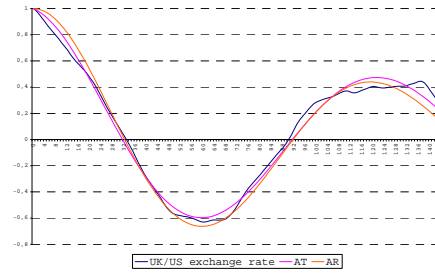
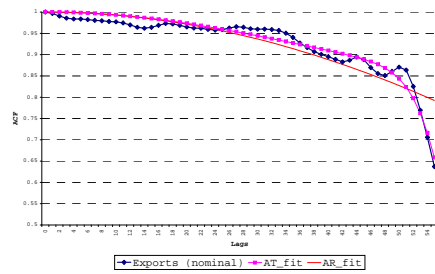
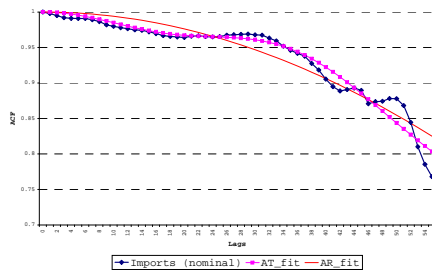
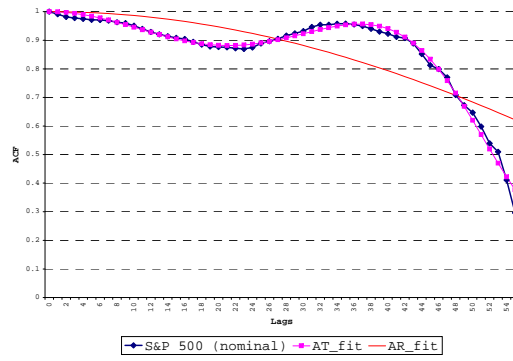
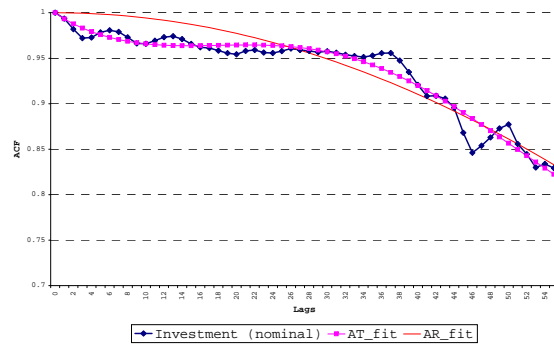
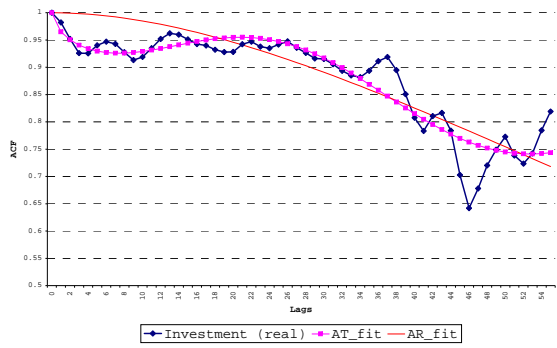
Second, we represent this evolution parametrically and give the following pictures, using ACFs as viewing glasses.

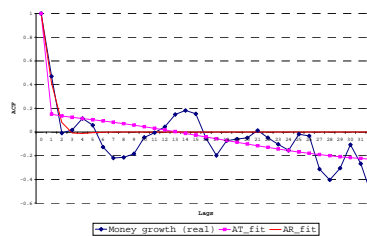
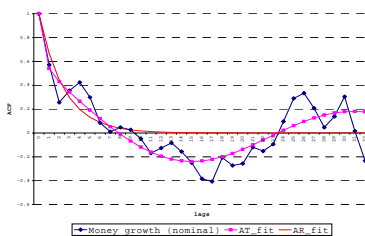
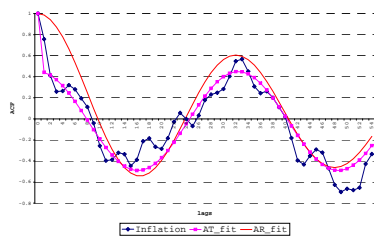
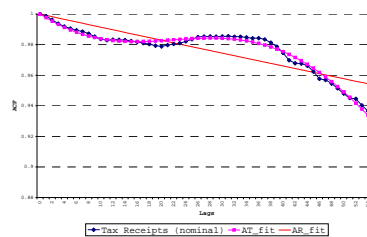
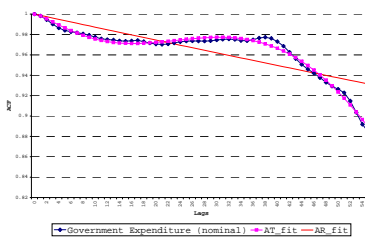
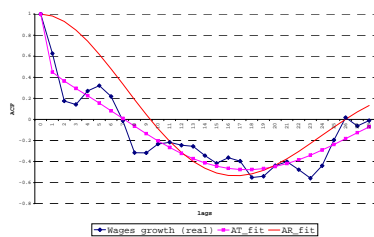
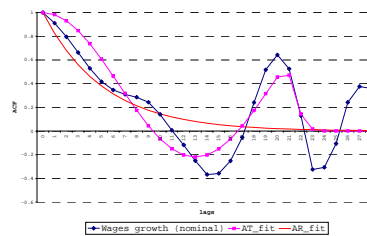
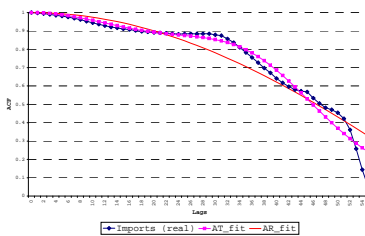
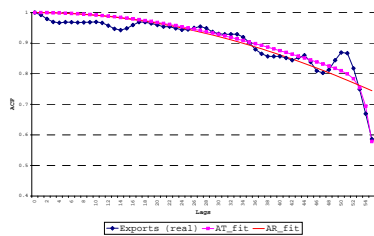
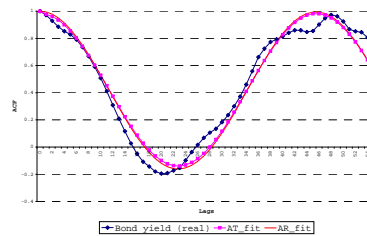
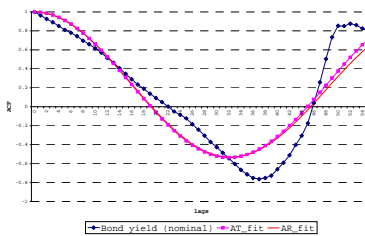
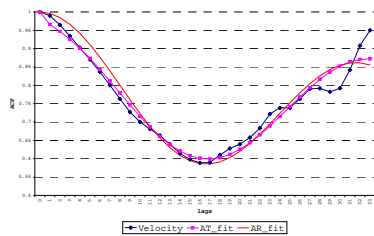
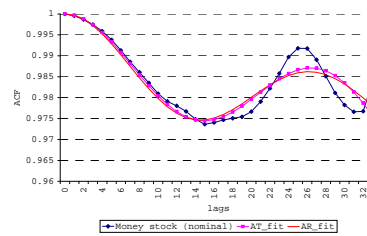
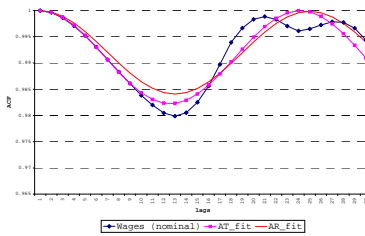
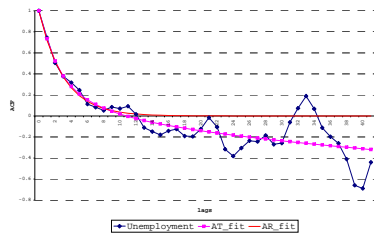
They compare the actual ACF to two possible models by fitting:

1. a formula similar to the one given at the end of the previous section;
2. the best $AR(p)$ model.









The dominance over traditional models remains even when we:

- take structural breaks into account;
- consider different sample periods.

The dominance over traditional models remains even when we:

- take structural breaks into account;
- consider different sample periods.

Standard econometric models ignore the nonlinearities and long-memory dynamics implied by these ACFs, thus erroneous conclusions arise.

The intermediate persistence of the cycles *and* the sudden turning points are completely missed by these standard models.

The dominance over traditional models remains even when we:

- take structural breaks into account;
- consider different sample periods.

Standard econometric models ignore the nonlinearities and long-memory dynamics implied by these ACFs, thus erroneous conclusions arise.

The intermediate persistence of the cycles *and* the sudden turning points are completely missed by these standard models.

Through the ACFs, we can see that:

1. impulses decay very slowly and these variables takes a long time to alter directions;
2. once they do, the change is abrupt and accelerates.

In the context of financial variables, these two features are the (short-run) momentum and (medium-run) cycles that are often noted empirically, but not captured within a single model.

In the context of financial variables, these two features are the (short-run) momentum and (medium-run) cycles that are often noted empirically, but not captured within a single model.

In the context of macroeconomic stabilization, the features imply that, if an intervention takes place (e.g. change in interest rates), it should:

1. occur as soon as possible to give time to the policy to operate;
2. impart a stimulus sufficient to achieve the objective, taking into account the increments due to the persistent dynamics; and
3. revert to a neutral stance well before its objective is achieved, letting the economy ease onto its intended path.

Unfortunately, counterexamples to this course of action still exist.