STYLIZED FACTS OF ECONOMIC METRICS: MODELING REAL AND VIRTUAL WORLDS

OLIVIER HEYNDRICKX

GHENT UNIVERSITY

Department of Physics and Astronomy
Faculty of Sciences

STYLIZED FACTS OF ECONOMIC METRICS: MODELING REAL AND VIRTUAL WORLDS

OLIVIER HEYNDRICKX
PROF. JAN RYCKEBUSCH
UNDER SUPERVISION OF ANDRES BELAZA

GHENT UNIVERSITY

Department of Physics and Astronomy
Faculty of Sciences

I would like to thank prof. Jan Ryckebusch for granting me the opportunity to write a thesis in the exiting new field of econophysics. Furthermore, Andres Belaza proved to be of indescribable help during the process, and together with Corneel Casert made the work enjoyable. In addition, I would like offer my sincere gratitude to *CCP Games* for providing market data of *EVE Online*, and Kevin Hoefman for working as a facilitator and providing valuable insight.
Finally, I wish to thank my parents, who supported me through an unconventional academic career and have always wanted the best for me.

ABSTRACT

The origin of stylized facts, being empirical features found in a wide range of financial markets, remains largely unknown. This thesis investigates four of them, regarding price log-returns: fat-tailed distributions, aggregational Gaussianity, absence of auto-correlation and volatility clustering.

First, the presence of these stylized facts is looked for in prices of a wide range of goods in the real world. Secondly, these markets are compared to those of the video game *EVE Online*. Finally, an agent-based model (ABM) is constructed based on the dynamics of *EVE*'s economy, and its price behavior is examined.

Stylized facts are found in the prices of many real goods, but remain absent in others. The same conclusion holds for the virtual world of *EVE*. The ABM also shows signs of the stylized facts, under the condition that a degree of randomness is added to its implementation.

ABSTRACT (NEDERLANDS)

In financiële markten zijn een aantal empirische vaststellingen in zulke mate prevalent dat ze als “gestyleerde feiten” beschouwd worden. De redenen voor hun aanwezigheid zijn nog grotendeels onbekend. Deze thesis onderzoekt vier van deze feiten die betrekking hebben tot prijsverschillen: het relatief vaker voorkomen van extreme waarden, aggregationele Gaussianiteit, afwezigheid van autocorrelaties, en geclusterde volatiliteit.

Als eerste wordt de prevalentie van deze gestyleerde feiten onderzocht in de prijzen van een diverse groep goederen in de echte wereld. Vervolgens worden deze vergeleken met de markten van de video game *EVE Online*. Tenslotte word een agent-based model (ABM) geconstrueerd, en de markten ervan worden op dezelfde manier verkend.

De gestyleerde feiten blijken aanwezig in de prijzen van veel van de goederen in de echte wereld, maar blijven afwezig in een beperkt aantal. Hetzelfde resultaat werd gevonden voor de goederen van de virtuele wereld. Het ABM toont ook tekens van de gestyleerde feiten, maar enkel indien er voldoende willekeurigheid aan het ABM werd toegevoegd.

PREFACE

What exactly is physics? If the reader's view is that it only relates to the study of matter, energy and force, then it will be difficult for it to encompass the subject of this thesis. However, placed in the larger framework of science, physics can be seen as the way of thought that provides us with methods to explain how the universe behaves. Patterns arising in various aspects of our lives are examined: early thermodynamics studied the behavior of gases in an attempt to increase the efficiency of steam engines, geophysics investigates the dynamics of plate tectonics, and modern chemical physicists look into the quantum mechanical behavior of chemical reactions. Yet, why should we limit the study of these patterns to the realm of the inanimate? Is a cell, the smallest building block of all known life, not subject to the same basic laws of physics that govern a rock? If the answer is yes, then could physics not stand to learn from the complex behavior that arises in it? Biophysics does precisely this: it seeks to find the physical underpinnings of biomolecular phenomena. It does this on all scales: from cells to large lifeforms and populations. Which brings us to the interdisciplinary subject at hand: econophysics. Groups of humans interact which each other, not unlike particles in a fluid, to form a complex system in which regularities can be found. From the interconnectedness between individual agents, collective behavior emerges. Econophysics approaches these economic systems from a physicist's viewpoint, and investigates its underlying mechanisms with a bottom-up approach.

The interdisciplinary aspect of this thesis is part of its appeal, but also presents additional challenges. An understanding of all relevant fields is required, and agent-based modeling in economics is itself so broad that it will be impossible to present a full summary of available research. I will thus have to limit myself to what I believe is essential for giving my model and the subsequent results a proper context. The exclusion of any specific subject by no means signifies that its impact on econophysics was less pronounced– only that I thought it not mandatory for building a minimal framework of understanding.

In chapter 1, I attempt to familiarize the reader with concepts and terms used in economics, and give a brief review of its history of modeling. In chapters 2 and 3, so-called stylized facts are investigated in real-world markets, and the market of an online video game: *EVE Online*. In chapter 4, I present an agent-based model inspired by *EVE Online*, and examine to what extent it replicates these stylized facts.

CONTENTS

1 INTRODUCTION

1.1 AN OVERVIEW OF ECONOMICS

In a world with limited supplies, the management of goods is a fundamental aspect of life. Humans are subject to this as much as ants that collect leaves to use as manure for underground fungi farms, or chimpanzees that trade sex for food. An *economy*, in its broadest sense, encompasses all practices, discourses, and material expressions associated with the production, use and management of resources.

For humans, these economies are contextualized by a larger framework of culture, education, history, social organization, political structure, legal systems, and so on. Many forms of economic systems appeared throughout human history; some examples of this are the inter-band bartering of nomadic hunter-gatherer groups, the need for management of food supplies in emerging agricultural-based settlements, and nation-focused mercantilism during the Renaissance.

As societies advanced, goods like salt and cattle proved to be less cumbersome methods of exchange than those in general bartering. From this, the concept of *money* evolved: anything generally accepted as either a medium of exchange, a measure of value, or a means of payment.

Nowadays, the most common form of economy that moves both money and physical goods around is the *market economy*. In such a system, decisions regarding investment, production and distribution are, under varying degrees of regulation by local governments, based on the interplay between *supply and demand*. The theory of supply and demand assumes people want to allocate their resources in the most efficient way possible. It states that if the price of a good goes up, the demand for that good will go down as less people can/want to afford it, yet its supply will go up as potential profits increase. In markets where the participants are free to compete against each other, the price of goods and services will then settle at a point where the quantity demanded (at the current price) will equal the quantity supplied (at the current price).

The crucial part of this theory lies in the assumed competition of self-interested market participants. As described by Adam Smith in *The Wealth Of Nations* (1776), a baker bakes bread out of self-interest, to earn enough money to feed his family and purchase the goods he wants as effectively as possible. The baker is still halted in raising the price of bread by the following loss of consumers to competing bakers on the market, and so a price equilibrium will be reached. In this way, as if by some *invisible hand of the market*, competing individuals in a market economy will, through self-interested behavior, take part in serving the needs of society as a whole.

In modern capitalist societies, trade is not limited to *physical assets* (e.g. bread), but includes *financial assets*. An example of such a financial asset is a *stock*. Stocks represents a financial investment in a company, and gives the owner of it (the *shareholder*) part ownership of that company. These stocks are traded on market types known as *financial markets* for prices that, like physical assets, are assumed to reflect the law of supply and demand.

All these parts form the foundation of current economies. *Economics*, then, is the field of science that deals with their theories, and with the behavior and interaction of agents in them. An interesting modern discussion regarding economic theory lies in the random vs. non-random nature of prices. Traced back as far as Regnault (1863) and more recently Kendall [KH53], it was noted that stock prices appeared to follow a random walk (Fig. 1). A random walk is one in which future steps or directions cannot be predicted on the basis of past actions. Applied to the stock market, it would mean that short-run changes in stock prices cannot be predicted. I will introduce the discussion through two books: *A Random Walk Down Wall Street* and a *A Non-Random Walk Down Wall Street*.

1. Burton Malkiel's *A Random Walk Down Wall Street* (1973)

   In 1965, Samuelson [Sam65] postulated that in a market that is informationally efficient– i.e. one where prices incorporate all available information– prices change randomly. This paradoxically-sounding statement assumes that individual agents seek to increase their profit at all times by exploiting patterns in prices. If the price of a good is higher than average, the agents will sell. If it is lower, they will buy. In doing so, they change the price patterns: selling lowers the price, and buying increases the price. In other words, the market equalizes, and the information about the patterns that agents possess becomes reflected in the

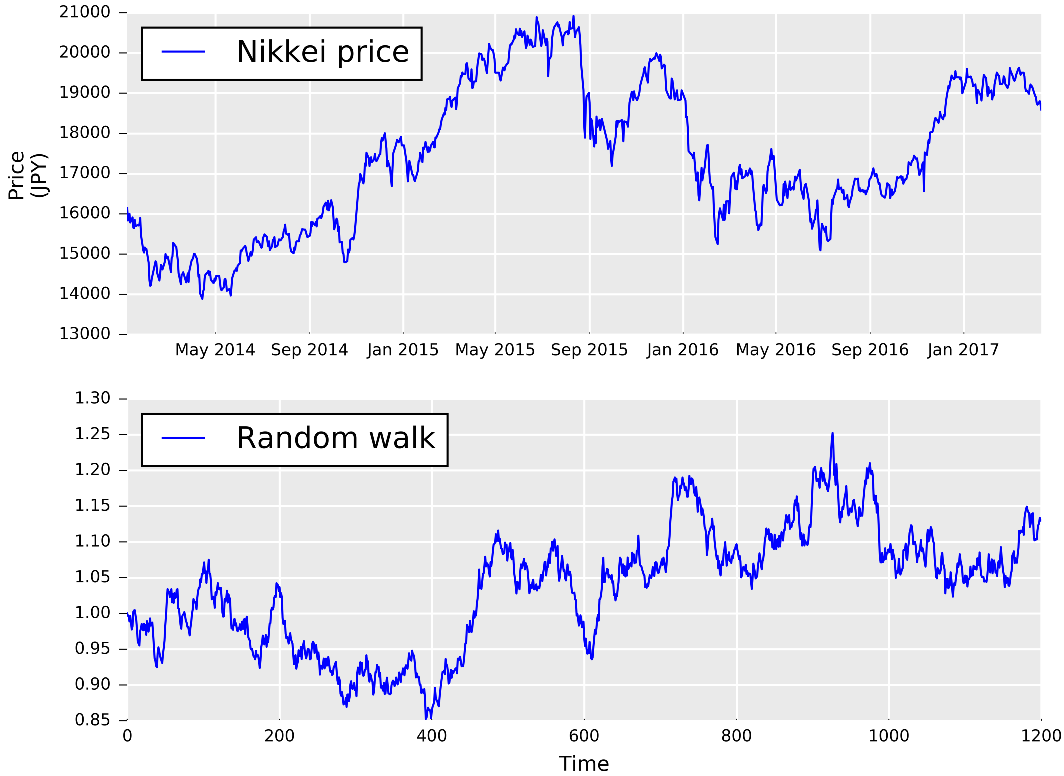

Figure 1: A comparison of the price of of the Nikkei stock market index on the Tokyo Stock Exchange (top) and a random walk (bottom).

market prices. If this incorporation of information happens instantly, then prices can be assumed to reflect all available information at all times. What causes price changes then? Only external factors remain, and as these are unpredictable and numerous, they are best modeled as random processes. As a result, the faster the market absorbs information through agent actions, the more *efficient* the market, and the more random its prices will appear. Fama [Fam+69; Fam70] worked in a similar direction as Samuelson and postulated the *Efficient-Market Hypothesis* (EMH).

Malkiel's 1973 book *A Random Walk Down Wall Street* [Mal73] gives a summary of this side of the discussion. According to the EMH or similar theories, the stock market adjusts so quickly and perfectly to new information that any person buying at current prices can do just as well as professionals.

If random walk theory were "taken to its logical extreme, it means that a blindfolded monkey throwing darts at a newspaper's financial pages could select a portfolio that would do just as well as one carefully selected by experts" [Mal73]. Malkiel didn't deny that price charts seemed to display some non-random patterns, but argued that after correcting for long-term inflation stock charts would be indistinguishable from a random walk.

2. Andrew Lo & Craig MacKinlay's *A Non-Random Walk Down Wall Street* (1999)

   The implied unpredictability of the market by the EMH sparked much debate. Countering Malkiel, Lo & MacKinlay published a collection of research papers arguing for a degree of predictability to financial markets as *A Non-Random Walk Down Wall Street* in 1999 [LM99]. Evidence against the EMH was found in so-called *anomalies*– empirical results that are inconsistent with maintained theories of asset-pricing behavior. Some of these anomalies, as listed by William Schzert [Sch02], are

   a) The size effect [Ban70; al15], a tendency of smaller companies to outperform larger ones. A possible explanation lies in different growth opportunities.

   b) The weekend effect [Fre80], a tendency of stock returns on Mondays to often be significantly lower than those of the previous working day. A possible explanation is that companies are more likely to release bad news on Fridays.

   c) The value effect [Bas77], a tendency of companies with high earnings-to-price ratios (E/P, relating share per earnings to share price) to have positive abnormal returns over the long term. A possible explanation is a general over-appreciation of growth stocks (low E/P ratio) over value stocks (high E/P ratio).

   Schzert mentioned that many of the well-known anomalies in the financial literature do not hold up in different sample periods [Sch02]. The size effect and the value effect, and to a lesser extent the weekend effect, seem to have somewhat disappeared after the papers that highlighted them were published. This could suggest that these anomalies were only statistical fluctuations, or that traders exploiting the anomalies caused them to disappear. These explanations are in favor of the EMH; a further discussion can be found in [Mal03].

There is no current consensus on the validity of the EMH. However, much of the opposition does not attack the randomness of markets, but merely the explanation the EMH gives for it. Voit [Voi01] called a random walk property of prices "the standard model of finance", and prices are assumed to follow a geometric Brownian motion in the central tool for pricing options: the Black-Scholes-Merton (BSM) model [BS73; Mer73].

With the EMH I conclude the overview of modern economics. In the next sections, I will introduce the research on economic systems, followed by a description of modeling in *econophysics*. Finally, the value of research on video games is discussed.

1.2 EARLY EXAMPLES OF ECONOMIC MODELING

1.2.1 *Lorenz, Gini, and representations of inequality*

Throughout history, inequality has been present in the management and allocation of goods, and distributions of income and wealth have been studied. In 1905, Lorenz [Lor05] criticized some of the methods to represent changing income distributions. These methods would only take either changes in income or changes in population into account. He proposed a graphical representation where the cumulative percent of the population from richest to poorest is plotted against the total income held by these percentages. Lorenz' original image can be seen in Fig 2. On such a *Lorenz curve*, a perfectly equal society will show up as a straight line, as each increment in population fraction gives an equally large increment in income. Increasing inequality will present itself as a Lorenz curve that deviates from this line. For wealth distributions, similar curves can be constructed.

Although Lorenz curves allow one to visually represent income distributions, they give no direct metric to discriminate between them. Working on numerical representations of inequality, Gini proposed several measures between 1908 and 1914. A summary of his efforts can be found in [CV12]. They culminated in the measure that is named after him: the *Gini coëfficient*, or simply the *Gini*. If A is the area between the line of equality and a

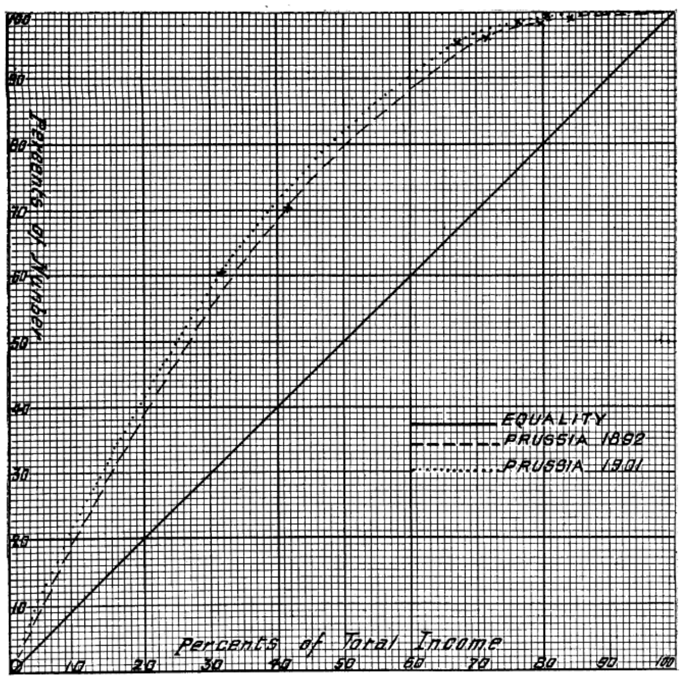

Figure 2: Lorenz' original depiction of a Lorenz curve, taken from [Lor05]. Modern Lorenz curves commonly their axes reversed.

Lorenz curve (see Fig. 3), and $A + B$ is the total area under the line of equality, then the Gini coëfficient is defined as:

$$\text{Gini} = \frac{A}{A + B}. \tag{1}$$

Gini coëfficients lie between zero (perfect equality) and one (all income gained by one person). As with Lorenz curves, the Gini is often calculated from wealth distributions and used to represent wealth inequality.

1.2.2 *Pareto and Pareto distributions*

In 1897, Pareto [Par96] noted some recurring features in the distributions of income. For many populations, plotting the logarithm of the number of incomes above a level, or $P(x)$, against the logarithm of that income level x produce a straight line. Such a distribution $P(x)$ is known as a *complimentary cumulative distribution function* (CCDF):

$$P(x) = \int_x^\infty p(x')\mathrm{d}x'. \tag{2}$$

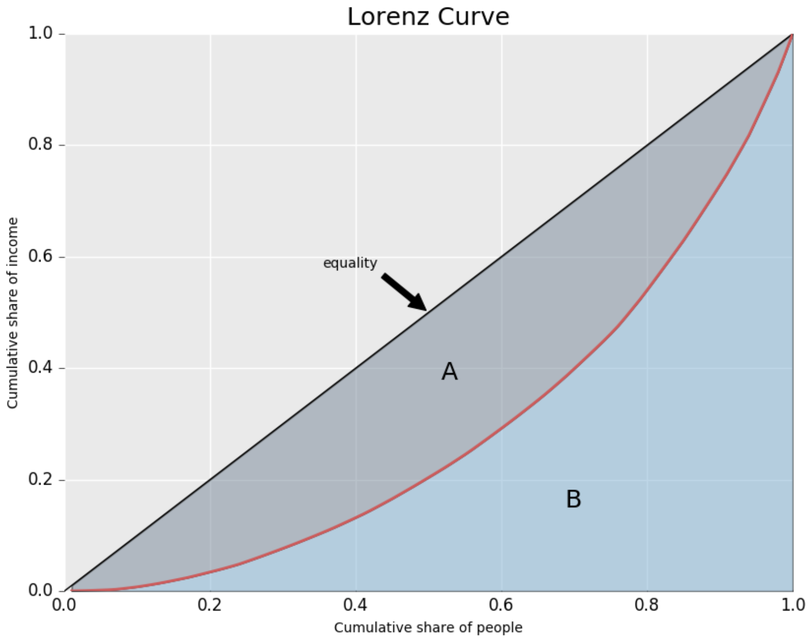

Figure 3: Graphical calculation of the Gini, with the red line indicating a sample Lorenz curve. The Gini is then calculated according to Eq. 1.

CCDFs are closely related to Lorenz curves L. If the *cumulative distribution function* $F(x) = 1 - P(x)$ with F^{-1} its inverse, then the Lorenz curve $L(x)$ can be written as

$$L(x) = \frac{\int_0^x F^{-1}(u)\mathrm{d}u}{\int_0^1 F^{-1}(u)\mathrm{d}u}. \tag{3}$$

A distribution that shows up as a straight line on log-log scales is known as a *power-law* (or in specific cases *Pareto distribution*) with coefficient α, and can be written as $\ln p(x)$ = -$\alpha \ln x$ + c, or

$$p(x) = Cx^{-\alpha} \tag{4}$$

with C = e^c and $x > 0$.

Pareto noticed power-law behavior not in income distributions, but in their CCDF. An interesting property of power-laws can however be highlighted here: for a power-law distribution, the CCDF is a power-law with coefficient $\alpha - 1$:

$$P(x) = C\int_x^\infty x'^{-\alpha}\mathrm{d}x' = \frac{C}{\alpha - 1}x^{-(\alpha-1)}. \tag{5}$$

Pareto's observations suggested a power-law with α =2.5, over its entire range (a *strong* Pareto distribution). He argued that the income distribution curve "varies very little in space and time; different peoples and different eras yield very similar curves. There is a remarkable stability in the form of this curve" [Par06]. Later findings (such as those by Gibrat, see next section) suggested that the power-law did not apply to the entire income distribution, but only the tail (a *weak* Pareto distribution).

The discovery of power-law behavior had lasting impact on other fields, including physics. Power-laws are e.g. closely related to scale invariance, a concept important in phase transitions: scaling the argument x in Eq. 4 only causes a proportionate scaling of Eq. 4

$$f(ax) = C(ax)^{-\alpha} \propto f(x). \tag{6}$$

An excellent discussion of the many interesting mathematical properties of power laws can be found in a review by Newman [New05].

1.2.3 *Gibrat's log-normal distribution*

In 1931, Gibrat [Gib31] suggested that Pareto's power-law income distributions only holds for the richer tail. He reached this conclusion by modeling income $I(t)$ at time t as an accumulation of random multiplicative changes $(1+\epsilon_i)$ (the law of proportionate effect):

$$I(t) = I(0)(1+\epsilon_1)(1+\epsilon_2)...(1+\epsilon_t), \tag{7}$$

recognizable as a Markov chain model. If the time period is small, ϵ_i can be assumed small, and approximating $\ln(1+t) \approx t$ for small t, Eq. 7 can be rewritten as

$$\ln \frac{I(t)}{I(0)} \approx \epsilon_1 + \epsilon_2 + ... + \epsilon_t. \tag{8}$$

For increments ϵ_i that are independent and normally distributed with mean μ and variance σ^2, the Central Limit Theorem states that $\ln \frac{I(t)}{I(0)}$ will approximate a normal (Gaussian) distribution with mean μt and variance $\sigma^2 t$. Then, $I(t)$ will follow a log-normal distribution [Sut97]

$$p(t) \sim \frac{1}{\sqrt{2\pi}\sigma t} e^{-\dfrac{(\ln t - \ln t_0)^2}{2\sigma^2}}. \tag{9}$$

On a log-log scale, the upper end of a log-normally distributed variable will appear as a straight line, similar to a power-law-distributed variable. To see this, we take the logarithm of Eq. 9

$$\ln p(t) \sim -\ln t - \ln\sqrt{2\pi}\sigma - \frac{(\ln t - \ln t_0)^2}{2\sigma^2}. \tag{10}$$

If σ is sufficiently large, the last term of Eq. 10 will be small for a large range of t values, giving a linear appearance for those ranges on a log-log scale. Thus, according to Gibrat, power-laws only correctly model the tail of income distributions, and the actual distribution more closely follows a log-normal.

However, Champernowne [Cha53] and Wold & Whittle [WW57] again found power laws. Is income power-law distributed, or does it follow a log-normal? What does this mean about the underlying generating process? Is there a different process going on in the tail of the distributions? Regardless of the answers, the power-law behavior of the tail of income distributions indicates that extreme cases, or *outliers*, are more frequently present than what could be explained by a Gaussian distribution. The power-law exponent α then gives an estimate of the probabilities of these extreme events.

In economics, the influence of power laws reaches beyond income distributions. For example, the size of firms appears to follow a power-law [Axt01], and an increased probability of outliers is also seen in the subject of the next section: changes in price.

1.3 THE LEGACY OF THE POWER LAW, AND STYLIZED FACTS

Mandelbrot [Man63] showed in 1963 that the distribution of differences in prices at high frequencies (daily data) for various assets was not Gaussian, but fat-tailed. As can be seen in Fig. 4, which shows the probability density function for daily price differences of cotton, a Gaussian fit describes the bulk of the data properly but underestimates the probability of outliers. If financial risk is to be properly estimated, these outliers cannot be ignored.

This fat-tailed nature of price differences is now known as a *stylized fact*, a term introduced by Nicholas Kaldor in 1957. Kaldor worked on modeling the process of growth in capitalist economies, and noted that such a model "must account for the

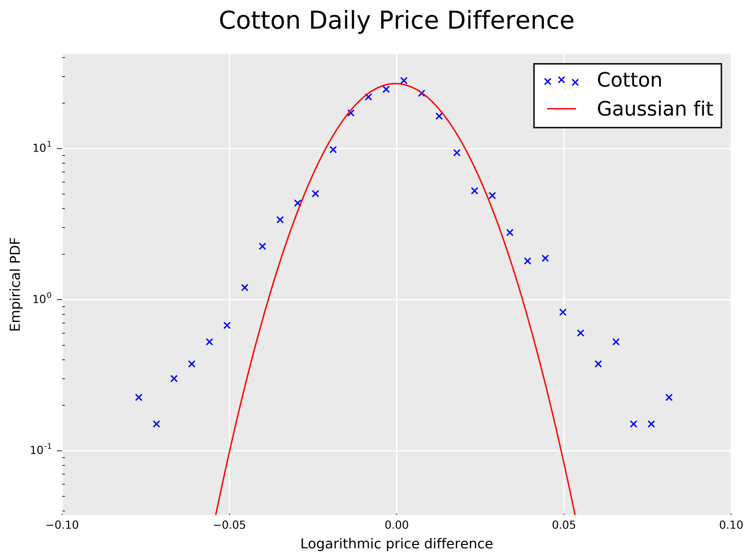

Figure 4: Empirical PDF of cotton daily futures price differences on the Chicago Mercantile Exchange. Data between 1972-11-16 and 2017-05-15, retrieved from [Qua17b].

remarkable historical constancies revealed by recent empirical investigations". These empirical constancies should, according to Kaldor, be the relevant starting point for economic modeling:

> "(...) Since facts, as recorded by statisticians, are always subject to numerous snags and qualifications, (...) the theorist, in my view, should be free to start off with a 'stylized' view of the facts – i.e. concentrate on broad tendencies, ignoring individual detail (...) and construct a hypothesis that could account for these 'stylized' facts" [Hag61].

In other words, stylized facts are robust empirical features that are seen throughout different markets and asset classes. Since Mandelbrot and Kaldor, many more of these stylized facts have been uncovered. The ones discussed by Chakraborti [Ani11] are as follows:

1. *Fat-tailed empirical distribution of returns*

   Economists often work with price differentials rather than prices. Defining the (simple) *return* of an asset with price p over a time period τ as

$$R_\tau(t) = \frac{p(t+\tau) - p(t)}{p(t)}, \tag{11}$$

and the *log-return* as

$$r_\tau(t) = \ln(p(t+\tau)) - \ln(p(t)), \tag{12}$$

it is clear that these are both dimensionless quantities appropriate to quantify relative price changes. For small changes, the simple and log-return are approximately equal. Log-returns tend to be the preferred quantity for technical analysis.

Mandelbrot [Man63] was among the first to note that the empirical distribution of financial returns was fat-tailed. Since then, many empirical studies (e.g. [Gop+98]) confirmed this.

2. *Aggregational Gaussianity*

   The fat tails of the first stylized fact appear when log-returns are calculated over short time periods, e.g. τ =1 day. However, it has been observed that as one increases τ, this feature becomes less pronounced, and the distribution approaches a Gaussian [KSY06]: so-called aggregational Gaussianity.

3. *Absence of auto-correlation of returns*

   The EMH assumes that price changes behave as a random walk. This would indicate that the autocorrelation of the log-returns, defined as $\rho(T) \sim \langle r_\tau(t+T) r_\tau(t) \rangle$ for a lag T, should be zero. For many markets, this appears to be the case [Pag96; Con01]: a positive change in price does not indicate that next change is more likely to be positive.

4. *Volatility clustering*

   The final stylized fact states that if one looks at absolute (or squared) returns, the auto-correlation function is significantly larger than zero and decays slowly. This means, in Mandelbrot's words, that "large changes tend to be followed by large changes of either sign, and small changes tend to be followed by small changes". Volatility, or the degree of variation of these price changes, is then clustered in time.

Whereas the third stylized fact can be explained with a random walk, the others don't fit the geometric Brownian motion that the BSM model is based on. It involves normally distributed returns, conflicting with the empirical evidence showing e.g. fat tails. As the stylized facts appear across a wide variety of markets, they could perhaps be caused by universal underlying principles. This is where econophysics enters the stage– after all, physicists are in the business of laws and invariances [Ric08].

1.4 THE EMERGENCE OF ECONOPHYSICS

As stated by Huber & Sornettte [HS16], there have been various collisions between economics and physics throughout their evolution. Starting with Adam Smith's market dynamics, driven by forces inspired by Newtonian mechanics, a notion of economic laws paralleling physical laws became entrenched in economics.

In particular, thermodynamics proved especially attractive, and the concept of an equilibrium state of economic activity was introduced by Marshall and Edgeworth through various works in the late 19th century. Similar to how Maxwell and Boltzmann abstracted the heterogeneity of particles and their microstates into a thermodynamic description, economic equilibrium theory reduced the heterogeneity of economic agents to a single representative agent. Boltzmann himself, in his arguments for using ratios and averages in kinetic gas theory, drew analogies between statistical physics and social statistics such as economics [Por86].

Over time, economics and physics became more rigid in their specializations, and a mainstream movement of the application of statistical physics to economics did not re-appear until the last few decades.

Driven by attempts to properly describe the stylized fact 'anomalies', the field of *econophysics* emerged in the 1990s– a synthesis of economics and physics, coined by Eugene Stanley. Similar to fields such as biophysics, which studies living organisms through their underlying obedience to the laws of physics, econophysics seeks to explain economic phenomena by applying tools and concepts from statistical and theoretical physics to them. From another direction, it provides the insight that dynamics of financial systems are perhaps best understood as emergent properties of a complex adaptive system.

The following sections will deal with two distinct types of models in econophysics. First statistical modeling will be discussed, which subdues individual characteristics of economic agents. Afterwards, agent-based modeling is introduced, aiming to integrate the learning and adaptive features of market participants.

1.5 DIRECT APPLICATIONS OF STATISTICAL MECHANICS

Statistical physics is, in a broad sense, a framework that allows systems consisting of many interdependent, interacting parts to be rigorously analyzed. According to econophysics, financial markets are a part of this class of systems. Their internal microscopic structure consists of 'economic particles': investors, traders, consumers, and so on. In the systems traditional statistical physics deals with, properties can be found that are invariant with respect to transformation of scale: *scaling laws*. These laws are viewed as emergent properties generated by the interactions of the microscopic subunits– via collective behavior. Perhaps, then, a similar reasoning could be applied to financial markets, and the previously mentioned stylized facts could be explained through it.

We have already seen an example of scale invariance earlier in this introduction: the power-law income distributions. Models such as those by Gibrat and Champernowne used a stochastic process to describe individual income or wealth, and can be seen as a one-body approach– income and wealth fluctuations are considered independently for each agent. One of the first truly novel directions of econophysics was the addition of two-body interactions, inspired by collisions in gas theory. After a short review on relevant aspects of statistical mechanics (see [YR09]), this will be introduced with a model by Drăgulescu & Yakovenko [DY00].

1.5.1 *Review of the Boltzmann-Gibbs distribution*

The Boltzmann-Gibbs distribution remains a pillar of modern statistical mechanics. Its derivation is as follows: consider N particles with total energy E. The energy axis can be divided into small intervals of size $\Delta\epsilon$, with N_k the number of particles with energy between ϵ_k and $\epsilon_k + \Delta\epsilon$. The ratio N_k/N then gives the probability of a particle to have energy ϵ_k, which is noted as P_k.

The multiplicity W, or the number of ways in which a a certain system state can be produced, is given by

$$W = \frac{N!}{N_1!N_2!N_3!\dots}. \tag{13}$$

The natural logarithm of the multiplicity is called the entropy $S = k \ln W$, where k is the Boltzmann constant. Using Stirling's approximation in the limit of large numbers, this can be approximated as

$$\frac{S}{N} \approx -\sum_k \frac{N_k}{N} \ln\left(\frac{N_k}{N}\right) = -\sum_k P_k \ln P_k. \tag{14}$$

It is then possible to find the highest multiplicity state, being the state with the highest entropy, using the method of Lagrange multipliers with the constraint that the total energy of the system $E = \sum_k N_k \epsilon_k$ is constant. The result of this is the exponential Boltzman-Gibbs distribution, stating that the probability $P(\epsilon)$ of finding a physical (sub)system in a state with energy ϵ is given by

$$P(\epsilon) = c\, e^{-\epsilon/kT} \tag{15}$$

with c a normalizing constant and T the temperature. The expectation value of any physical variable x can then be obtained as

$$\langle x \rangle = \frac{\sum_k x_k e^{-\epsilon_k/kT}}{\sum_k e^{-\epsilon_k/kT}} \tag{16}$$

where the sum is taken over all states of the system. Temperature is related to the average energy per particle: $T \sim \langle \epsilon \rangle$.

1.5.2 *Statistical mechanics of money*

The above derivation is general, and its main constraint is the presence of a conserved property. It is not unreasonable that the Boltzmann-Gibbs distribution would apply to other statistical systems, such as an economy consisting of many interacting participants, as long as a conserved property is present. Drăgulescu & Yakovenko [DY00] suggested that for economies, this conserved quantity is money m. Economic participants are by law not permitted to create or destroy money, and the total amount remains the same after every transaction between participants i and j:

$$\begin{cases} m_i' = m_i - \Delta m, \\ m_j' = m_j + \Delta m. \end{cases} \tag{17}$$

This local conservation of money allowed them to make an analogy with the energy transfer between molecular collisions in gas,

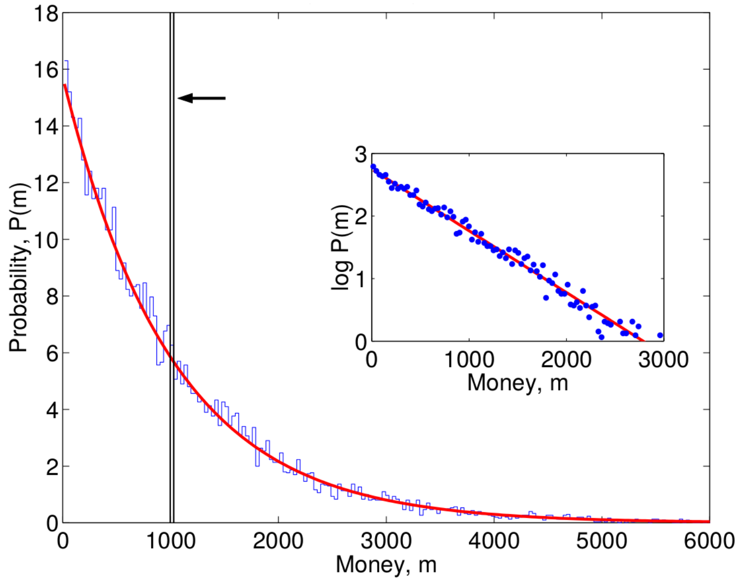

Figure 5: Probability distribution of money $P(m)$ obtained in simulations by Drăgulescu & Yakovenko, adapted from [DY00]. Red curves are Boltzmann-Gibbs distributions. The vertical line indicates the initial distribution of money.

where energy is conserved. Central banks or governments can inject money into an economy, but this is analogous to an influx of energy from external sources such as the Earth receiving energy from the Sun. As long as the rate of money influx is slow compared to relaxation processes in the economy, the system can be considered in a quasi-stationary statistical equilibrium.

Drăgulescu & Yakovenko go on to argue that the stationary distribution of money $P(m)$ should then be of the form of Eq. 15:

$$P(m) = c\, e^{-m/kT_m} \tag{18}$$

with T_m the 'money temperature', defined as the average amount of money per participant: $T = \langle m \rangle$. To verify this, they performed simulations starting from uniform distributions of money with varying rules for the size of money transfer Δm, and found the Boltzmann-Gibbs distribution in all of them. An example result is seen in Fig 5.

When multiplicative money exchanges are considered rather than additive, the probability distribution deviates from the Boltzmann-

Gibbs distribution for lower values, and appears closer (but not exactly equal) to a Gamma distribution:

$$P(m) = cm^{\beta} e^{-m/kT}, \tag{19}$$

differing from Eq. 15 by the power-law prefactor m^{β}. Examples of this can be found in [IKR98; CC00]. Chatterjee [CKM04] added a randomly distributed saving factor, and found power-law tails originating from the participants who hoard money and do not give it back.

For direct comparison with empirical data, the distribution of money in the real world would need to be known. This could be approximated from balances on bank accounts– no such studies appear to be available at the time. As some interplay between theory and empirical data is desired, I will shift to more verifiable regions: distributions of wealth and income.

1.5.3 *Statistical mechanics of wealth and income*

Data on the distribution of wealth is rather limited. In the real world, income is routinely reported by individuals to the government due to their relation to taxation methods, but wealth rarely is. The available data on wealth tends to come from countries where all assets must be reported in the event of a death for the purpose of inheritance tax. This gives us a wealth distribution of dead people, which through adjustment procedures based on age, gender, and so on, can be used to infer a wealth distribution of the entire population.

In statistical physics, it is known that identical molecules in a gas can spontaneously develop a widely unequal distribution of energies as a result of random energy transfers in molecular collisions. By analogy, very unequal wealth distributions might spontaneously develop in an economic system as a result of random interactions between economic participants, which could explain economic inequality [Yak12].

Wealth w has various definitions, but can be seen as the sum of a person's money and the monetary value of all his assets. An approximation of these monetary values are the prices p. The wealth of a participant is then

$$w_i = m_i + \sum_k p_k \tag{20}$$

where the sum is taken over all assets. Total wealth $W = \sum w$ is generally not conserved as prices p_k vary over time– an important difference from the money transfers in the previous section.

Wealth of a participant can change without any transactions, and a transaction does not change the wealth of a participant as he gains as much in value as he loses in money. Thus, the redistribution of wealth in these models is directly related to price fluctuations.

Silver, Slud & Takamoto [SST02] implemented a version of such a model and found, analytically and through simulations, that the stationary distribution of wealth $P(w)$ was best described by a Gamma distribution. Chatterjee & Chakrabarti [CC06] found a similar Gamma distribution shape for fixed saving propensities. Several authors have proposed models where wealth evolves over time (e.g. [BM00]), giving rise to power-law tails. This could suggest that the presence of a power-law is a nonequilibrium effect that requires constant growth or inflation of the economy, and disappears for a closed system with conservation laws.

Regarding income, data is more easily available. Silva and Yakovenko [SY05] examined U.S. income data between 1983-2001, finding two distinct regions as seen in Fig. 6. The majority of the population could be described by the Boltzmann-Gibbs distribution, whereas the top few percent follow a power-law. They go on to describe this two-part structure as 'thermal' and 'superthermal', and relate it to thermal equilibrium in statistical systems. A discussion of further modeling of these income distributions can be found in [YR09].

In summary, approaching economics from a statistical mechanics viewpoint has provided valuable insights. The stylized facts however seem to be harder to reproduce using generalized particle behavior. The next section will introduce a different approach, taking the heterogeneity of market participants into account.

1.6 AGENT-BASED MODELING (ABM)

In 1953, Champernowne [Cha53] said that

> "The forces determining the distribution of incomes in any community are so varied and complex, and interact and fluctuate so continuously, that any theoretical model must either be unrealistically simplified or hopelessly complicated."

As income (and wealth) distributions arise from the actions and interactions of individual people, any perfect description of such a distribution would need to take these interactions into account.

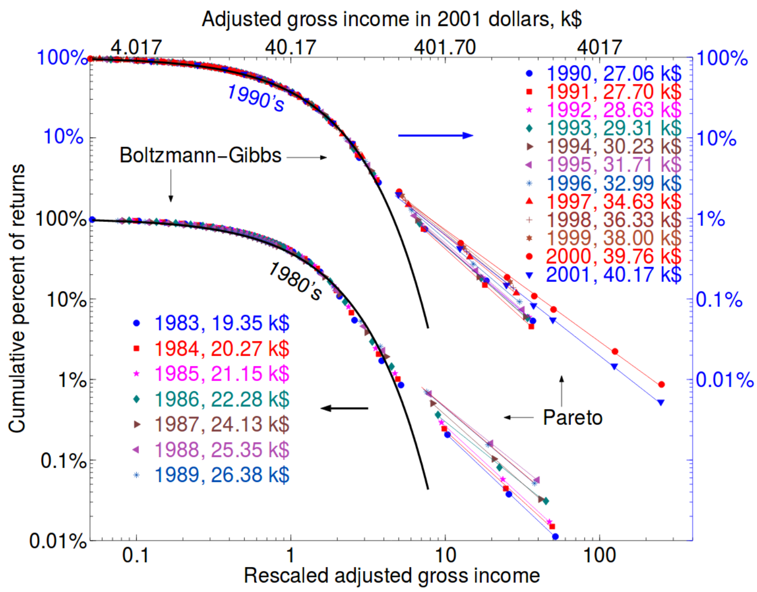

Figure 6: CCDF of annual personal U.S. income, taken from [SY05]. Two regions are visible: one following a Boltzmann-Gibbs distribution, the other following a Pareto distribution (power-law).

This is not easily done using theoretical stochastic models, and *Agent-Based Modeling* (ABM) attempts to address some of these limitations.

1.6.1 *Complexity and complex systems*

ABM is often mentioned with *complex(ity)*: a term that proves to be somewhat hard to describe. Whatever complexity may be, *complex systems* are supposed to have it, so the question can be redirected to what a complex system is. According to Simon [Sim96],

> "(...) in such systems the whole is more than the sum of the parts in the weak but important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is no trivial matter to infer the properties as a whole."

Dean Rickles [Ric08] extracts a triplet of characteristics for a complex system:

1. The system is made up of many subunits.

2. These subunits are, at least some of the time, interdependent.

3. The interaction between these subunits is, at least some of the time, nonlinear.

with additionally

4. (*adaptive* complex system) The subunits modify their properties and behavior with respect to a changing environment, generating new systemic properties that reflect the changed environment.

5. (*self-organizing* adaptive complex system) The subunits modify their properties and behavior with respect to those of of the system they form– a 'downward causation' exists from the systemic properties to the subunits' properties.

At first glance, economic systems are eligible to be categorized as a complex system: multiple agents of different types (traders, producers, firms, etc.) compete (interact) with each other and in doing so generate properties that cannot be explained by the parts themselves (stylized facts). These properties, arising in the system but not possessed by the individual members, are *emergent properties*. Explaining the economic stylized facts by labeling them as emergent properties in a complex system is, under the Rickles's characteristics, reasonable.

However, some doubt still exists on whether economic systems are *truly* complex. Durlauf [Dur05] argues that there is not enough evidence to reject alternative approaches, and says that the scaling law research is consistent with complex system models but the evidence is far from decisive and is amenable to alternative interpretations. Pisarenko & Sornette [PS06] show that a power-law model of prices at best provides an approximation, and no complexity can directly be inferred from it. Financial markets could still be complex systems, but such claims should be treated with scrutiny. An agent-based modeling approach could provide a way to investigate this conjectured complexity.

1.6.2 *Agent-based models*

How, then, does ABM differ from more classical approaches? Unlike particle systems like idealized gases, economic systems consist of participants that are extremely heterogeneous. These participants can differ in behavioral rules, view of the external world, decision-making, extent of memory, and so on. An *agent* tries to capture this diversity, and an agent-based model refers to a model in which the dynamic processes of agent interaction are simulated over time.

Similar to 'complexity', the term 'agent' tends to be hard to define. According to Macal [MN09], the definition of an agent can include the following:

- An agent is an identifiable, discrete, or modular, individual with a set of characteristics and rules governing its behaviors and decision-making capability. Agents are self-contained. The discreteness requirement implies that an agent has a boundary and one can easily determine whether something is part of an agent, is not part of an agent, or is a shared characteristic.

- An agent is autonomous and self-directed. An agent can function independently in its environment and in its interactions with other agents for the limited range of situations that are of interest.

- An agent is social, interacting with other agents. Agents have protocols for interaction with others, e.g. for communication. Agents have the ability to recognize and distinguish the traits of other agents.

- An agent is situated, living in an external environment wherein the agent interacts with other agents.

- An agent is goal-directed, having goals to achieve (not necessarily objectives to maximize) with respect to its behaviors. This allows an agent to compare the outcome of its behavior to the goals it is trying to achieve.

- An agent is flexible, having the ability to learn and adapt its behaviors based on experience. This requires some form of memory. An agent may have rules that modify its behavior.

These properties may not all be present, depending on the specific purpose of the agent-based model. He labels those agents *proto-agents*: "agents who miss one or more of the characteristics noted above but to which the characteristics can easily be added without modification to the structure of the model".

1.6.3 *Challenges with ABM*

Adding complicated behavioral rules may lead to a better prediction of the stylized facts. However, it is important for agent-based models to not just replicate features of real markets, but also to show which aspects of the model may have caused them. Overly complicated models may make it difficult to determine which underlying mechanisms cause the stylized facts to appear. In addition, many ABMs only replicate the stylized facts in a very specific and limited region of the model's parameters– e.g. the number of agents [CPZ01]. Are these limitations of the models, or do real market dynamics also evolve to the specific region which generates the stylized facts?

General issues regarding ABM were summarized by Windrum, Fagiolo & Alessio Moneta [WFM07]. They listed four key problems:

1. The neoclassical community has developed a core set of theoretical models and applied these to a range of research areas; the ABM community has not.

2. ABMs have different theoretical content and seek to explain widely varying phenomena, with little in-depth research being done to compare and evaluate their relative explanatory performance.

3. No standard techniques for constructing and analyzing ABMs exist.

4. The validation of ABMs involves strong assumptions regarding unknown processes that generate empirical data.

Amilon [Ami08] provided an example of these problems. Using maximum likelihood techniques, he found that the presence of the stylized facts in some ABMs was highly dependent on how noise was implemented in agent behavior. This indicates that certain model assumptions could have significant impact on their ability to reproduce empirical data. Windrum, Fagiolo & Moneta conclude with a call for generalized protocols regarding analysis and parameter calibration in order to address these issues. However, despite the possible flaws, the complex system approach of ABM in econophysics can still provide valuable insights.

This concludes the introduction on economics and econophysics. Much of this thesis will deal with the examination of stylized facts in various markets, and the creation of an ABM. Some of the markets examined will come from virtual rather than real worlds, and for this reason, a final introductory discussion regarding video games is presented in the next section.

1.7 VIDEO GAMES AS A RESEARCH LABORATORY

Games played on electronic devices, or *video games*, are an important segment of the modern entertainment industry. For example, in 2016, *League of Legends* logged more than 100 million players per month [Kol13], and the total revenue of MMOGs (massively multiplayer online game) exceeded US$20 billion [Sup17]. A subtype of the MMOG is the *MMORPG*, or massive multiplayer online role-playing game. These are internet-based games which are played by a large number of players at the same time, all accessing the same *virtual world*. Players engage in numerous activities such as moving around, communicating, purchasing/producing/consuming goods, and fighting one another. These virtual worlds can be extremely complex, with some of them having more inhabitants than small countries [Cas02]. Players are often permitted to trade goods, and economies emerge. How similar are these virtual economies compared to real ones?

In the real world, a person might have no direct use for a diamond. That person could still be willing to pay a large sum for it, because the value of the diamond was determined by the market as whole, based on more than its direct impact on that person's life. The same can be said for a virtual object in an MMOG. The assessment of economic value is made through willingness to pay– with currency, time, or effort. In video games, the scarcity of certain virtual goods makes them valuable, with players investing hours of time to obtain them. This investment can be felt as very real. Despite being intrinsically worthless, virtual goods have value to those who trade them, and in this way differ little from e.g. the stock market. However, a few considerations have to be made.

1.7.1 *How games differ from the real world*

People play video games because they get some form of satisfaction from it. Besides possible addictive tendencies, nothing is forcing them to play the game, and they can start or stop playing whenever they want. In this way, player participation might not be driven by the same motivations as economic participation in the real world. This could have an impact on the emerging market structures.

Additionally, video games have a distinct encoded nature. In contrast to the physical universe, a virtual world can be manipulated by the owners of the game, usually without cost. Goods can be added or removed, game mechanics such as the ease of production can be adjusted, and so on. For example, this allows the creators of the game to control prices, something which real governments tend to avoid. Coding errors can also be present. When discovered and abused, these can severely impact the game's economy, as seen by the 20% inflation in a single day after a duplication error was found in the MMORPG *Everquest II* [Alp15].

Unlike the real world, players in virtual economies generally cannot misrepresent assets. Forged goods can't exist unless specifically encoded. Other forms of fraud can however be present, and little deters players from this 'criminal' behavior. An example is charging ten times as much for a specific good, and counting on the buyer not noticing the difference due to the visual representation of in-game currency. Real world policies regarding anti-trust rules are also absent, allowing players to corner the market of certain goods, creating artificial prices that would not form in the real world.

Many more differences between real and virtual economies exist, depending on the specific game in question. There is then no reason to a priori assume that they should behave similarly. Virtual economies should still be investigated however, as any similarities/differences could tell us more about the underlying dynamics that give rise to the emergent complex behavior.

1.7.2 *Insights from virtual worlds*

Complex video games are a relatively new phenomena, and the research on virtual worlds has remained limited. Lofgren & Fefferman [LF07] compared the unintended spread of a virtual disease in the MMORPG *World of Warcraft* to real world infections, arguing for the use of video games as a testing ground for epidemic control measures. Kim, Keegan, Park, & Oh [Kim+15] investigated player behavior in the video game *League of Legends* and found that individual player proficiency increased team performance more than team congruency, having implications on team building in the real world. Fuchs & Thurner [FT14] studied the MMORPG *PARDUS*, and found a correlation between political status and wealth, in addition to finding wealth distributions that were comparable to those in the real world, as seen in Fig. 7. Szell et. al. [Sze+12] also used *PARDUS* data, and found that the in-game motion of players was strongly shaped by the presence of socio-economic areas. Castronova et. al. [Cas+09], as a last example, looked at the MMORPG *Everquest II* and found that the aggregate economical behavior was comparable to that of the real world.

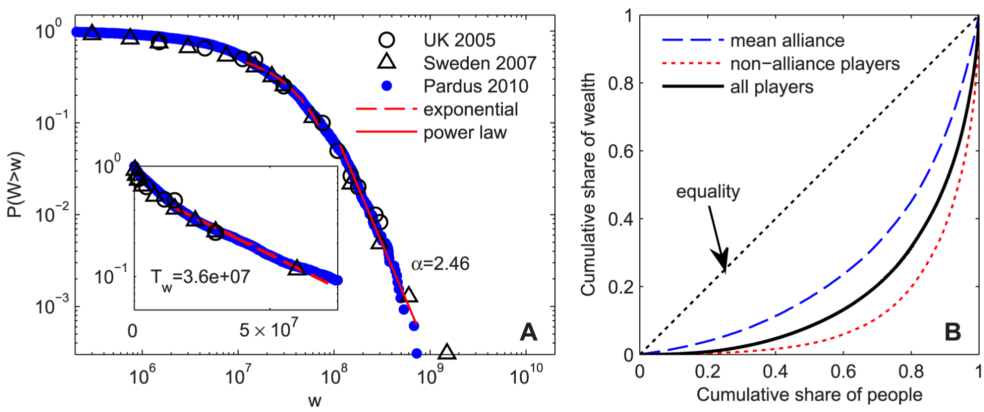

Figure 7: Wealth distribution and Lorenz curve comparing the UK, Sweden and the game *PARDUS*, taken from [FT14]. The left image shows a comparable distribution of wealth in the real and virtual world.

Most of the available literature appears to focus on behavioral aspects, rather than economic analysis. One of the goals of this thesis is to compare the general behavior of MMORPG *EVE Online*'s economy with that of the real world. This will be done through first exploring the stylized facts in the real world, and then using the same methods to look for them in *EVE Online*. Afterwards, an agent-based model will be introduced based on the dynamics of *EVE Online*, and I will investigate its emerging economy in a similar manner.

2 STYLIZED FEATURES OF REAL-WORLD FINANCIAL MARKETS

In this chapter, a framework is built for comparing stylized facts in real and virtual markets. The methodology to visualize a selection of stylized facts is first explained, illustrated with daily frequency data of the NASDAQ stock price. Afterwards, it is applied to a wide range of markets.

2.1 METHODOLOGY

The focus on this chapter is on the distribution of price log-returns, defined by Eq. 12 in the introduction:

$$r_\tau(t) = \ln(p(t+\tau)) - \ln(p(t)).$$

If the price $p(t)$ at some point in time is zero, the logarithm is not defined, and the (log-)return can't be calculated. A similar problem arises when no price was recorded at all– e.g. financial stock markets record no prices during the weekends. There are several possibilities to handle missing data, each with their own disadvantages. The majority of data I used was of daily frequency, with weekend gaps in real world data. Log-returns across weekends were of a similar size as those calculated across weekdays. For this reason, I ignored the gaps and calculated the log-returns as if the data after them immediately followed the last logged value.

2.1.1 *Fat-tailed empirical distribution of log-returns*

The first stylized fact states that the log-returns of prices are fat-tailed: the probability of finding extreme events is higher than what could be expected if they were to follow a Gaussian distribution. To visualize this, I calculated and binned the log-returns to form a probability density function (PDF). I then fitted a Gaussian distribution to the data using the method of least squares. For daily frequency NASDAQ data, the result can be seen in Fig. 8. Fat-tailed behavior is clearly visible.

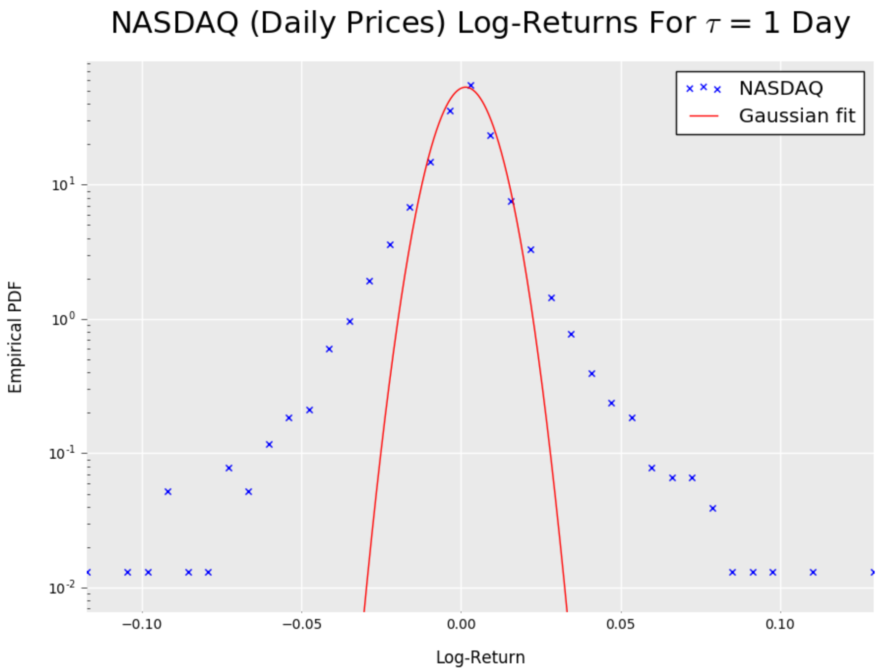

Figure 8: Empirical PDF of NASDAQ log-returns. Log-returns were calculated using daily frequency data between 1971-02-05 and 2017-04-05, retrieved from YAHOO finance.

To quantify the fat tails, the kurtosis k (the fourth standardized moment) is often used in literature. For a sample $x_1, x_2, ..., x_n$, this can be estimated by

$$k = (n-1)\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2}. \tag{21}$$

The kurtosis of a normal distribution is 3. A distribution is defined as *leptokurtic*, or heavy tailed, if its kurtosis is larger than 3.

As mentioned in the introduction, the heavy tails of log-return distributions are sometimes approximated as power-laws. For a power-law $p(x) = Cx^{-\alpha}$ starting at x_{min} however, the fourth moment is [New05]

$$\begin{aligned} \langle x^4 \rangle &= \int_{x_{min}}^{\infty} x^4 p(x) dx = C \int_{x_{min}}^{\infty} x^{-\alpha+4} dx \\ &= \frac{C}{5-\alpha}\left[x^{-\alpha+4}\right]_{x_{min}}^{\infty}, \end{aligned} \tag{22}$$

which becomes infinite if $\alpha \leq 4$. This implies that, while the kurtosis for a particular sample may be relatively small, it occasionally takes on a huge value, causing the kurtosis on average to diverge. This tells us that the kurtosis is not a well defined quantity for power-laws with $\alpha \leq 4$.

In the markets I investigated for this thesis, fat tails occasionally exhibited power-law exponents exceeding 4. However, as I was more interested in the presence of fat tails rather than their precise behavior, no rigorous comparison between possible distributions was done to determine if the a power-law was indeed the best fit. I then cautiously followed the literature in using kurtosis as a marker for non-Gaussianity.

2.1.2 *Aggregational normality*

The second stylized fact states that whereas the log-returns show fat tails, this feature decreases as the time period τ over which the log-returns are calculated increases. To visualize this, I first determined the log-return empirical PDFs as described in the previous subsection for different τ. I then placed the individual PDFs side by side, as seen in Fig. 9 for daily frequency NASDAQ data. Fat tails are seen for short τ, and as τ increases the distribution becomes more Gaussian.

After this, I standardized the log-returns by using the parameters (μ_i, σ_i) of the Gaussian distribution fitted to them

$$r_\tau(t) \rightarrow \frac{r_\tau(t) - \mu_i}{\sqrt{\sigma_i^2}}. \tag{23}$$

I then plotted the CCDFs on a single image, on which changes in fat-tailed behavior become more clear. For daily frequency NASDAQ data, this is shown in Fig. 10.

Finally, I quantified the fat-tailedness by calculating the kurtosis for each of the PDFs and plotting the *excess* kurtosis $k - 3$ in function of τ. For comparison purposes, a measure whose change in value is less susceptible to extreme outliers than kurtosis was also calculated: the Hogg coefficient [KW04]

$$Hogg = \frac{U_\alpha - L_\alpha}{U_\beta - L_\beta}, \tag{24}$$

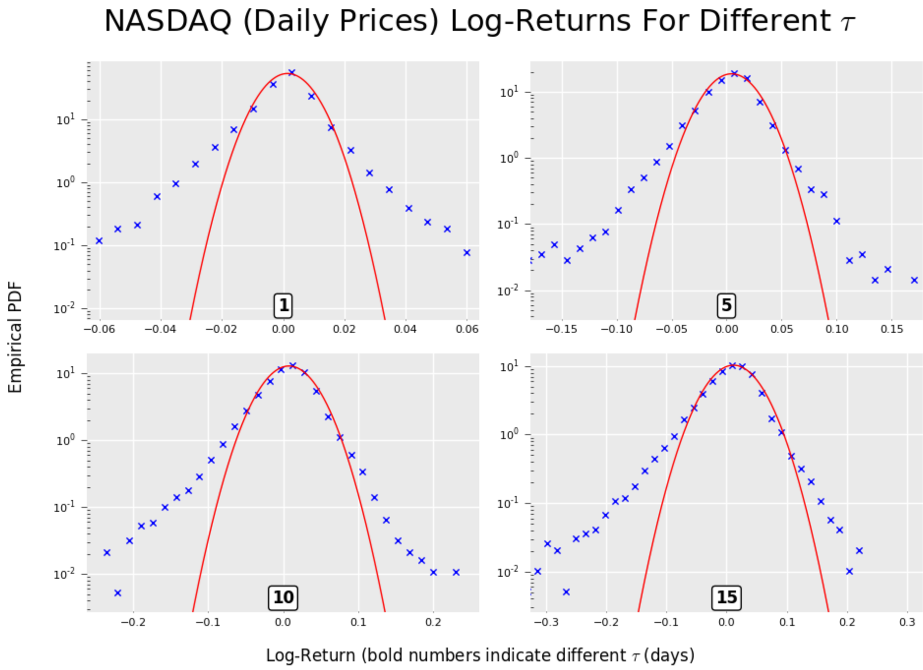

Figure 9: Empirical PDF of NASDAQ log-returns, for several τ indicated by bold numbers. Log-returns were calculated using daily frequency data between 1971-02-05 and 2017-04-05, retrieved from YAHOO finance.

where U_α (L_α) is the average of the upper (lower) α quantiles defined as

$$\begin{aligned} U_\alpha &= \frac{1}{\alpha}\int_{1-\alpha}^{1} CDF^{-1}(y)dy \\ L_\alpha &= \frac{1}{\alpha}\int_{0}^{\alpha} CDF^{-1}(y)dy, \end{aligned} \tag{25}$$

for $0 \le \alpha \le 1$. The Hogg coefficient for a Gaussian distribution is 2.59; like kurtosis, the *excess* Hogg coefficient $Hogg - 2.59$ will be used to look for leptokurtic behavior.

Nothing a priori dictates that the PDFs of price log-returns should exhibit only one peak. Multimodal distributions could cause a misinterpretation of the value of the kurtosis. For this reason, I calculated Sarle's bimodality coefficient [Pfi+13]:

$$b = (g^2+1)\left(k-3+\frac{3(n-1)^2}{(n-2)(n-3)}\right)^{-1} \tag{26}$$

where n is the sample size, g is the sample skewness, and k is the sample kurtosis. The value of this coefficient lies between

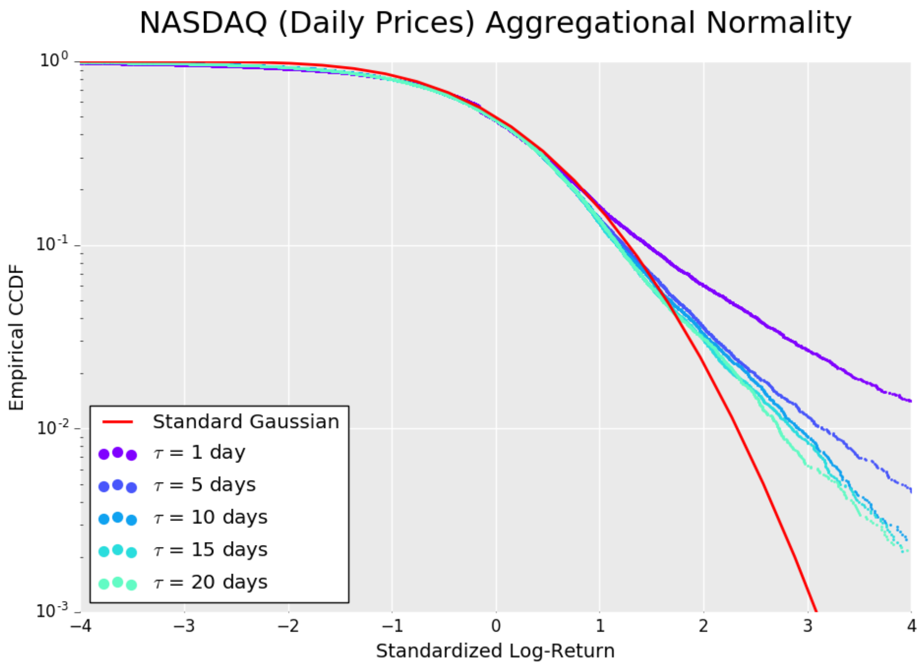

Figure 10: Empirical CCDF of NASDAQ log-returns, for several τ. Log-returns were calculated using daily frequency data between 1971-02-05 and 2017-04-05 retrieved, from YAHOO finance.

0 and 1, and increases as the multimodality of the distribution increase.

Excess kurtosis, Hogg coefficient and Sarle's bimodality coefficient are then plotted against τ. For daily frequency NASDAQ data, this is shown in Fig. 11.

2.1.3 *Absence of autocorrelation of log-returns*

The third stylized fact states that there is no evidence of autocorrelation between log-returns. The Pearson correlation coefficient for two datasets with n values $\{x_1 \ldots x_n\}$, $\{y_1 \ldots y_n\}$ is defined as

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{27}$$

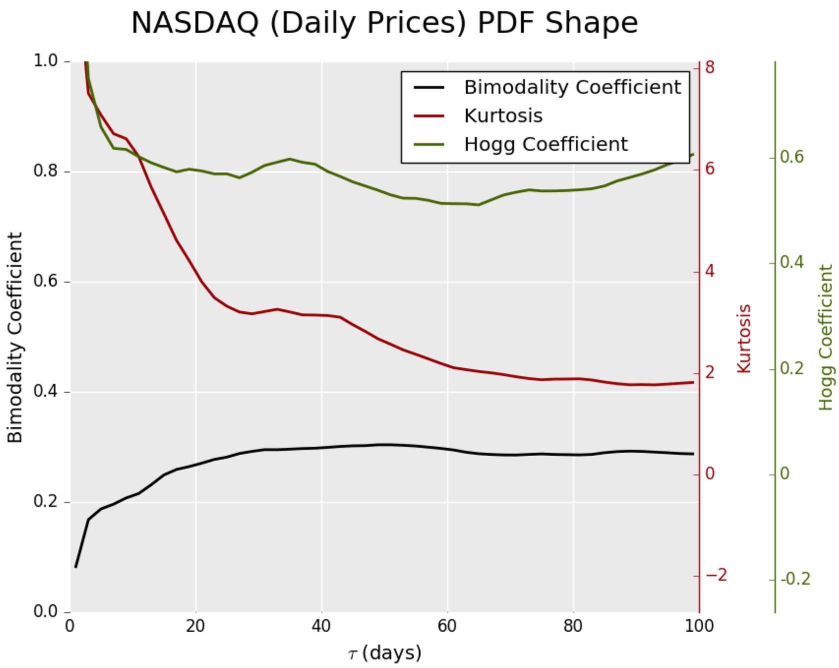

Figure 11: Sarle's bimodiality coefficient, kurtosis and Hogg coefficient of NASDAQ log-returns, plotted against τ. Log-returns were calculated using daily frequency data between 1971-02-05 and 2017-04-05, retrieved from YAHOO finance.

where $\bar{x}, \bar{y}$ are the sample means. For two datasets from the same time series, shifted over a time T, the cross-correlation at lag T is defined as

$$\rho(T) = \frac{\sum_{i=1}^{n-T}(x_i - \bar{X})(x_{i+T} - \bar{Y})}{\sqrt{\sum_{i=1}^{n-T}(x_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n-T}(x_{i+T} - \bar{Y})^2}} \tag{28}$$

where $\bar{X}$ and $\bar{Y}$ are means of datasets X = $\{x_1 \ldots x_{n-T}\}$ and Y = $\{x_{1+T} \ldots x_n\}$. To investigate the third stylized fact, I constructed correlograms, plotting the sample autocorrelation $\rho(T)$ against the time lag T over which the autocorrelation is calculated.

In order to reject the null hypothesis that $\rho(T)$ is zero, confidence intervals were created using Bartlett's formula [Pec]

$$\mathrm{CI} = \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{1}{N}(1 + 2\sum_{i=1}^{k} r_i^2)} \tag{29}$$

where $z_{1-\frac{\alpha}{2}}$ indicates the quantile function of the normal distribution.

The resulting correlogram for daily frequency NASDAQ data can be seen in Fig. 12. The values of $\rho(T)$ quickly become indistinguishable from zero.

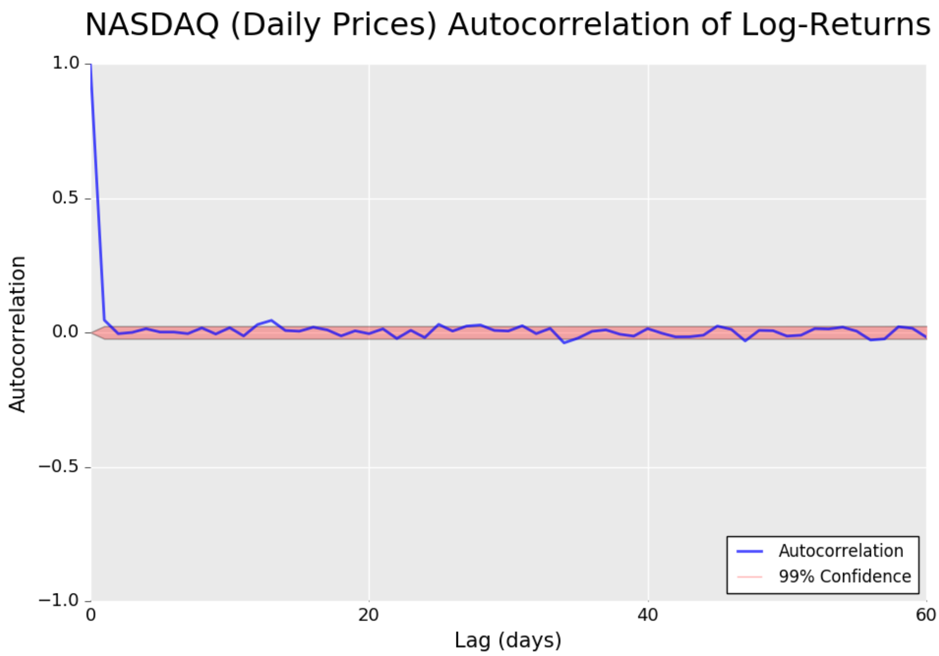

Figure 12: Correlogram of NASDAQ log-returns. Shaded area indicates a 99% confidence interval, calculated using Bartlett's formula. Log-returns were calculated using daily frequency data between 1971-02-05 and 2017-04-05, retrieved from YAHOO finance.

2.1.4 *Volatility clustering*

The fourth stylized fact states that even though the log-returns are not autocorrelated, absolute log-returns do exhibit an autocorrelation function that is significantly larger than zero and decays slowly. To show this, correlograms were constructed using the method described in the previous segment. The resulting correlogram for daily frequency NASDAQ data can be seen in Fig. 13: volatility clustering is clearly present.

2.2 RESULTS

The above methods provide a visual way to look for the presence of stylized facts. For a variety of markets, composite im-

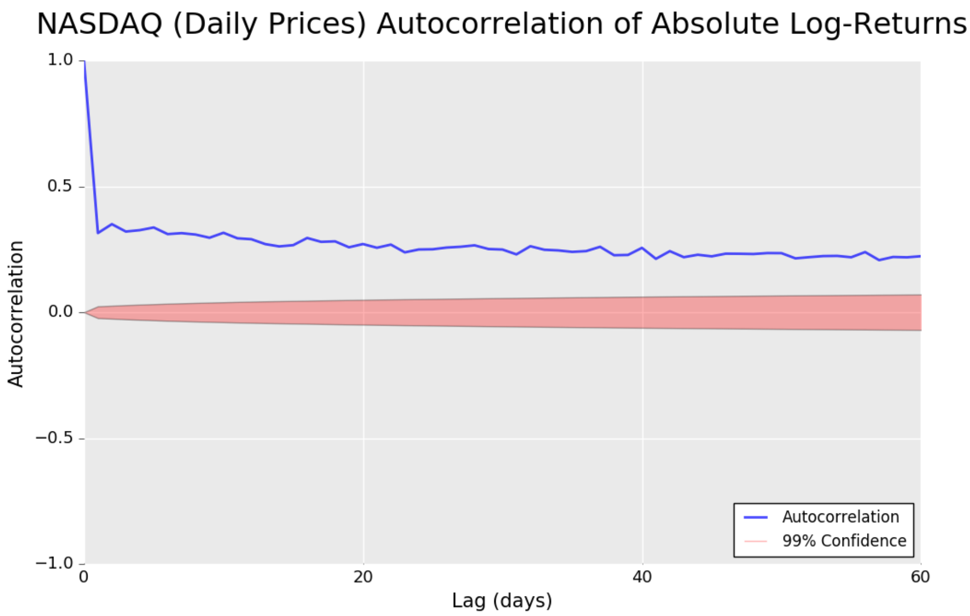

Figure 13: Correlogram of NASDAQ absolute log-returns. Shaded area indicates a 99% confidence interval, calculated using Bartlett's formula. Log-returns were calculated using daily frequency data between 1971-02-05 and 2017-04-05, retrieved from YAHOO finance.

ages were created, and a selection will be discussed here. The remaining ones can be found in the Appendix. On such a figure, subfigure A) shows the price over the entire range of available data. Subfigures B) and C) list the correlograms, where the stylized facts suggest that B) should be zero and C) should show a slowly decaying autocorrelation to indicate volatility clustering. Subfigures D-F) show the PDFs and aggregational normality indicators. Fat tails should be visible in the upper section of D), and E-F) should show a decrease of the fat-tailed behavior over τ.

As mentioned in the introduction, the assumption that price changes are to some extent random is one of the foundations of the BSM model of price returns. This model assumes geometrical Brownian motion (GBM), which can be modeled as [Sig13]

$$p(t_{i+1}) = p(t_i)e^{\sigma\sqrt{t_{1+1}-t_i}N(0,1)+\mu(t_{1+1}-t_i)},\ t \geq 0 \tag{30}$$

where μ is the percentage drift, σ is the percentage volatility, and $N(0,1)$ is a random number from a standard Gaussian distribution. For comparison, similar composite images were created for GBM processes.

- Crude oil (Fig. 14): daily WTI crude oil prices. The absolute log-returns show a significant autocorrelation, whereas the raw log-returns do not. Fat tails are clearly present in the first row of subfigure D). The positive log-returns quickly approach a Gaussian distribution, whereas the negative ones exhibit fat tails for a longer period. This can be explained by the presence of a sudden large drop in price, i.e. the 2008 and 2014 crashes, visible in A). Log-returns calculated using data across this period will be large and negative.

- Aluminum (Fig. 15): daily aluminum prices on the London Metal Exchange. These prices exhibit much less of the stylized facts: the raw log-returns show no autocorrelation, but neither do the absolute ones. The PDFs also show no fat tails, as can be seen in D), E), and low values of the excess kurtosis in F).

- Lumber (Fig. 16): daily lumber futures prices on the Chicago Mercantile Exchange. Futures are contracts where the price is determined in the present, for delivery of the goods in the future. Long-term price histories can be constructed by chaining together individual short-term futures contracts, providing an estimate for the daily price. For lumber futures, all stylized facts appear to be present.

- Cattle (Fig. 17): daily live cattle futures prices on the Chicago Mercantile Exchange. Although the fat tails in D) appear to be limited, all stylized facts are arguably present.

- Rubber (Fig. 18): daily natural rubber futures prices on the Shanghai Futures Exchange. Subfigure C) shows no clear volatility clustering. In addition, whereas E) and F) would suggest aggregational gaussianity, the PDFs in D) exhibit multimodal behavior that starts as early as $\tau = 2$ days. This could indicate an additional underlying process, and makes it hard to determine if fat-tailed behavior is ever present.

- GBM (Fig. 19): GBM process with $\mu = 0.0001$ and $\sigma = 0.01$. The price pattern in A) is very similar to real price patterns and no autocorrelation of the log-returns is visible in B), but C-F) show no sign of the other stylized facts.

2.3 CONCLUSION

A few general observations can be made. First, the stylized facts are present in many markets that were investigated, but not all of them. How robust are these features then? Do certain markets lack the required underlying mechanisms for them to emerge, or do they suppress them in some way?

Second, the general behavior of the excess kurtosis in F) is similar to that of the excess Hogg coefficient, and seems to be a good proxy for the behavior of fat tails. However, some caution should be exercised, as seen in the results for natural rubber in Fig. 18. The bimodality coefficient appears to move inversely to the kurtosis, and a decrease in kurtosis may indicate multimodal behavior rather than a decrease of fat tails.

Finally, GBM processes as described by Eq. 30 don't reproduce stylized facts. The BSM can account for this by e.g. assuming the values of μ and σ as varying over time, but the base random process is not sufficient to fully replicate market behavior. The question arises if whether this is a sufficiently strong argument against the validity of the random walk theories such as the EMH.

This concludes the examination of some markets in the real world. In the next chapter, I move on to the economy of a virtual one: *EVE Online*.

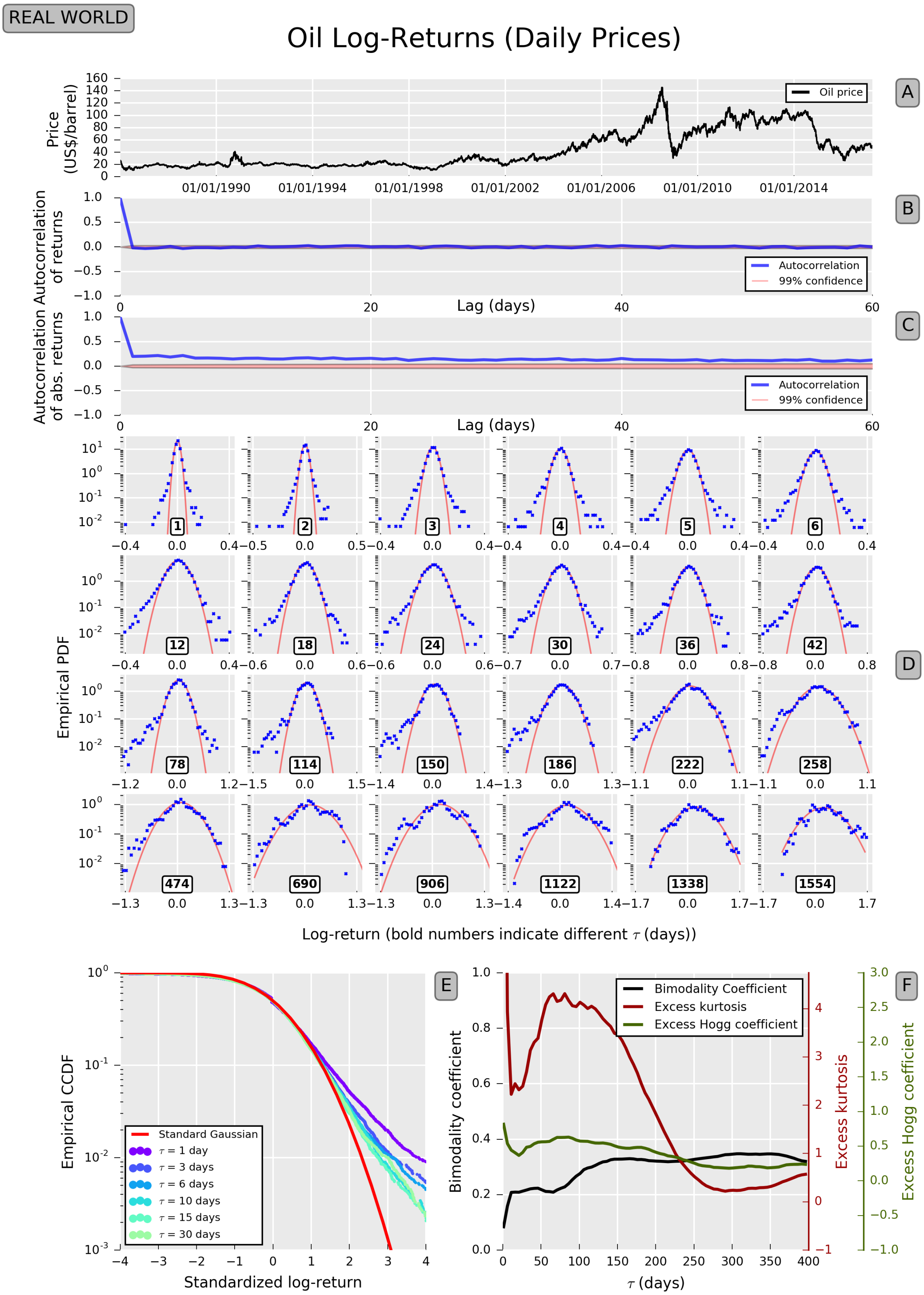

Figure 14: West Texas Intermediate (WTI) daily crude oil price between 1986-01-02 and 2017-04-03, retrieved from [Dat17].

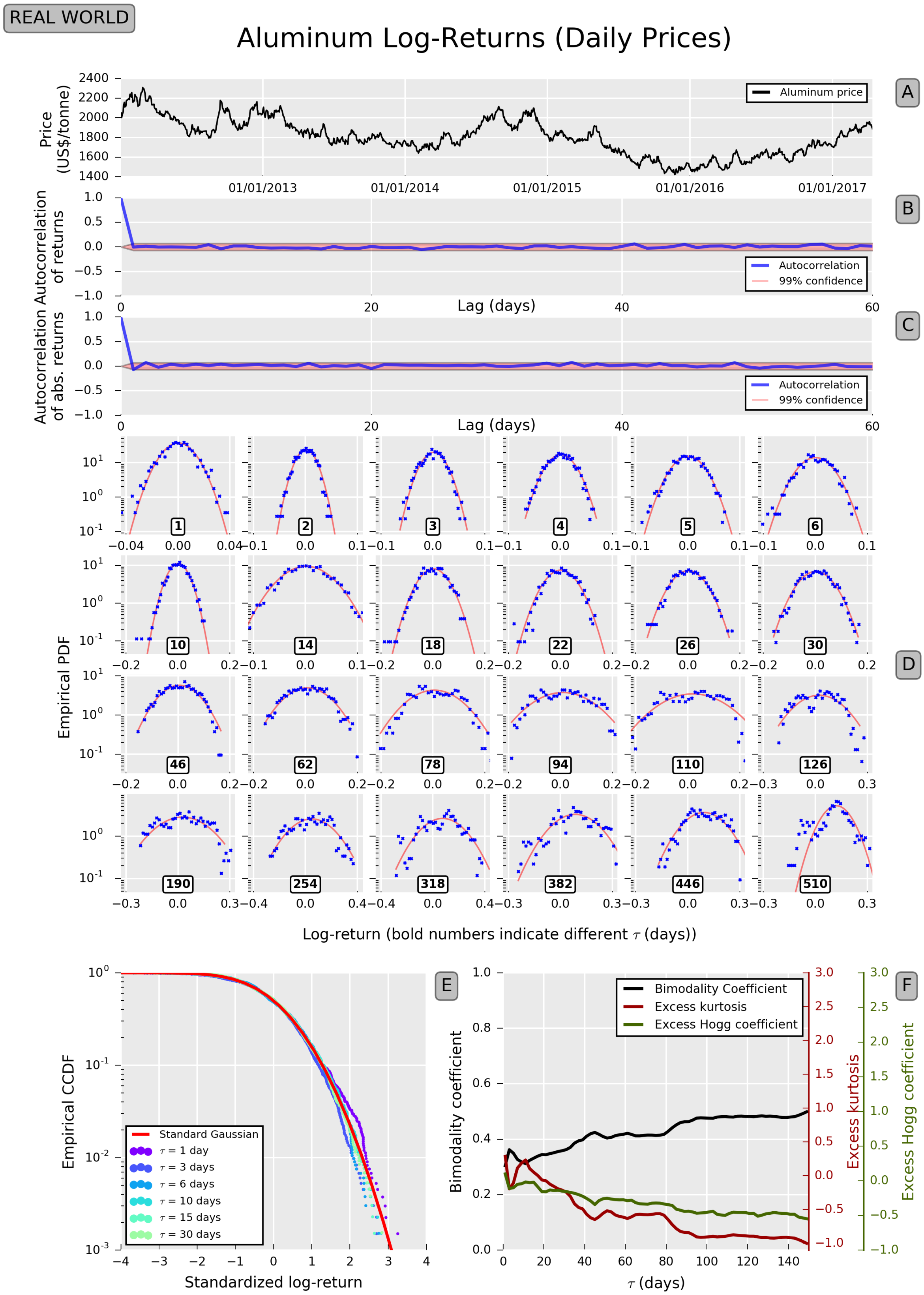

Figure 15: Daily aluminum price on the London Metal Exchange between 2012-01-03 and 2017-04-13, retrieved from [Qua17b].

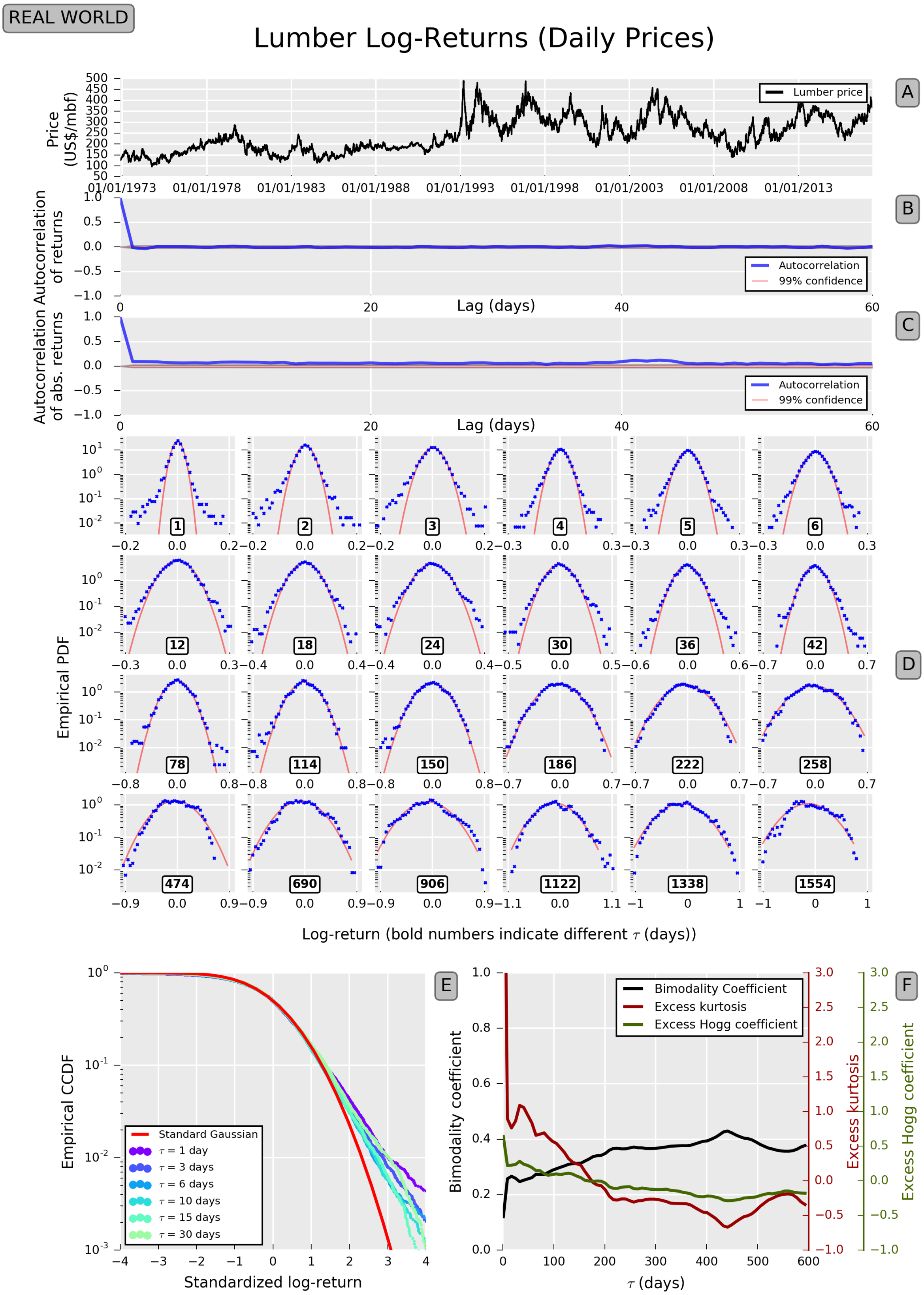

Figure 16: Daily lumber futures price on the Chicago Mercantile Exchange between 1972-11-16 and 2017-05-15, retrieved from [Qua17b].

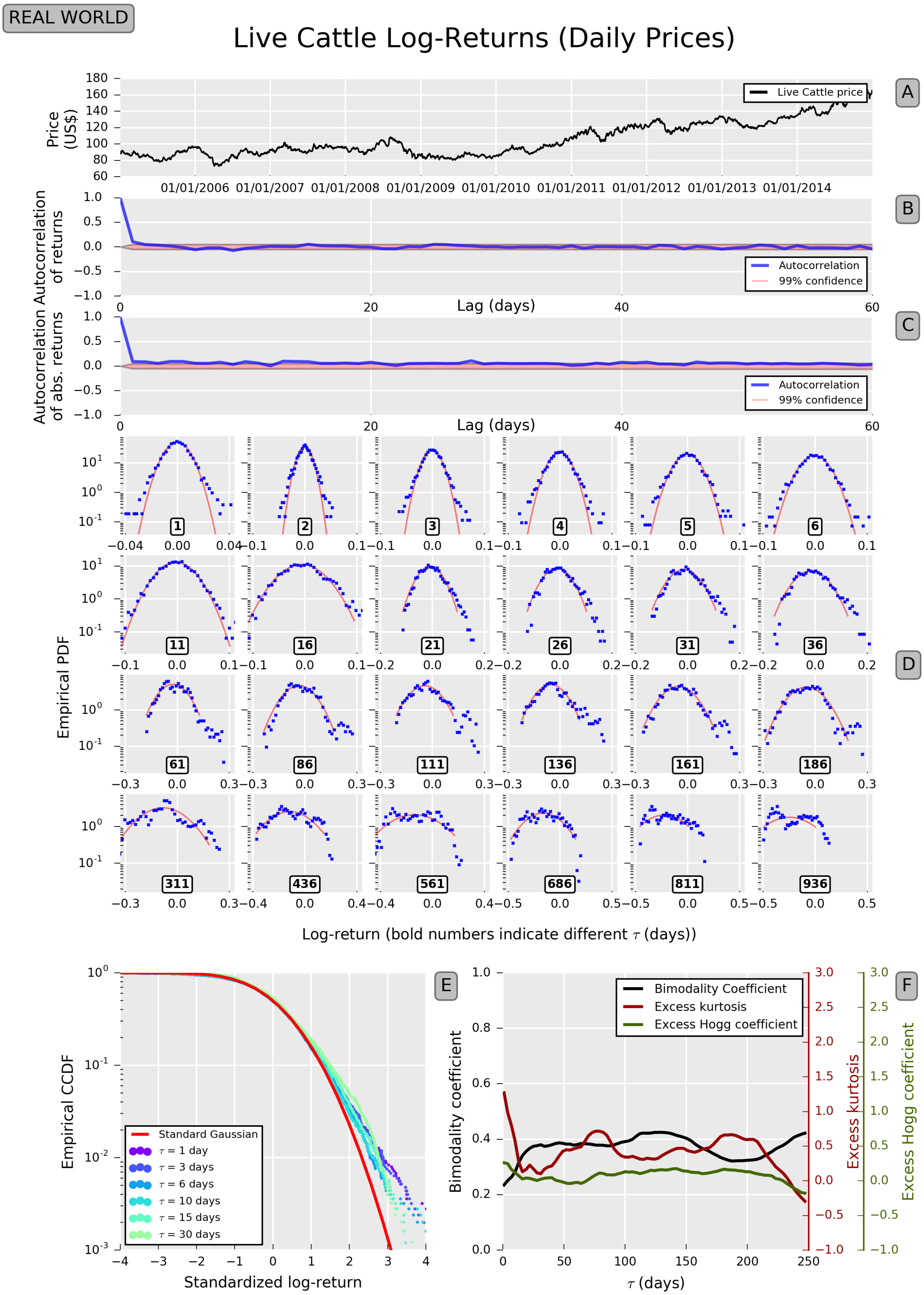

Figure 17: Daily live cattle futures price on the Chicago Mercantile Exchange between 2005-01-03 and 2014-12-31, retrieved from [Qua17b].

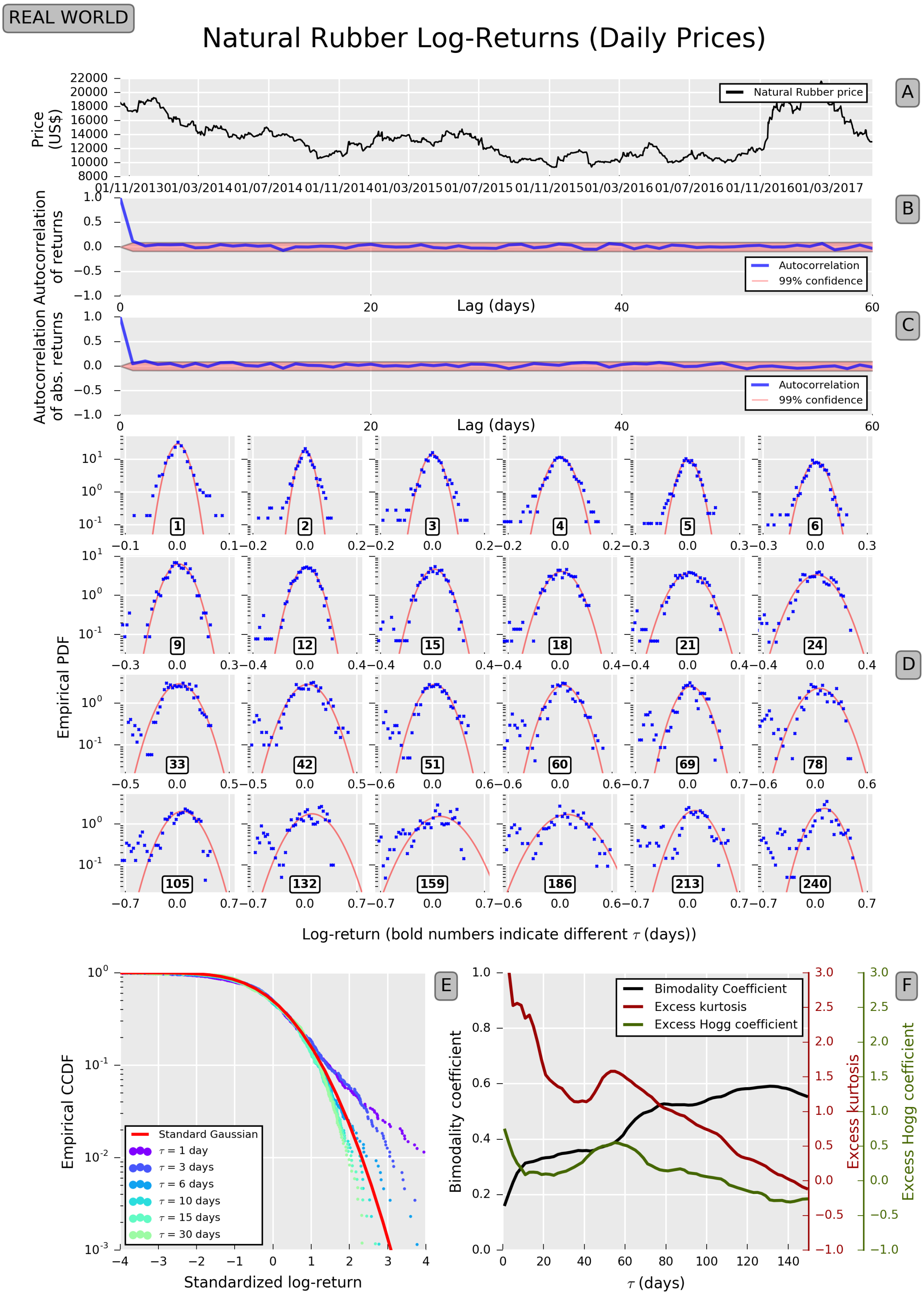

Figure 18: Daily natural rubber futures price on the Shanghai Futures Exchange between 2013-10-16 and 2017-05-15, retrieved from [Qua17b].

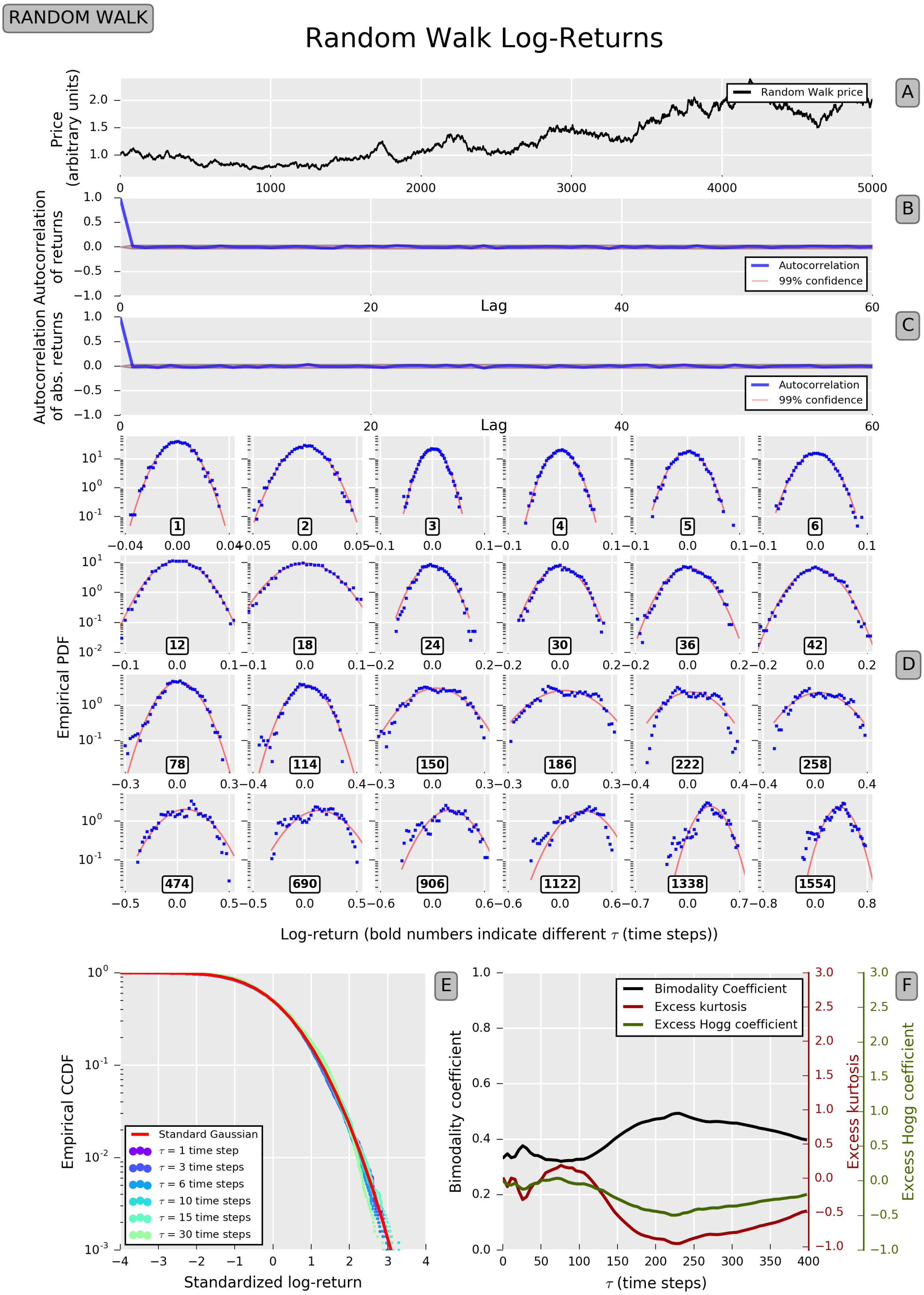

Figure 19: Geometric Brownian motion as described by Eq. 30, for 5000 steps with $\mu = 0.0001$ and $\sigma = 0.01$.

3 STYLIZED FACTS IN THE VIRTUAL WORLD OF EVE

Having explored some stylized facts in real markets, I will now look at the virtual world of the MMORPG *EVE Online* (or simply *EVE*). In *EVE*, players act as pilots in a futuristic world. They fly ships across space, completing various objectives. Some example objectives are fighting others, transporting goods, creating ships, and mining ores. In-game goods are traded on a virtual market in search of profit, measured in the in-game currency *ISK* (InterStellar Kredit). The entire production chain of goods is in the hands of the players, who analyse *EVE*'s markets with as much scrutiny as investors in the real world. In this chapter, I investigate whether the prices in this virtual world behave similarly to those of the real one with respect to the stylized facts. I conclude with a short discussion on the wealth distribution of *EVE*.

3.1 STYLIZED FACTS

As discussed in the introduction, the price behavior of goods in video games may depend on factors specific to the game. For *EVE*, prices are set by the players, and the developers generally do not interfere. However, these prices may be indirectly affected by modifications of the game known as *expansions*. These expansions include new game areas and objects, changes in material requirements for existing goods, and so on. An analogy in the real world would be the discovery of a new technology such as fracking, which drastically alters the supply of a good. In the composite images for *EVE* markets, these expansions are shown by numbered lines in subfigures A).

The virtual goods I examined were four basic ships and the ores required to make them. Market data was provided by *EVE*'s developer *CCP Games* [Gam17]. As with the real world markets I will discuss a selection of them, and the remaining results can be found in the Appendix. For ship data, an initial cleaning phase removed single-day price spikes that more than doubled the average price. These can be attributed to wrongly matched buy/sell orders, and don't reflect the general market prices.

- Tritanium (Fig. 20): daily tritanium ore price. Subfigure A) shows a general price behavior not unlike aluminum in the previous chapter. Some inflation is present, which can be seen in D) for $\tau = 1554$ days, giving a PDF centered around a positive value. All four stylized facts appear to be present.

- Pyerite (Fig. 21): daily pyerite ore price. If one only looks at B-C) and F), all stylized facts appear present. However, A) shows that pyerite was much more impacted by the release of expansions. For example, A.6) marks the release date of *Revelations I*, which reduced the supply of pyerite, increasing its price. Log-returns calculated using data before and inside this plateau are large and positive, giving the appearance of a positive shoulder in D) with $\tau = 24$ days. The following drop in price at A.7) happens more slowly, and so negative shoulders in the log-returns only appear at larger τ.

  Additionally, in 2012 at A.16-18), a series of game changes were implemented which resulted in large increases of ore and ship prices. This splits A) up in two regions. The higher average price in the second region results in a larger number of positive log-returns. This gives D) a multimodal appearance for high τ. For example, at $\tau = 1122$ days, an expected peak centered at zero is accompanied by a positive peak for log-returns calculated using data across the two regions.

- Condor (Fig. 22): daily Condor ship price. As all ship prices examined showed large increases during 2012, data was limited to either before or after this period. For Condor prices before 2012, all four stylized facts are visible. Subfigure A) shows flattened regions, where the price appears to stay constant. This is reflected in D) with $\tau = 1$ day, where the probability of finding log-returns equal to zero is increased. These give much weight to the Gaussian fit, which can be seen in the kink of the standardized CCDF in E).

- Slasher (Fig. 23): daily Slasher ship price starting 2013. The stylized facts in subfigures B) and C) are present, but fat-tailed behavior of the PDFs is much less clear. In addition, E) and F) seem to indicate that the fat tails increase for higher τ.

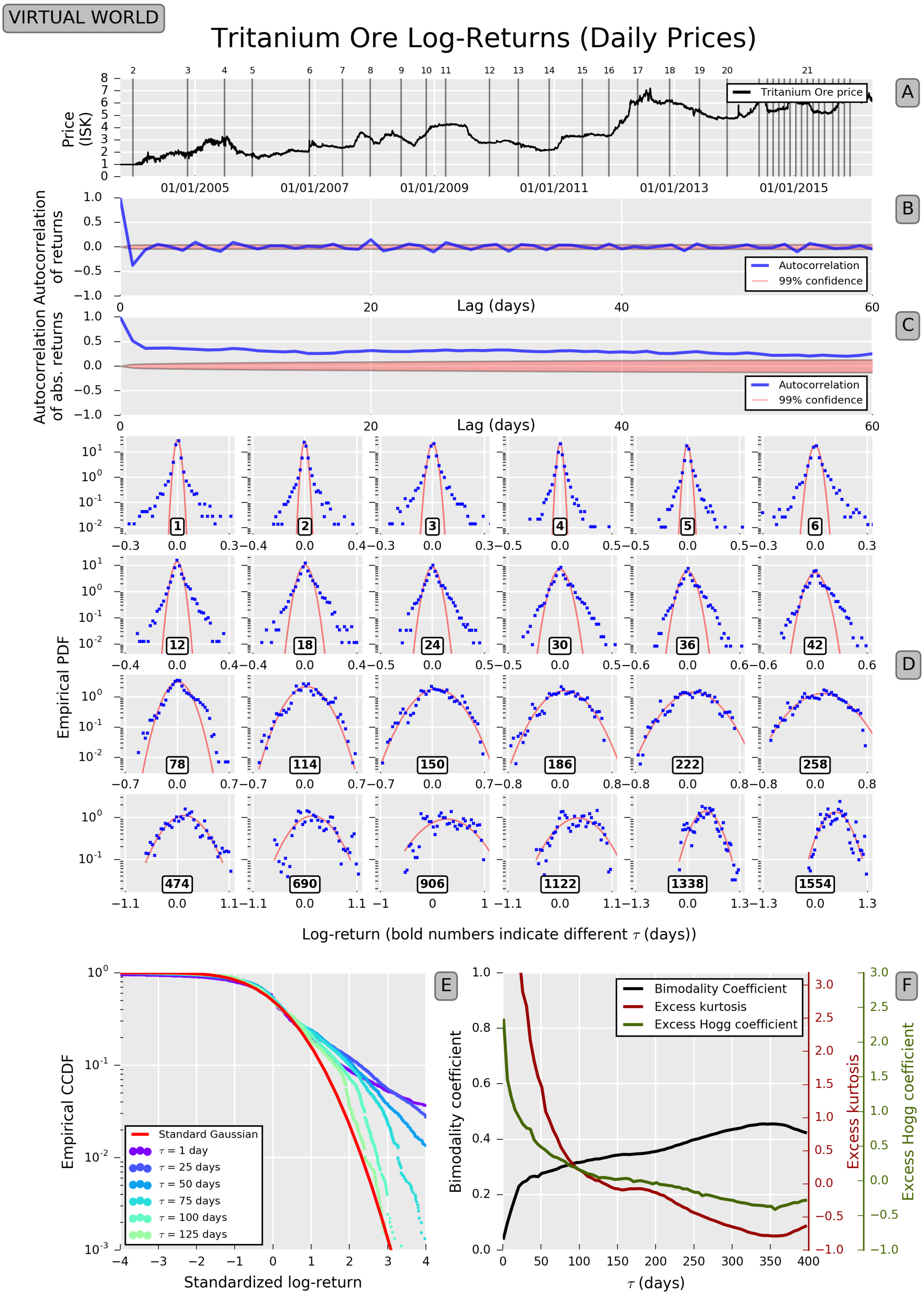

Figure 20: Tritanium ore daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17].

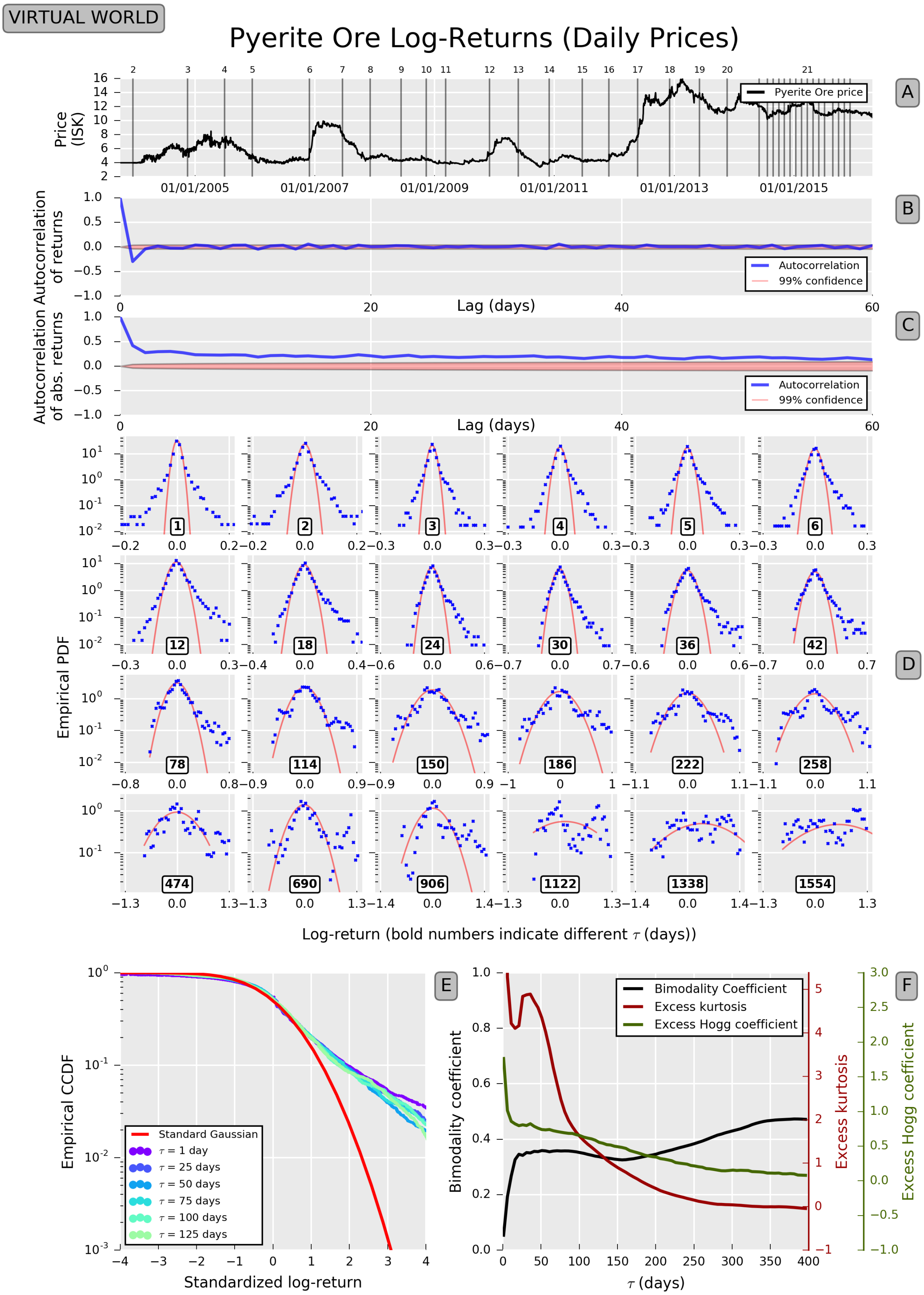

Figure 21: Pyerite ore daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17].

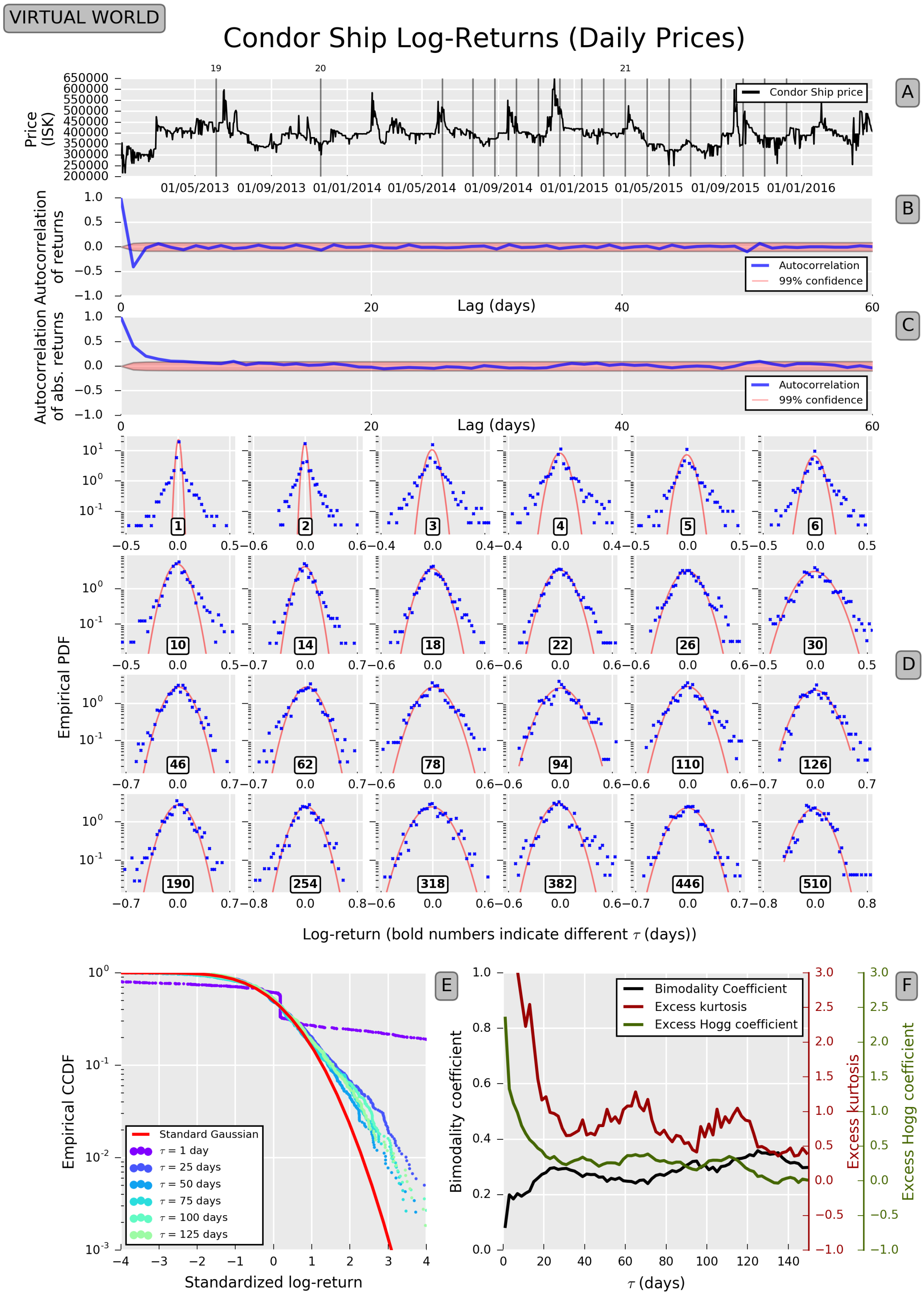

Figure 22: Atron ship daily price in *EVE* between 2013-01-01 and 2016-04-23, courtesy of [Gam17].

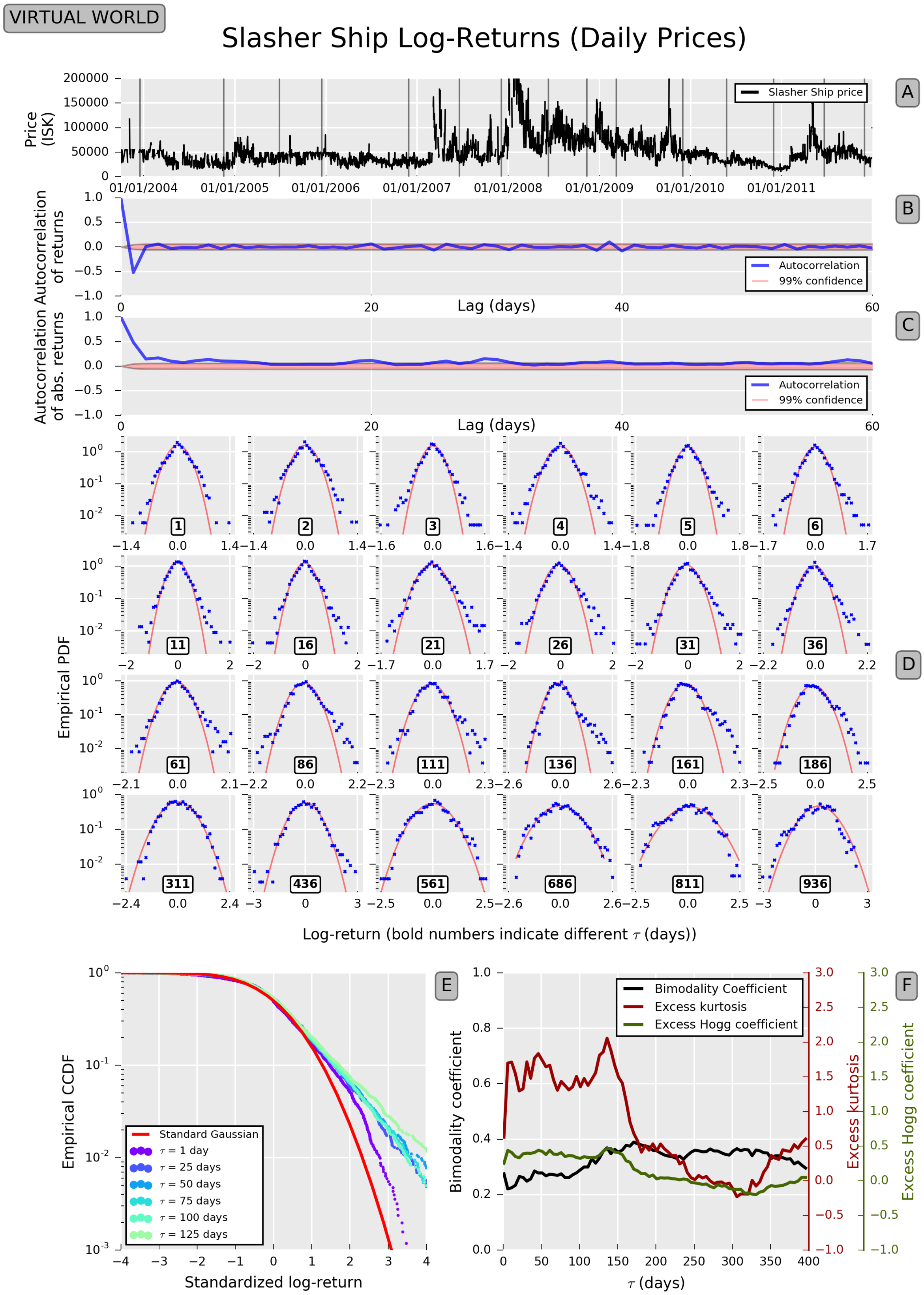

Figure 23: Slasher ship daily price in *EVE* between 2003-10-01 and 2012-01-01, courtesy of [Gam17].

3.2 WEALTH DISTRIBUTION

In addition to examining price behavior in *EVE*, a rough estimate of its wealth distribution was created. Plans exist to add direct wealth tracers to the database provided by [Gam17], but at the time of writing, these are not available. Wealth must then be estimated through a proxy. As players invest much of their *ISK* in ships and other goods, a proper proxy should include all of these. The database did not allow the tracking of inventories, and only player *ISK* could be obtained, giving a rough estimate of wealth. More rigorous investigation of the wealth distribution will have to wait until better data can be obtained.

The Lorenz curve of the *ISK* distribution of active players on 2016-04-27 can be seen in Fig 24. For comparison, Fig. 25 shows real world Ginis of the year 2000. *EVE* exhibits a highly unequal distribution of wealth with a Gini of 0.90, exceeding that of any country in the real world. Namibia comes closest, with a Gini of 0.847. A common factor between these two is the absence of a welfare structure, where wealth is spread out through progressive taxation and social services.

As a remark, the Gini of the world being higher than that of its individual countries is not a mistake. If the inhabitants of two countries possess the same amount of money as their countrymen, but the first ones own \$10 and the second ones \$1000, the inequality of the individual countries will be low but their combined inequality will be high.

The CCDF of the *ISK* distribution can be seen in Fig. 26. An exponential and log-normal were fitted to the bulk of the distribution (90%) and a power-law to the tail (10%), shifted for clarity. A log-normal seems to be a superior fit, but no rigorous analysis was done to investigate this as better data is expected to become available in the near future. This small discussion on *EVE*'s wealth distribution was only included to argue that the virtual economy may be comparable to real ones in more than its price behavior.

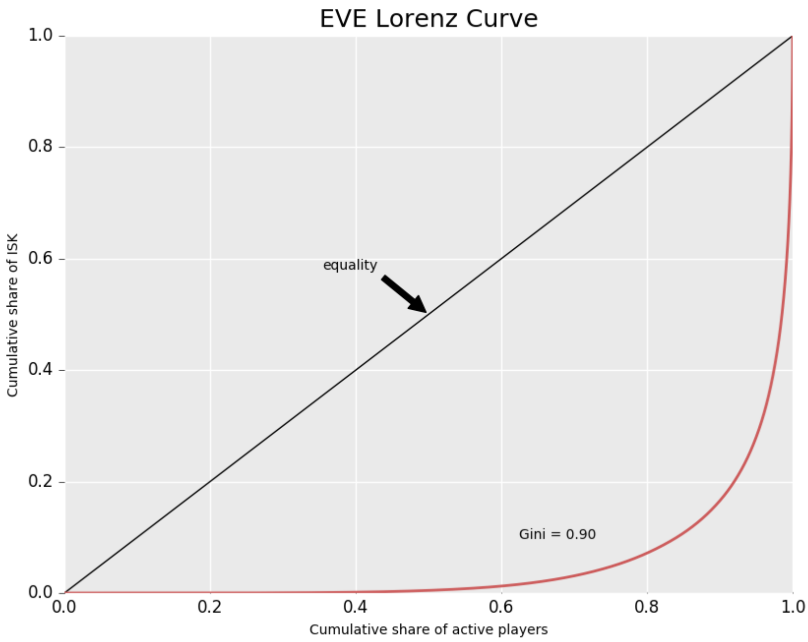

Figure 24: Lorenz curve of *ISK* distribution in *EVE* on 2016-04-27. Data courtesy of [Gam17].

3.3 CONCLUSION

To conclude, despite being a free market with minimum 'government' regulation, the prices of goods in *EVE* behave similar to those in the real world. Stylized facts are often present, but not always. These results strengthen the claim that real and virtual economies are comparable, and that perhaps results from research in one could extend to the other. As virtual worlds offer opportunities for experimentation which would be impossible in the real world, this field deserves continued exploration.

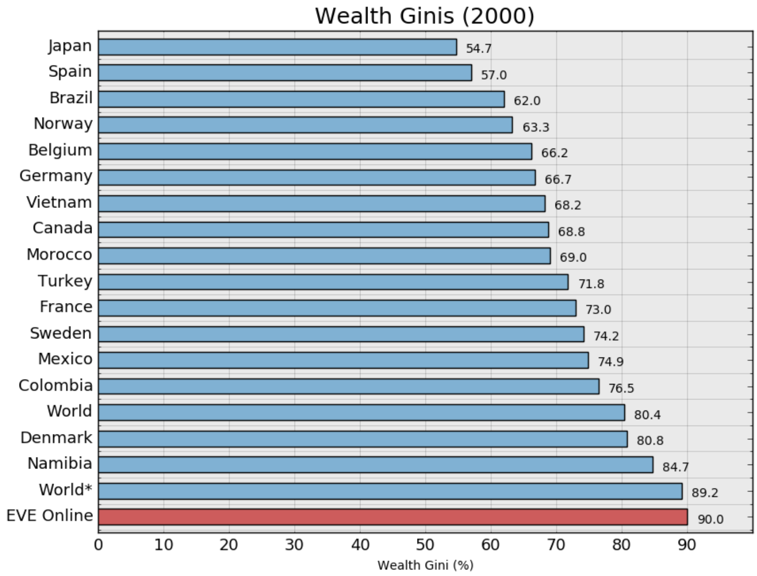

Figure 25: Wealth Ginis of a selection of countries of the year 2000. World* is calculated using purchasing parity power dollars, rather than conventional US$ exchange rates. Data from [Dav+09].

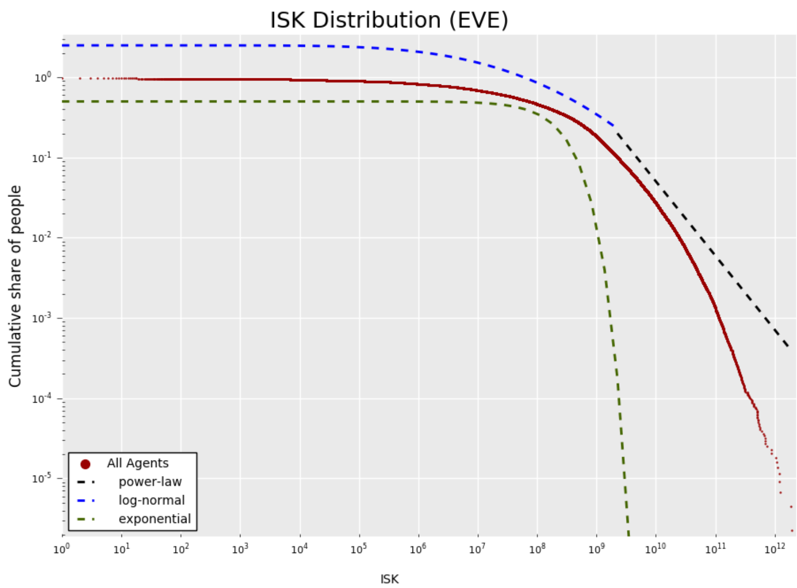

Figure 26: CCDF of *EVE*'s wealth distribution. A log-normal and exponential were fitted to the bulk 90%, and a power-law with $\alpha = 1.92$ to the bottom 10%.

4 REPLICATING EVE MARKET STRUCTURE THROUGH ABM

In the introduction, some examples of ABMs were given that aim to replicate economies. These ABMs generally use multiple agent types representing traders with different behavior. In the final part of this thesis, I present an ABM with an alternative approach, based on the game dynamics of *EVE*.

4.1 DESIGN PHILOSOPHY OF THE MODEL

A classical breakdown of economic activity distinguishes three sectors:

- primary: the retrieval and production of raw materials, e.g. corn or iron,
- secondary: the transformation of raw or intermediate materials into goods, e.g. steel into cars or textiles into clothing,
- and tertiary: the supplying of services, e.g. baby-sitting or banking.

A raw good such as iron flows through the three sectors by being mined and refined into steel (primary sector), being transformed into a car (secondary sector), which is then sold to and used by a laborer (tertiary sector).

In *EVE*, a similar distinction can be made. All *EVE*'s players need a ship to travel around the fictional universe, and the production of these ships from raw minerals (ore) plays a big part in its economy. Though far from the only option a player in *EVE* has to make money, a reasonable classification of player jobs is

- miners: extracting ore from the game,
- industrialists: converting ores into ships,
- consumers: engaging in missions the game provides or in *Player-Versus-Player* (PVP) activities, destroying ships in the process.

I built an ABM based on this classification, with agents shifting between the three sectors. Rather than focusing on e.g. 'fundamental' price and historical price based agents [CPZ01], this model focuses on the creation and consumption of goods, and the market that emerges from agents trying to maximize their profit. The model is then driven by the law of supply and demand on one side, with agents having the desire to possess a ship, and profit maximization on the other. Some potential benefits of this approach include:

- The life cycle of a good is similar to one in the real world, where goods are created from raw materials and later consumed. This cycle may be of importance in the emergence of stylized facts in certain markets.
- If stylized facts emerge, they are shown not to be limited to ABMs which limit themselves to different types of financial traders.
- In addition to following market prices, this approach allows us to later on use the model to investigate the impact of several parameters on the wealth distributions between different agent jobs. This could then be compared to the wealth distributions of the real world.

The initial goal of the model is to create a stable market by mimicking *EVE* player behavior. If an equilibrium is reached, the price behavior can be studied where stylized facts may or may not be present.

In the following sections, the implementation of the model is explained in greater detail and its robustness is tested. Subsequently, the ABM's market is investigated in in the same way as those of real and virtual markets in the previous chapters..

4.2 IMPLEMENTATION OF THE MODEL

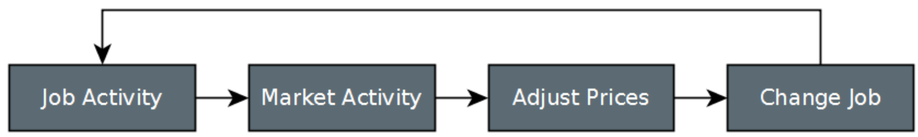

Figure 27: Flowchart depicting general agent activity.

Consider a set of N agents who wish to maximize their profit. At any time, an agent has one specific job, providing the agent with either ores (miner), tools of production (industrialist), or a source of cash income (consumer). The ores are used in the production of ships, which occasionally get destroyed. The source of cash income for consumers mimics tasks that players in *EVE* can complete in order to gain money.

At every timestep $t = 1, 2, \ldots$, an agent undergoes the following steps:

1. The agent practices his job,
2. The agent attempts to sell (buy) goods on the market,
3. The agent adjusts his prices to sell (buy) goods,
4. The agent potentially changes jobs if he expects higher profits elsewhere.

Fig. 27 shows a flowchart of this general behavior. In the following subsections, I explain the four steps in detail.

4.2.1 *Job activity*

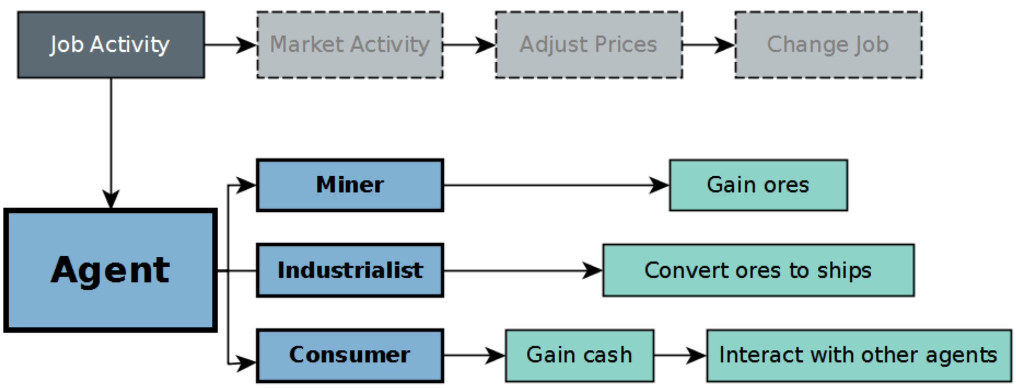

Figure 28: Flowchart depicting agent job activity.

Fig. 28 depicts basic job activity in the model. As a miner, an agent receives a set amount of ores Λ_{ore}^{inc} every time step. As an industrialist, an agent converts up to Λ_{ship}^{inc} ores in ships per turn, with a 1-1 ratio. As a consumer, an agent receives Λ_{cash}^{inc} money. In addition, consumers are allowed to interact with other agents.

They do so with a probability of ξ_{int}^{cons} per time step, consuming theirs or their target's ship in the process.

$$Job = \begin{cases} \mathrm{Ores}(t+1) = \mathrm{Ores}(t) + \Lambda_{ore}^{inc} & \text{if miner,} \\ \begin{cases} \mathrm{Ores}(t+1) = \mathrm{Ores}(t) - \min(\mathrm{Ores}(t), \Lambda_{ship}^{inc}) \\ \mathrm{Ships}(t+1) = \mathrm{Ships}(t) + \min(\mathrm{Ores}(t), \Lambda_{ship}^{inc}) \end{cases} & \text{if industrialist,} \\ \begin{cases} \mathrm{Cash}(t+1) = \mathrm{Cash}(t) + \Lambda_{cash}^{inc} \\ \text{Agent interaction with probability } \xi_{int}^{cons} \end{cases} & \text{if consumer.} \end{cases} \tag{31}$$

4.2.2 *Market activity*

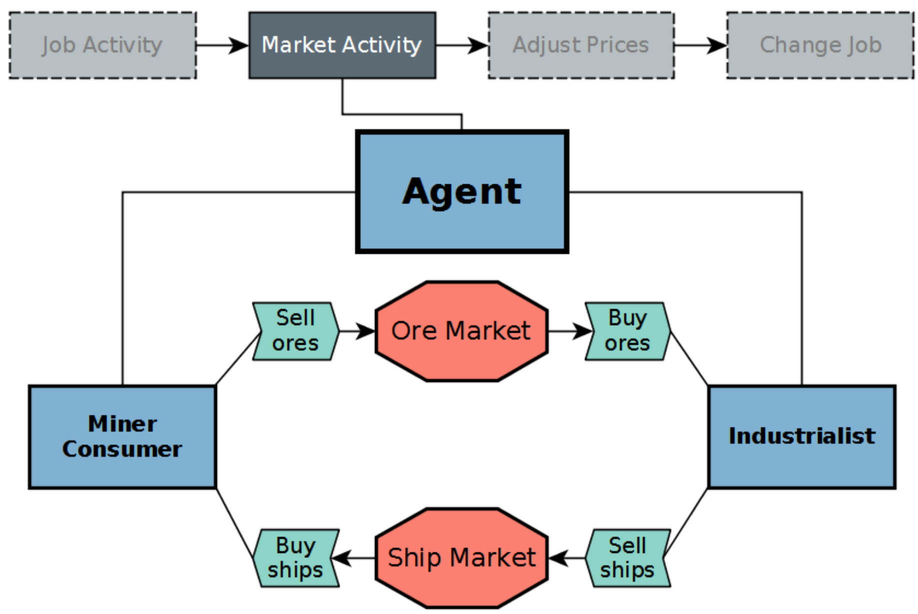

Figure 29: Flowchart depicting market activity of agents.

Agents participate in the market by placing orders that last for τ_{order} time steps. Orders can take two forms: *buy orders*, a request to buy a certain amount of goods for a certain price, and *sell orders*, a request to sell a certain amount of goods for a certain price. An agent seeking to sell a good will first look at all available buy orders on the market for that same good. If buy orders are present with a higher price than his sell price $p^{sell}(t)$, the agent will fulfill the most expensive one. If no such order is available, the agent will place a sell order for $p^{sell}(t)$. A similar reasoning holds for placing buy orders at price $p^{buy}(t)$, fulfilling the cheapest available sell orders first if their price is lower than $p^{buy}(t)$. Agents will only place an order after attempting to fulfill available complimentary orders. This implementation of market activity replicates the one of *EVE*.

In the model, two types of goods are traded: ores and ships. Ores are traded in a continuous manner, whereas ships are discrete. Miners sell all ores in their possession, and buy ships if they have none. Consumers, potentially having ores left if they were a miner in the previous time steps, do the same. Industrialists buy ores, and sell the ships they create. This market activity is shown in Fig. 29.

4.2.3 *Price adjustment*

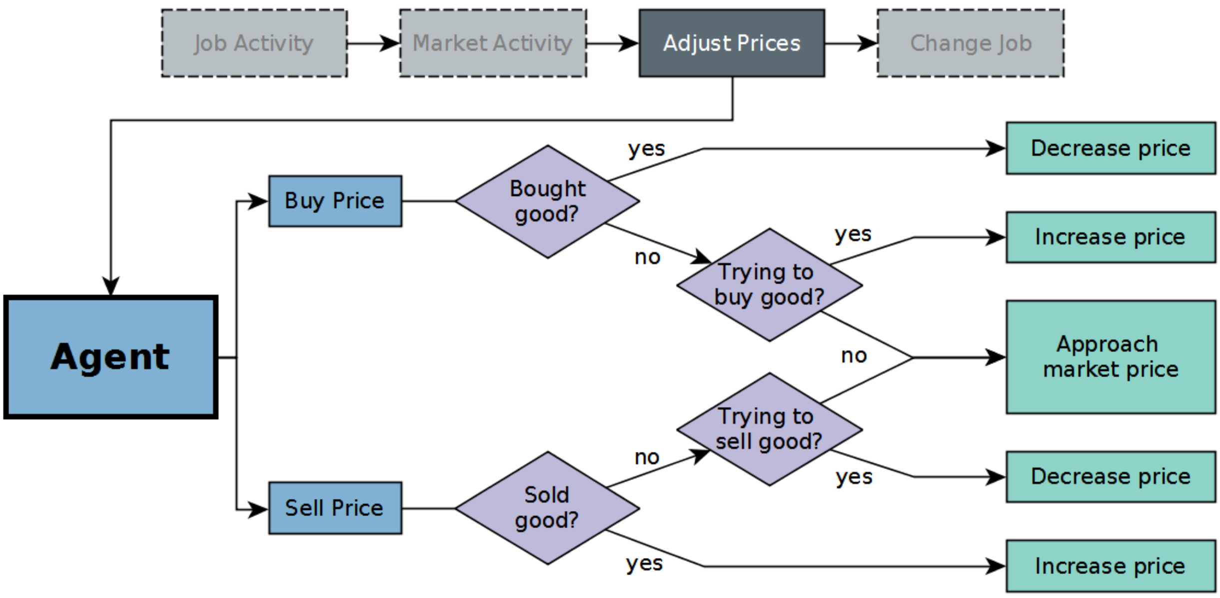

Figure 30: Flowchart depicting agent buy and sell price modification.

An agent keeps four prices $p^{sell,buy}_{ore,ship}(t)$ in his memory, at which he will place sell (buy) orders for ores (ships). Every time step, these prices are multiplied or divided by a factor $\rho > 1$, in such a way that the agent maximizes his profit.

For sell prices of a good,

- If the agent sold a good in the previous time step, he will increase his $p^{sell}(t)$ in an attempt to increase his profit.
- If the agent sold nothing but has outstanding sell orders, $p^{sell}(t)$ is too high and he will decrease it.
- If none of the above are true, the agent approaches the cheapest sell order remaining on the market. If no sell orders remain, the agent will approach the most expensive buy order remaining on the market. If no such orders are available either, the agent will decrease his sell price.

For buy prices, a similar reasoning holds,

- If the agent bought a good, he will decrease $p^{buy}(t)$ in an attempt to lower his expenses.
- If the agent bought nothing but has outstanding buy orders, $p^{buy}(t)$ is too low and he will increase it.
- If none of the above are true, the agent approaches the most expensive buy order remaining on the market. If no buy orders remain, the agent will approach the cheapest sell order remaining on the market. If no such orders are available either, the agent will increase his buy price.

This is represented by Fig. 30. Only local information is used in the adjustment of an agent's prices: his own price sales history, and the price extrema of the openly accessible market.

4.2.4 *Job change*

After having practiced a job for τ_{job} time steps, the agent will start to potentially change jobs. He does so by comparing his current profit with the expected profit of other jobs every time step.

The agent's current profit $\bar{P}_k(t)$ takes all gains and losses of cash into account, and is calculated as a moving average over a time interval $t - \tau_{job}, ..., t - 1, t$.

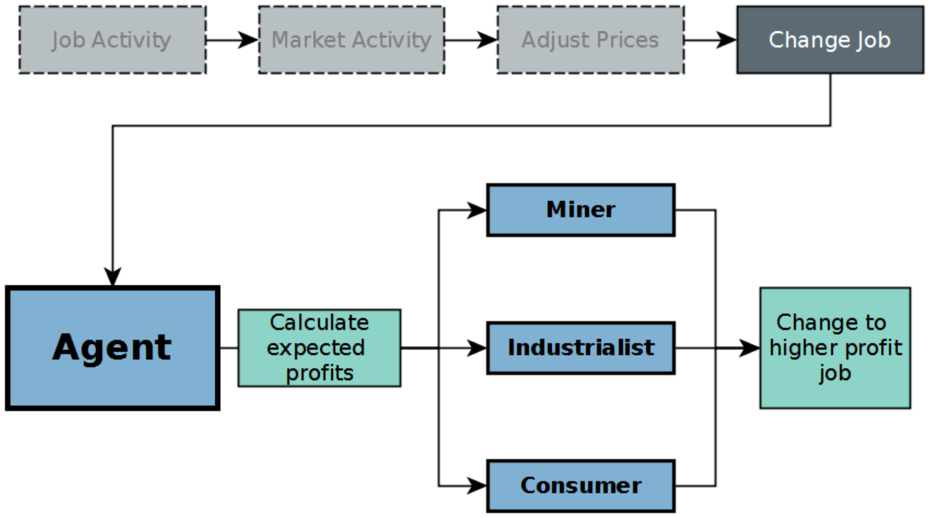

Figure 31: Flowchart depicting agent job change.

The expected profit $P_{j,exp}(t+1)$ is calculated using the agent's current prices and his expected income of goods:

$$P_j^{exp}(t+1) = \begin{cases} p_{ore}^{sell}(t) \cdot \Lambda_{ore}^{inc} & \text{for } j = \text{miner,} \\ \left(p_{ship}^{sell}(t) \ - \ p_{ore}^{buy}(t)\right) \cdot \Lambda_{ship}^{inc} & \text{for } j = \text{for industrialist,} \\ \Lambda_{cash}^{inc} & \text{for } j = \text{for consumer.} \end{cases} \tag{32}$$

These expected profits are then multiplied by a factor μ_j. The value of μ_j is built up when the agent practices job j, and accounts for all potential differences between the expected profit calculation and actual job profit. In the base model, these differences come from mismatched supply and demand.

$$\begin{aligned} &P_j^{exp}(t+1) \to \mu_j \cdot P_j^{exp}(t+1) \\ &\text{with } \mu_j = \left\langle \frac{\bar{P}_j}{P_j^{exp}(t+1)} \right\rangle_j \end{aligned} \tag{33}$$

where $\langle \cdot \rangle_j$ indicates the moving average over time steps where an agent practiced job j. To limit the amount of parameters of the model, the time over which this moving average is taken is set to τ_{job}.

After calculating the expected profits for other jobs $P^{exp}_{l\neq k}(t+1)$, the agent compares them to his current profit $\bar{P}_k(t)$. Should $P^{exp}_{l\neq k}(t+1) > \bar{P}_k(t)$, the agent will potentially change from job k to l. The probability of this change grows exponentially to a maximum value of ξ_{change} , shown in Eq. 34.

$$Pr(k \to l) = \begin{cases} \frac{\xi_{change}}{e^{A \cdot B} - 1} \cdot e^{A \cdot (P^{exp}_{l\neq k}(t+1) - \bar{P}_k(t))} & \text{if } P^{exp}_{l\neq k}(t+1) - \bar{P}_k(t) < B \\ \xi_{change} & \text{otherwise,} \end{cases} \tag{34}$$

where A determines the strength of the exponential increase, and B determines at which profit difference the maximum value ξ_{change} is reached. In the model, both were set to 2.

4.2.5 *Additional details*

Consumers receive cash every time step, which could cause constant inflation of prices. To prevent this, a cash sink may be required. Inspired by video games, this was implemented here in the form of agents 'quitting'. At every time step, an agent has a probability ξ_{quit} to quit, replacing him with a new agent one. The starting job of the new agent is randomly chosen, and its starting prices lie anywhere between the current orders on the market.

The agents as described above fulfill the requirements mentioned in section 1.6.2. In the next section, the equilibrium state of the model and its resulting price behavior are examined.

4.2.6 *Summary of symbols and parameters*

- N: number of agents in the simulation
- T: total time steps of the simulation
- t: current time step
- τ: time difference, or *lag*
- j : an agent's job (miner, industrialist or consumer)
- $n_j(t) = \dfrac{N}{N_j(t)}$, $j = m, i, c$: fraction of miners, industrialists and consumers at time step t
- $p^{sell,buy}_{ore,ship}(t)$: an agent's individual sell (buy) price for ores (ships) at time step t
- $\bar{p}_{ore,ship}(t)$: the market's average price of all sold ores (ships) at time step t
- $P_j(t)$, $j = m, i, c$: actual profit of an agent at time step t with job j
- $P^{exp}_j(t)$, $j = m, i, c$: expected profit of an agent at time step t for job j

Non-arbitrary parameters

- Λ^{inc}_{ore}: a miner's ore income per time step
- Λ^{inc}_{cash}: a consumer's cash income per time step
- Λ^{inc}_{ship}: an industrialist's ship production per time step
- ξ^{cons}_{int}: a consumer's random interaction chance per time step
- ρ: an agent's multiplicative price change per time step
- τ_{job}: amount of time steps over which actual profit is averaged, and after which an agent can change jobs
- τ_{order}: duration of a market order in time steps
- ξ_{quit}: an agent's agent quit chance per time step

4.3 BASE MODEL RESULTS

In this discussion, the *base model* will refer to the implementation described above, with the following parameters:

- $n_j(t=0) = \begin{cases} 1 & \text{for } j = m \\ 0 & \text{otherwise,} \end{cases}$
- $\rho = 1.003$,
- $\Lambda^{inc}_{ore} = \Lambda^{inc}_{cash} = \Lambda^{inc}_{ship} = \zeta^{cons}_{int} = 1$,
- $\tau_{job} = 50$,
- $\tau_{order} = 5$,
- $\zeta_{change} = 0.1$,
- $\zeta_{quit} = 0.001$.

The results of such a run can be seen in Fig. 32 for $N = 10000$ and $T = 20000$. On these images, subfigure A) shows the number of agents with a specific job, and B) and C) show the average prices of completed sell orders for ores and ships per time step. After an initial thermalizing phase, the system reaches a dynamic equilibrium at approximately 3000 time steps.

The ore and ship prices in B) and C) are not comparable to those of *EVE* or the real world. A high degree of periodicity is present and the system switches between states with different agent job fractions. The model behaves according to changes in supply and demand, caused by the interplay between the agent job distribution and prices. If the demand for e.g. ships is too high, industrialists will increase their sell price. As prices go up, expected profit of industrialists will go up and agents start to change to that job. At a certain point, supply exceeds demand, and prices will start decreasing. This cycle repeats itself and leads to periodic price behavior.

This periodicity is reflected in the markers for stylized facts seen in Fig. 33. Subfigures B) and C) show autocorrelations of both normal and absolute log-returns that are much larger than zero. The cause of this can be seen in Fig. 34, where the log-returns show periodic behavior. This is a direct result of the regular increasing and decreasing of prices. Upward slopes of the general price motif generate positive log-returns, whereas downward slopes generate negative ones.

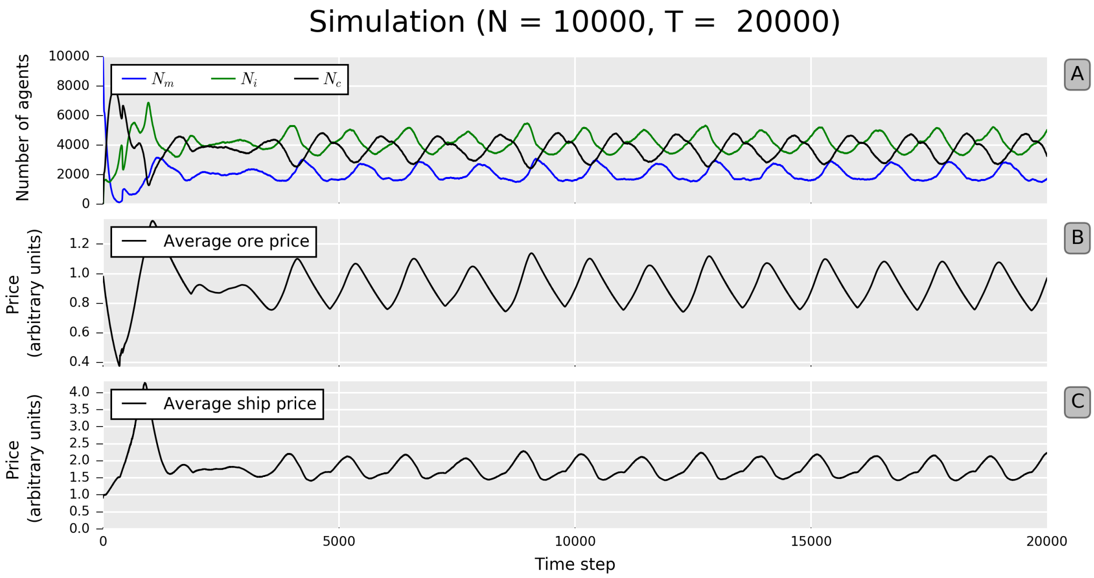

Figure 32: Results of the base model as described by 4.3. Subfigure A) shows the distribution of agent jobs; B) and C) show the average prices of completed sell orders for ores and ships.

The PDFs in D) exhibit bimodal behavior, as seen in the bimodality coefficient in F). Like the autocorrelations, this results from the periodic nature of the prices. Prices move up (down) in consistent patterns and the probability of finding positive (negative) log-returns increases.

In the next section, the model is tested for robustness with regards to a variety of parameters. Afterwards, the ABM is modified as to better approximate real markets.

4.4 ROBUSTNESS OF THE MODEL

Ideally, the equilibrium state of the ABM should not be highly dependent on the arbitrary choices made in the base implementation. As long as the values of these parameters are within reason, their impact should remain limited and predictable. The following lists a number of simulations that tested the general behavior of the model.

4.4.1 *Number of agents*

As mentioned in section 1.6.2, some ABMs only operate properly within a certain range of agents. The equilibrium arising in the

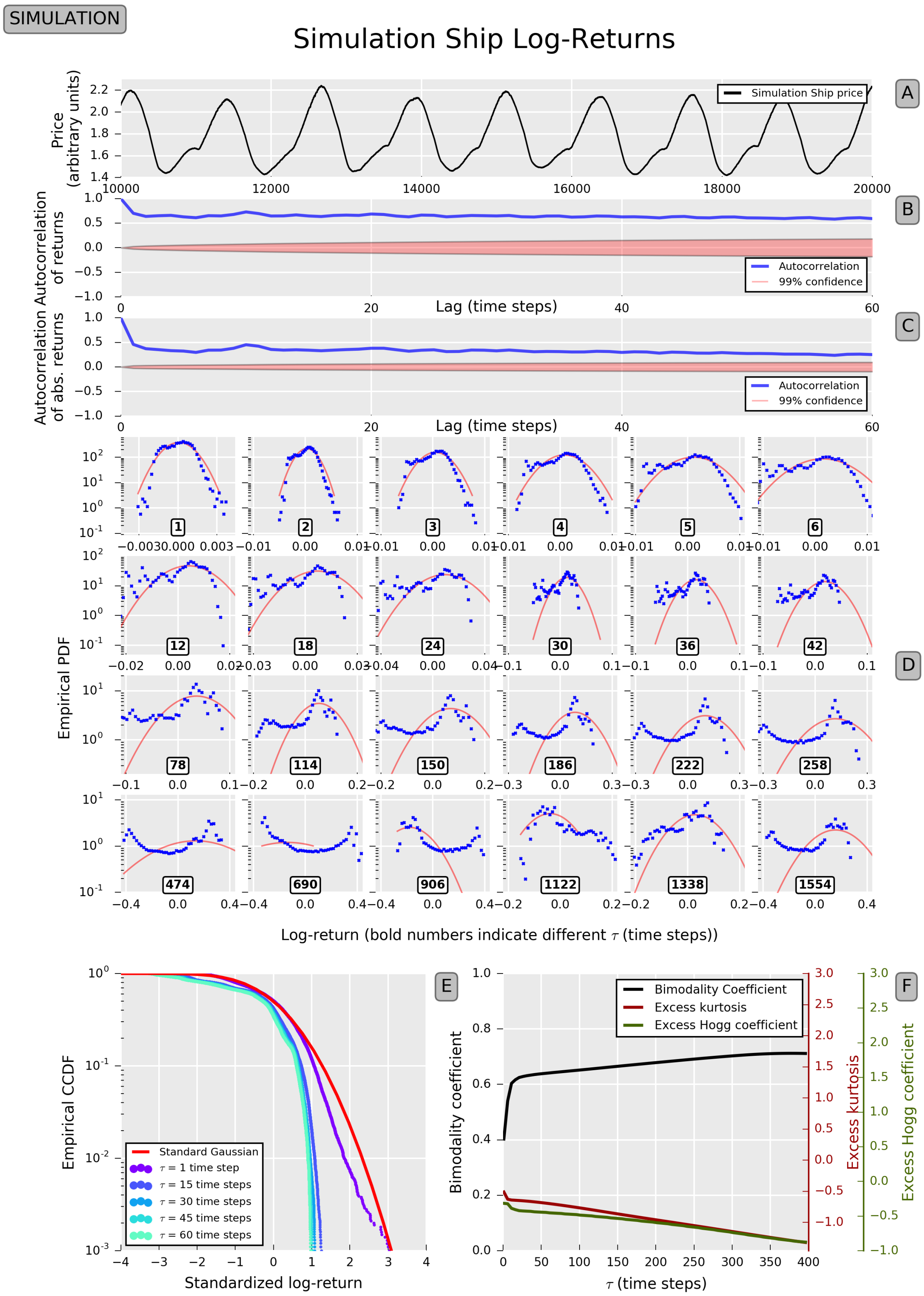

Figure 33: Stylized facts of the base model as described by 4.3.

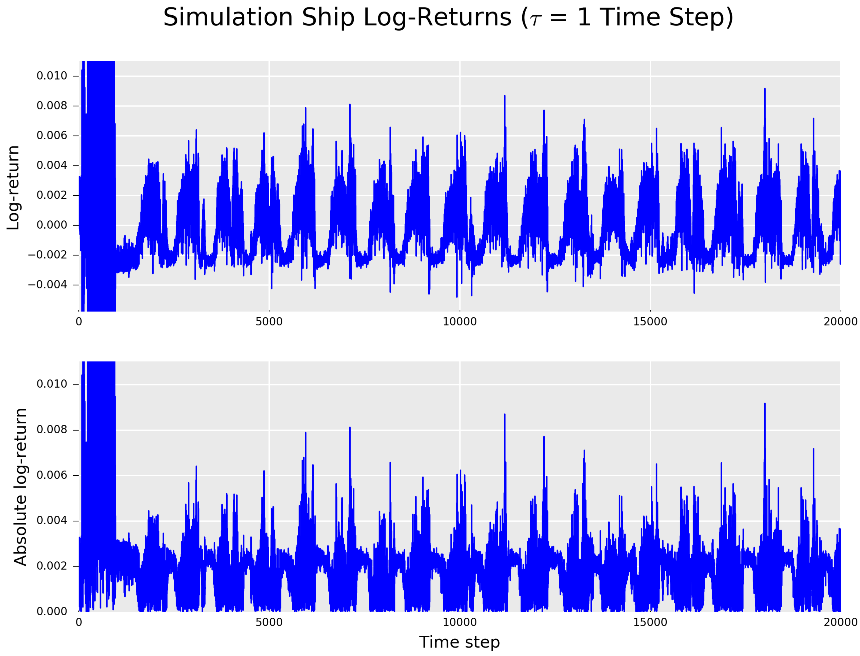

Figure 34: Log-returns of base model as described by 4.3.

ABM of this thesis should be unaffected by the agent count. A higher number of agents will increase demand for goods, but also their supply. Only when the amount of agents drop below a reasonable number will the market break down as a stable supply of goods can no longer be guaranteed.

Fig. 35 shows simulation results for N = 100 (1), 1000 (2) and 10000 (3) agents. Their equilibrium states are comparable. As expected, the fluctuations are larger for a smaller system size. The decreasing trend in subfigure 1C) is a statistical fluctuation, and does not indicate a significant difference from the base model results.

4.4.2 *Starting distribution of agent jobs*

The starting distribution of agent jobs may prevent the model from reaching a stable equilibrium. Fig. 35 shows a comparison between three extreme starting situations: all miners (1), all industrialists (2) and all consumers (3). The lower prices at the start of 2C) are seen in other simulations as well (e.g. Fig. 35 1C) and do not indicate any significant difference from the base model results. The ABM reaches the same equilibrium in all cases.

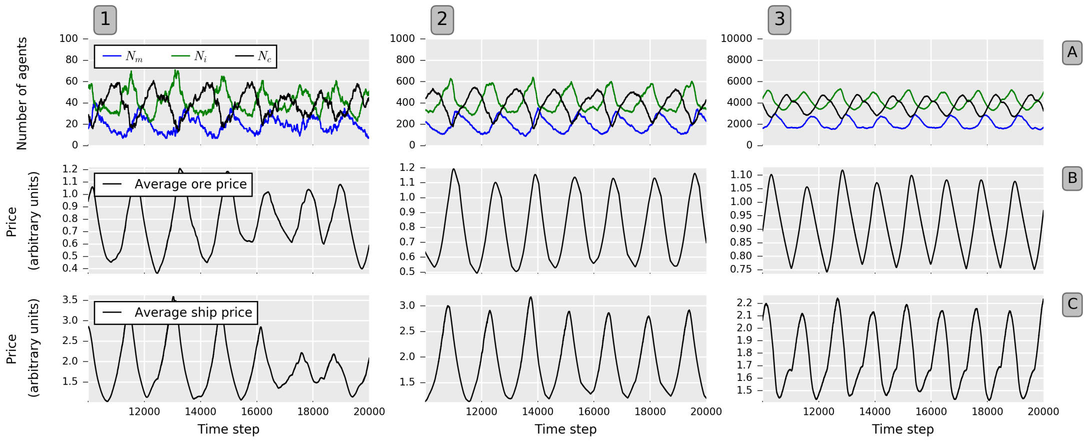

Figure 35: Comparison of base model implementations with varying number of agents. Subfigure 1-3) show simulations with respectively N = 100, 1000 and 10000.

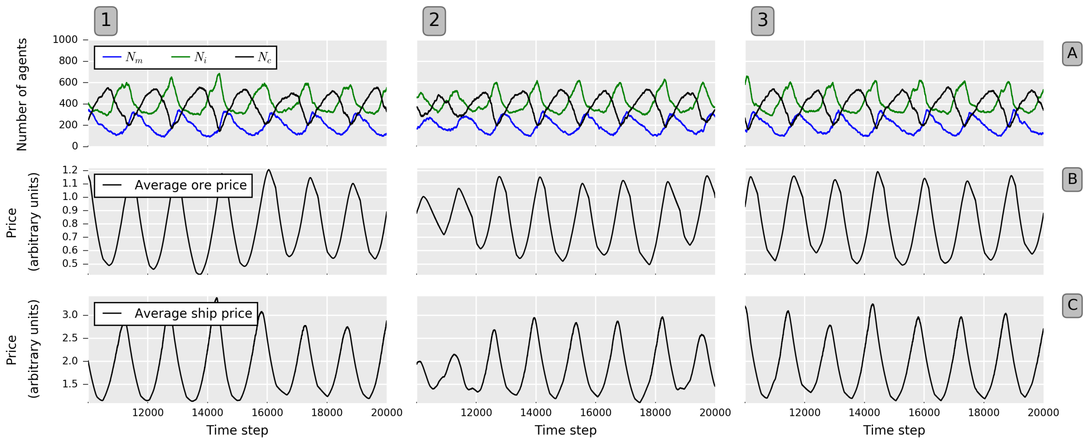

Figure 36: Comparison of base model implementations with changed starting agent distributions. Subfigure 1-3) respectively start with all miners, all industrialists and all consumers.

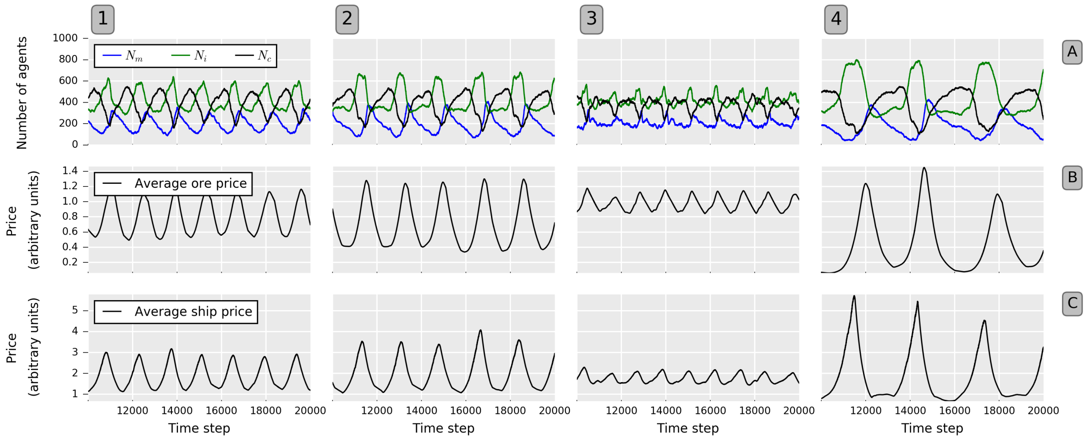

Figure 37: Comparison of base model implementations with changes to the arbitrarily chosen parameter values of Eq. 34. Subfigure 1) shows the base model, 2) increases the strength of the exponential increase A to 3, 3) lowers the profit difference B at which ξ_{change} is reached to 1, and 4) lowers the maximum change probability ξ_{change} to 0.01.

4.4.3 *Job change parameters*

The probability to change jobs was implemented with an exponential function described in Eq. 34. Fig. 37 shows some variations of its parameters: the maximum probability to change per turn ξ_{change} and the scaling factors A and B. A higher exponential increase A in 2) has no significant impact on price behavior when compared to the base model 1). In 3), agents change at lower profits, which decreases the magnitude of price fluctuations as the changes in supply and demand happen more rapidly. In 4) agents are slower to move between jobs, causing the opposite effect: price fluctuations are larger, and their periodicity decreases. In summary, the general behavior of the model is not altered by changing these parameters.

4.4.4 *Price change*

In Fig. 38, a comparison can be seen between different values of price change ρ. The base model implemented $\rho = 1.003$ whereas for subfigures 1-4), this value is respectively 1.03, 1.005, 1.001 and 1.0003. The effect is as expected. For large ρ, changes in supply and demand quickly inflate the prices, and agents frequently switch jobs. When ρ is small, agents are given more time to adapt

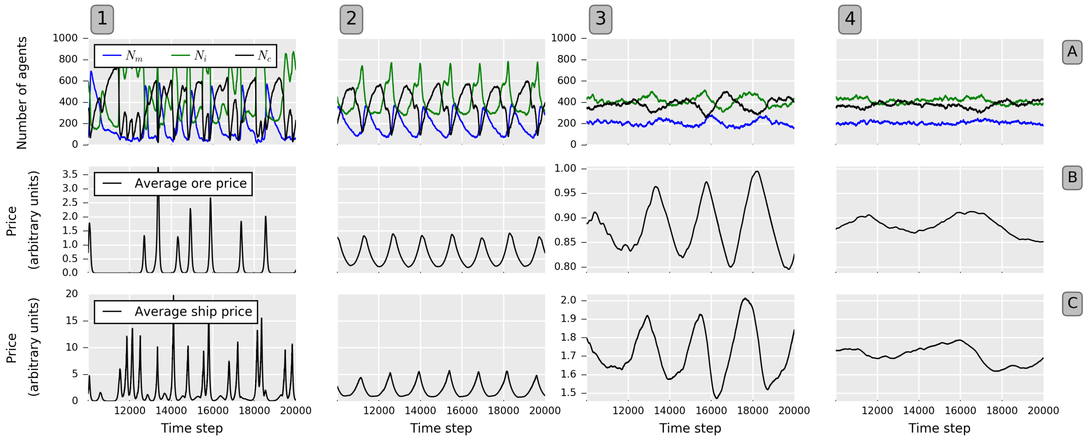

Figure 38: Comparison of base model implementations with changed values of ρ. The base model uses a $\rho = 1.003$; subfigures 1-4) use respectively 1.03, 1.005, 1.001 and 1.0003.

to small profit increases. This results in a more stable agent job distribution, and a slower market.

4.4.5 *Order duration*

As market order cannot be updated, the duration of orders may affect the general behavior of the model. Once a sell or buy order is placed, it will either be fulfilled at its initial price, or expire after τ_{order} time steps.

Placing a sell order at a high price removes the goods from an agent's inventory. If τ_{order} is small, the agent will quickly recover these goods and re-place them for his adjusted price. Upon reaching a suitable sell price, he can sell these accumulated goods all at once. In contrast, for large τ_{order}, expired orders are returned and re-placed incrementally later on. These may lead to different volumes being sold at different time steps, affecting the average price.

Ongoing buy orders on the other hand will prevent agents from placing additional buy orders at higher prices, as they have no desire risk the purchase of excess goods. A small τ_{order} will result in the buy orders reflecting the agent's buy price at all times–similarly to sell orders. For large τ_{order} however, the agent will continuously decrease his price: no goods are being purchased, but he cannot change his buy order. When the order eventually expires, a new buy order will be placed for a substantially lower price.

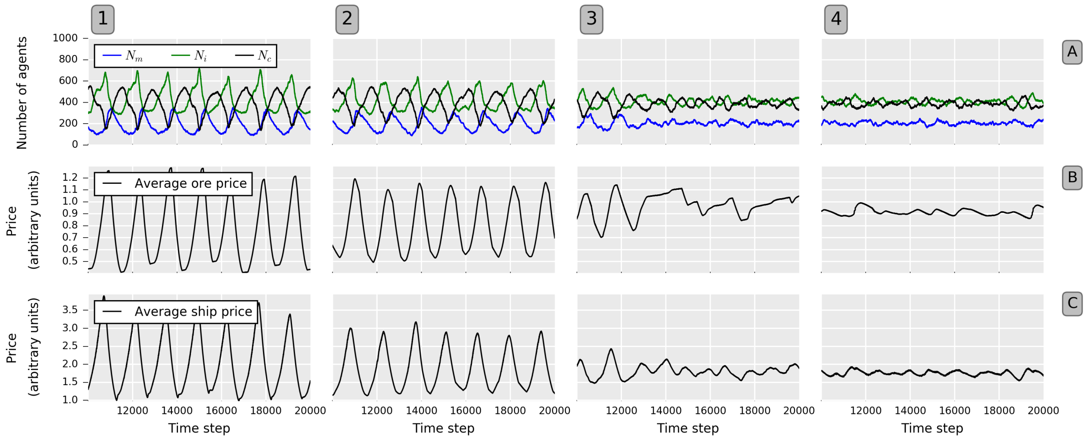

Figure 39: Comparison of base model implementations with changed values of τ_{order}. Subfigure 1) lowers τ_{order} to 2 timesteps, 2) shows the base model with τ_{order} = 5, 3-4) increase τ_{order} to respectively 10 and 25.

The combination of these effects make the impact of changes in τ_{order} difficult to quantify. Fig. 39 shows simulation results for τ_{order} of respectively 2, 5, 10 and 25. Higher values of τ_{order} appear to decrease the price fluctuations and stabilize the agent job distribution. This could indicate that when agents change to jobs with a higher profit, initial price values prevent them from immediately impacting the market. When they return to the market after their order expires, heavily adjusted prices may then cause a mean-reversing effect.

4.5 FINE-TUNING THE MODEL

The base model seems to be robust to changes in its parameters. However, the resulting price behavior is too periodic to be comparable to that of real markets. Several methods to introduce randomness to the system were investigated, and in this section I argue how their implementation might lead to a better approximation of real markets.

4.5.1 *Uncertainty of target price*

In the base model, agents not participating in a type of market activity will approach the price extrema of orders still on the market. A realistic agent might not be interested in keeping track of such a market, which introduces uncertainty to this target price.

A price uncertainty of x on price p causes an agent to see that price at any value between $p \cdot x$ and p/x. For sell prices, the probability is skewed to the increasing side: agents prefer to sell for more (potentially increasing their profit) rather than less. The opposite holds for buy prices.

Fig. 40 shows the results for various degrees of uncertainty of the target price. Prices are confined within a smaller region and the agent job distribution is constant. A possible explanation of this is the following. When prices are increasing due to a change in supply and demand, agents that change their job with perfect information will have a price close to that of their new competitors, and they will contribute to the general behavior of the price. With imperfect information, some agents will arrive at the new job with higher prices, and others with lower ones. Those with optimistic prices (e.g. high sell prices) will have no impact on the market as their orders will not be fulfilled. Those with pessimistic prices will steal customers from the old competitors. This counteracts the change in price that led them to change jobs in the first place, resulting in a mean-reversing effect that stabilizes the market. As prices fluctuate less, so does the expected profit of agents, and their job distribution remains more constant.

In summary, the base model can be said to exhibit a form of herding behavior, where all agents aim for the same price. Adding uncertainty limits this, reducing the regularity and size of price changes.

4.5.2 *Random price changes*

The agents in the base model are perfectly rational, increasing or decreasing their price based on sales made at every time step. Real agents may be more fickle in their behavior, adjusting prices randomly rather than based on their history. This was implemented by changing an agent's prices at random with probability ξ_p every time step, besides the normal history-based change.

Fig. 41 shows the different model equilibriums for several values of ξ_p. The effect of random price adjustment is similar to that of price uncertainty in the previous section. Both implementations result in agents occasionally acting against their best interest, which slows down changes in average prices. The agents

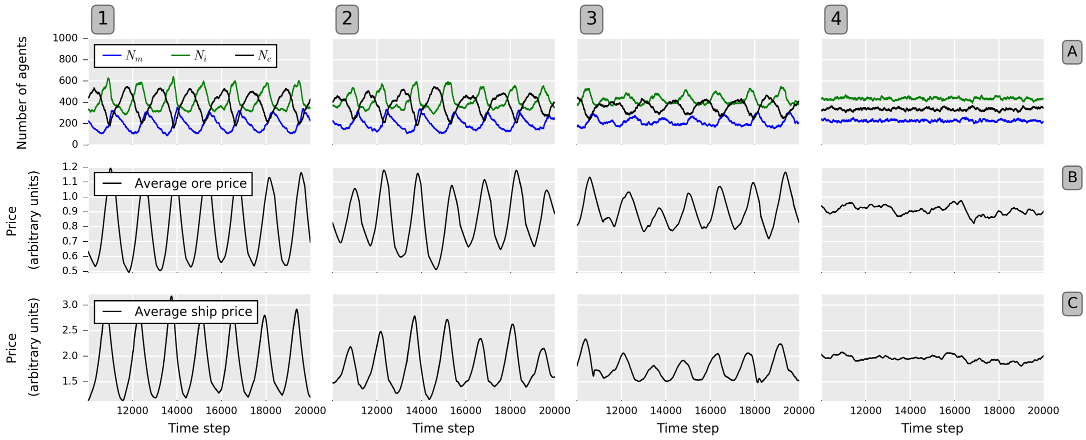

Figure 40: Comparison of base model implementations with added target price uncertainty. Subfigure 1) shows the base model, 2-4) add price uncertainties of respectively 1.1, 1.5 and 2.

are given more time to adapt to changes in supply and demand, and less extreme prices are seen.

4.5.3 *Successive shocks*

Finally, I apply shocks to the system, defined as an abrupt change in a parameter that slowly decays over time. These replicate external events that impact the market in some way. In the real world, an example could be a temporary conflict affecting the supply of a good. To mimic this, Λ_{ore}^{inc} was modified at random time steps by a randomly chosen amount, lasting for a randomly chosen duration. These random values were drawn from a distribution of the form

$$A \cdot (-\ln x + 1)^{\pm B} \tag{35}$$

with A and B scaling parameters and x indicating uniform numbers between zero and one. The resulting shocks have a magnitudes that are generally confined between 0.5 and 2, for durations between 75 and 250 time steps.

Results of a simulation applying these shocks with a probability of 0.004 per time step are shown in Fig. 42. The shocks do not appear to affect the general price motif, but they increase the volatility during certain intervals. As they alter the supply of ores which becomes reflected in the prices, this is an expected result.

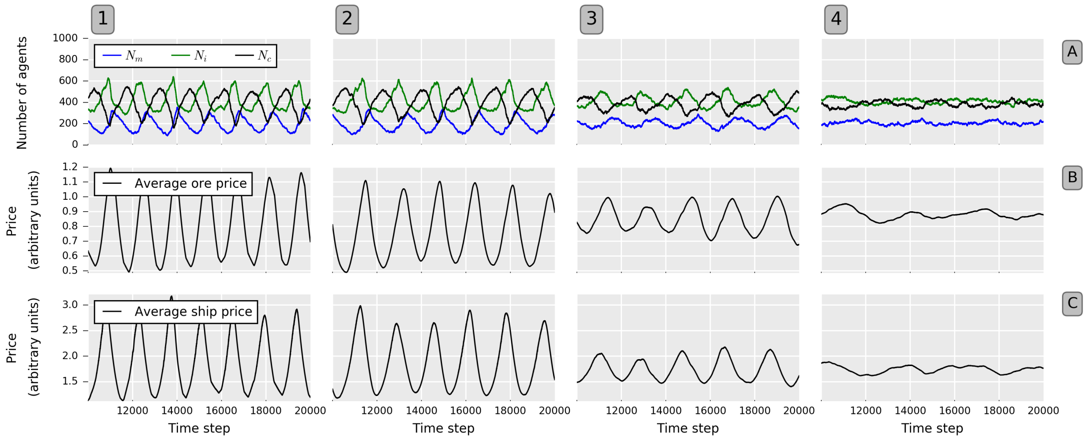

Figure 41: Comparison of base model implementations with added random price changes. Subfigure 1) shows the base model, 2-4) add random change probabilities of respectively 0.1, 0.5 and 0.75 per time step.

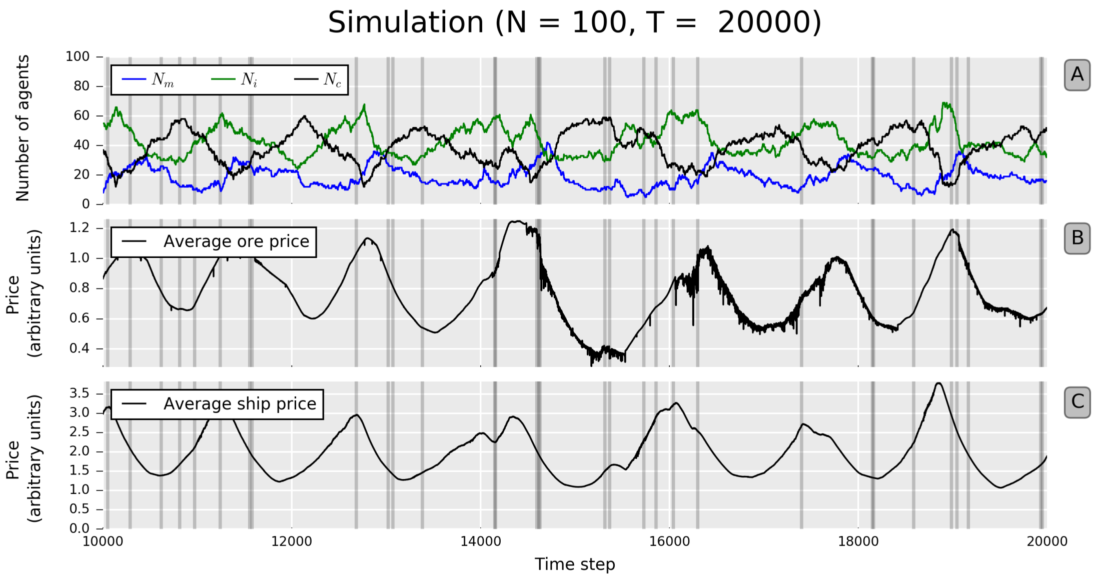

Figure 42: Results of the base model with added ore income shocks with a probability of 0.004 per time step, indicated by the vertical lines.

4.5.4 *Impact on stylized facts*

The previously discussed addition to the model may improve its ability to generate stylized facts. Price uncertainty and random price changes reduced the large patterns of price fluctuations, which affects the autocorrelation of the log-returns. Successive shocks give rise to larger fluctuations in the prices, and could generate fat-tailed behavior.

Fig. 43 shows the stylized facts for ores of a simulation that combined a price uncertainty of 2 and a $\xi_p = 0.25$ with random Λ_{ore}^{inc} shocks. The autocorrelation in subfigure B) is slightly negative for shorter lags for reasons that are not directly clear, but it quickly becomes indistinguishable from zero. The autocorrelation of the absolute log-returns in C) remains significantly larger than zero: clustered increases in volatility are introduced by the Λ_{ore}^{inc} shocks. In D-F), fat-tailed behavior is present which decreases as τ increases. Fig. 44 shows that for log-returns of ship prices the stylized facts are less clear, but still visible to some degree.

4.6 DISCUSSION AND FUTURE PROSPECTS

In summary, the base model gives rise to a dynamic equilibrium with prices reflecting supply and demand of goods. The model appears robust to changes in many of the chosen parameters, whose impact on general price behavior is as expected.

Periodic changes in the ABM's supply and demand result in price patterns that appear different from those of real markets. High autocorrelations are present in the log-returns, indicating alternating periods of increasing and decreasing prices. This is also visible in the PDFs of the log-returns, showing bimodal rather than (fat-tailed) Gaussian behavior.

To better approximate real markets, imperfect information, irrationality and random external influences were added to the model. These were implemented in the form of uncertainty with regards to market prices, random price change behavior and shocks. Their addition caused the emergence of stylized facts in both ore and ship prices.

Random GBM as seen in section 2.2 does not exhibit stylized facts, but some degree of randomness appears necessary to generate them in this model. These results cautiously suggest that real markets may operate under a certain balance between randomness and the law of supply and demand. A highly ran-

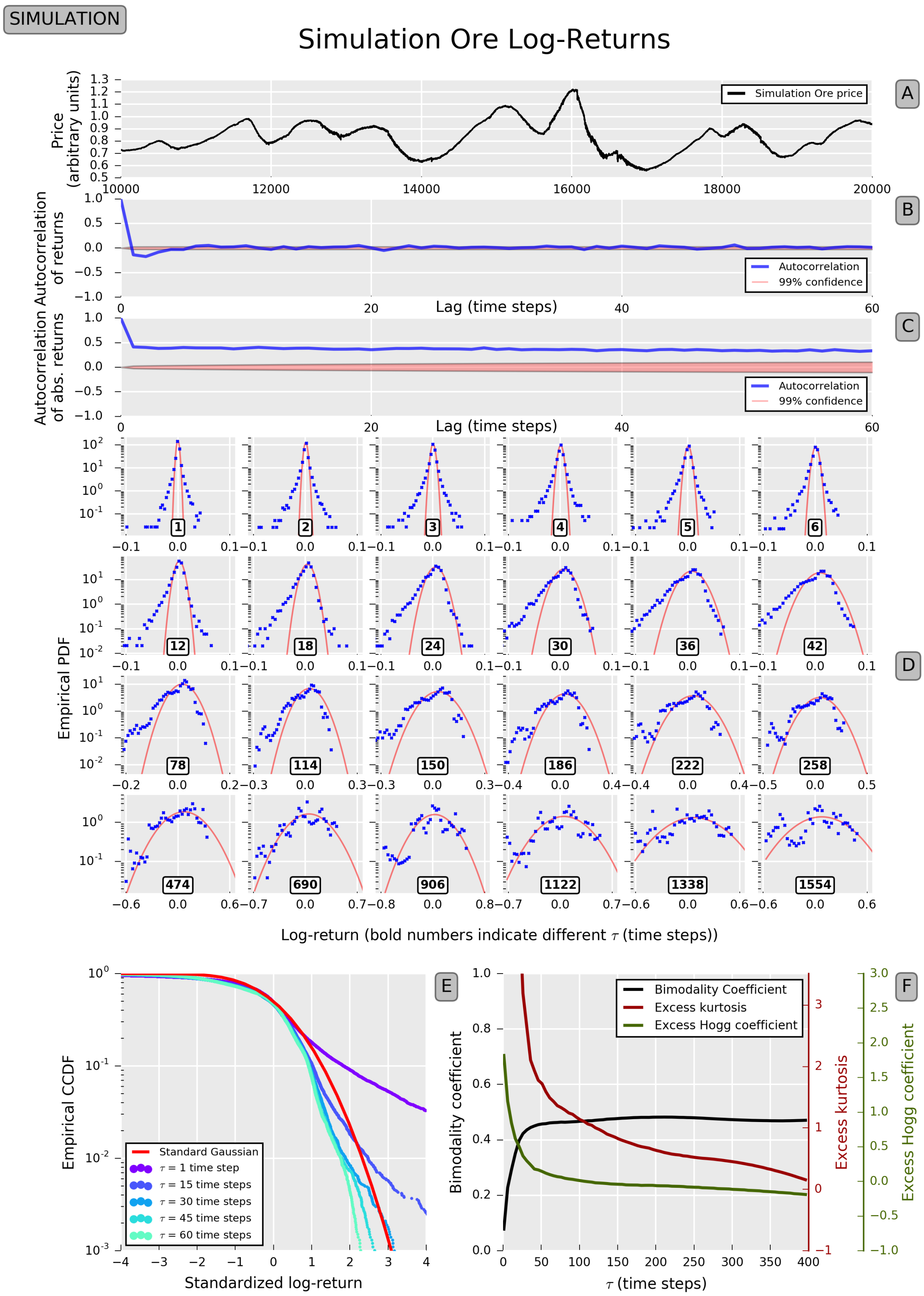

Figure 43: Stylized facts of ore prices in the base model with added price uncertainty of 2, $\xi_p = 0.25$ and random ore income shocks. The strength and duration of each shock was chosen randomly.

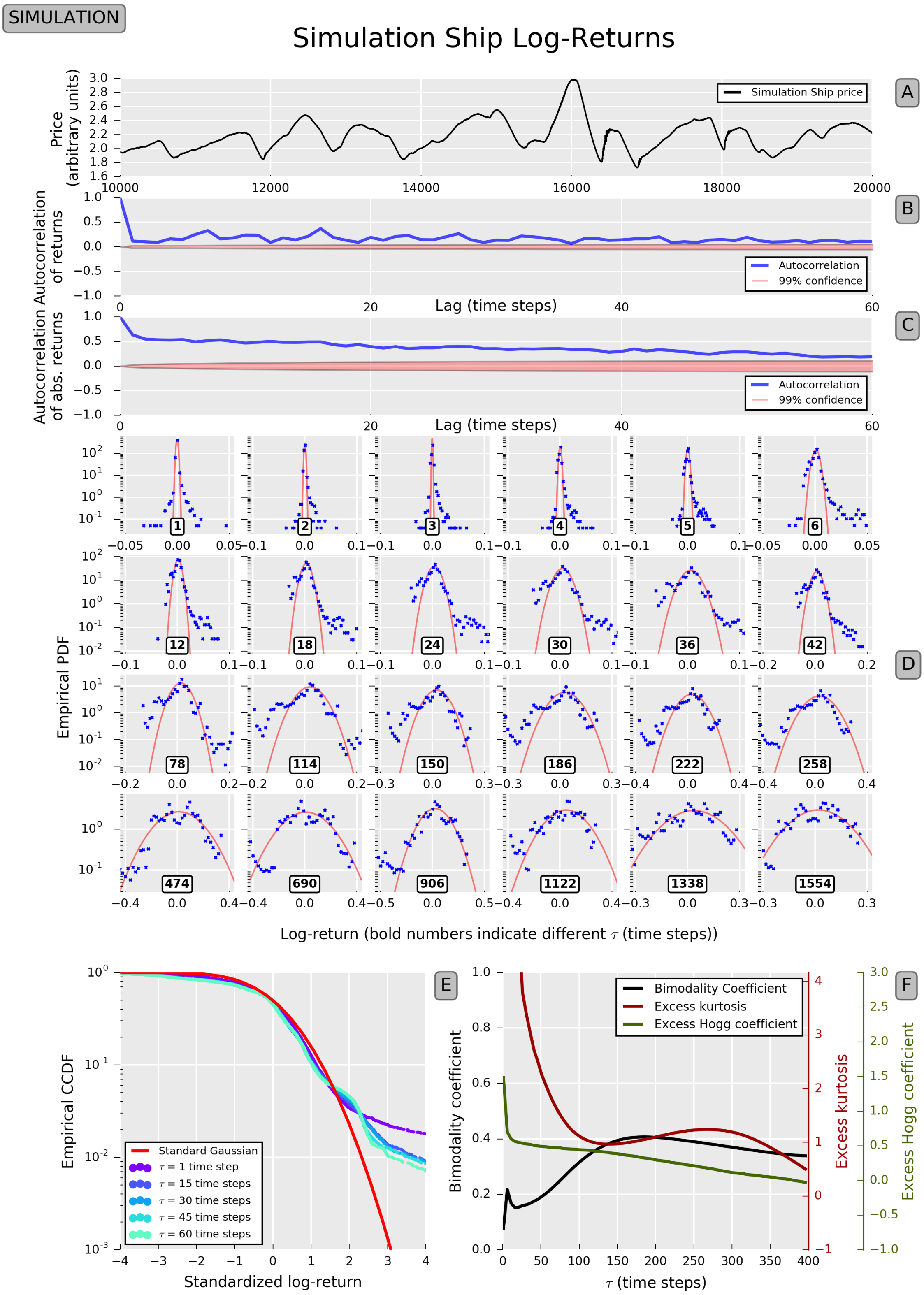

Figure 44: Stylized facts of ship prices in the base model with added price uncertainty of 2, $\xi_p = 0.25$ and random ore income shocks. The strength and duration of each shock was chosen randomly.

dom market may be too unstable, whereas too little randomness would allow participants to abuse patterns to increase their profit. Perhaps markets self-organize to a region where the interplay of both effects gives rise to stylized facts.

However, the EMH mentioned in the introduction states that if the market participants possess more accurate information, the price behavior will appear more random. This is not seen in the ABM that was presented in this thesis. The base model exhibited less random behavior with better information when compared to section 4.5.1. As the model is built using basic assumptions of profit maximization and the law of supply and demand, it remains an open question why its results are contrary to what the EMH suggests.

The ABM created for this thesis was designed to be easily adaptable, and could play a role in further research. Possible additional modifications include:

- Adjusting the price changing behavior of agents as to better approximate that of real market participants.
- Increasing the learning behavior of agents, allowing them to adopt different strategies in order to maximize their profit.
- Implementing additional job change functions in addition to the profit-driven one.
- Making the income of agents dependent on proficiency or time spent at a certain job.
- Expanding the trading mechanisms to allow for direct exchange of goods in addition to through the market.
- Adding more job types, such as traders that attempt to buy low and sell high.
- Increasing the social aspect of agents by allowing them to set up supply chains, form firms, or form loyal bonds with others.
- Adding movement to the model with resource gain depending on the location.

These may or may not be required to create the empirical features seen in real markets. However, as ABMs become more complex, specifying the underlying mechanics that give rise to the emergent features becomes more difficult. If the origin of the stylized facts is to be determined, the simplest approaches ought to be exhausted first. It is my hope that the model presented in this thesis, mimicking the market structure of *EVE* using the law of supply and demand, added to the discussion in a meaningful way.

5 CONCLUSION

The first goal of this thesis was to investigate the prevalence of economic stylized facts. In specific, the fat-tailed behavior, aggregational Gaussianity, volatility clustering and absence of autocorrelation of log-returns were studied. Time series of daily frequency prices were obtained for goods ranging from stocks to raw metals and food. Most of these markets exhibited the stylized facts to some degree– but not all. This may indicate that the underlying mechanisms that generate the facts are not present in some markets, or that additional factors prevent them from appearing.

I then turned to the virtual world of *EVE Online*. *EVE*'s economic structure is unlike that of most real countries and its markets could behave differently. However, similarly to real markets, stylized facts appear with varying consistency. This supports the claim that economies in real and virtual worlds are comparable. Perhaps in the future, video games could serve as a testing ground for economic concepts with social impacts that are difficult to predict, such as basic income.

Finally, an Agent-Based Model inspired by the dynamics of *EVE* was constructed. The base implementation resulted in a dynamic equilibrium with periodic price behavior. The ABM was then tested for robustness under change of its parameters. This had no unexpected impact on the model dynamics. Several implementations were then added to increase randomness in the system. These did not impede the model's ability to reach an equilibrium, but resulted in price behavior that more closely resembled that of real markets.

The base implementation of the model showed no sign of the stylized facts introduced in chapter 1, and observed in real and virtual markets as discussed in chapters 2 and 3. Under additional randomness, implemented in the form of imperfect information, irrational behavior and shocks, they seemed to appear in the ABM. This could mean that a certain degree of noise plays an important role in the emergence of stylized facts.

A APPENDIX

This appendix includes the results for real and virtual markets that were investigated but not discussed in the thesis. All use daily frequency data, whose source will be mentioned in the figure caption. Data gaps (weekends in the real world) were ignored– Friday and Monday were considered successive days. Unless specified in the figure caption, no data cleaning took place.

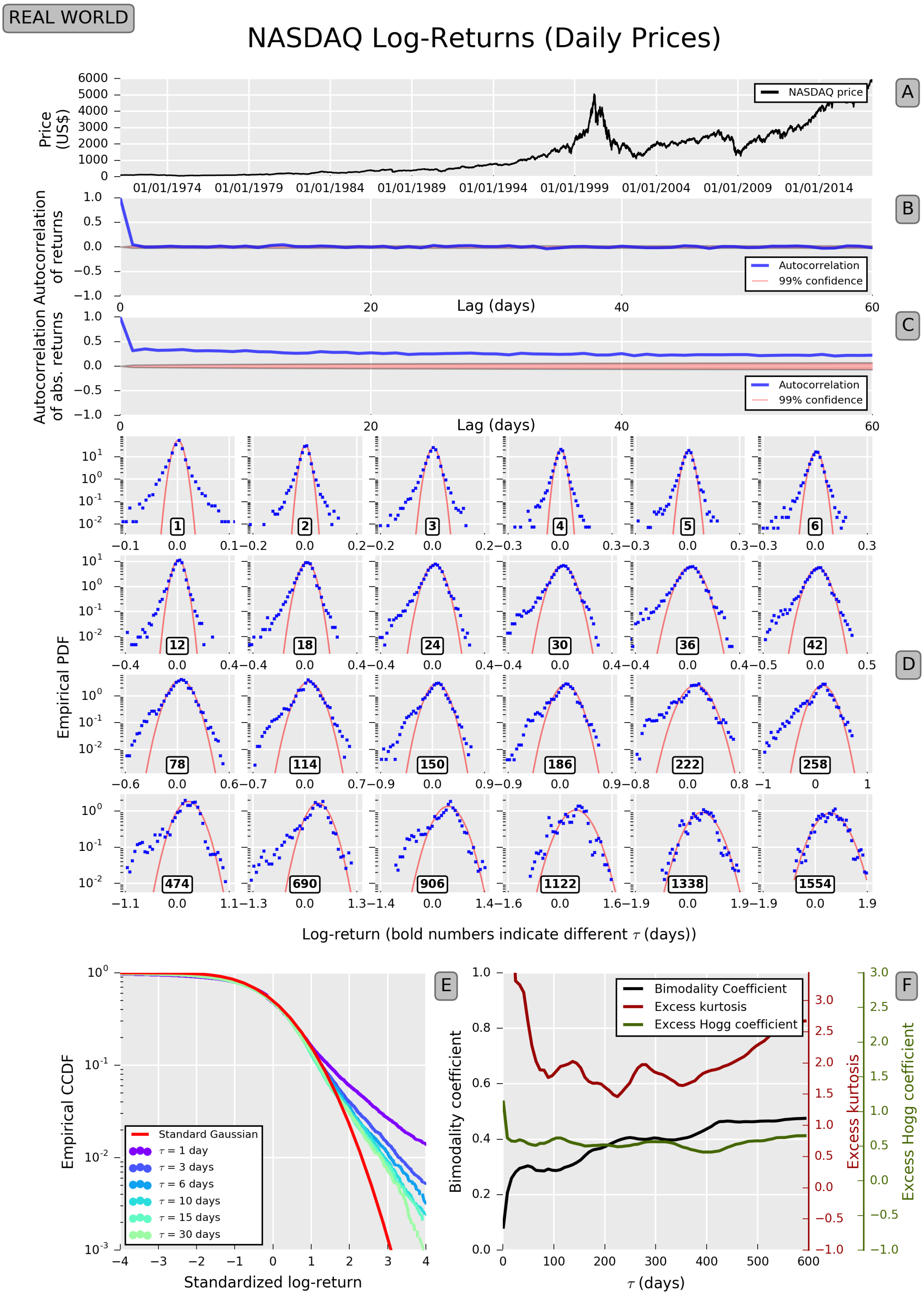

Figure 45: Daily NASDAQ opening price between 1971-02-05 and 2017-02-02, retrieved from [Qua17a].

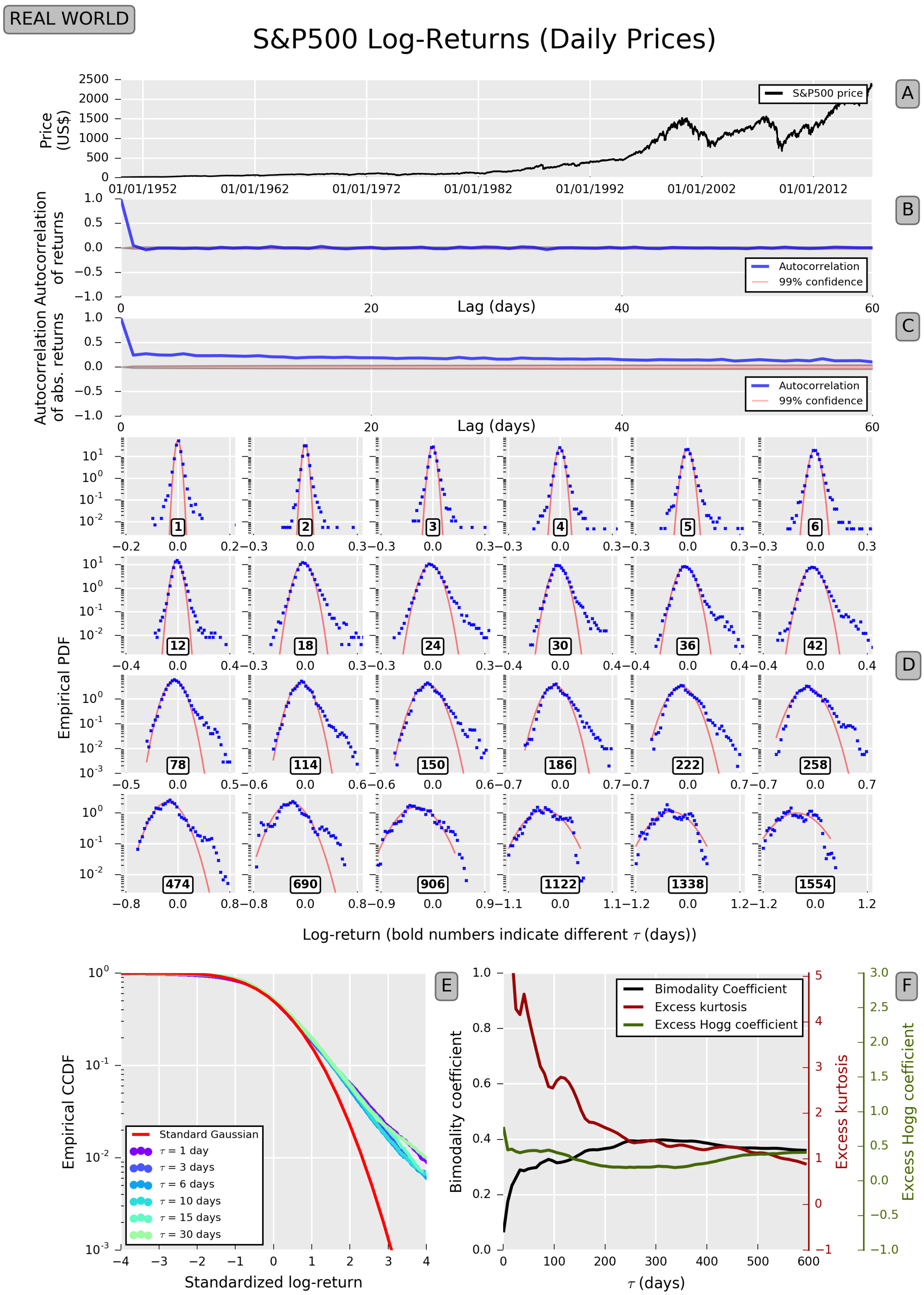

Figure 46: Daily S&P opening price between and 1950-01-03 and 2017-04-13, retrieved from [Qua17a].

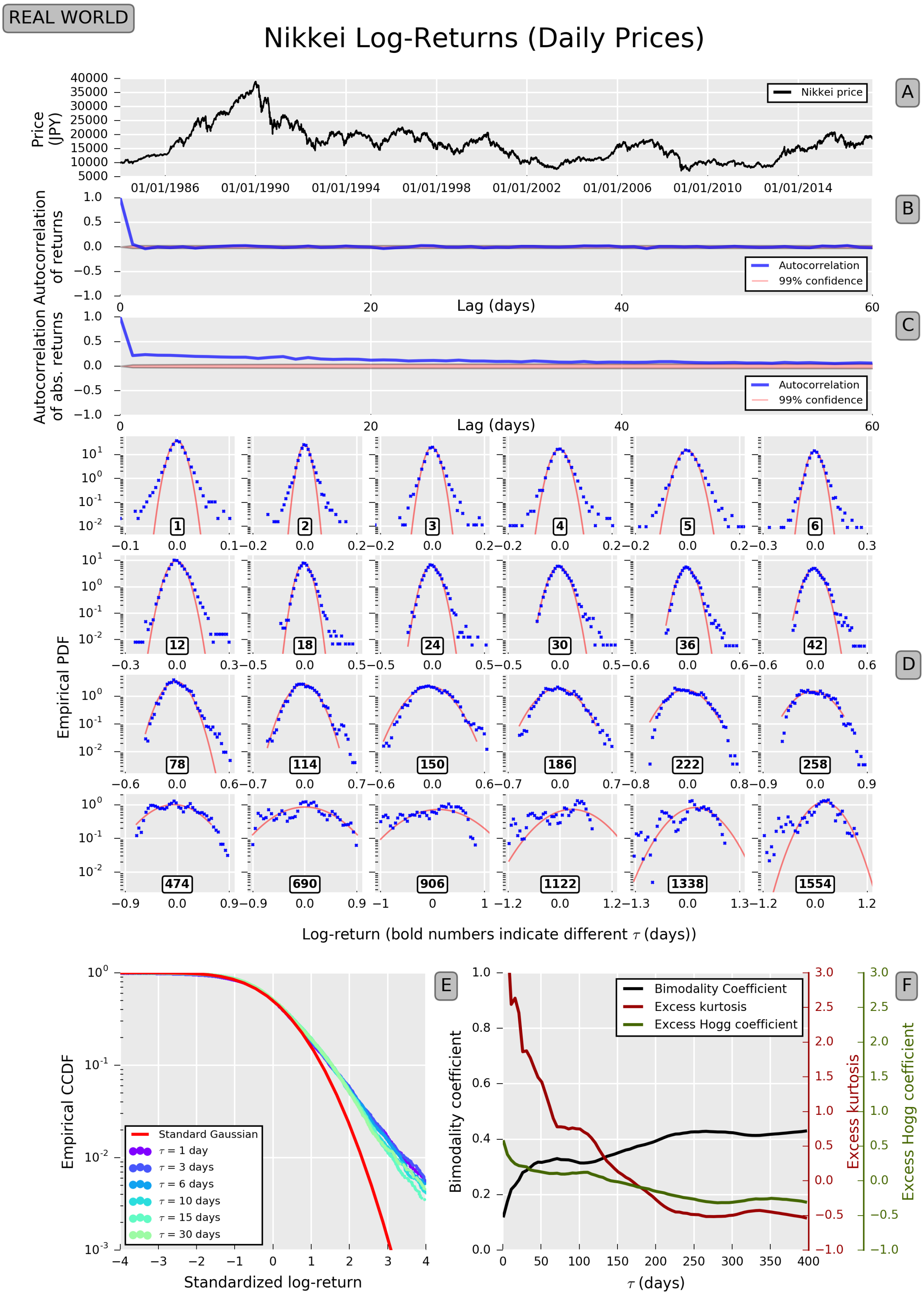

Figure 47: Daily Nikkei 225 opening price between 1984-01-04 and 2017-04-12, retrieved from [Qua17a].

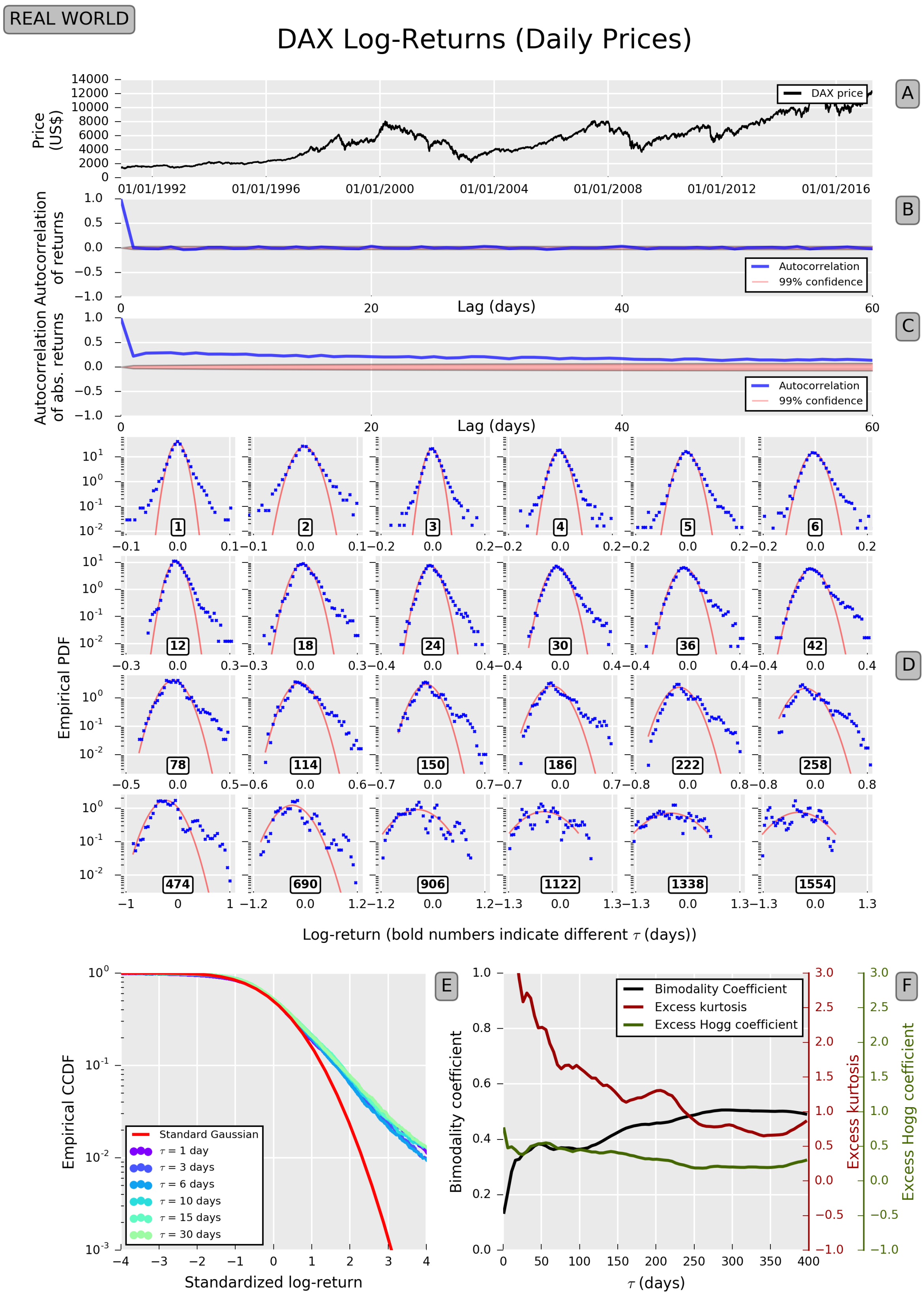

Figure 48: Daily DAX opening price between 1990-11-26 and 2017-04-12, retrieved from [Qua17a].

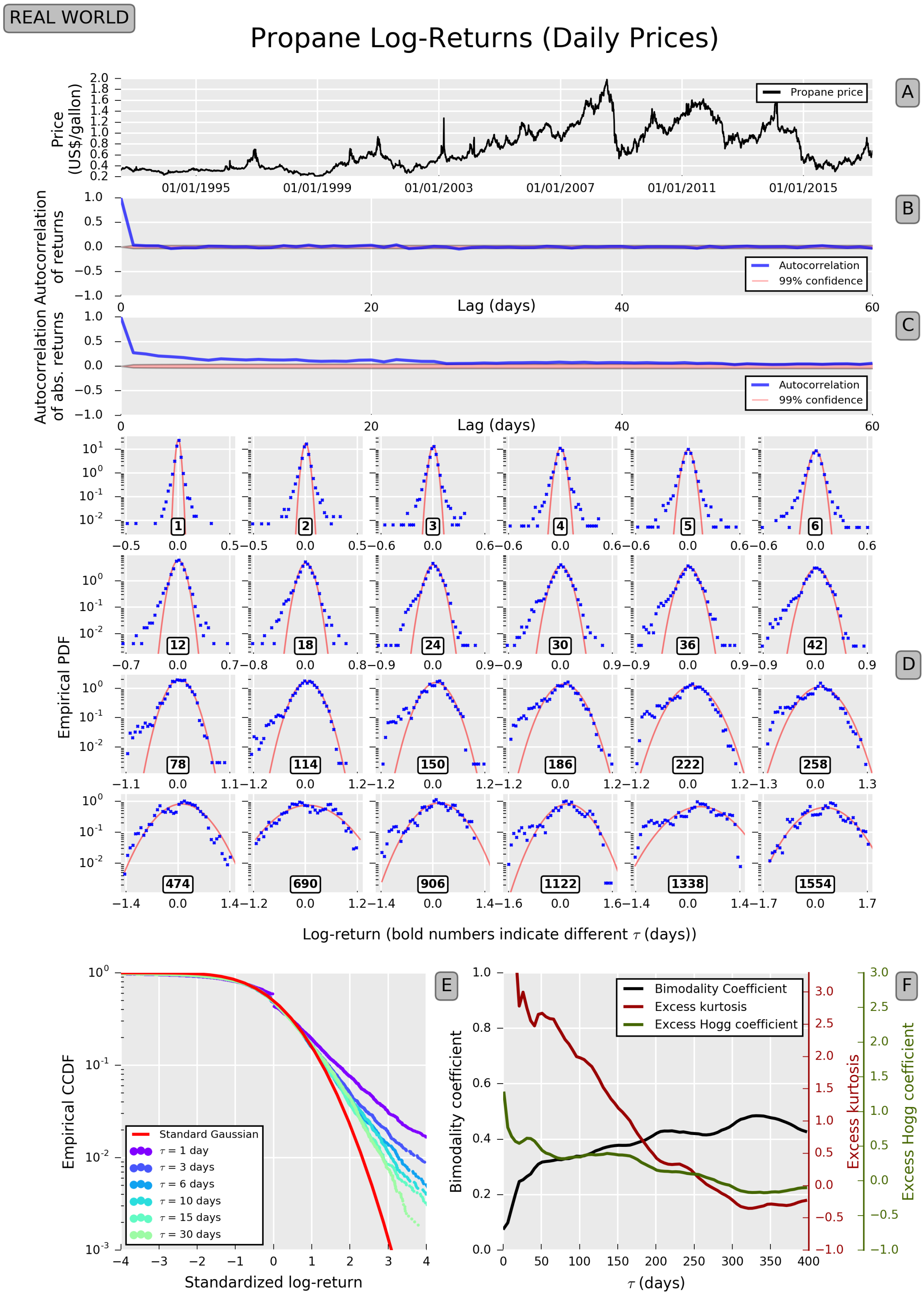

Figure 49: Daily Mont Belvieu propane price between 1992-07-09 and 2017-04-10, retrieved from [Dat17].

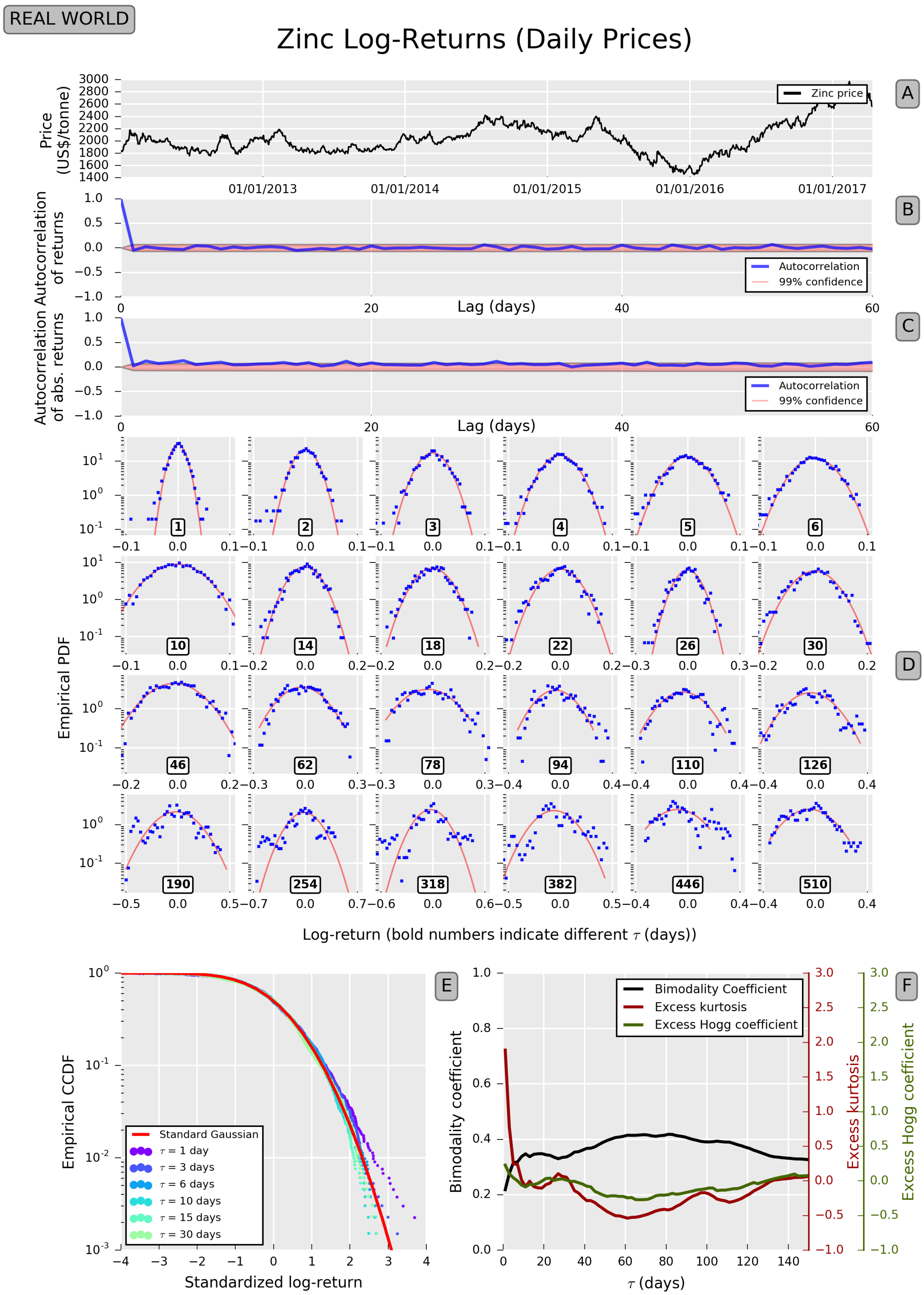

Figure 50: Daily zinc price on the London Metal Exchange between 2012-01-03 and 2017-04-13, retrieved from [Qua17b].

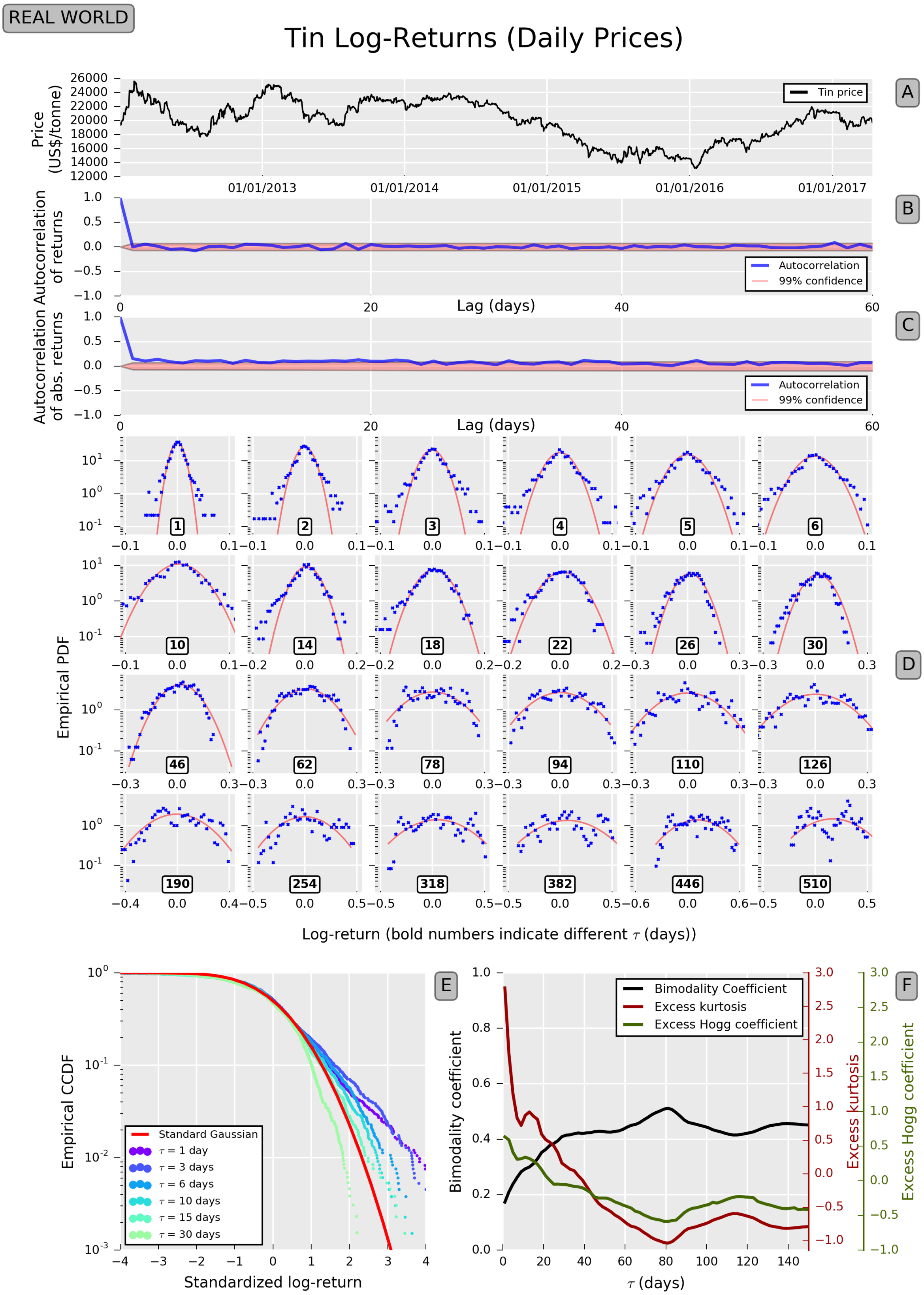

Figure 51: Daily tin price on the London Metal Exchange between 2012-01-03 and 2017-04-13, retrieved from [Qua17b].

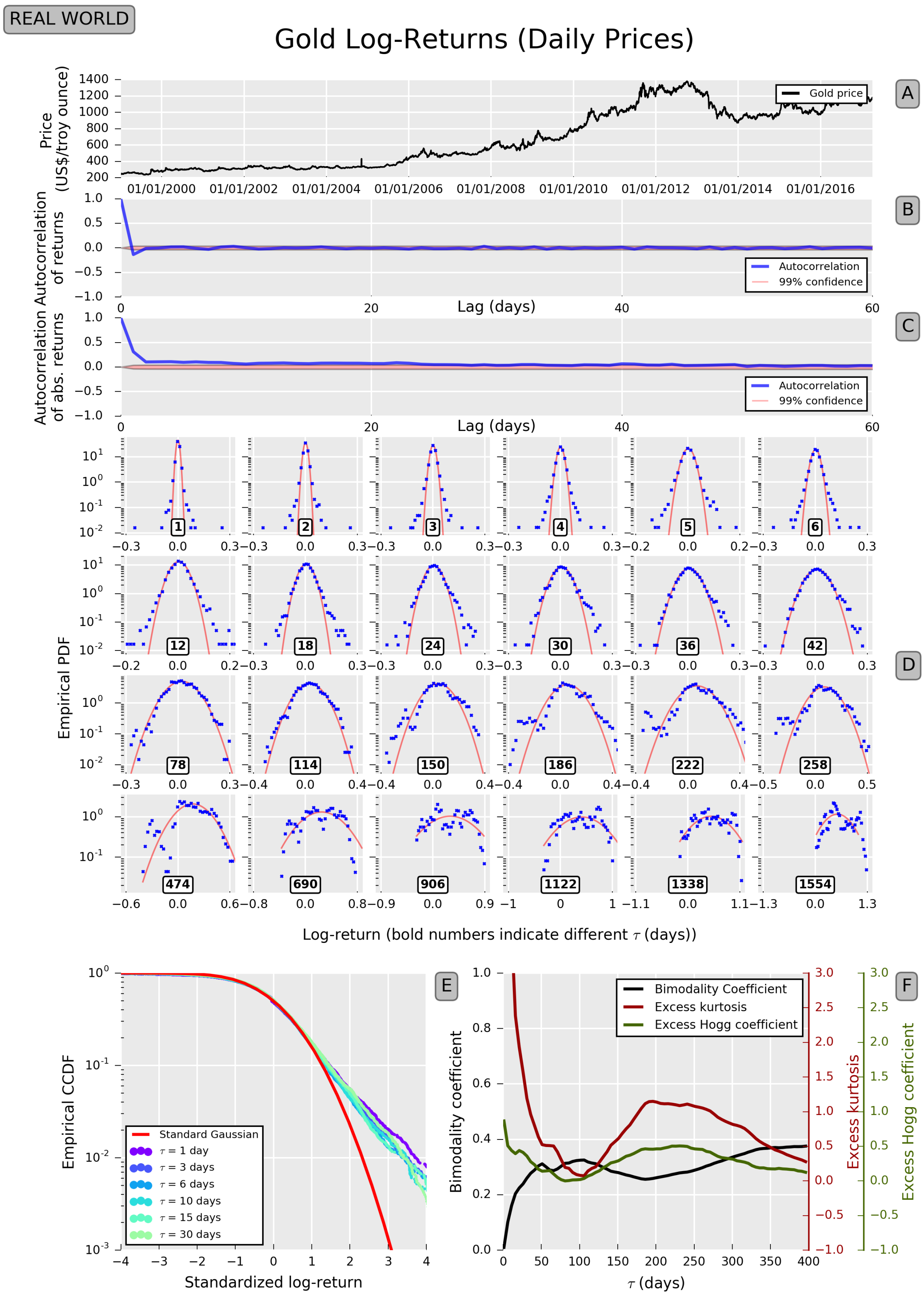

Figure 52: Daily 10:30 A.M. (London time) gold fixing price on the London Bullion Market between 1999-01-04 and 2017-04-05, retrieved from [Dat17]. No post-processing.

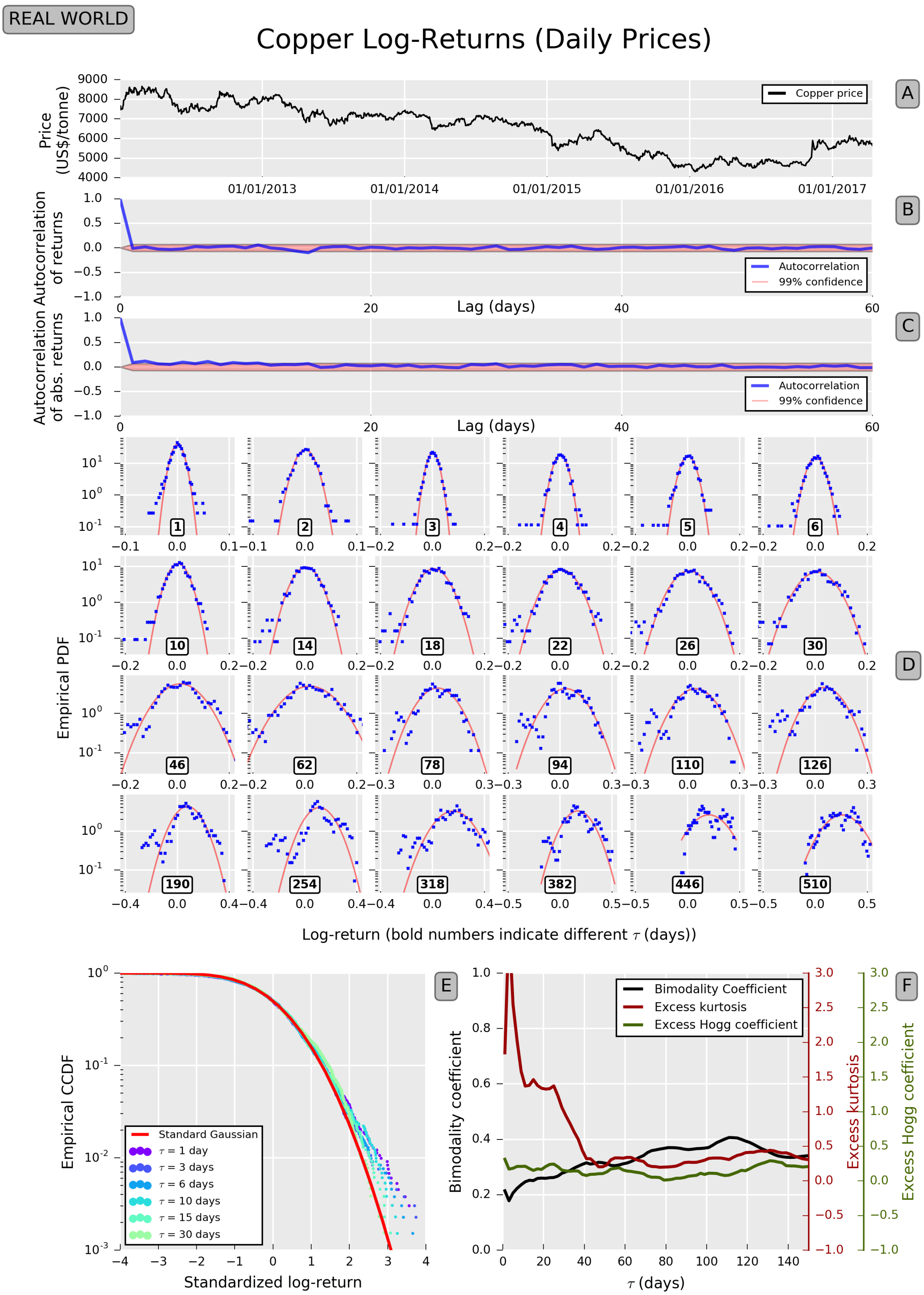

Figure 53: Daily copper price on the London Metal Exchange between 2012-01-03 and 2017-04-13, retrieved from [Qua17b].

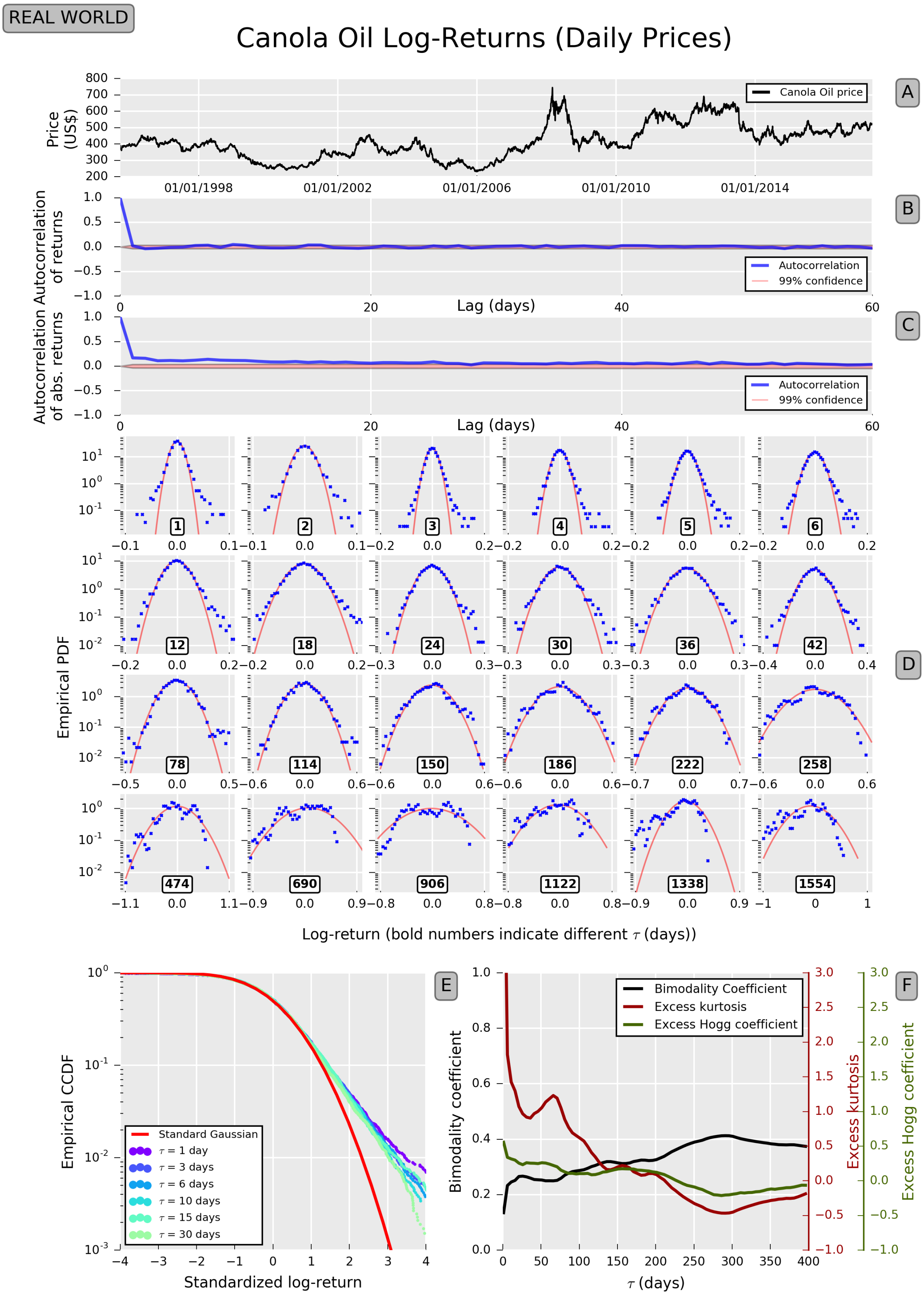

Figure 54: Daily canola oil futures price on the Chicago Mercantile Exchange between 1995-09-28 and 2017-05-15, retrieved from [Qua17b].

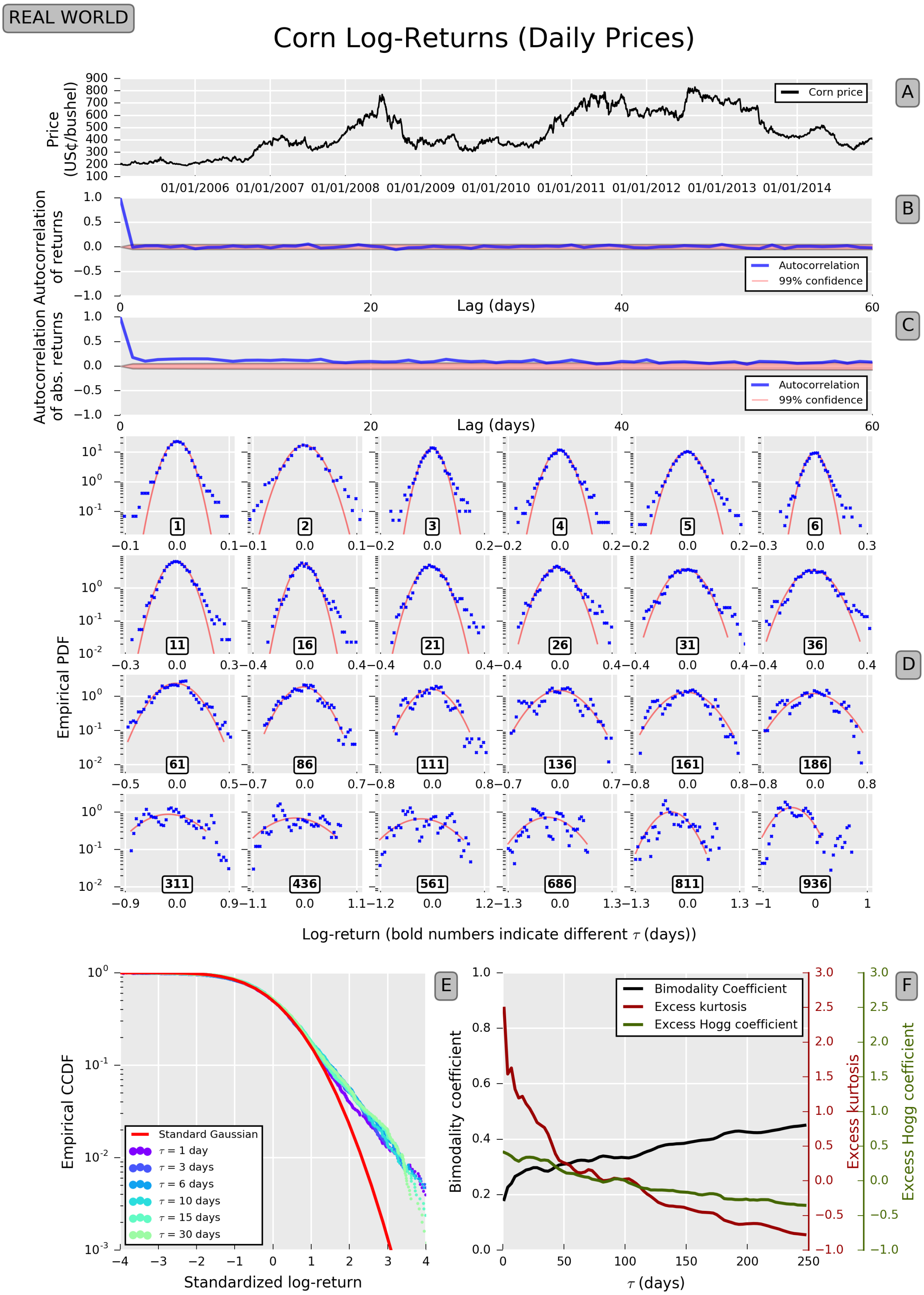

Figure 55: Daily corn futures price on the Chicago Mercantile Exchange between 2005-01-03 and 2014-12-31, retrieved from [Qua17b].

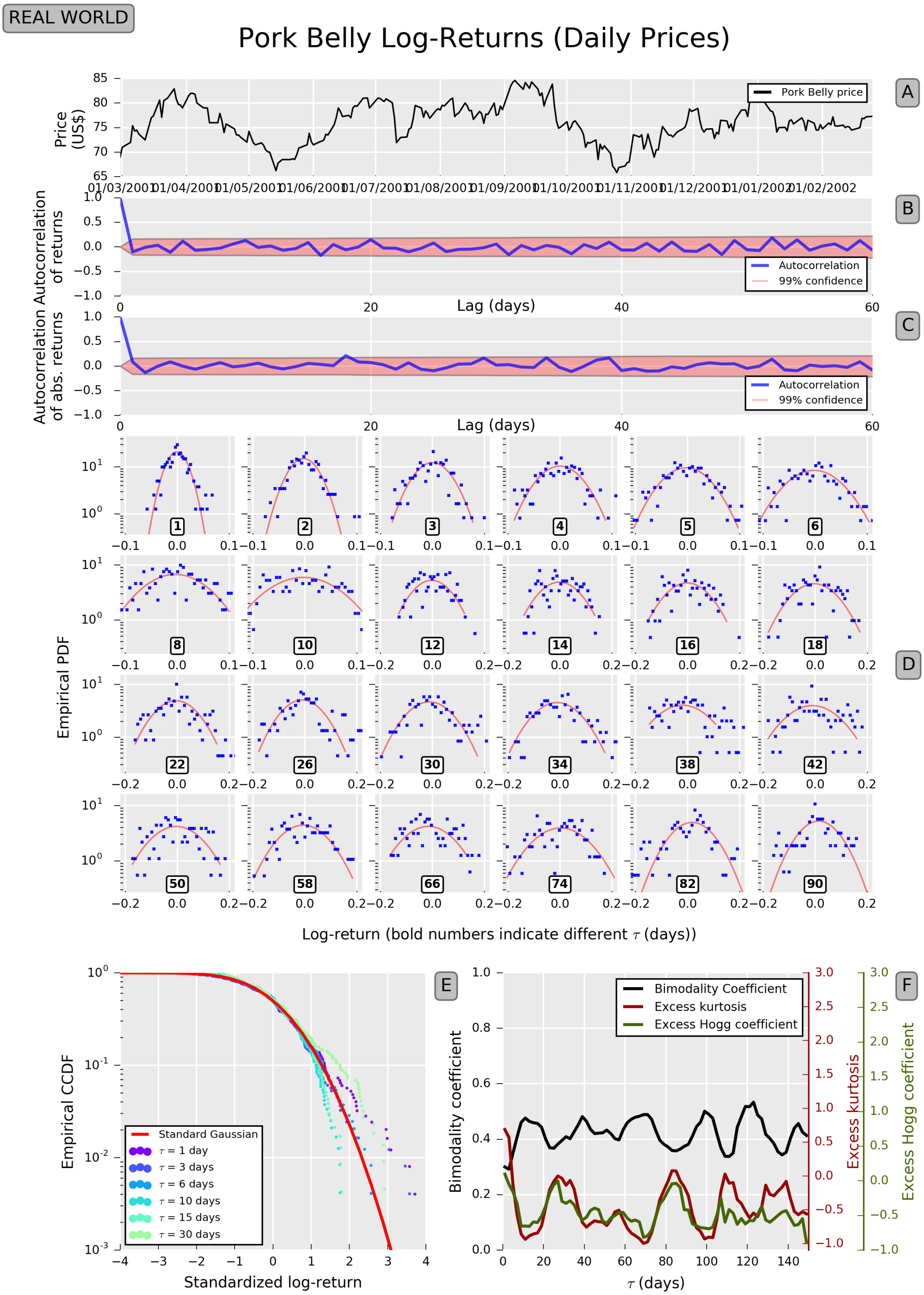

Figure 56: Daily pork belly oil futures price on the Chicago Mercantile Exchange between 2001-02-28 and 2002-02-25, retrieved from [Qua17b].

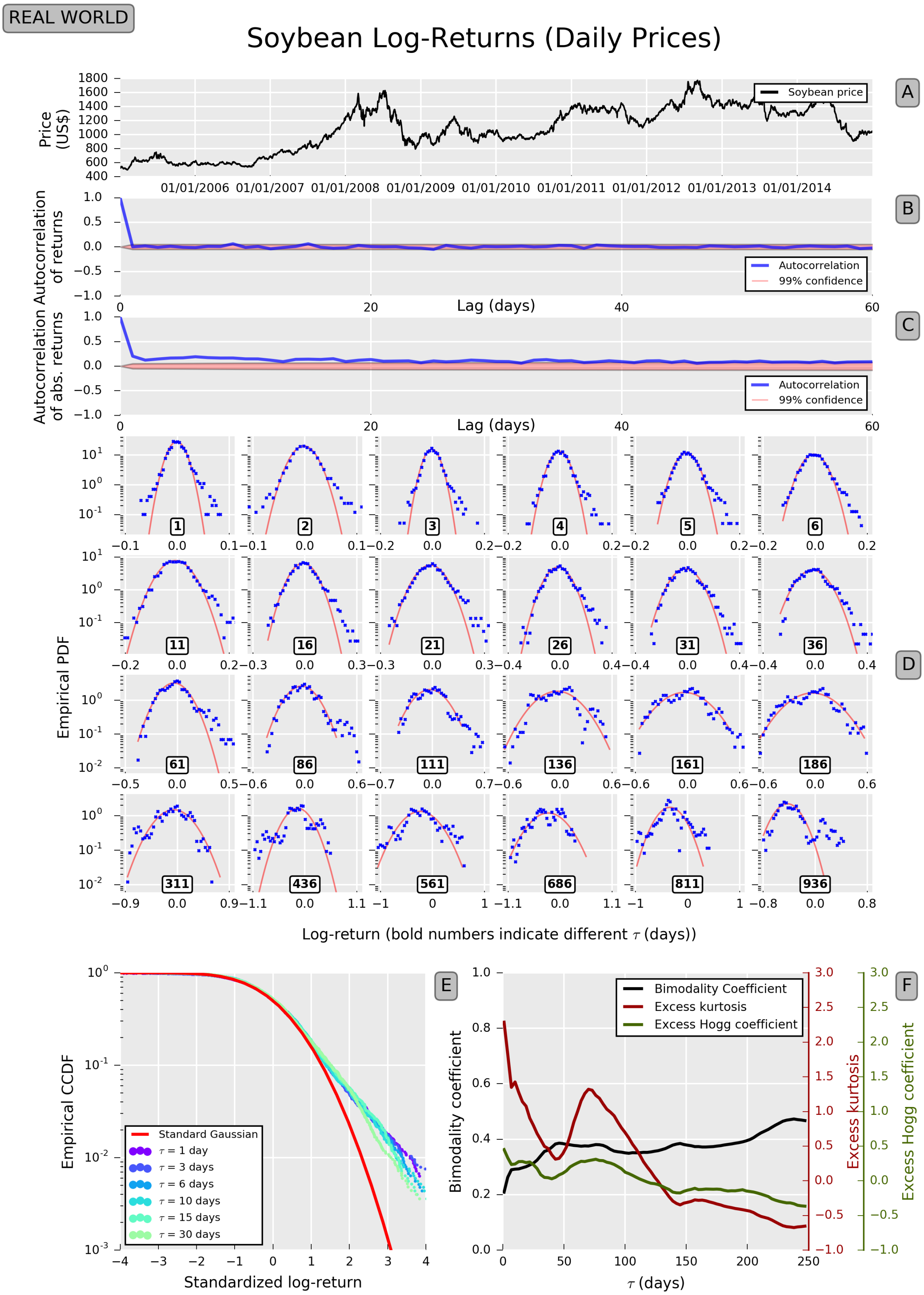

Figure 57: Daily soybean futures price on the Chicago Mercantile Exchange between 2005-01-03 and 2014-12-31, retrieved from [Qua17b].

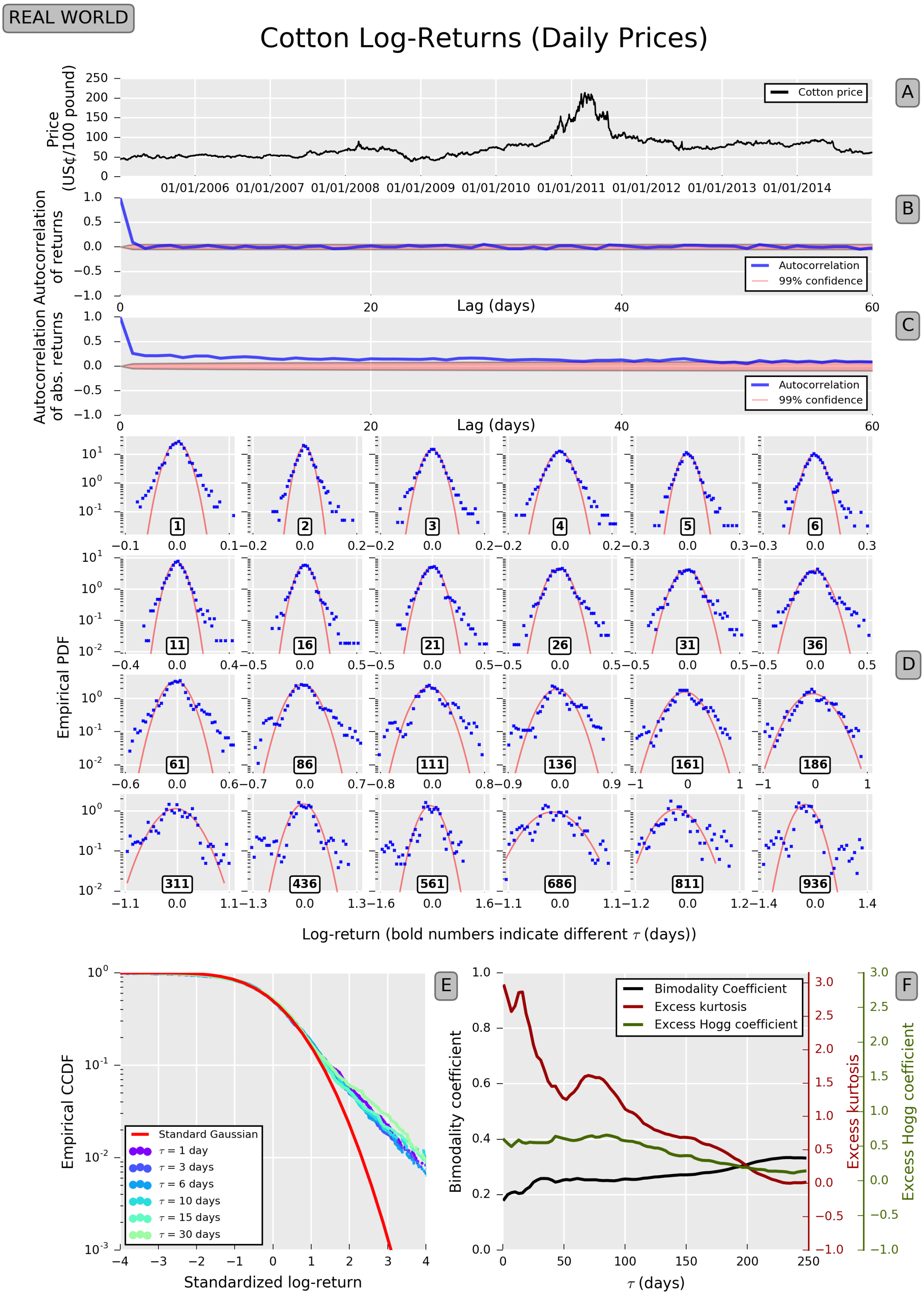

Figure 58: Daily cotton futures price on the Chicago Mercantile Exchange between 2005-01-03 and 2014-12-31, retrieved from [Qua17b].

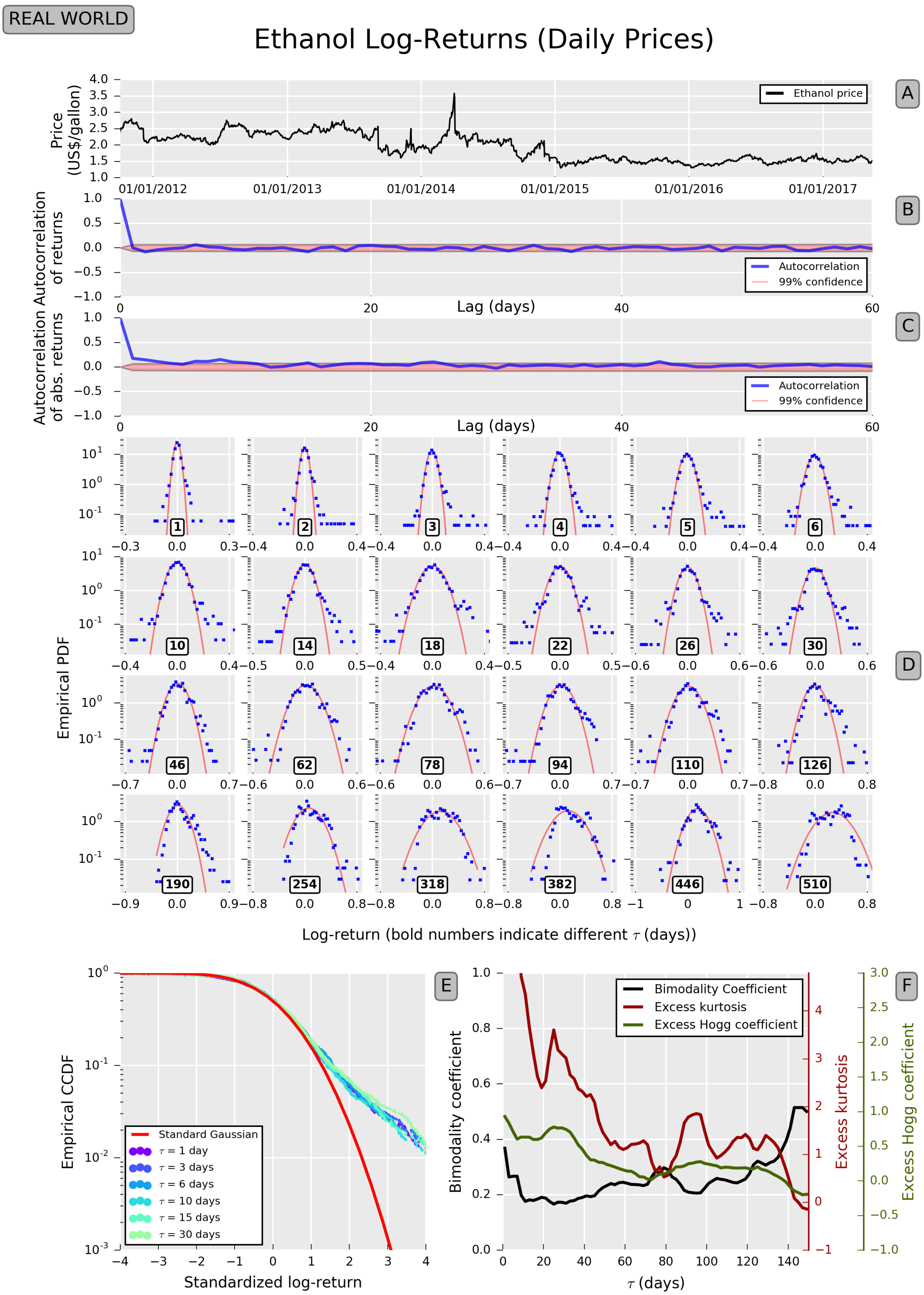

Figure 59: Daily ethanol futures price on the Chicago Mercantile Exchange between 2011-10-03 and 2017-05-15, retrieved from [Qua17b].

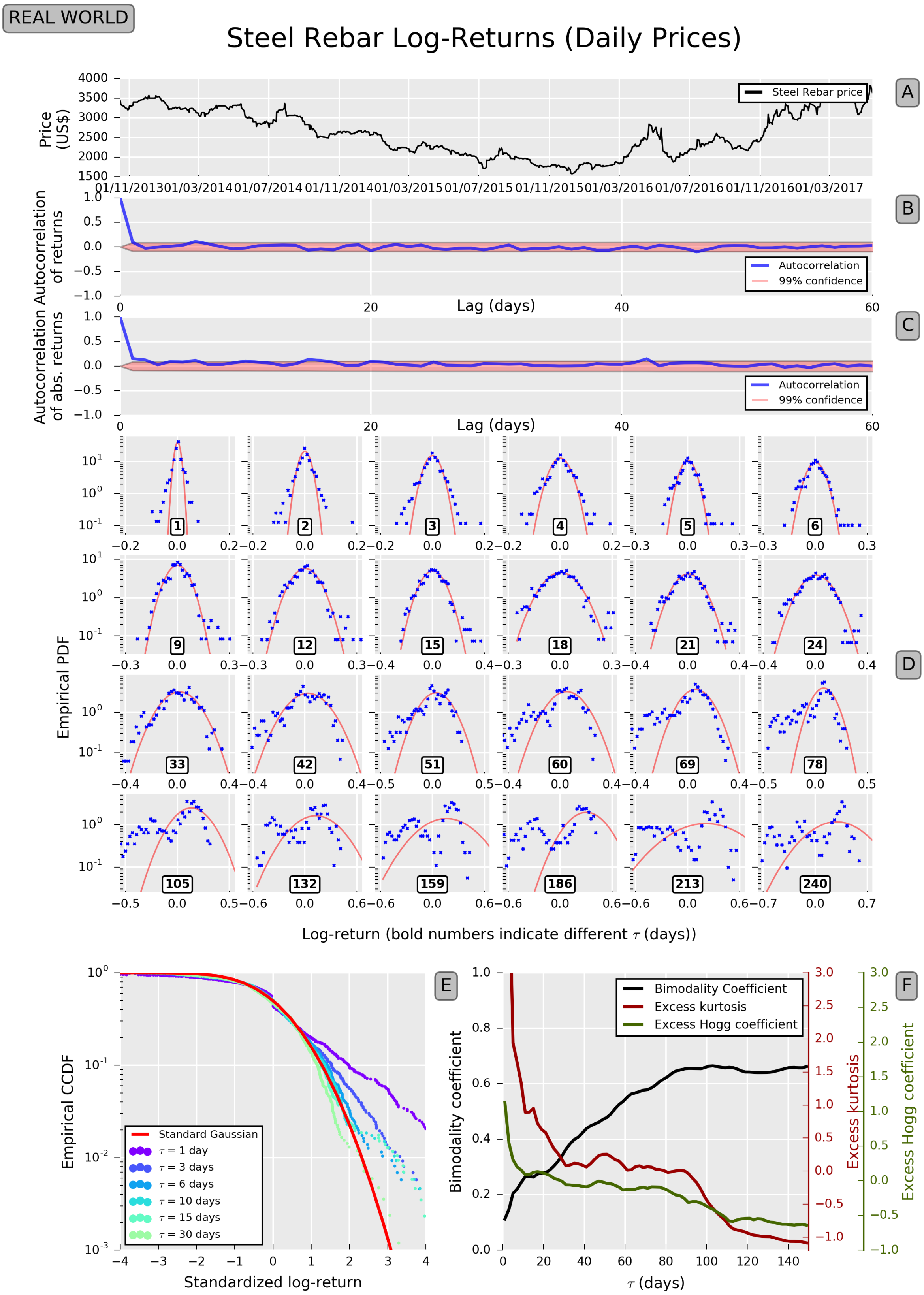

Figure 60: Daily steel rebar futures price on the Shanghai Futures Exchange between 2013-10-16 and 2017-05-15, retrieved from [Qua17b].

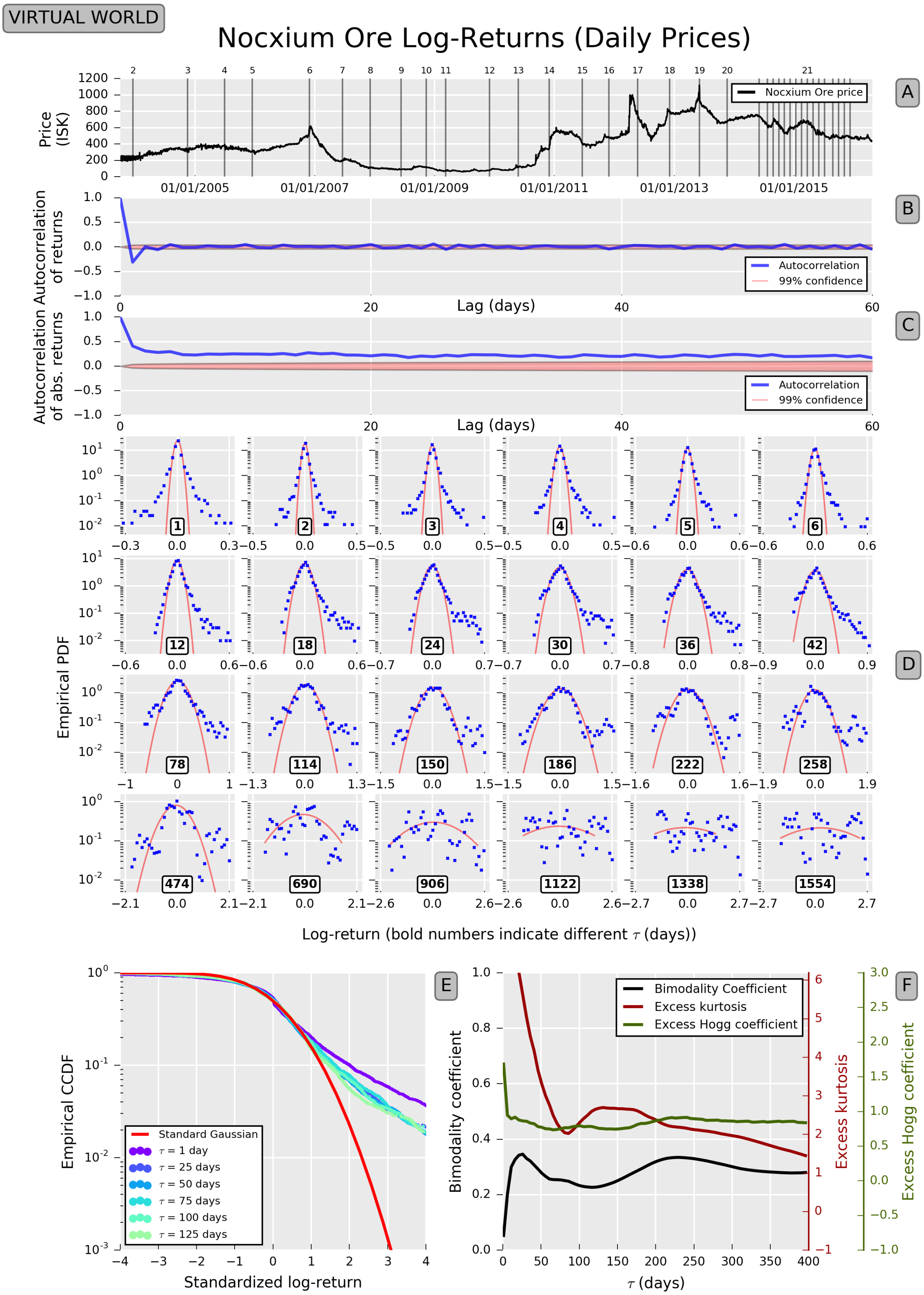

Figure 61: Nocxium ore daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17].

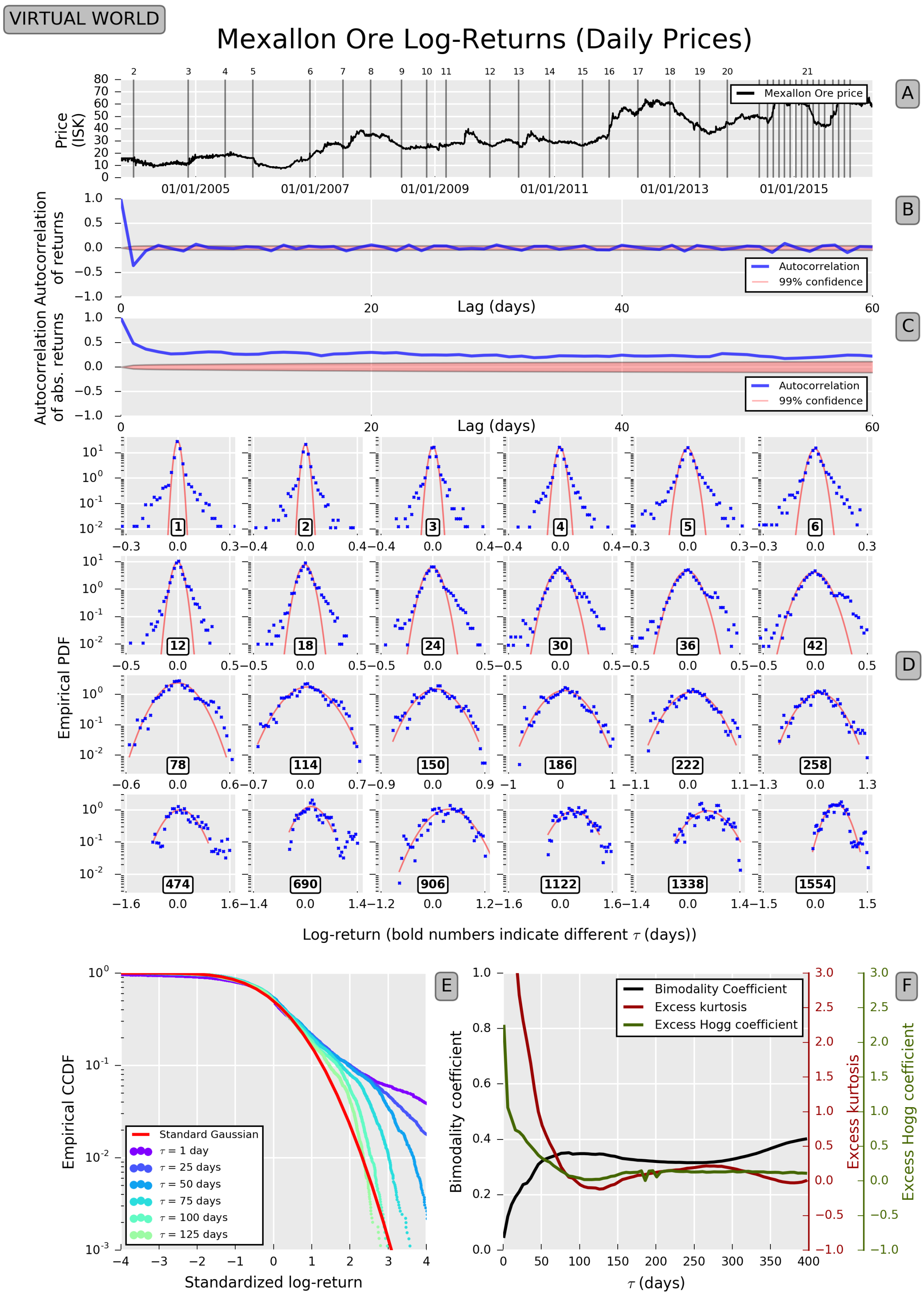

Figure 62: Mexallon ore daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17].

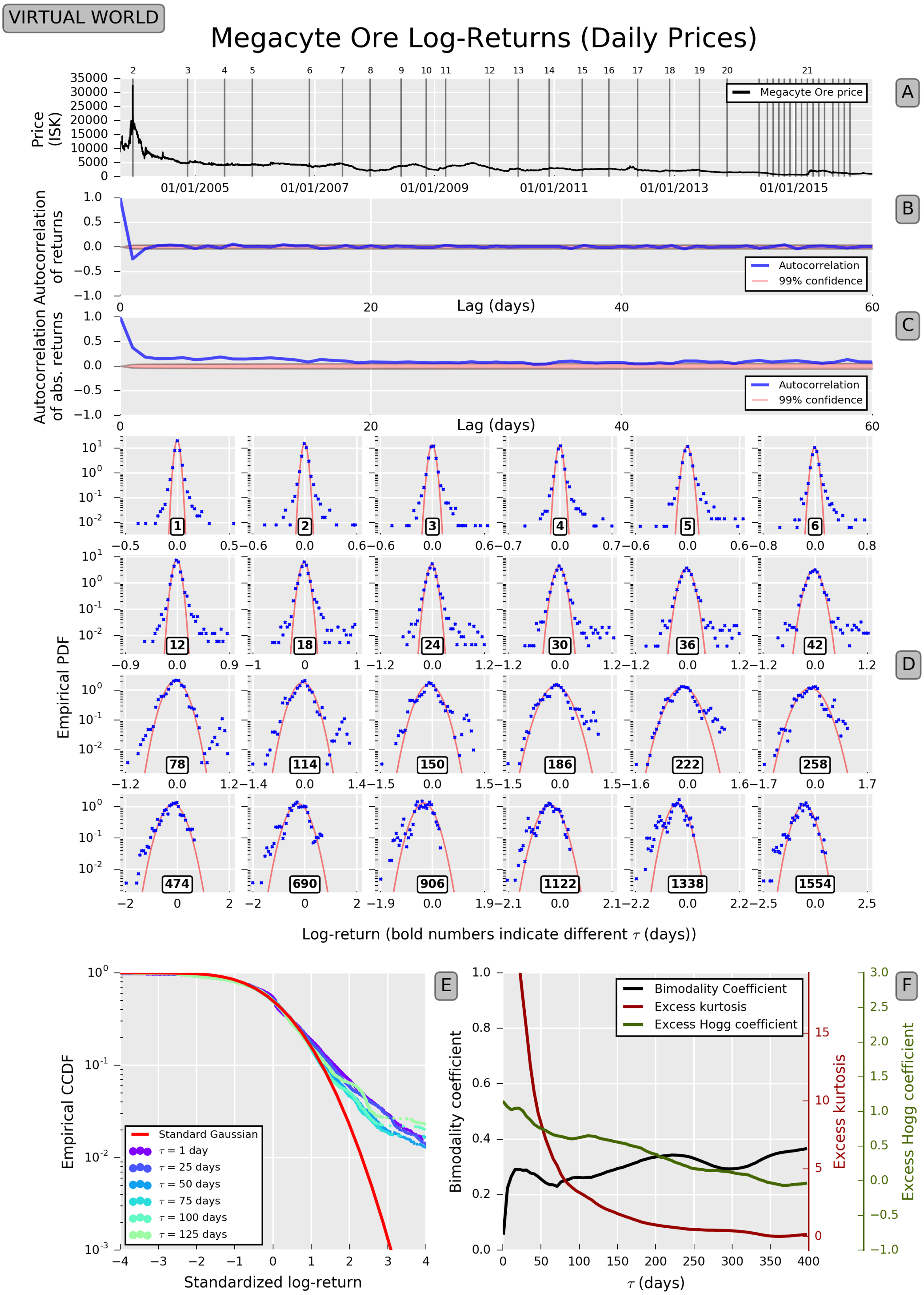

Figure 63: Megacyte ore daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17].

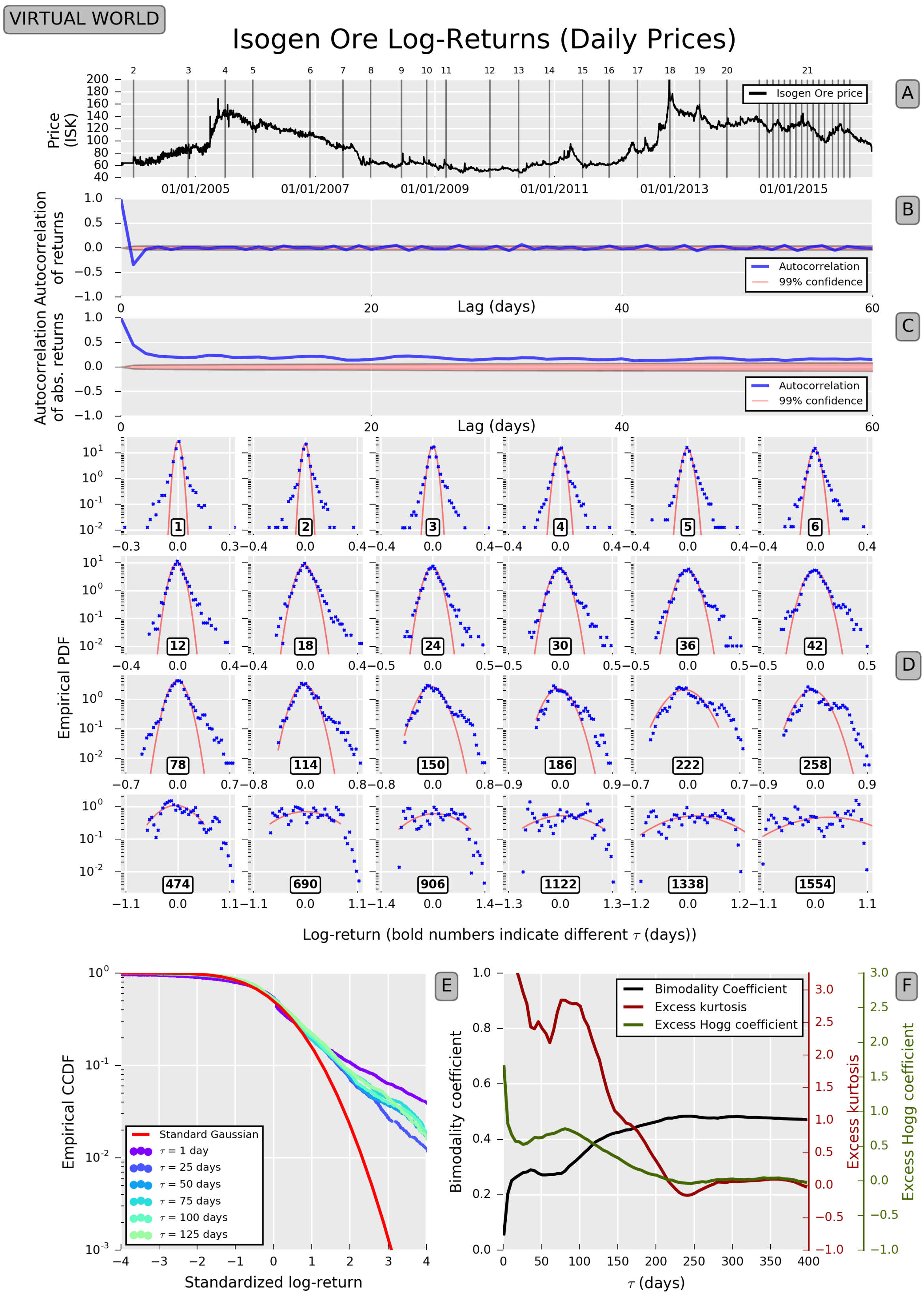

Figure 64: Isogen ore daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17].

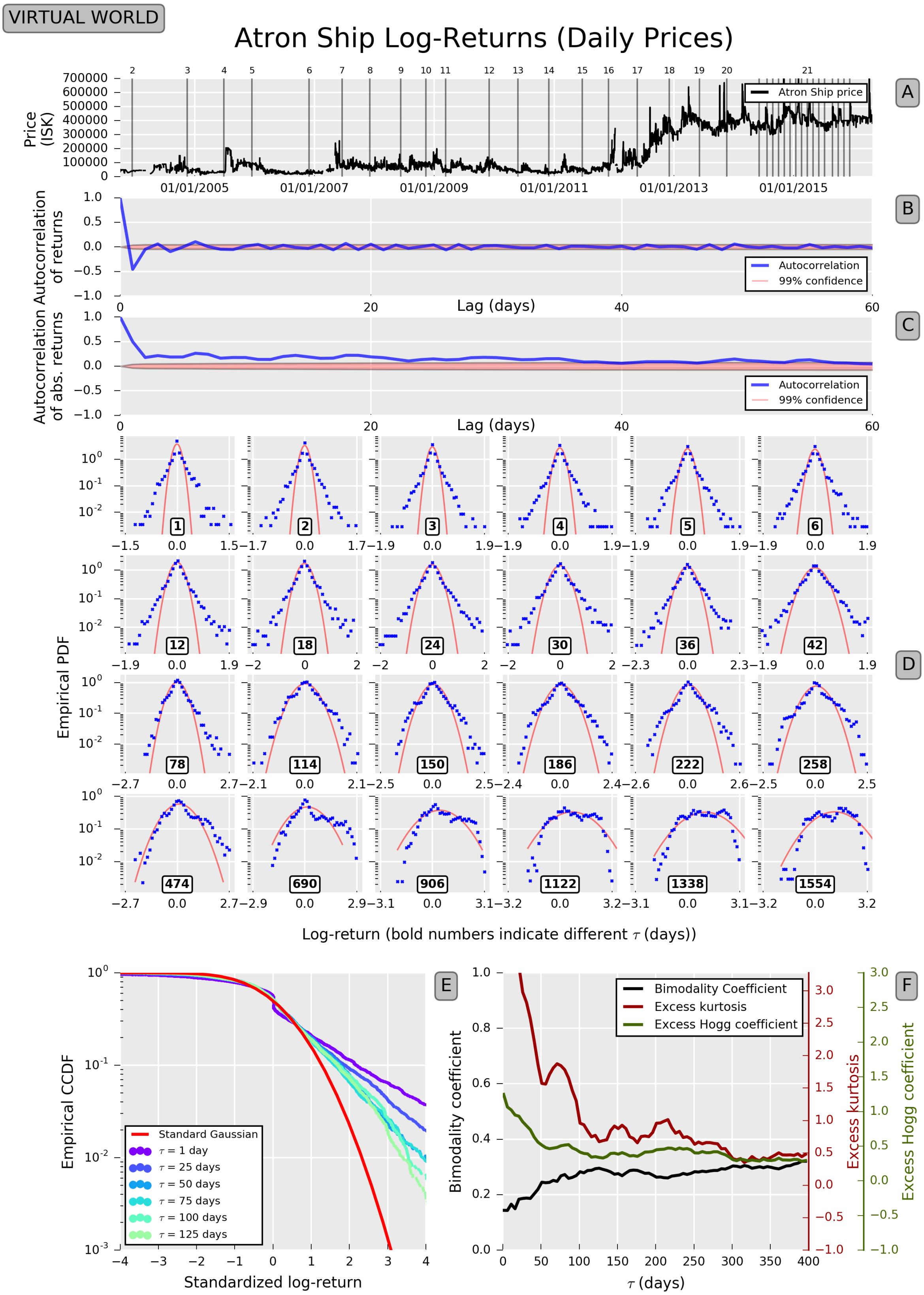

Figure 65: Atron ship daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17]. Peaks removed.

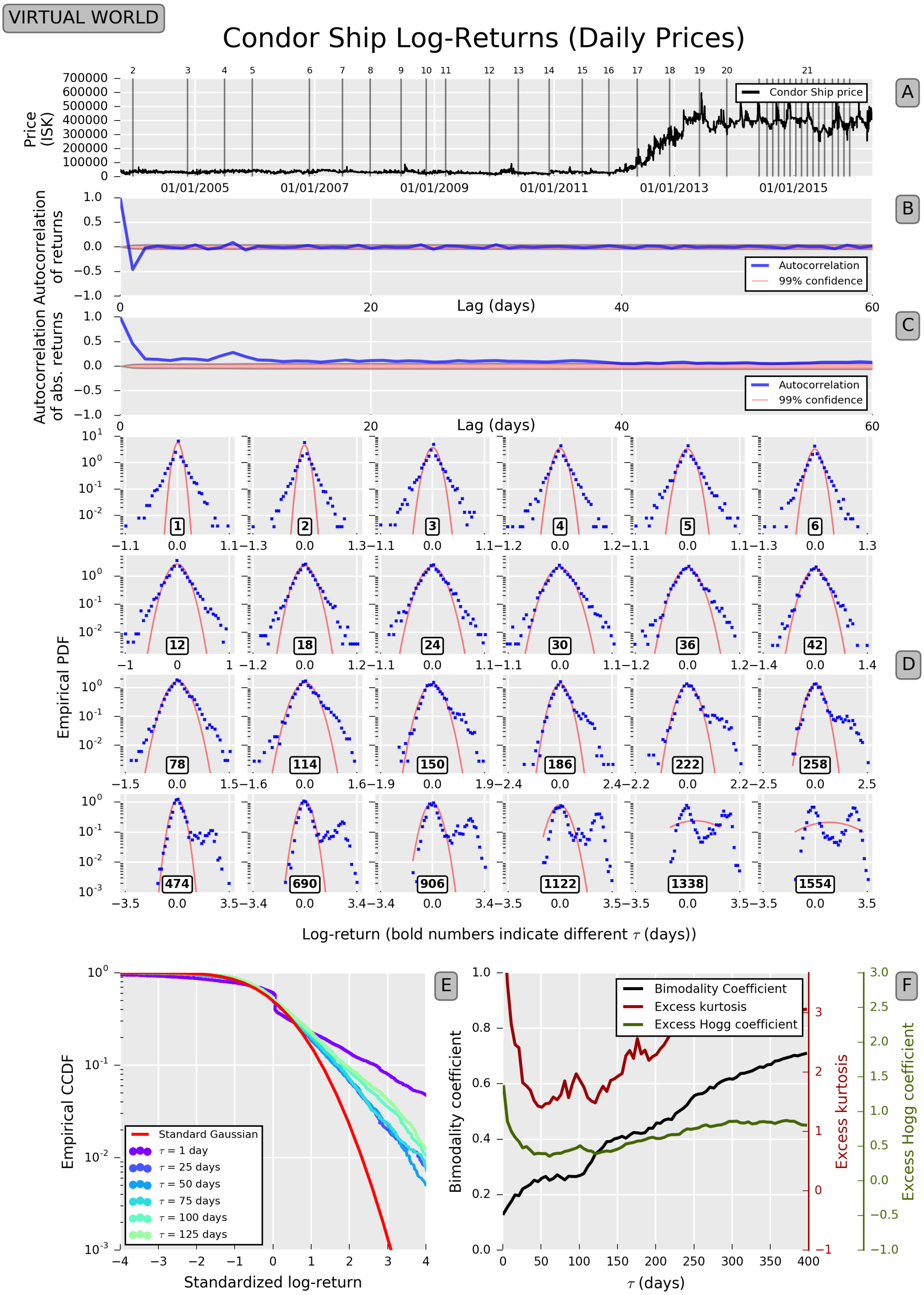

Figure 66: Condor ship daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17]. Peaks removed.

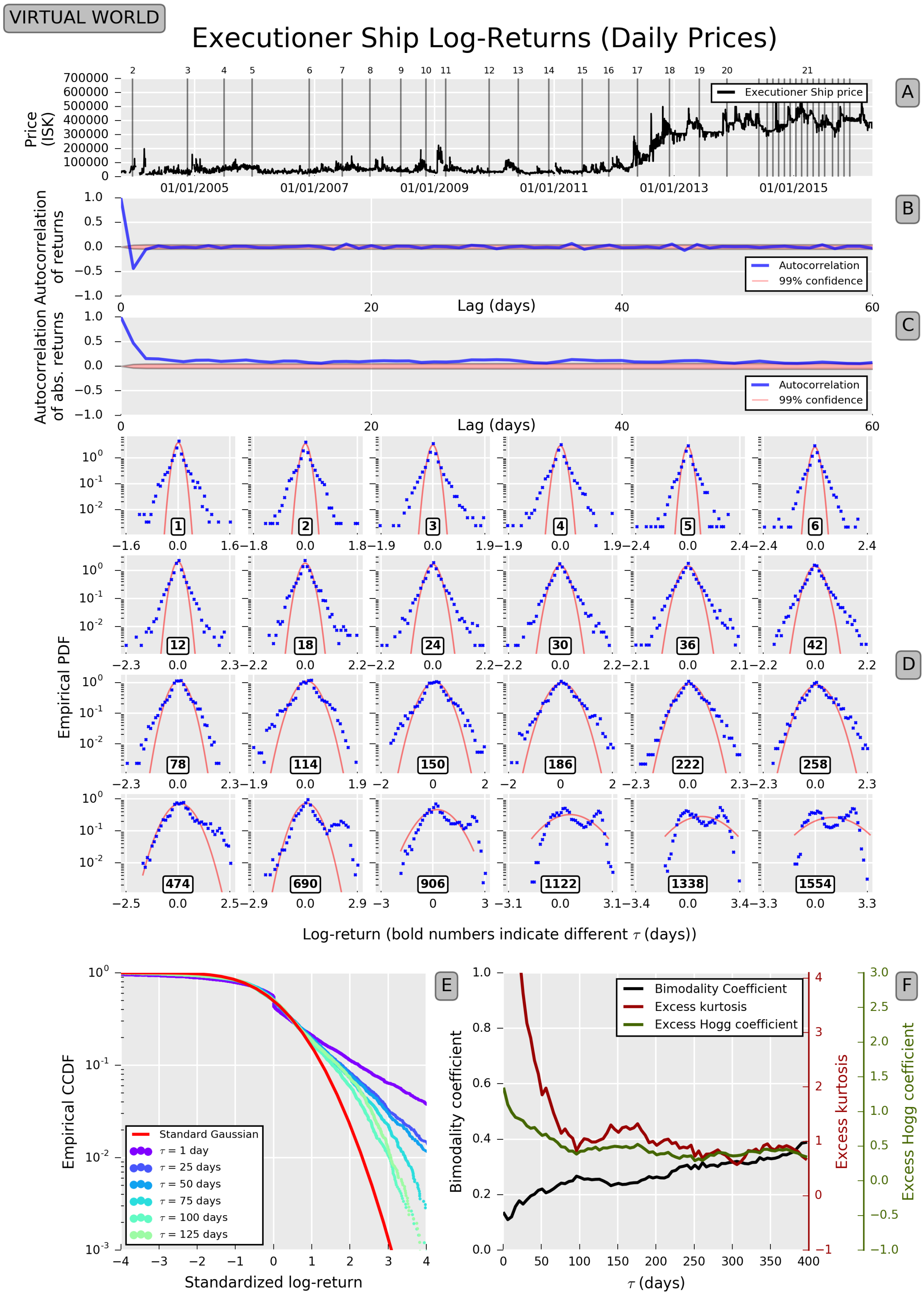

Figure 67: Executioner ship daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17]. Peaks removed.

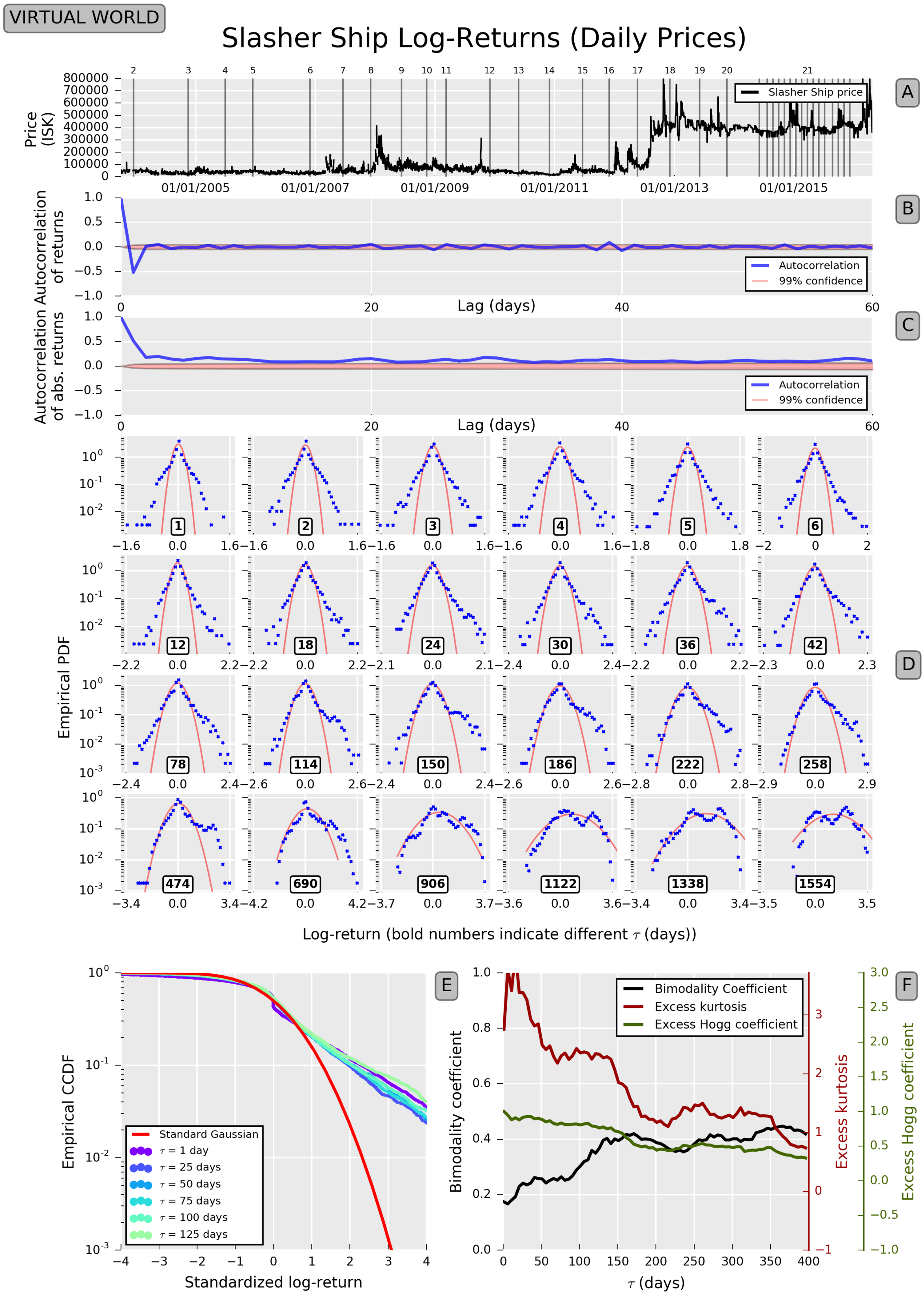

Figure 68: Slasher ship daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17]. Peaks removed.

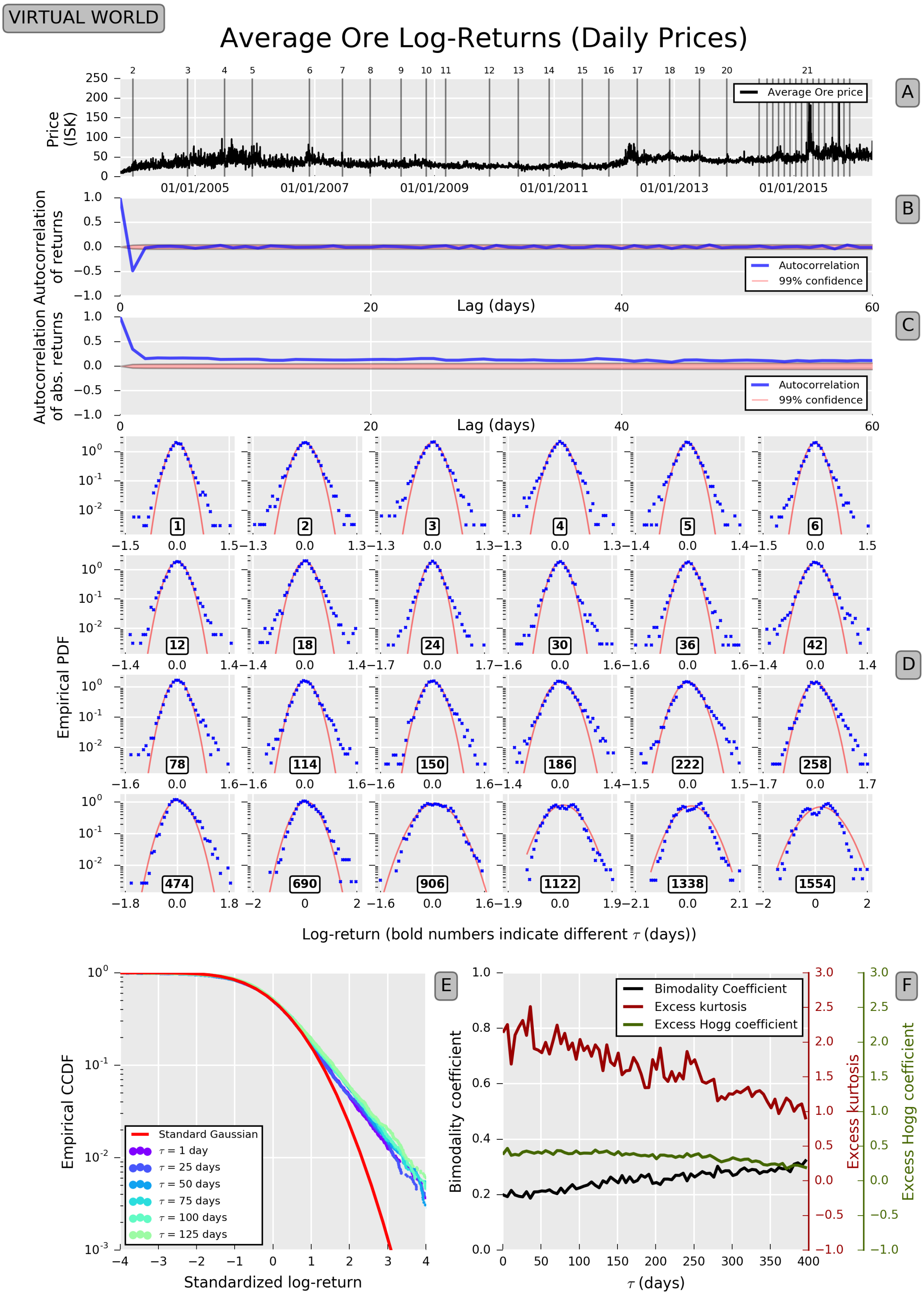

Figure 69: Average ore (tritanium, pyerite, noxcium, mexallon, megacyte, isogen) daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17].

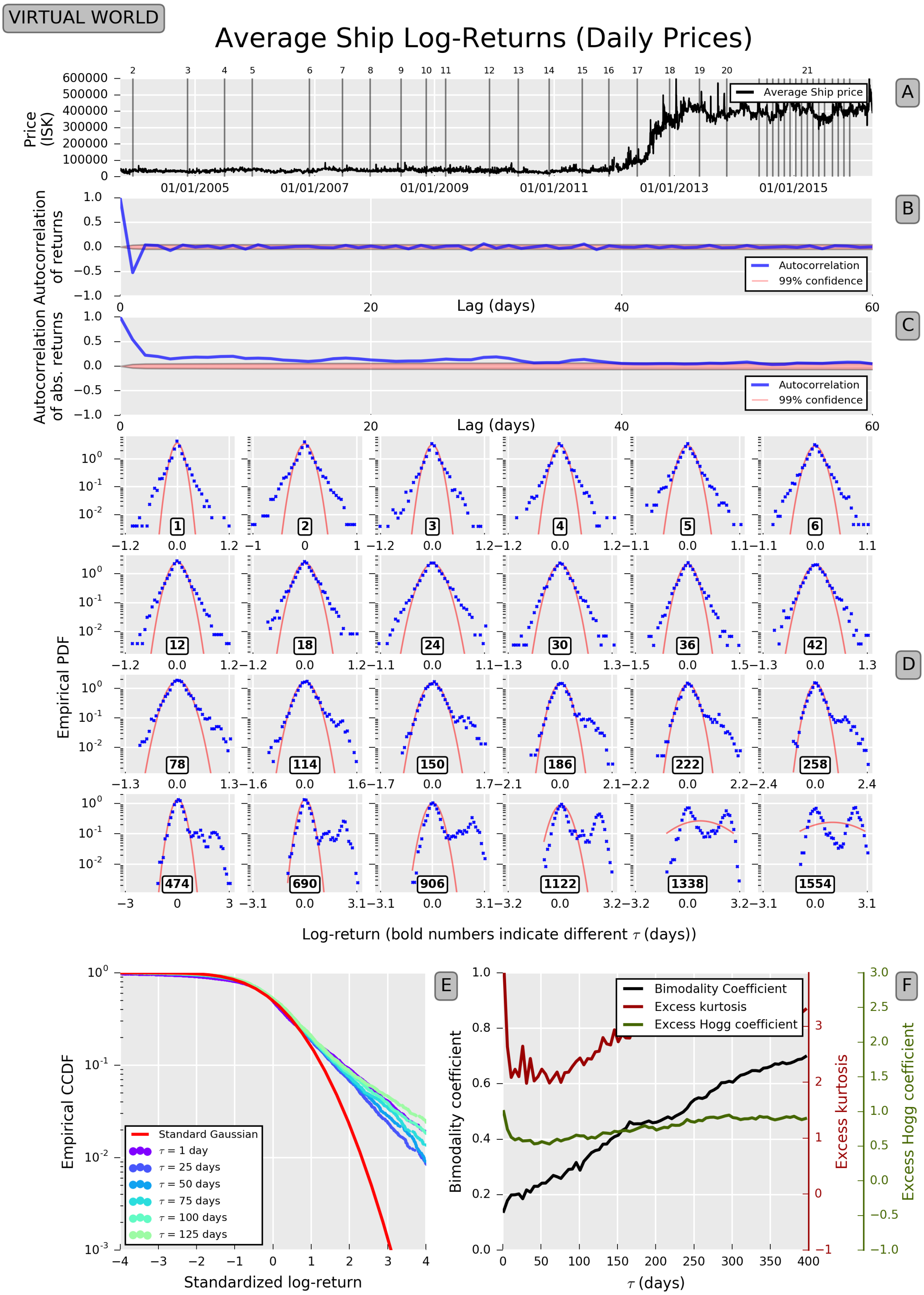

Figure 70: Average ship (atron, slasher, executioner and condor) daily price in *EVE* between 2003-10-01 and 2016-04-23, courtesy of [Gam17]. Peaks removed.

BIBLIOGRAPHY

[al15] C. Asness et. al. *Size Matters, If You Control Your Junk*. Working Paper. Apr. 2015. DOI: 10.2139/ssrn.2553889. URL: https://ssrn.com/abstract=2553889.

[Alp15] Alphameridian. *Everquest II Economy Hacked. Don't Grow Too Much Money on Trees Too Fast*. https://arstechnica.com/gaming/2005/08/948/. Accessed: 2017-05-12. Dec. 8, 2015.

[Ami08] Henrik Amilon. "Estimation of an adaptive stock market model with heterogeneous agents." In: *Journal of Empirical Finance* 15.2 (2008), pp. 342–362. ISSN: 0927-5398. DOI: https://doi.org/10.1016/j.jempfin.2006.06.007. URL: http://www.sciencedirect.com/science/article/pii/S0927539807000606.

[Ani11] et. al. Anirban Chakraborit. "Econophysics: Empirical facts and agent-based models." In: *Quantative Finance* 11 (2011), pp. 1013–1041.

[Axt01] Robert L. Axtell. "Zipf Distribution of U.S. Firm Sizes." In: *Science* 293.5536 (Sept. 2001), pp. 1818–1820. ISSN: 1095-9203. DOI: 10.1126/science.1062081. URL: http://dx.doi.org/10.1126/science.1062081.

[Ban70] R. Banz. "The Relationship Between Return and Market Value Of Common Stocks." In: *Journal of Financial Economics* 9.1 (1970), pp. 3–18. URL: http://www.sciencedirect.com/science/article/pii/0304405X81900180.

[Bas77] S. Basu. "Investment Performance of Common Stocks In Relation To Their Price-Earnings Ratios: A Test Of The Efficient Market Hypothesis." In: *The Journal of Finance* 32.3 (1977), pp. 663–682. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1977.tb01979.x. URL: http://dx.doi.org/10.1111/j.1540-6261.1977.tb01979.x.

[BS73] Fischer Black and Myron Scholes. "The Pricing of Options and Corporate Liabilities." In: *Journal of Political Economy* 81.3 (1973), pp. 637–54. URL: http://EconPapers.repec.org/RePEc:ucp:jpolec:v:81:y:1973:i:3:p:637-54.

[BM00] J.-P. Bouchaud and M. Mezard. "Wealth condensation in a simple model of economy." In: *eprint arXiv:cond-mat/0002374* (Feb. 2000). eprint: cond-mat/0002374.

[Cas02] Edward Castronova. *On Virtual Economies*. Working Paper 752. July 2002. URL: https://ssrn.com/abstract=338500.

[Cas+09] Edward Castronova, Dmitri Williams, Cuihua Shen, Rabindra Ratan, Li Xiong, Yun Huang, and Brian Keegan. "As real as real? Macroeconomic behavior in a large-scale virtual world." In: *New Media & Society* 11.5 (2009), pp. 685–707. DOI: 10.1177/1461444809105346. eprint: http://dx.doi.org/10.1177/1461444809105346. URL: http://dx.doi.org/10.1177/1461444809105346.

[CV12] Lidia Ceriani and Paolo Verme. "The origins of the Gini index: extracts from Variabilità e Mutabilità (1912) by Corrado Gini." In: *The Journal of Economic Inequality* 10.3 (2012), pp. 421–443.

[CC00] A. Chakraborti and B.K. Chakrabarti. "Statistical mechanics of money: how saving propensity affects its distribution." In: *European Physical Journal B* 17 (Aug. 2000), pp. 167–170. DOI: 10.1007/s100510070173. eprint: cond-mat/0004256.

[Cha53] D.G. Champernowne. "A Model of Income Distribution." In: *The Economic Journal* 63.250 (1953), pp. 318–351.

[CC06] A. Chatterjee and B. K. Chakrabarti. "Kinetic market models with single commodity having price fluctuations." In: *The European Physical Journal B - Condensed Matter and Complex Systems* 54.3 (2006), pp. 399–404. ISSN: 1434-6036. DOI: 10.1140/epjb/e2007-00011-1. URL: http://dx.doi.org/10.1140/epjb/e2007-00011-1.

[CKM04] A. Chatterjee, B. K. Chakrabarti, and S.S. Manna. "Pareto law in a kinetic model of market with random saving propensity." In: *Physica A Statistical Mechanics and its Applications* 335 (Apr. 2004), pp. 155–163. DOI: 10.1016/j.physa.2003.11.014. eprint: cond-mat/0301289.

[Con01] R. Cont. "Empirical properties of asset returns: stylized facts and statistical issues." In: *Quantitative Finance* 1.2 (2001), pp. 223–236. DOI: 10.1080/713665670. eprint: http://dx.doi.org/10.1080/713665670. URL: http://dx.doi.org/10.1080/713665670.

[CPZ01] M. Cristelli, L. Pietronero, and A. Zaccaria. "Critical Overview of Agent-Based Models for Economics." In: *ArXiv e-prints* (2001).

[Dat17] St. Louis Fed. Reserve Economic Data. *FRED Economic Data*. https://fred.stlouisfed.org/. Accessed: 2017-04-01. 2017.

[Dav+09] James B. Davies, Susanna Sandstrom, Anthony B. Shorrocks, and Edward N. Wolff. *The Level and Distribution of Global Household Wealth*. Working Paper 15508. National Bureau of Economic Research, Nov. 2009. DOI: 10.3386/w15508. URL: http://www.nber.org/papers/w15508.

[DY00] A. Drăgulescu and V.M. Yakovenko. "Statistical mechanics of money." In: *The European Physical Journal B - Condensed Matter and Complex Systems* 17.4 (2000), pp. 723–729. ISSN: 1434-6036. DOI: 10.1007/s100510070114. URL: http://dx.doi.org/10.1007/s100510070114.

[Dur05] Steven N. Durlauf. "Complexity and Empirical Economics." In: *The Economic Journal* 115.504 (2005), F225–F243. ISSN: 00130133, 14680297. URL: http://www.jstor.org/stable/3590439.

[Fam70] Eugene F. Fama. "Efficient Capital Markets: A Review of Theory and Empirical Work." In: *The Journal of Finance* 25.2 (1970), pp. 383–417. ISSN: 00221082, 15406261. URL: http://www.jstor.org/stable/2325486.

[Fam+69] Eugene F. Fama, Lawrence Fisher, Michael C. Jensen, and Richard Roll. "The Adjustment of Stock Prices to New Information." In: *International Economic Review* 10.1 (1969), pp. 1–21. ISSN: 00206598, 14682354. URL: http://www.jstor.org/stable/2525569.

[Fre80] Kenneth R. French. "Stock returns and the weekend effect." In: *Journal of Financial Economics* 8.1 (1980), pp. 55–69. ISSN: 0304-405X. DOI: http://dx.doi.org/10.1016/0304-405X(80)90021-5. URL: http://www.sciencedirect.com/science/article/pii/0304405X80900215.

[FT14] B. Fuchs and S. Thurner. "Behavioral and Network Origins of Wealth Inequality: Insights from a Virtual World." In: *PLoS ONE* 9.10 (Aug. 2014), e103503. DOI: 10.1371/journal.pone.0103503. arXiv: 1403.6342 [physics.soc-ph].

[Gam17] CCP Games. *EVE online*. www.eveonline.com. Accessed: 2017-05-26. 2003-2017.

[Gib31] R. Gibrat. *Les inégalités économiques*. Théses. Recueil Sirey, 1931. URL: https://books.google.be/books?id=m9fuoAEACAAJ.

[Gop+98] P. Gopikrishnan, M. Meyer, L.A.N. Amaral, and H.E. Stanley. "Inverse cubic law for the distribution of stock price variations." In: *The European Physical Journal B - Condensed Matter and Complex Systems* 3.2 (1998), pp. 139–140. ISSN: 1434-6036. DOI: 10.1007/s100510050292. URL: http://dx.doi.org/10.1007/s100510050292.

[Hag61] D.C. Hagued. *The Theory of Capital*. 1st ed. Palgrave Macmillan UK, 1961, pp. 177–178. ISBN: 978-1-349-08452-4.

[HS16] T. A. Huber and D. Sornette. "Can there be a physics of financial markets? Methodological reflections on econophysics." In: *European Physical Journal Special Topics* 225 (Dec. 2016). DOI: 10.1140/epjst/e2016-60158-5.

[IKR98] S. Ispolatov, P.L. Krapivsky, and S. Redner. "Wealth distributions in asset exchange models." In: *The European Physical Journal B - Condensed Matter and Complex Systems* 2.2 (1998), pp. 267–276. ISSN: 1434-6036. DOI: 10.1007/s100510050249. URL: http://dx.doi.org/10.1007/s100510050249.

[KH53] M. G. Kendall and A. Bradford Hill. "The Analysis of Economic Time-Series-Part I: Prices." In: *Journal of the Royal Statistical Society. Series A (General)* 116.1 (1953), pp. 11–34. ISSN: 00359238. URL: http://www.jstor.org/stable/2980947.

[Kim+15] J. Kim, B. C. Keegan, S. Park, and A. Oh. “The Proficiency-Congruency Dilemma: Virtual Team Design and Performance in Multiplayer Online Games.” In: *ArXiv e-prints* (Dec. 2015). arXiv: 1512.08321 [cs.HC].

[KW04] Tae-Hwan Kim and Halbert White. “On more robust estimation of skewness and kurtosis.” In: *Finance Research Letters* 1.1 (2004), pp. 56–73. ISSN: 1544-6123. DOI: https://doi.org/10.1016/S1544-6123(03)00003-5. URL: http://www.sciencedirect.com/science/article/pii/S1544612303000035.

[KSY06] Ken Kiyono, Zbigniew R. Struzik, and Yoshiharu Yamamoto. “Criticality and Phase Transition in Stock-Price Fluctuations.” In: *Phys. Rev. Lett.* 96 (6 Feb. 2006), p. 068701. DOI: 10.1103/PhysRevLett.96.068701. URL: https://link.aps.org/doi/10.1103/PhysRevLett.96.068701.

[Kol13] Phil Kollar. *The past, present and future of League of Legends studio Riot Games*. https://www.polygon.com/2016/9/13/12891656/the-past-present-and-future-of-league-of-legends-studio-riot-games. Accessed: 2017-05-12. 2016-09-13.

[LM99] Andrew W. Lo and A. Craig MacKinlay. *A Non-Random Walk Down Wall Street*. Princeton University Press, 1999. ISBN: 9780691092560. URL: http://www.jstor.org/stable/j.ctt7tccx.

[LF07] E. Lofgren and N. Fefferman. “The Untapped Potential of Virtual Game Worlds to Shed Light on Real World Epidemics.” In: *The Lancet Infectious Diseases* 7 (2007). DOI: 10.1016/S1473-3099(07)70212-8.

[Lor05] M. O. Lorenz. “Methods of Measuring the Concentration of Wealth.” In: *Publications of the American Statistical Association* 9.70 (1905), pp. 209–219. DOI: 10.1080/15225437.1905.10503443. eprint: http://www.tandfonline.com/doi/pdf/10.1080/15225437.1905.10503443. URL: http://www.tandfonline.com/doi/abs/10.1080/15225437.1905.10503443.

[MN09] Charles M. Macal and Michael J. North. “Agent-based Modeling and Simulation.” In: *Winter Simulation Conference*. WSC ’09. Winter Simulation Conference, 2009, pp. 86–98. ISBN: 978-1-4244-5771-7. URL: http://dl.acm.org/citation.cfm?id=1995456.1995474.

[Mal73] Burton Malkiel. *A Random Walk Down Wall Street*. Z.Z. Norton and Company, Inc., 1973. ISBN: 0-393-06245-7.

[Mal03] Burton G. Malkiel. "The Efficient Market Hypothesis and Its Critics." In: *Journal of Economic Perspectives* 17.1 (Mar. 2003), pp. 59–82. DOI: 10.1257/089533003321164958. URL: http://www.aeaweb.org/articles?id=10.1257/089533003321164958.

[Man63] Benoit Mandelbrot. "The Variation of Certain Speculative Prices." In: *The Journal of Business* 36 (1963). URL: http://EconPapers.repec.org/RePEc:ucp:jnlbus:v:36:y:1963:p:394.

[Mer73] Robert Merton. "An Intertemporal Capital Asset Pricing Model." In: *Econometrica* 41.5 (1973), pp. 867–87. URL: http://EconPapers.repec.org/RePEc:ecm:emetrp:v:41:y:1973:i:5:p:867-87.

[New05] MEJ Newman. "Power laws, Pareto distributions and Zipf's law." In: *Contemporary Physics* 46.5 (2005), pp. 323–351. DOI: 10.1080/00107510500052444. eprint: http://www.tandfonline.com/doi/pdf/10.1080/00107510500052444. URL: http://www.tandfonline.com/doi/abs/10.1080/00107510500052444.

[Pag96] Adrian Pagan. "The econometrics of financial markets." In: *Journal of Empirical Finance* 3.1 (1996), pp. 15–102. URL: http://EconPapers.repec.org/RePEc:eee:empfin:v:3:y:1996:i:1:p:15-102.

[Par96] V. Pareto. *Cours d'économie politique professé á l'Université de Lausanne*. F. Rouge, 1896. URL: https://books.google.be/books?id=eIAOnQEACAAJ.

[Par06] Vilfredo Pareto. *Manual of Political Economy*. [1971] Translated by Schwier Ann. Augustus M. Kelley, 1906. ISBN: 0678008817.

[Pec] Branko Pecar. *Associations Between and Within The Time series*. http://pecar-uk.com/Autocorrelations.pdf. Accessed: 2017-05-16.

[Pfi+13] Roland Pfister, Katharina A Schwarz, Markus Janczyk, Rick Dale, and Jon Freeman. "Good things peak in pairs: a note on the bimodality coefficient." In: *Frontiers in psychology* 4 (2013), p. 700.

[PS06] V. Pisarenko and D. Sornette. "New statistic for financial return distributions: Power-law or exponential?" In: *Physica A Statistical Mechanics and its Applications* 366 (July 2006), pp. 387–400. DOI: 10.1016/j.physa.2005.10.015. eprint: physics/0403075.

[Por86] Theodore M. Porter. *The Rise of Statistical Thinking, 1820-1900*. Princeton, NJ, USA: Princeton University Press, 1986. ISBN: 0691084165.

[Qua17a] Quandl. *Yahoo! Finance*. https://finance.yahoo.com/. Accessed: 2017-04-01. 2007-2017.

[Qua17b] Quandl. *Core Financial Data*. https://www.quandl.com/. Accessed: 2017-04-01. 2011-2017.

[Ric08] Dean Rickles. *Econophysics and Financial Market Complexity*. To appear in J. Collier and C. Hooker (eds.), Handbook of the Philosophy of Science, Vol.10: Philosophy of Complex Systems. North Holland: Elsevier. Jan. 2008. URL: http://philsci-archive.pitt.edu/3851/.

[Sam65] P. Samuelson. "Proof That Properly Anticipated Prices Fluctuate Randomly." In: *Industrial Management Review* 6.2 (1965), pp. 41–50. URL: https://www.ifa.com/media/images/pdf%7B%5C%%7D20files/samuelson-proof.pdf.

[Sch02] G. William Schwert. *Anomalies and Market Efficiency*. Working Paper 9277. National Bureau of Economic Research, Oct. 2002. DOI: 10.3386/w9277. URL: http://www.nber.org/papers/w9277.

[Sig13] Karl Sigman. *Simulating Brownian motion (BM) and geometric Brownian motion (GBM)*. http://www.columbia.edu/~ks20/4404-Sigman/4404-Notes-sim-BM.pdf. Accessed: 2017-05-26. 2013.

[SY05] A. C. Silva and V. M. Yakovenko. "Temporal evolution of the thermal and superthermal income classes in the USA during 1983 2001." In: *EPL (Europhysics Letters)* 69 (Jan. 2005), pp. 304–310. DOI: 10.1209/epl/i2004-10330-3. eprint: cond-mat/0406385.

[SST02] Jonathan Silver, Eric Slud, and Keiji Takamoto. "Statistical Equilibrium Wealth Distributions in an Exchange Economy with Stochastic Preferences." In: *Journal of Economic Theory* 106.2 (2002), pp. 417–435. URL: http://EconPapers.repec.org/RePEc:eee:jetheo:v:106:y:2002:i:2:p:417-435.

[Sim96] Herbert A. Simon. *The Sciences of the Artificial (3rd Ed.)* Cambridge, MA, USA: MIT Press, 1996. ISBN: 0-262-69191-4.

[Sup17] Superdata. *The MMO & MOBA Games Market Report, 2016*. https://www.superdataresearch.com/market-data/mmo-market/. Accessed: 2017-05-12. 2017.

[Sut97] John Sutton. "Gibrat's legacy." In: *Journal of economic literature* 35.1 (1997), pp. 40–59.

[Sze+12] M. Szell, R. Sinatra, G. Petri, S. Thurner, and V. Latora. "Understanding mobility in a social petri dish." In: *Scientific Reports* 2, 457 (June 2012), p. 457. DOI: 10.1038/srep00457. arXiv: 1112.1220 [physics.soc-ph].

[Voi01] J. Voit. *The Statistical Mechanics of Financial Markets*. Physics and astronomy online library. Springer, 2001. ISBN: 9783540414094. URL: https://books.google.be/books?id=TV-IQgAACAAJ.

[WFM07] Paul Windrum, Giorgio Fagiolo, and Alessio Moneta. "Empirical Validation of Agent-Based Models: Alternatives and Prospects." In: *Journal of Artificial Societies and Social Simulation* 10.2 (2007), p. 8. ISSN: 1460-7425. URL: http://jasss.soc.surrey.ac.uk/10/2/8.html.

[WW57] H. O. A. Wold and P. Whittle. "A Model Explaining the Pareto Distribution of Wealth." In: *Econometrica* 25.4 (1957), pp. 591–595.

[YR09] V. M. Yakovenko and J. B. Rosser Jr. "Colloquium: Statistical mechanics of money, wealth, and income." In: *Reviews of Modern Physics* 81 (Oct. 2009), pp. 1703–1725. DOI: 10.1103/RevModPhys.81.1703. arXiv: 0905.1518 [q-fin.ST].

[Yak12] V.M. Yakovenko. "Applications of statistical mechanics to economics: Entropic origin of the probability distributions of money, income, and energy consumption." In: *ArXiv e-prints* (Apr. 2012). arXiv: 1204.6483 [q-fin.ST].