Ergodicity Economics

Ole Peters and Alexander Adamou

LONDON MATHEMATICAL LABORATORY

2018/06/25 at 20:51:03

Contents

1 Populations

The previous chapter developed a model of individual behaviour based on an assumed dynamic imposed on wealth. If we know the stochastic process that describes individual wealth, then we also know what happens at population level – each individual is represented by a realisation of the process, and we can compute the dynamics of wealth distributions. We answer questions about inequality and poverty in our model economy. It turns out that our decision criterion generates interesting emergent behaviour – cooperation, the sharing and pooling of resources, is often time-average growth optimal. This provides answers to the puzzles of why people cooperate, why there is an insurance market, and why we see socio-economic structure from the formation of firms to nation states with taxation and redistribution systems.

1.1 Every man for himself

{section:Every_man}

We have seen that risk aversion constitutes optimal behaviour under the assumption of multiplicative wealth growth and over time scales that are long enough for systematic trends to be significant. In this chapter we will continue to explore our null model, GBM. By "explore" we mean that we will let the model generate its world – if individual wealth was to follow GBM, what kind of features of an economy would emerge? We will see that cooperation and the formation of social structure also constitute optimal behaviour.

GBM is more than a random variable. It's a stochastic process, either a set of trajectories or a family of time-dependent random variables, depending on how we prefer to look at it. Both perspectives are informative in the context of economic modelling: from the set of trajectories we can judge what is likely to happen to an individual, *e.g.* by following a single trajectory for a long time; while the PDF of the random variable $x(t^*)$ at some fixed value of t^* tells us how wealth is distributed in our model.

We use the term wealth distribution to refer to the density function $\mathcal{P}_x(x)$ (not to the process of distributing wealth among people). This can be interpreted as follows: if I select a random individual (each individual with uniform probability $\frac{1}{N}$), the probability of the selected individual having wealth greater than x is given by the CDF $F_x(x) = \int_x^\infty \mathcal{P}_x(s)\, ds$. In a large population of N individuals, $\Delta x \mathcal{P}_x(x) N$ is the approximate number of individuals who have wealth between x and $x + \Delta x$. Thus, a broad wealth distribution with heavy tails indicates greater wealth inequality.

Examples:

- Under perfect equality everyone would have the same, meaning that the wealth distribution would be a Dirac delta function centred at the sample mean of x, that is
$$\mathcal{P}_x(x) = \delta(x - \langle x \rangle_N); \tag{1}$$
- Maximum inequality would mean that one individual owns everything and everyone else owns nothing, that is
$$\mathcal{P}_x(x) = \frac{N-1}{N}\delta(x-0) + \frac{1}{N}\delta(x - N\langle x \rangle_N). \tag{2}$$

1.1.1 Log-normal distribution

{section:Log-normal_wealt

At a given time, t, GBM produces a random variable, $x(t)$, with a log-normal distribution whose parameters depend on t. (A log-normally distributed random variable is one whose logarithm is a normally distributed random variable.) If each individual's wealth follows GBM,

$$dx = x(\mu dt + \sigma dW), \tag{3}$$ {eq:GBM}

with solution

$$x(t) = x(0) \exp\left[\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W(t)\right], \tag{4}$$ {eq:GBM_sol}

then we will observe a log-normal distribution of wealth at each moment in time:

$$\ln x(t) \sim \mathcal{N}\left(\ln x(0) + \left(\mu - \frac{\sigma^2}{2}\right) t, \sigma^2 t\right). \tag{5}$$ {eq:lognormal}

It will be convenient hereafter to assume the initial condition $x(0) = 1$ (and, therefore, $\ln x(0) = 0$) unless otherwise stated.

Note that the variance of $\ln x(t)$ increases linearly in time. We will develop an understanding of this shortly. As we will see, it indicates that any meaningful measure of inequality will grow over time in our simple model. To see what kind of a wealth distribution (Eq. 5) is, it is worth spelling out the lognormal PDF:

$$\mathcal{P}_x(x) = \frac{1}{x\sqrt{2\pi\sigma^2 t}} \exp\left(-\frac{[\ln x - (\mu - \frac{\sigma^2}{2})t]^2}{2\sigma^2 t}\right). \tag{6}$$ {eq:PDFx}

This distribution is the subject of a wonderful book [2], sadly out-of-print now. We will find it useful to know a few of its basic properties. Of particular importance is the expected wealth under this distribution. This is

$$\langle x(t) \rangle = \exp(\mu t) \tag{7}$$ {eq:exp_x}

or, equivalently, $\ln \langle x(t) \rangle = \mu t$. We could confirm this result by calculating $\langle x(t) \rangle = \int_0^\infty s\mathcal{P}_x(s)\, ds$, but this would be laborious. Instead we use a neat trick, courtesy of [42, Chapter 4.2], which will come in handy again in Sec. **??**. To compute moments, $\langle x^n \rangle$, of stochastic differential equations for x, like (Eq. 1.1.2), we find ordinary differential equations for the moments, which we know how to solve. For the first moment we do this simply by taking expectations of both sides of (Eq. 1.1.2). The noise term vanishes to turn the SDE for x into an ODE for $\langle x \rangle$:

$$\begin{aligned} \langle dx \rangle &= \langle x(\mu dt + \sigma dW) \rangle & (8) \\ d\langle x \rangle &= \langle x \rangle \mu dt + \sigma \overbrace{\langle dW \rangle}^{=0} & (9) \\ &= \langle x \rangle \mu dt. & (10) \end{aligned}$$

This is a very simple first-order linear differential equation for the expectation value of x. Its solution with initial condition $x(0) = 1$ is (Eq. 7).

For $\mu > 0$ the expected wealth grows exponentially over time, as do its population median and variance:

$$\begin{aligned} \text{median}[x(t)] &= \exp[(\mu - \sigma^2/2)t]; & (11) \\ \text{var}[x(t)] &= \exp(2\mu t)[\exp(\sigma^2 t) - 1]. & (12) \end{aligned}$$ {eq:median_x} {eq:var_x}

1.1.2 Two growth rates

{section:two_rates}

We will recap briefly here one of our key ideas, covered in detail in Sec. **??**, that the ensemble average of all possible trajectories of GBM grows at a different (faster) rate from that achieved by a single trajectory almost surely in the long-time limit. Understanding this difference was the key to developing a coherent theory of individual decision-making. We will see here that it is also crucial in understanding how wealth becomes distributed in a population of individuals whose wealths follow and, in particular, how we can measure the inequality in such a distribution. {eq:GBM}

From (Eq. **??**), we recall that the growth rate of the expected wealth is

$$g_{\langle\rangle} = \frac{d \ln \langle x \rangle}{dt} = \mu, \tag{13}$$

while, from (Eq. ??), the time-average growth rate of wealth is

$$\overline{g} = \frac{d \langle \ln x \rangle}{dt} = \mu - \frac{\sigma^2}{2}. \tag{14}$$

1.1.3 Measuring inequality

{section:Inequality_measu

In the case of GBM we have just seen how to compute the exact full wealth distribution $\mathcal{P}$. This is interesting but often we want only summary measures of the distribution. One such summary measure of particular interest to economists is inequality. How much inequality is there in a distribution like (Eq. 5)? And how does this quantity increase over time under GBM, as we have suggested?

Clearly, to answer these questions, we must quantify "inequality". In this section, and also in [1], we develop a natural way of measuring it, which makes use of the two growth rates we identified for the non-ergodic process. We will see that a particular inequality measure, known to economists as Theil's second index of inequality [76], is the difference between typical wealth (growing at the time-average growth rate) and average wealth (growing at the ensemble-average growth rate) in our model. Thus, the difference between the time average and ensemble average, the essence of ergodicity breaking, is the fundamental driver of the dynamics of inequality.

The two limits of inequality are easily identified: minimum inequality means that everyone has the same wealth, and maximum inequality means that one individual has all the wealth and everyone else has nothing. (This assumes that wealth cannot become negative.) Quantifying inequality in any other distribution is reminiscent of the gamble problem. Recall that for gambles we wanted make statements of the type "this gamble is more desirable than that gamble". We did this by collapsing a distribution to a scalar. Depending on the question that was being asked the appropriate way of collapsing the distribution and the resulting scalar can be different (the scalar relevant to an insurance company may not be relevant to an individual). In the case of inequality we also have a distribution – the wealth distribution – and we want to make statements of the type "this distribution is more unequal than that distribution". Again, this is done by collapsing the distribution to a scalar, and again many different choices of collapse and resulting scalar are possible. The Gini coefficient is a particularly well-known scalar of this type, the 80/20 ration is another, and many other measures exist.

In this context the expectation value is an important quantity. For instance, if everyone has the same wealth, everyone will own the average $\forall i, x_i = \langle x \rangle_N$, which converges to the expectation value for large N. Also, whatever the distribution of wealth, the total wealth is $N \langle x \rangle_N$ which converges to $N \langle x \rangle$ as N grows large. The growth rate of the expectation value, $g_{\langle\rangle}$, thus tells us how fast the average wealth and the total population wealth grow with probability one in a large ensemble. The time-average growth rate, $\overline{g}$, on the other hand, tells us how fast an individual's wealth grows with probability one in the long run. If the typical individual's wealth grows at a lower rate than the expectation value of wealth then there must be atypical individuals with very large wealths that account for the difference. This insight suggests the following measure of inequality.

<u>Definition</u> **Inequality, J, is the quantity whose growth rate is the**

difference between expectation-value and time-average growth rates,

$$\frac{dJ}{dt} = g_{\langle\rangle} - \overline{g}. \tag{15}$$ {eq:dJ}

Equation (15) defines the dynamic of inequality, and inequality itself is found by integrating over time:

$$J(t) = \int_0^t ds[g_{\langle\rangle}(s) - \overline{g}(s)]. \tag{16}$$ {eq:J}

This definition may be used for dynamics other than GBM. Whatever the wealth dynamic, typical minus average growth rates are informative of the dynamic of inequality. Within the GBM framework we can write the difference in growth rates as

$$\frac{dJ}{dt} = \frac{d \ln \langle x \rangle}{dt} - \frac{d \langle \ln x \rangle}{dt} \tag{17}$$ {eq:J_dyn}

and integrate over time to get

$$J(t) = \ln \langle x \rangle - \langle \ln x \rangle \,. \tag{18}$$

This quantity is known as the mean logarithmic deviation (MLD) or Theil's second index of inequality [76]. This is rather remarkable. Our general inequality measure, (Eq. 16), evaluated for the specific case of GBM, turns out to be a well-known measure of inequality that economists have identified independently, without considering non-ergodicity and ensemble average and time average growth rates. Merely by insisting on measuring inequality well, Theil used the GBM model without realising it!

Substituting the known values of the two growth rates into (Eq. 15) and integrating, we can evaluate the Theil inequality as a function of time:

$$J(t) = J(0) + \frac{\sigma^2}{2} t. \tag{19}$$ {eq:J_t}

Thus we see that, in GBM, our measure of inequality increases indefinitely.

1.1.4 Wealth condensation

{section:condensation}

The log-normal distribution generated by GBM broadens indefinitely, (Eq. 12). Likewise, the inequality present in the distribution – measured as the time-integrated difference between ensemble and time average growth rates – grows continually. A related property of GBM is the evolution towards wealth condensation. Wealth condensation means that a single individual will own a non-zero fraction of the total wealth in the population in the limit of large N, see *e.g.* [13]. In the present case an arbitrarily large share of total wealth will be owned by an arbitrarily small share of the population.

One simple way of seeing this is to calculate the fraction of the population whose wealths are less than the mean, *i.e.* $x(t) < \exp(\mu t)$. To do this, we define a new random variable, $z(t)$, whose distribution is the standard normal:

$$z(t) \equiv \frac{\ln x(t) - (\mu - \sigma^2/2)t}{\sigma t^{1/2}} \sim \mathcal{N}(0,1). \tag{20}$$

We want to know the mass of the distribution with $\ln x(t) < \mu t$ or, equivalently, $z < \sigma t^{1/2}/2$. This is

$$\Phi\left(\frac{\sigma t^{1/2}}{2}\right), \tag{21}$$

where Φ is the CDF of the standard normal distribution. This fraction tends to one as $t \to \infty$.

1.1.5 Rescaled wealth

{section:rescaled}

Over the decades, economists have arrived at many inequality measures, and have drawn up a list of conditions that particularly useful measures of inequality satisfy. Such measures are called "relative measures" [72, Appendix 4], and J is one of them. One of the conditions is that inequality measures should not change when x is divided by the same factor for everyone. Since we are primarily interested in inequality in this section it is useful to remove absolute wealth levels from the analysis and study an object called the rescaled wealth.

Definition The rescaled wealth,

$$y = \frac{x}{\langle x \rangle_N}, \tag{22}$$

{eq:rescaled}

is the proportion of the population-average wealth owned by an individual.

This quantity is useful, for instance because its numerical value does not depend on the currency used, it is a dimensionless number. Thus if my rescaled wealth, $y = 1/2$, this means that my wealth is half the average wealth, irrespective of whether I measure wealth in Kazakhstani Tenge or in Swiss Francs. For a large population, (Eq. 17) may be expressed in terms of y as $\frac{dJ}{dt} = -\frac{d\langle \ln y \rangle}{dt}$.

1.1.6 Power law resemblance

{section:power_law}

It is an established empirical observation [55] that the upper tails of real wealth distributions look more like a power law than a log-normal. Our trivial model does not strictly reproduce this feature, but it is instructive to compare the lognormal distribution to a power-law distribution. A power law PDF has the asymptotic form

$$\mathcal{P}_x(x) = x^{-\alpha}, \tag{23}$$

{eq:power_law}

for large arguments x. This implies that the logarithm of the PDF is proportional to the logarithm of its argument, $\ln \mathcal{P}_x(x) = -\alpha \ln x$. Plotting one against the other will yield a straight line, the slope being the exponent $-\alpha$.

Determining whether an empirical observation is consistent with such behaviour is difficult because the behaviour is to be observed in the tail (large x) where data are, by definition, sparse. A quick-and-dirty way of checking for possible power-law behaviour is to plot an empirical PDF against its argument on log-log scales, look for a straight line, and measure the slope. However, plotting any distribution on any type of scales results in some line. It may not be a straight line but it will have some slope everywhere. For a known distribution (power law or not) we can interpret this slope as a local apparent power-law exponent.

What is the local apparent power-law exponent of a log-normal wealth distribution near the expectation value $\langle x\rangle = \exp(\mu t)$, *i.e.* in the upper tail where approximate power law behaviour has been observed empirically? The logarithm of (Eq. 6) is

$$\begin{aligned}\ln \mathcal{P}(x) = & -\ln\left(x\sqrt{2\pi\sigma^2 t}\right) - \tfrac{[\ln x-(\mu-\frac{\sigma^2}{2})t]^2}{2\sigma^2 t} & (24)\\ = & -\ln x - \tfrac{\ln(2\pi\sigma^2 t)}{2} - \tfrac{(\ln x)^2-2(\mu-\frac{\sigma^2}{2})t\ln x+(\mu-\frac{\sigma^2}{2})^2 t^2}{2\sigma^2 t}. & (25)\end{aligned}$$

Collecting terms in powers of $\ln x$ we find

$$\ln \mathcal{P}(x) = -\frac{(\ln x)^2}{2\sigma^2 t} + \left(\frac{\mu}{\sigma^2} - \frac{3}{2}\right)\ln x - \frac{\ln(2\pi\sigma^2 t)}{2} - \frac{(\mu - \frac{\sigma^2}{2})^2 t}{2\sigma^2} \qquad (26)$$

with local slope, *i.e.* apparent exponent,

$$\frac{d\ln \mathcal{P}(x)}{d\ln x} = -\frac{\ln x}{\sigma^2 t} + \frac{\mu}{\sigma^2} - \frac{3}{2}. \qquad (27)$$

Near $\langle x\rangle$, $\ln x \sim \mu t$ so that the first two terms cancel approximately. Here the distribution will resemble a power-law with exponent $-3/2$ when plotted on doubly logarithmic scales. This is consistent with exponents between -1.6 and -1.2 measured from wealth data. (The distribution will also look like a power-law where the first term is much smaller than the others, *e.g.* where $\ln x \ll \sigma^2 t$.) We don't believe that such empirically observed power laws are merely a manifestation of this mathematical feature. Important real-world mechanisms that broaden real wealth distributions, *i.e.* concentrate wealth, are missing from the null model. However, it is interesting that the trivial model of GBM reproduces so many qualitative features of empirical observations.

1.2 Finite populations

{section:finite_populatio

So far we have considered the properties the random variable, $x(t)$, generated by GBM at a fixed time, t. Most of the mathematical objects we have discussed are, strictly speaking, relevant only in the limit $N \to \infty$, where N is the number of realisations of this random variable. For example, the expected wealth, $\langle x(t)\rangle$, is the limit of the sample mean wealth

$$\langle x(t)\rangle_N \equiv \frac{1}{N}\sum_{i=1}^{N} x_i(t), \qquad (28)$$ {eq:sample}

as the sample size, N, grows large. In reality, human populations can be very large, say $N \sim 10^7$ for a nation state, but they are most certainly finite. Therefore, we need to be diligent and ask what the effects of this finiteness are. In particular, we will focus on the sample mean wealth under GBM. For what values of μ, σ, t, and N is this well approximated by the expectation value? And when it is not, what does it resemble?

1.2.1 Sums of lognormals

{section:sketch}

In [62] we studied the sample mean of GBM, which we termed the "partial ensemble average" (PEA). This is the average of N independent realisations the random variable $x(t)$, (Eq. 28). Here we sketch out some simple arguments about how this object depends on N and t.

Considering the two growth rates in Sec. 1.1.2, we anticipate the following tension:

A) for large N, the PEA should resemble the expectation value, $\exp(\mu t)$;

B) for long t, all trajectories in the sample – and, therefore, the sample mean – should grow like $\exp[(\mu - \sigma^2/2)t]$.

Situation A – when a sample mean resembles the corresponding expectation value – is known in statistical physics as "self-averaging." A simple strategy for estimating when this occurs is to look at the relative variance of the PEA,

$$R \equiv \frac{\mathrm{var}(\langle x(t) \rangle_N)}{\langle \langle x(t) \rangle_N \rangle^2}. \tag{29}$$

To be explicit, here the $\langle \cdot \rangle$ and $\mathrm{var}(\cdot)$ operators, without N as a subscript, refer to the mean and variance over all possible PEAs. The PEAs themselves, taken over finite samples of size N, are denoted $\langle \cdot \rangle_N$. Using standard results for the mean and variance of sums of independent random variables and inserting the results in (Eq. 7) and (Eq. 12), we get

$$R(N) = \frac{e^{\sigma^2 t} - 1}{N}. \tag{30}$$

If $R \ll 1$, then the PEA will likely be close to its own expectation value, which is equal to the expectation value of the GBM. Thus, in terms of N and t, $\langle x(t) \rangle_N \approx \langle x(t) \rangle$ when

$$t < \frac{\ln N}{\sigma^2}. \tag{31}$$ {eq:short_t}

This hand-waving tells us roughly when the large-sample – or, as we see from (Eq. 31), short-time – self-averaging regime holds. A more careful estimate of the cross-over time in (Eq. 47) is a factor of 2 larger, but the scaling is identical.

For $t > \ln N/\sigma^2$, the growth rate of the PEA transitions from μ to its $t \to \infty$ limit of $\mu - \sigma^2/2$ (Situation B). Another way of viewing this is to think about what dominates the average. For early times in the process, all trajectories are close together and none dominate the PEA. However, as time goes by the distribution broadens exponentially. Since each trajectory contributes with the same weight to the PEA, after some time the PEA will be dominated by the maximum in the sample,

$$\langle x(t) \rangle_N \approx \frac{1}{N} \max_{i=1}^{N} \{x_i(t)\}, \tag{32}$$

as illustrated in Fig. 1.

Self-averaging stops when even the "luckiest" trajectory is no longer close to the expectation value $\exp(\mu t)$. This is guaranteed to happen eventually because the probability for a trajectory to reach $\exp(\mu t)$ decreases towards zero as t grows. We know this from Sec. 1.1.4. Of course, this takes longer for larger samples, which have more chances to contain a lucky trajectory.

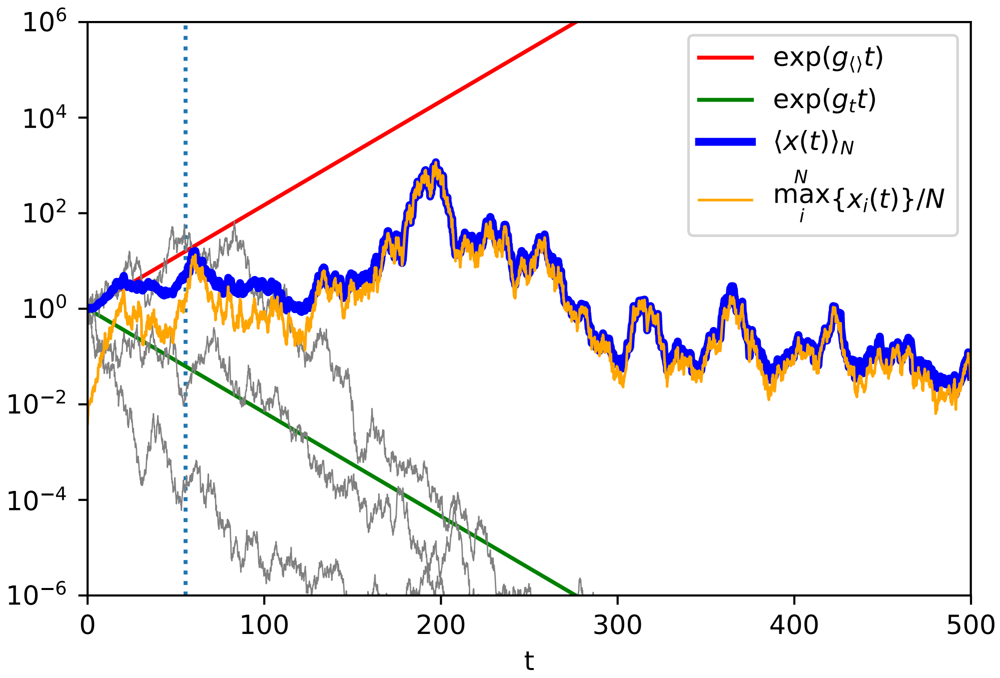

Figure 1: PEA and maximum in a finite ensemble of size $N = 256$. **Red line:** expectation value $\langle x(t)\rangle$. **Green line:** exponential growth at the time-average growth rate. In the $T \to \infty$ limit all trajectories grow at this rate. **Yellow line:** contribution of the maximum value of any trajectory at time t to the PEA. **Blue line:** PEA $\langle x(t)\rangle_N$. **Vertical line:** Crossover – for $t > t_c = \frac{2\ln N}{\sigma^2}$ the maximum begins to dominate the PEA (the yellow line approaches the blue line). **Grey lines:** randomly chosen trajectories – any typical trajectory soon grows at the time-average growth rate. **Parameters:** $N = 256$, $\mu = 0.05$, $\sigma = \sqrt{0.2}$. {fig:trajectories}

In [62] we analysed PEAs of GBM analytically and numerically. Using (Eq. 4) the PEA can be written as

$$\langle x\rangle_N = \frac{1}{N}\sum_{i=1}^{N} \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_i(t)\right], \qquad (33) \quad \texttt{\{eq:PEA\}}$$

where $\{W_i(t)\}_{i=1\ldots N}$ are N independent realisations of the Wiener process. Taking the deterministic part out of the sum we re-write (Eq. 33) as

$$\langle x\rangle_N = \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)t\right] \frac{1}{N}\sum_{i=1}^{N} \exp\left(t^{1/2}\sigma\xi_i\right), \qquad (34) \quad \texttt{\{eq:PEA\_2\}}$$

where $\{\xi_i\}_{i=1\ldots N}$ are N independent standard normal variates.

We found that typical trajectories of PEAs grow at $g_{\langle\rangle}$ up to a time t_c that is logarithmic in N, meaning $t_c \propto \ln N$. This is consistent with our analytical sketch. After this time, typical PEA trajectories begin to deviate from expectation-value behaviour, and eventually their growth rate converges to g_t. While the two limiting behaviours $N \to \infty$ and $t \to \infty$ can be computed exactly, what happens in between is less straightforward. The PEA is a random object outside these limits.

A quantity of crucial interest to us is the exponential growth rate experienced by the PEA,

$$g_{\text{est}}(t, N) \equiv \frac{\ln(\langle x(t)\rangle_N) - \ln(x(0))}{t - 0} = \frac{1}{t}\ln(\langle x(t)\rangle_N). \qquad (35) \quad \texttt{\{eq:gest\}}$$

In [62] we proved that the $t \to \infty$ limit for any (finite) N is the same as for the case $N = 1$,

$$\lim_{t\to\infty} g_{\text{est}}(t, N) = \mu - \frac{\sigma^2}{2} \qquad (36) \quad \texttt{\{eq:gest\_2\}}$$

for all $N \geq 1$. Substituting (Eq. 34) in (Eq. 35) produces

$$\begin{aligned} g_{\text{est}}(t, N) &= \mu - \frac{\sigma^2}{2} + \frac{1}{t}\ln\left(\frac{1}{N}\sum_{i=1}^{N}\exp(t^{1/2}\sigma\xi_i)\right) && (37) \\ &= \mu - \frac{\sigma^2}{2} - \frac{\ln N}{t} + \frac{1}{t}\ln\left(\sum_{i=1}^{N}\exp(t^{1/2}\sigma\xi_i)\right). && (38) \quad \texttt{\{eq:gest\_4\}} \end{aligned}$$

We didn't look in [62] at the expectation value of $g_{\text{est}}(t, N)$ for finite time and finite samples, but it's an interesting object that depends on N and t but is not stochastic. Note that this is not g_{est} of the expectation value, which would be the $N \to \infty$ limit of (Eq. 35). Instead it is the $S \to \infty$ limit,

$$\langle g_{\text{est}}(t, N)\rangle = \frac{1}{t}\langle \ln(\langle x(t)\rangle_N)\rangle = f(N, t), \qquad (39) \quad \texttt{\{eq:gest\_3\}}$$

where, as previously, $\langle\cdot\rangle$ without subscript refers to the average over all possible samples, *i.e.* $\lim_{S\to\infty}\langle\cdot\rangle_S$. The last two terms in (Eq. 38) suggest an exponential relationship between ensemble size and time. The final term is a tricky stochastic object on which the properties of the expectation value in (Eq. 39) will hinge. This term will be the focus of our attention: the sum of exponentials of normal random variates or, equivalently, log-normal variates.

1.2.2 The random energy model

Since the publication of [62] we have learned, thanks to discussions with J.-P. Bouchaud, that the key object in (Eq. 38) – the sum log-normal random variates – has been of interest to the mathematical physics community since the 1980s. The reason for this is Derrida's random energy model [25, 26].

It is defined as follows. Imagine a system whose energy levels are $2^K = N$ normally-distributed random numbers, ξ_i (corresponding to K spins). This is a very simple model of a disordered system, such as a spin glass, the idea being that the system is so complicated that we "give up" and simply model its energy levels as realisations of a random variable. (We denote the number of spins by K and the number of resulting energy levels by N, while Derrida uses N for the number of spins). In this model The partition function is then

$$Z = \sum_{i=1}^{N} \exp\left(\beta J \sqrt{\frac{K}{2}} \xi_i\right), \qquad (40) \quad \{\text{eq:Z}\}$$

where the inverse temperature, β, is measured in appropriate units, and the scaling in K is chosen so as to ensure an extensive thermodynamic limit [25, p. 79]. J is a constant that will be determined below. The logarithm of the partition function gives the Helmholtz free energy,

$$\begin{aligned} F &= -\frac{\ln Z}{\beta} && (41) \\ &= -\frac{1}{\beta} \ln\left[\sum_{i=1}^{N} \exp\left(\beta J \sqrt{\frac{K}{2}} \xi_i\right)\right]. && (42) \quad \{\text{eq:F}\} \end{aligned}$$

Like the growth rate estimator in (Eq. 35), this involves a sum of log-normal variates and, indeed, we can rewrite (Eq. 38) as

$$g_{\text{est}} = \mu - \frac{\sigma^2}{2} - \frac{\ln N}{t} - \frac{\beta F}{t}, \qquad (43) \quad \{\text{eq:gest\_5}\}$$

which is valid provided that

$$\beta J \sqrt{\frac{K}{2}} = \sigma t^{1/2}. \qquad (44) \quad \{\text{eq:map}\}$$

Equation (44) does not give a unique mapping between the parameters of our GBM, (σ, t), and the parameters of the REM, (β, K, J). Equating (up to multiplication) the constant parameters, σ and J, in each model gives us a specific mapping:

$$\sigma = \frac{J}{\sqrt{2}} \quad \text{and} \quad t^{1/2} = \beta\sqrt{K}. \qquad (45) \quad \{\text{eq:choice\_1}\}$$

The expectation value of g_{est} is interesting. The only random object in (Eq. 43) is F. Knowing $\langle F \rangle$ thus amounts to knowing $\langle g_{\text{est}} \rangle$. In the statistical mechanics of the random energy model $\langle F \rangle$ is of key interest and so much about it is known. We can use this knowledge thanks to the mapping between the two problems.

Derrida identifies a critical temperature,

$$\frac{1}{\beta_c} \equiv \frac{J}{2\sqrt{\ln 2}}, \qquad (46) \quad \{\text{eq:beta\_c}\}$$

above and below which the expected free energy scales differently with K and β. This maps to a critical time scale in GBM,

$$t_c = \frac{2 \ln N}{\sigma^2}, \tag{47}$$ {eq:t_c}

with high temperature ($1/\beta > 1/\beta_c$) corresponding to short time ($t < t_c$) and low temperature ($1/\beta < 1/\beta_c$) corresponding to long time ($t > t_c$). Note that t_c in (Eq. 47) scales identically with N and σ as the transition time, (Eq. 31), in our sketch.

In [25], $\langle F \rangle$ is computed in the high-temperature (short-time) regime as

$$\begin{aligned} \langle F \rangle &= E - S/\beta && (48) \\ &= -\frac{K}{\beta}\ln 2 - \frac{\beta K J^2}{4}, && (49) \end{aligned}$$ {eq:F_2}

and in the low-temperatures (long-time) regime as

$$\langle F \rangle = -KJ\sqrt{\ln 2}. \tag{50}$$ {eq:F_3}

Short time

We look at the short-time behavior first (high $1/\beta$, (Eq. 49)). The relevant computation of the entropy S in [25] involves replacing the number of energy levels $n(E)$ by its expectation value $\langle n(E) \rangle$. This is justified because the standard deviation of this number is $\sqrt{n}$ and relatively small when $\langle n(E) \rangle > 1$, which is the interesting regime in Derrida's case.

For spin glasses, the expectation value of F is interesting, supposedly, because the system may be self-averaging and can be thought of as an ensemble of many smaller sub-systems that are essentially independent. The macroscopic behavior is then given by the expectation value.

Taking expectation values and substituting from (Eq. 49) in (Eq. 43) we find

$$\langle g_{\text{est}} \rangle^{\text{short}} = \mu - \frac{\sigma^2}{2} + \frac{1}{t}\frac{KJ^2}{4T^2}. \tag{51}$$ {eq:gest_6}

From (Eq. 44) we know that $t = \frac{KJ^2}{2\sigma^2 T^2}$, which we substitute, to find

$$\langle g_{\text{est}} \rangle^{\text{short}} = \mu, \tag{52}$$ {eq:gest_7}

which is the correct behavior in the short-time regime.

Long time

Next, we turn to the expression for the long-time regime (low temperature, (Eq. 50)). Again taking expectation values and substituting, this time from (Eq. 50) in (Eq. 43), we find for long times

$$\langle g_{\text{est}} \rangle^{\text{long}} = \mu - \frac{\sigma^2}{2} - \frac{\ln N}{t} + \sqrt{\frac{2 \ln N}{t}}\,\sigma, \tag{53}$$ {eq:gest_8}

which has the correct long-time asymptotic behavior. The form of the correction to the time-average growth rate in (Eq. 53) is consistent with [62] and [67], where it was found that approximately $N = \exp(t)$ systems are required for ensemble-average behavior to be observed for a time t, so that the parameter

$\ln N/t$ controls which regime dominates – if the parameter is small, then (Eq. 53) indicates that the long-time regime is relevant.

Figure 2 is a direct comparison between the results derived here, based on [25], and numerical results using the same parameter values as in [62], namely $\mu = 0.05, \sigma = \sqrt{0.2}, N = 256$ and $S = 10^5$.

Notice that $\langle g_{\text{est}} \rangle$ is not the (local) time derivative $\frac{\partial}{\partial t} \langle \ln(\langle x \rangle_N) \rangle$, but a time-average growth rate, $\left\langle \frac{1}{t} \ln\left(\frac{\langle x(t) \rangle_N}{\langle x(0) \rangle_N} \right) \right\rangle$. In [62] we used a notation that we've stopped using since then because it caused confusion – $\langle g \rangle$ there denotes the growth rate of the expectation value, which is not the expectation value of the growth rate.

It is remarkable that the expectation value $\langle g_{\text{est}}(N,t) \rangle$ so closely reflects the median, $q_{0.5}$, of $\langle x \rangle_N$, in the sense that

$$q_{0.5}(\langle x(t) \rangle_N) \approx \exp\left(\langle g_{\text{est}}(N,t) \rangle t\right). \qquad (54)$$ {eq:quant_ave}

In [61] it was discussed in detail that $g_{\text{est}}(1,t)$ is an ergodic observable for (Eq. 1.1.2), in the sense that $\langle g_{\text{est}}(1,t) \rangle = \lim_{t\to\infty} g_{\text{est}}$. The relationship in (Eq. 54) is far more subtle. The typical behavior of GBM PEAs is complicated outside the limits $N \to \infty$ or $t \to \infty$, in the sense that growth rates are time dependent here. This complicated behavior is well represented by an approximation that uses physical insights into spin glasses. Beautiful!

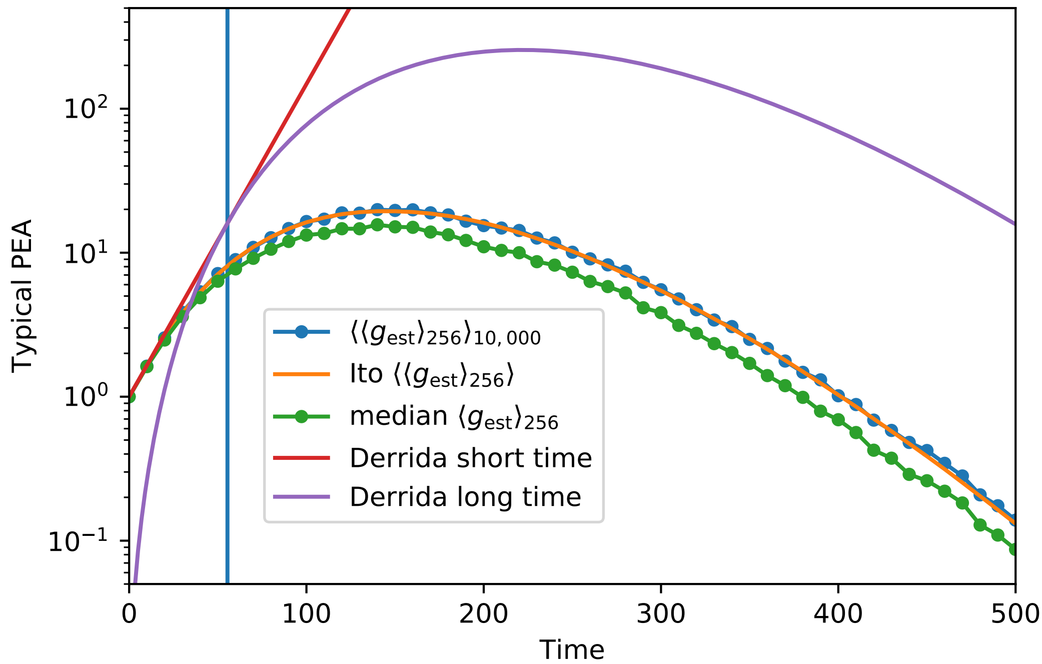

Figure 2: Lines are obtained by exponentiating the various exponential growth rates. **Blue line:** $\langle\langle g_{\text{est}}\rangle_{256}\rangle_{10,000}$ is the numerical mean (approximation of the expectation value) over a super-ensemble of $S = 10,000$ samples of g_{est} estimated in sub-ensembles of $N = 256$ GBMs each. **Green line:** median in a super-ensemble of S samples of g_{est}, each estimated in sub-ensembles of size N. **Yellow line:** (Eq. **??**) is an exact expression for $d\langle\ln\langle x\rangle_N\rangle$, derived using Itô calculus. We evaluate the expression by Monte Carlo, and integrate, $\langle\ln\langle x\rangle_N\rangle = \int_0^t d\langle\ln\langle x\rangle_N\rangle$. Exponentiation yields the yellow line. **Red line:** short-time behavior, based on the random energy model, (Eq. 52). **Purple line:** long-time behavior, based on the random energy model, (Eq. 53). **Vertical line:** Crossover between the regimes at $t_c = \frac{2\ln N}{\sigma^2}$, corresponding to $\beta_c = \frac{2(\ln 2)^{1/2}}{J}$. **Parameters:** $N = 256$, $S = 10,000$, $\mu = 0.05$, $\sigma = \sqrt{0.2}$. {fig:1}

List of Symbols

δ Most frequently used to express a difference, for instance δx is a difference between two wealths x. It can be the Kronecker delta function, a function of two arguments with properties $\delta(i,j) = 1$ if $i = j$ and $\delta(i,j) = 0$ otherwise. It can also be the Dirac delta function of one argument, $\int f(x)\delta(x - x_0)dx = f(x_0)$.

δt A time interval corresponding to the duration of one round of a gamble or, mathematically, the period over which a single realisation of the constituent random variable of a discrete-time stochastic process is generated..

Δ Difference operator, for instance Δv is a difference of two values of v, for instance observed at two different times.

η Langevin noise with the properties $\langle \eta \rangle = 0$ and $\langle \eta(t_1)\eta(t_2) \rangle = \delta(t_1 - t_2)$.

$g_{\mathbf{est}}$ Growth rate estimator for finite time and finite ensemble size.

$g_{\langle\rangle}$ Exponential growth rate of the expectation value.

$\overline{g}$ Time-average exponential growth rate.

$\mathcal{N}$ Normal distribution, $x \sim \mathcal{N}(\langle p \rangle, \mathrm{var}(p))$ means that the variable p is normally distributed with mean $\langle p \rangle$ and variance $\mathrm{var}(p)$..

p Probability, p_i is the probability of observing event i in a realization of a random variable.

$\mathcal{P}$ Probability density function.

t Time.

T Number of sequential iterations of a gamble, so that $T\delta t$is the total duration of a repeated gamble..

v Stationarity mapping function, so that $v(x)$ has stationary increments.

var Variance.

W Wiener process, $W(t) = \int_0^t dW$ is continuous and $W(t) \sim \mathcal{N}(0, \overline{g})$.

x Wealth.

ξ A standard normal variable, $\xi \sim \mathcal{N}(0,1)$.

References

[1] A. Adamou and O. Peters. Dynamics of inequality. *Significance*, 13(3):32–35, 2016.

[2] J. Aitchison and J. A. C. Brown. *The lognormal distribution.* Cambridge University Press, 1957.

[3] A. B. Atkinson. The timescale of economic models: How long is the long run? *The Review of Economic Studies*, 36(2):137–152, 1969.

[4] G. I. Barenblatt. *Scaling*. Cambridge University Press, 2003.

[5] C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers*. Springer, New York, 1978.

[6] J. Benhabib, A. Bisin, and S. Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1):123–157, 2011.

[7] Y. Berman. Understanding the mechanical relationship between inequality and intergenerational mobility. Available at: http://papers.ssrn.com/abstract=2796563, 2017.

[8] Y. Berman, E. Ben-Jacob, and Y. Shapira. The dynamics of wealth inequality and the effect of income distribution. *PloS one*, 11(6):e0199999, 2016.

[9] Y. Berman and Y. Shapira. Revisiting r > g—the asymptotic dynamics of wealth inequality. *Physica A: Statistical Mechanics and its Applications*, 467:562–572, 2017.

[10] D. Bernoulli. Specimen Theoriae Novae de Mensura Sortis. Translation "Exposition of a new theory on the measurement of risk" by L. Sommer (1954). *Econometrica*, 22(1):23–36, 1738.

[11] T. Bewley. The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory*, 16(2):252–292, 1977.

[12] J.-P. Bouchaud. On growth-optimal tax rates and the issue of wealth inequalities. *http://arXiv.org/abs/1508.00275*, August 2015.

[13] J.-P. Bouchaud and M. Mézard. Wealth condensation in a simple model of economy. *Physica A*, 282(4):536–545, 2000.

[14] J. Bricker, A. Henriques, J. Krimmel, and J. Sabelhaus. Measuring income and wealth at the top using administrative and survey data. *Brookings Papers on Economic Activity*, pages 261–331, 2016.

[15] M. Brzezinski. Do wealth distributions follow power laws? evidence from 'rich lists'. *Physica A: Statistical Mechanics and its Applications*, 406:155–162, 2014.

[16] T. W. Burkhardt. The random acceleration process in bounded geometries. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(07):P07004, 2007.

[17] J. Y. Campbell. *Financial decisions and markets, A course in asset pricing*. Princeton University Press, 2017.

[18] M. Corak. Income inequality, equality of opportunity, and intergenerational mobility. *The Journal of Economic Perspectives*, 27(3):79–102, 2013.

[19] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley & Sons, 1991.

[20] F. A. Cowell. Piketty in the long run. *The British Journal of Sociology*, 65(4):708–720, 2014.

[21] J. C. Cox, J. E. Ingersoll, and S. A. Ross. An intertemporal general equilibrium model of asset prices. *Econometrica*, 53(2):363–384, 1985.

[22] M. De Nardi. Quantitative models of wealth inequality: A survey. 2015. Working Paper 21106, NBER.

[23] M. De Nardi, G. Fella, and F. Yang. Piketty's book and macro models of wealth inequality. 2015. Working Paper 21730, NBER.

[24] M. A. B. Deakin. G. I. Taylor and the Trinity test. *Int. J. Math. Edu. Sci. Tech.*, 42(8):1069–1079, 2011.

[25] B. Derrida. Random-energy model: Limit of a family of disordered models. *Phys. Rev. Lett.*, 45(2):79–82, July 1980.

[26] B. Derrida. Random-energy model: An exactly solvable model of disordered systems. *Phys. Rev. B*, 24:2613–2626, September 1981.

[27] K. Devlin. *The unfinished game.* Basic Books, 2008.

[28] A. Drăgulescu and V. M. Yakovenko. Exponential and power-law probability distributions of wealth and income in the united kingdom and the united states. *Physica A: Statistical Mechanics and its Applications*, 299(1):213–221, 2001.

[29] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.*, IV(17):549–560, 1905.

[30] R. P. Feynman. *The Feynman lectures on physics.* Addison-Wesley, 1963.

[31] X. Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1(1):255–294, 2009.

[32] X. Gabaix, J.-M. Lasry, P.-L. Lions, and B. Moll. The dynamics of inequality. *Econometrica: Journal of the Econometric Society*, 84(6):2071–2111, 2016.

[33] J. M. Harrison. *Brownian motion of performance and control.* Cambridge University Press, 2013.

[34] H. Hinrichsen. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Adv. Phys.*, 49(7):815–958, 2000.

[35] J. C. Hull. *Options, Futures, and Other Derivatives.* Prentice Hall, 6 edition, 2006.

[36] K. Itô. Stochastic integral. *Proc. Imperial Acad. Tokyo*, 20:519–524, 1944.

[37] C. I. Jones. Pareto and piketty: The macroeconomics of top income and wealth inequality. *The Journal of Economic Perspectives*, 29(1):29–46, 2015.

[38] R. Kaas, M. Goovaerts, J. Dhaene, and M. Denuit. *Modern Actuarial Risk Theory.* Springer, 2 edition, 2008.

[39] B. Kaymak and M. Poschke. The evolution of wealth inequality over half a century: The role of taxes, transfers and technology. *Journal of Monetary Economics*, 77:1–25, 2016.

[40] J. L. Kelly Jr. A new interpretation of information rate. *Bell Sys. Tech. J.*, 35(4):917–926, July 1956.

[41] O. S. Klass, O. Biham, M. Levy, O. Malcai, and S. Solomon. The forbes 400 and the pareto wealth distribution. *Economics Letters*, 90(2):290–295, 2006.

[42] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23. Springer Science & Business Media, 1992.

[43] W. Kopczuk. What do we know about the evolution of top wealth shares in the united states? *The Journal of Economic Perspectives*, 29(1):47–66, 2015.

[44] W. Kopczuk and E. Saez. Top wealth shares in the united states, 1916–2000: Evidence from estate tax returns. *National Tax Journal*, 57(2):445–487, 2004.

[45] P. S. Laplace. *Théorie analytique des probabilités.* Paris, Ve. Courcier, 2 edition, 1814.

[46] K. Liu, N. Lubbers, W. Klein, J. Tobochnik, B. Boghosian, and H. Gould. The effect of growth on equality in models of the economy. *arXiv*, 2013.

[47] Z. Liu and R. A. Serota. Correlation and relaxation times for a stochastic process with a fat-tailed steady-state distribution. *Physica A: Statistical Mechanics and its Applications*, 474:301–311, 2017.

[48] J. Marro and R. Dickman. *Nonequilibrium Phase Transitions in Lattice Models.* Cambridge University Press, 1999.

[49] J. E. Meade. *Efficiency, Equality and The Ownership of Property.* Allen & Unwin, London, UK, 1964.

[50] R. Mehra and E. C. Prescott. The equity premium puzzle. *J. Monetary Econ.*, 15:145–161, 1985.

[51] K. Menger. Das Unsicherheitsmoment in der Wertlehre. *J. Econ.*, 5(4):459–485, 1934.

[52] P. R. Montmort. *Essay d'analyse sur les jeux de hazard.* Jacque Quillau, Paris. Reprinted by the American Mathematical Society, 2006, 2 edition, 1713.

[53] H. Morowitz. *Beginnings of cellular life.* Yale University Press, 1992.

[54] J. A. Nelder and R. Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.

[55] M. E. J. Newman. Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.*, 46(5):323–351, 2005.

[56] V. Pareto. *Cours d'économie Politique.* F. Rouge, Lausanne, Switzerland, 1897.

[57] O. Peters. Optimal leverage from non-ergodicity. *Quant. Fin.*, 11(11):1593–1602, November 2011.

[58] O. Peters. The time resolution of the St Petersburg paradox. *Phil. Trans. R. Soc. A*, 369(1956):4913–4931, December 2011.

[59] O. Peters and A. Adamou. Stochastic market efficiency. *SFI working paper 13-06-022*, 2013.

[60] O. Peters and A. Adamou. Rational insurance with linear utility and perfect informatio n. *arXiv:1507.04655*, July 2015.

[61] O. Peters and M. Gell-Mann. Evaluating gambles using dynamics. *Chaos*, 26:23103, February 2016.

[62] O. Peters and W. Klein. Ergodicity breaking in geometric Brownian motion. *Phys. Rev. Lett.*, 110(10):100603, March 2013.

[63] T. Piketty. *Capital in the twenty-first century.* Harvard University Press, 2014.

[64] T. Piketty and E. Saez. A theory of optimal inheritance taxation. *Econometrica: Journal of the Econometric Society*, 81(5):1851–1886, 2013.

[65] T. Piketty and G. Zucman. Capital is back: Wealth-income ratios in rich countries, 1700-2010. *The Quarterly Journal of Economics*, 129(3):1255–1310, 2014.

[66] Quandl. Dow Jones Industrial Average. http://www.quandl.com/data/BCB/UDJIAD1-Dow-Jones-Industrial-Average, 2016. Accessed: 04/19/2016.

[67] S. Redner. Random multiplicative processes: An elementary tutorial. *Am. J. Phys.*, 58(3):267–273, March 1990.

[68] J.-V. Rios-Rull and M. Kuhn. 2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics. *Quarterly Review*, (April):1–75, 2016.

[69] E. Saez and G. Zucman. Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. Technical report, National bureau of economic research, 2014.

[70] P. Samuelson. What classical and neoclassical monetary theory really was. *Canad. J. Econ.*, 1(1):1–15, February 1968.

[71] P. Samuelson. St. Petersburg paradoxes: Defanged, dissected, and historically described. *J. Econ. Lit.*, 15(1):24–55, 1977.

[72] A. Sen. *On Economic Inequality*. Oxford: Clarendon Press, 1997.

[73] J. E. Stiglitz. Distribution of income and wealth among individuals. *Econometrica: Journal of the Econometric Society*, 37(3):382–397, 1969.

[74] G. Taylor. The formation of a blast wave by a very intense explosion. II. the atomic explosion of 1945. *Proc. R. Soc. A*, 201:175–186, 1950.

[75] The World Wealth and Income Database. Usa top 10% and 1% and bottom 50% wealth shares, 1913–2014. http://wid.world/data/, 2016. Accessed: 12/26/2016.

[76] H. Theil. *Economics and information theory*. North-Holland Publishing Company, 1967.

[77] E. O. Thorp. *Beat the dealer*. Knopf Doubleday Publishing Group, 1966.

[78] G. E. Uhlenbeck and L. S. Ornstein. On the theory of the brownian motion. *Physical Review*, 36(5):823–841, Sep 1930.

[79] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier (Amsterdam), 1992.

[80] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188, 1977.

[81] P. Vermeulen. How fat is the top tail of the wealth distribution? *Review of Income and Wealth*, 2017.

[82] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, 1944.

[83] W. A. Whitworth. *Choice and chance*. Deighton Bell, 2 edition, 1870.

[84] E. N. Wolff. Household wealth trends in the united states, 1983–2010. *Oxford Review of Economic Policy*, 30(1):21–43, 2014.