Johannes Gutenberg-University Mainz
Bachelor of Science in Wirtschaftswissenschaften

Wealth distributions

Summer 2022

Niklas Scheuer (lecture) and (tutorials)

www.macro.economics.uni-mainz.de

April 21, 2022

Contents

Johannes Gutenberg-University Mainz
Bachelor of Science in Wirtschaftswissenschaften

Wealth distributions

Summer 2022

Niklas Scheuer (lecture) and (tutorials)

www.macro.economics.uni-mainz.de
April 21, 2022

1 Organizational matters and the idea of the lecture

1.1 Organization

- Office: 01-142 (old ReWi or ReWi II), first floor on the left
- Appointments: only by e-mail (scheuer@uni-mainz.de)
- Course material:
  - Tutorials:
    - 6 normal tutorials
    - Available at least 1 week before on the *webpage*
  - Lectures: available on *webpage*
    - 6 normal lectures
    - 1 guest lecture
    - 4 lectures on Python and how to use it
- Assessment: 1-hour examination-closed book

1.2 Motivation

- We focus on wealth explicitly
- We will see
  - facts about distributions→data and empirical work
  - how to analyze them→models
- How to use/interpret measures of inequality
  - What can we do about wealth inequality in general and where does it come from?
  - How can we *justify* what we see or can't we
  - How can we *change* what we see

- Why do we want to focus on wealth and not other measures such as income or human capital
- Firstly, the name of the lecture implies wealth
- But more important and reasonably:
  - Wealth is more unequally distributed than income generating more public interest
  - Wealth gives ultimately more analyzing possibilities (opinion)
- Reliable data on wealth, however, is rather rare

Johannes Gutenberg-University Mainz
Bachelor of Science in Wirtschaftswissenschaften

Wealth distributions

Summer 2022

Niklas Scheuer (lecture) and (tutorials)

www.macro.economics.uni-mainz.de
April 21, 2022

Part I
Facts about wealth

2 Tools for descriptive statistics

- First we need to understand distributions themselves before talking about wealth specifically
- Probability distributions are described by functions, we consider:
  - Probability density functions (henceforth, pdf)
  - Cumulative distribution function (henceforth, cdf)
- Generally, a distribution function F links every realization x of random variable X according to some probability $\Pr(X \leq x)$
- Depending on X being discrete or continuous different pdfs and cdfs emerge

- This section is structured as follows:
  - Discrete random variable
  - Continuous random variable
    - CDF
    - PDF
    - Example
- Literature for this section:
  - Caputo et al. (2009) for a german textbook
  - Beaumont (2005, chapter 4) for an english textbook

2.1 Discrete random variable

- What is a discrete random variable
  - Possible number of realizations is discrete (i.e. countable), e.g. Rolling a die
  -

Number on die	x	1	2	3	4	5	6
Probability to occur	$\Pr(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

- Connection to distributions: How are the probabilities distributed?
  - One can see, that the probabilities are evenly distributed over all possible outcomes
  - We face a discrete uniform distribution

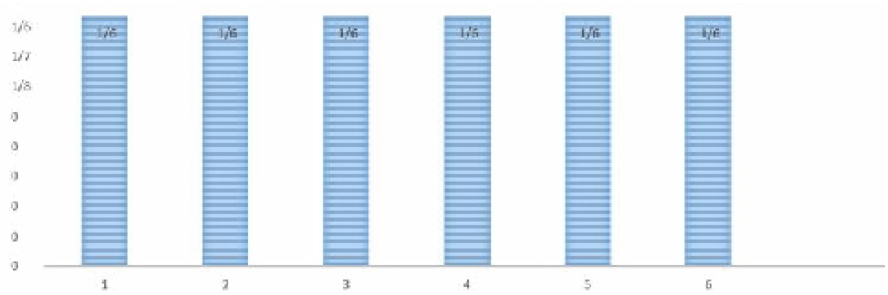

Figure 1 *Uniform distribution for probabilities of rolling a die*

- Expressed in a function (here also called the cumulative distribution function)

$$F(X) = \Pr(X < x) = \sum_{x_i < x} P(X = x_i) \tag{2.1}$$

- Other examples for discrete random variables:
  - IQ
  - Tossing a coin
  - and so on
- Other Examples of discrete distributions:
  - Binomial distribution
  - Poisson distribution
  - Bernoulli distribution

2.2 Continuous random variable

- More prominent
- But: More difficult to handle
- Examples:
  - a random person's height or weight
  - speed of a bird
- Predictions about exact values are impossible
- Technically, we do not have a continuous variable in real life
- We scale such that every variable ultimately is discrete

- An Example: Consider the following data on wealth

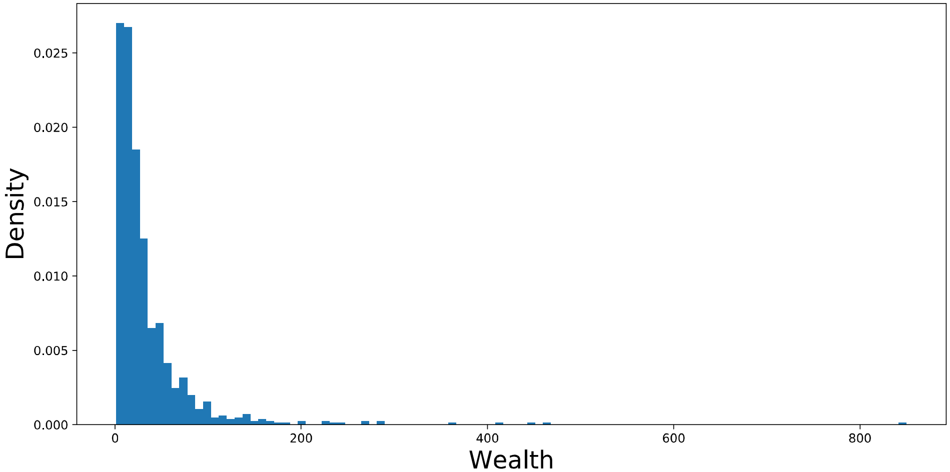

Figure 2 *Histogram of a wealth distribution*

- We see the density of households on the y-axis
- Holding wealth given by the x-axis
- This is a discrete variable in reality, i.e. we can count wealth in distinct elements
- Histograms are not so helpful in theory because:
  - Depending on size they may lead you to tell a different story
  - The smaller the step size, the thinner the bars
  - Eventually, we obtain a density/distribution function

- Hence, we ’transform’ this into

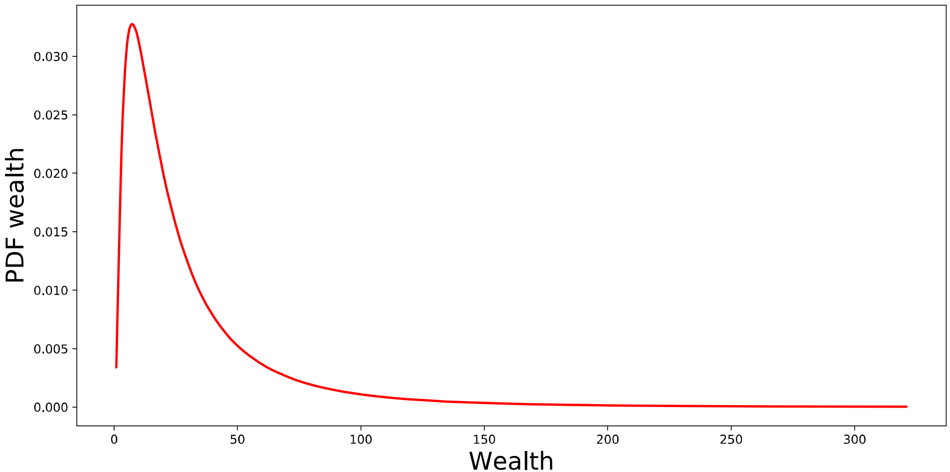

Figure 3 *Probability density function without bars*

- The thinner the bars, the less you will see separate bars
- Ultimately, we ’laid’ a pdf over a histogram

- Examples for continuous probability distributions:
  - Normal distribution
  - Lognormal distribution
  - Exponential distribution
  - Pareto distribution

2.2.1 CDF

- **C**umulative **D**istribution **F**unction $F(x)$
- Gives us the function with which
  - a distribution can be described and
  - the probability of a continuous random variable taking on a realization below or above a certain value can be characterized
- Connection to probability:
$$F(x) = \Pr(X < x) \tag{2.2}$$
- The probability of a continuous random variable taking on an exact value is 0
$$\Pr(X = x) = 0$$
- We further see that
$$\Pr(X > x) = 1 - F(x)$$
- What does this mean?

- Consider for instance the cdf of an exponential distribution

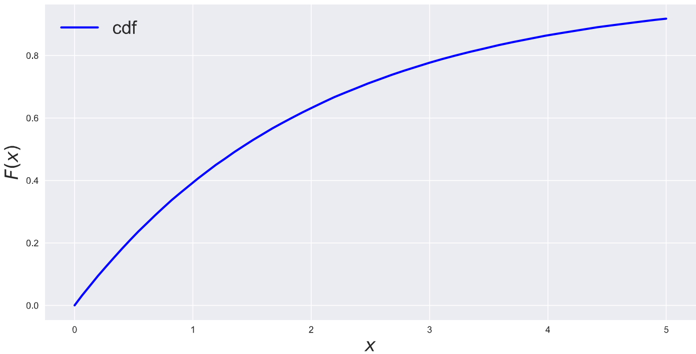

Figure 4 *cdf of an exponential distribution*

- Two helpful characteristics:
  - You can immediately state the probability
  - Cannot take on a value above 1
- But: Where does $F(x)$ come from?

2.2.2 PDF

- $F(x)$ is the result of integrating the **P**robability **D**ensity **F**unction f

$$\Pr(X < x) =: F(x) = \int_{-\infty}^{x} f(u)\, du \tag{2.3}$$

  or

$$\Pr(X > x) = 1 - F(x) = 1 - \int_{-\infty}^{x} f(u)\, du$$

- Integrating the pdf yields the cdf or deriving the cdf with respect to the state variable yields the pdf

- Consider for instance the pdf of an exponential distribution

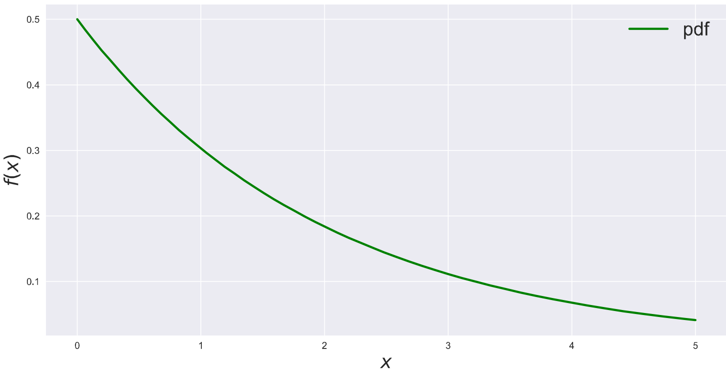

Figure 5 *pdf of an exponential distribution*

- In terms of probability:
  - The area under the pdf for a given support is the probability
  - The function itself can take on values above 1!

2.2.3 An example: The exponential distribution

- The cdf reads $1 - e^{-\lambda x} \implies$ pdf reads $\lambda e^{-\lambda x}$
- A graphic realization of the cdf and pdf of a exponential distribution

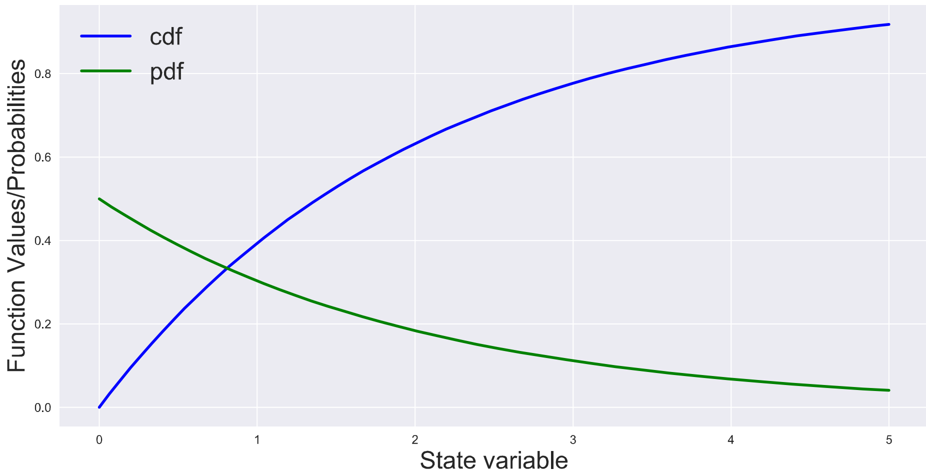

Figure 6 *CDF and PDF of an exponential distribution*

- Whether discrete or continuous, we can analyze a distribution with respect to many properties such as the mean, the variance etc.
- A property which attracts a lot of attention is for instance: Inequality
- Therefore we introduce prominent measures allowing us to analyze how unequal a distribution is
- Such measures are:
  - Lorenz curve and Gini coefficient
  - Ratios:
    - 90-10 ratio
    - Interquartile range
    - Other percentile ratios
    - Top shares
  - Coefficient of variation (CV)

2.3 Lorenz curve and Gini coefficient

- What was the Gini coefficient again?

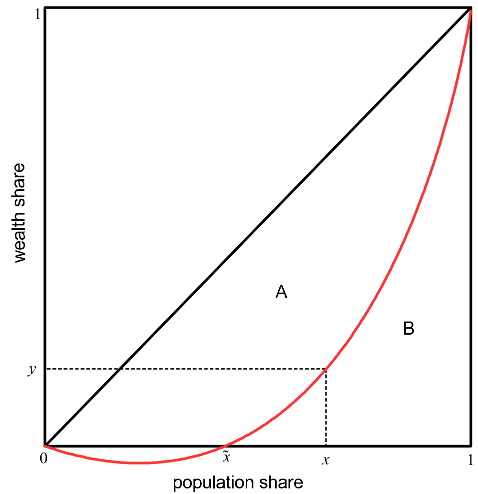

Figure 7 *Ilustrating the Gini coefficient for a (red) Lorenz curve with negative values*

- The equations behind the figure
  - The Gini coefficient is given by $G = \frac{A}{A+B}$
  - The areas B and A are given by, respectively,

$$B = \int_{\tilde{x}}^{1} y dx, \qquad A = \frac{1}{2} - B - \int_{0}^{\tilde{x}} y dx = \frac{1}{2} - \int_{\tilde{x}}^{1} y dx - \int_{0}^{\tilde{x}} y dx,$$

where $\tilde{x}$ satisfies $y(\tilde{x}) = 0$ and where $\int_0^{\tilde{x}} y dx < 0$ which requires the minus sign in the expression for A. Thus,

$$G = \frac{\frac{1}{2} - \int_0^1 y dx}{\frac{1}{2} - \int_0^{\tilde{x}} y dx}.$$

- What was the Lorenz curve again?
  - Let there be a continuous random variable a standing for levels of wealth
  - The pdf is $f(a)$ over the support $[a_{\min}, a_{\max}]$
  - The Lorenz curve tells us that $x\%$ of the population hold $y\%$ of total wealth
  - The share of population holding $\tilde{a}$ or less is given by

$$x(\tilde{a}) = \int_{a_{\min}}^{\tilde{a}} f(a)da,$$

    where $\tilde{a} \in [a_{\min}, a_{\max}]$.
  - The wealth share owned by x is given by

$$y(\tilde{a}) = \frac{\int_{a_{\min}}^{\tilde{a}} af(a)da}{\int_{a_{\min}}^{a_{\max}} af(a)da}$$

  - The Lorenz curve is then constructed by mapping x into y (see figure above)
  - For each $x \in [0, 1]$ there is one and only one value $y \in [0, 1]$
  - Thus, y is a function of x.

2.4 Ratios

2.4.1 Quantiles

- You can segment the distribution into *quantiles*
- Technically: Quantile function = inverse of CDF

$$y(p) \equiv F^{-1}(p), F(y) \equiv \Pr(Y \leq y) \tag{2.4}$$

- Various ways, e.g.
  - four equal parts, for which you refer to every part as *quartile*
  - five equal parts, for which you refer to every part as *quintile*
  - ten equal parts, for which you refer to every part as *decile*
  - one hundred equal parts, for which you refer to every part as *percentile*
- This can be visualized as follows:

- Imagine a normal distribution

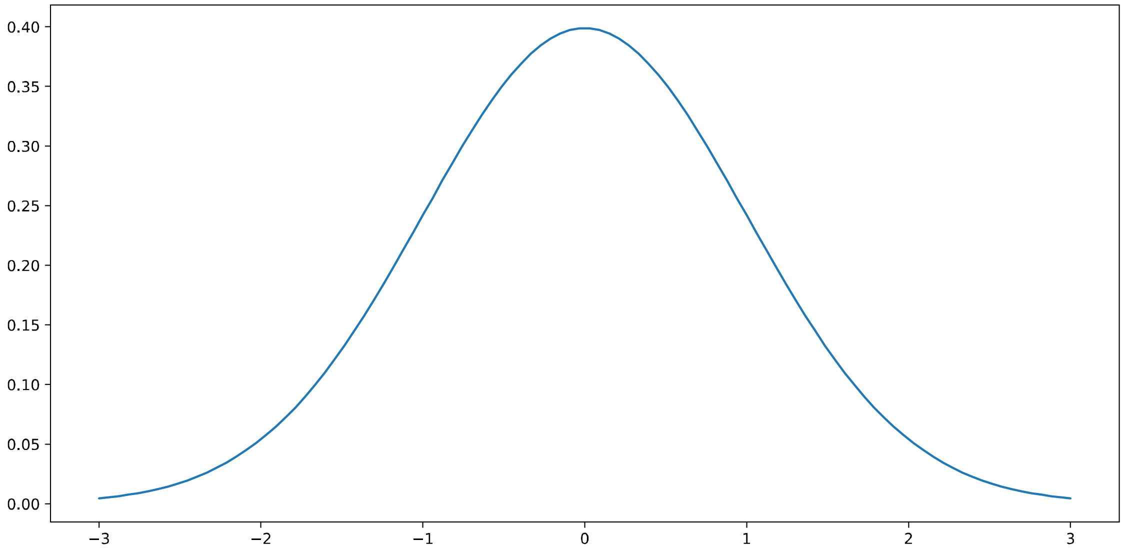

Figure 8 *Example of a normal distribution with* $\mu = 0$ *and* $\sigma = 1$

- Now imagine we were to determine the quartiles? How does it look like

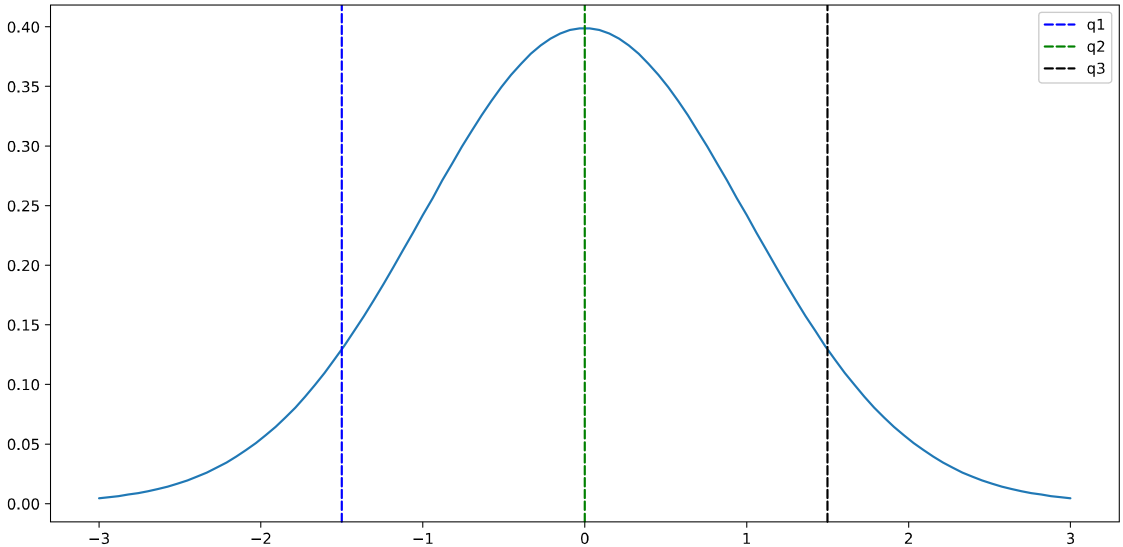

Figure 9 *Example of a normal distribution with $\mu = 0$ and $\sigma = 1$ including quartiles*

2.4.2 Top Shares

- Something which is frequently done: Comparing percentiles
- This shows how equal or unequal the distribution is
- How can we imagine this?
- Assume we have the following wealth distribution

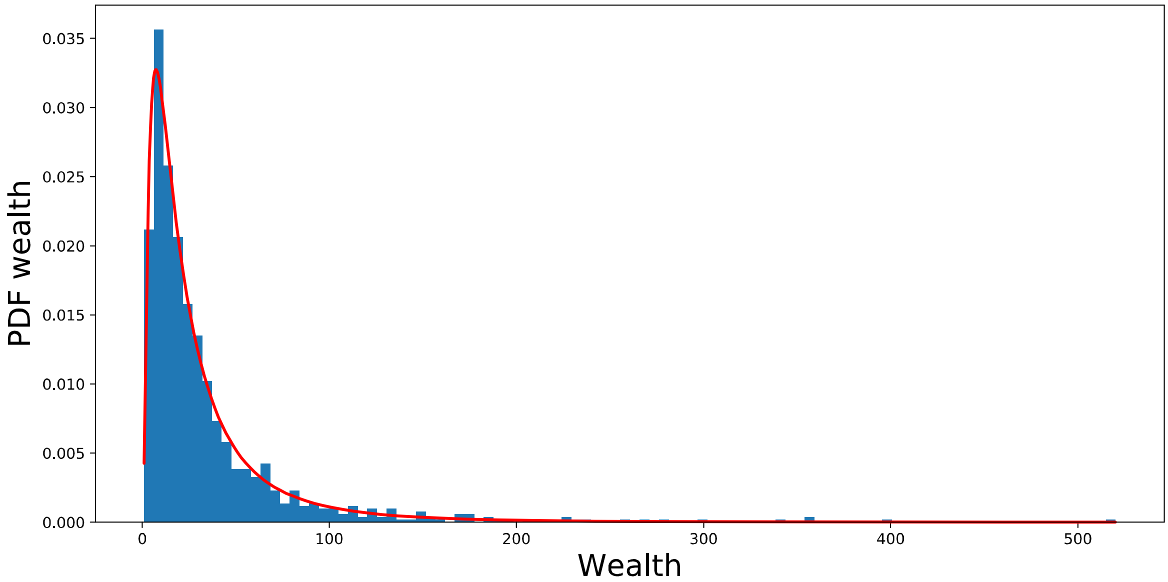

Figure 10 *Wealth distribution*

- Now we want to plot a vertical line showing the top percent wealth share
- One can also look at the 10% top wealth shares or 0.1% or whatever share you desire

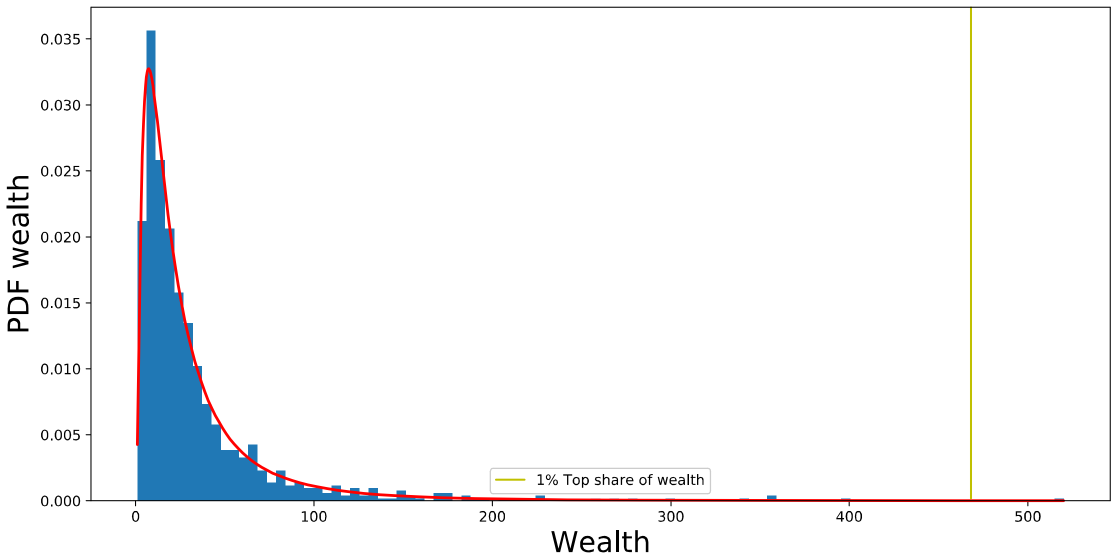

Figure 11 *Wealth distribution including top 1% share of wealth holders*

3 Some empirical facts

3.1 The world as a whole

- Oxfam numbers

3.2 The United States

3.2.1 An Introduction

- See a video on the wealth distribution in the USA
- What kind of data do we have? How does the wealth distribution look like?
- What do we see in terms of inequality?

3.2.2 The wealth distribution of the US

- See the evolution of the wealth distribution of a cohort in the USA
  - National Longitudinal Survey of Youth (born between 1957-64)
  - Cohort originally included 12,686 respondents ages 14-22 when first interviewed in 1979
  - https://www.nlsinfo.org/content/cohorts/nlsy79

• How do wealth distributions look like?

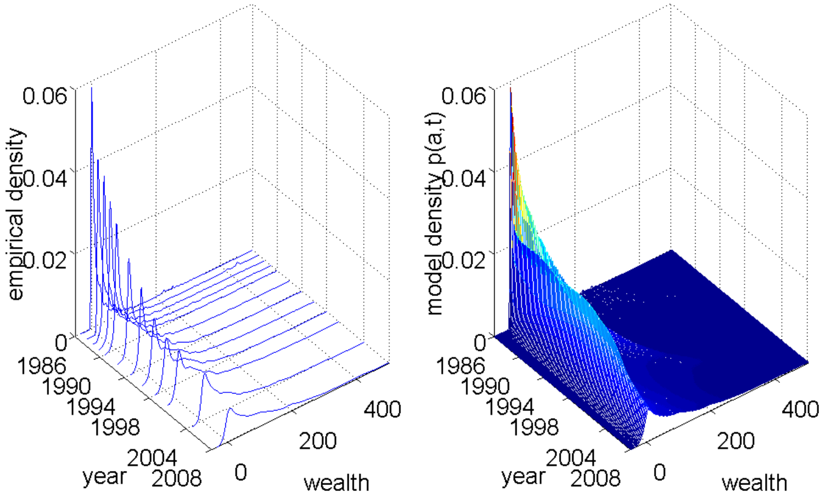

Figure 12 *The density of wealth for the NLSY 79 cohort (left figure)*

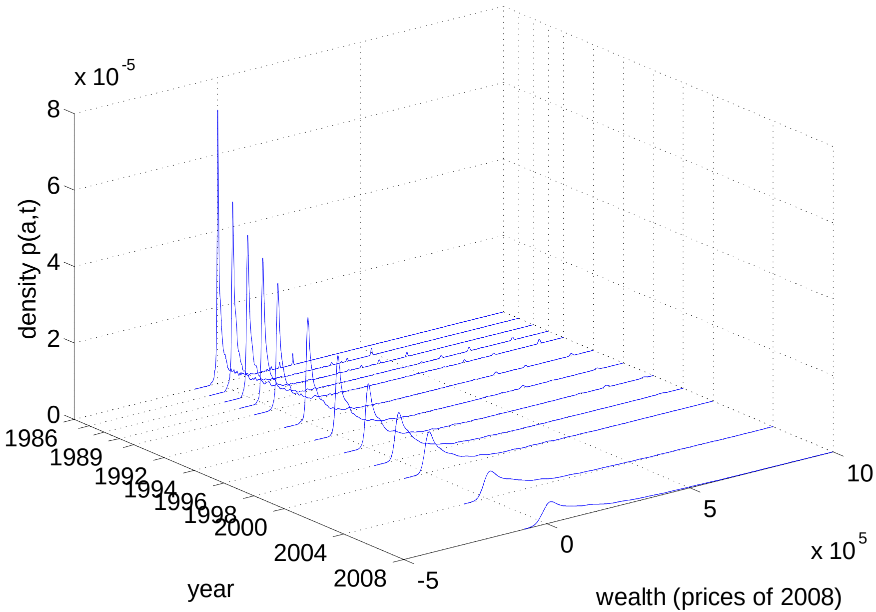

Figure 13 *Empirical wealth distribution of the 1979 cohort in the NLSY from 1986 to 2008*

- What do we see?
- Well both figures show the wealth distribution over time for one cohort
- The shape of every wealth distribution is very convincing
- The initial distribution has a relatively high concentration for low levels of wealth
- Over the years this concentration shrinks, put differently, wealth is distributed more unequally across individuals
- In order to see this, we have to consider the GINI coefficient
- We see:

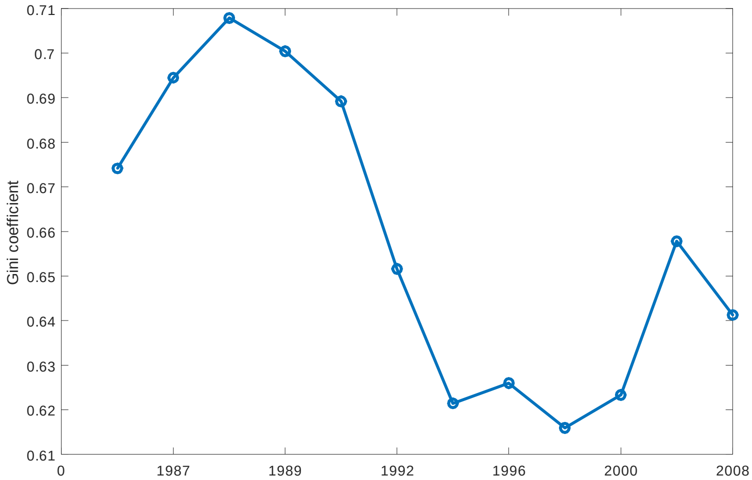

Figure 14 *GINI coefficient NLSY*

- Mass gets pushed either
  - further left (poor people become poorer) or
  - further right (rich people become richer)

- Especially the last point has been of interest in recent years

3.2.3 Do the rich get richer - an outlook

- Saez & Zucman (2016) about wealth inequality looking at top shares:

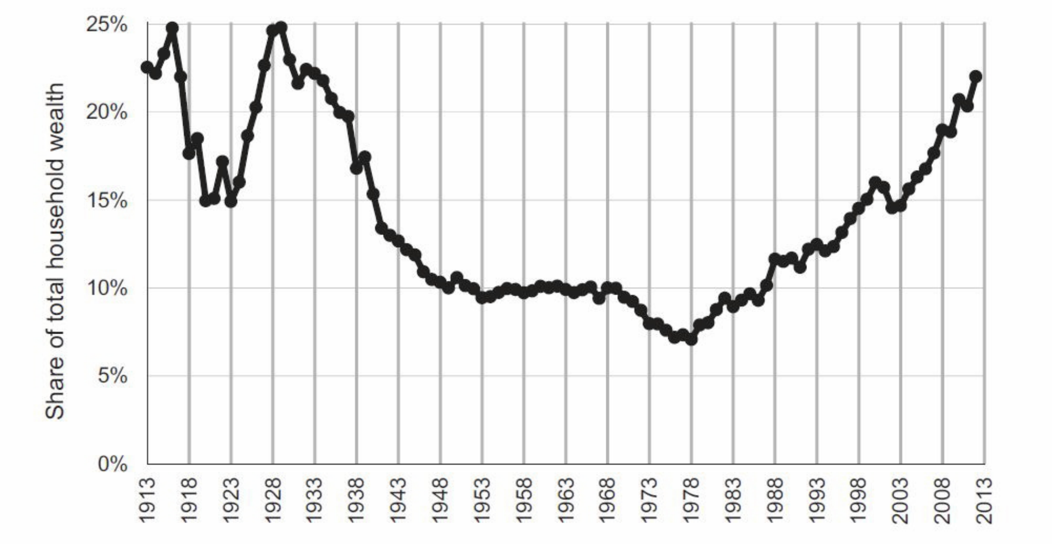

Figure 15 *Top 0.1 % wealth share in the US, 1913-2012 (Saez and Zucman, 2016)*

- Share of total household wealth owned by the richest 0.1% of families in the United States from 1913 to 2012
- What do we see?
  - This figure tells us the share of total household wealth (y-axis) the 0.1% richest households in the US held over the years (x-axis)
  - In the very beginning it rises slowly but then decreases from 1918 until 1923 substantially
  - Reasons: Spanish flu, World War I
  - Then we see a surge until 1928, to which we generally refer to a the Roaring Twenties
  - After 1928 a decrease in inequality took place until 1978
  - Reasons: Great depression, World War II, Marshall plan, oil crises in the 70s
  - Then we clearly see an increase over the years until now

3.3 Wealth in Germany

- Is there data on wealth in Germany?
- Where does that data come from?
  - Taxes?
  - Surveys?
- What does the data tell us about wealth in Germany?
  - What is the average wealth or median wealth of the population?
  - What about the top-shares (top 10 %, top 1 %)?
- What about bequests?
- Bequest volume in Germany (Grabka and Tiefensee, DIW, 2017)
  - Every year individuals leave bequests for their offspring
  - This contributes to wealth of the offspring

- Data on Germany origins from tax records of the bequest tax tells us that about 400 Billion Euro are bequeathed every year in Germany
  - Estimates only
  - Why?
- In order to avoid paying taxes individuals gift money over the years to their offspring (See Destatis data and also Ryglewski (2019))
  - However, officially this is subject of taxation as well (see the german law on what defines gifting)
  - Except when you own a family business and bequeath it to your offspring (no taxes required at all)
- Impact on wealth inequality?
  - Grossmann and Kirch (2016) view bequathing as contributing to more inequality
    - They inter alia suggest to tax bequests with a rate of 100 %
    - Speaking pragmatically, the current system favours rich heirs over talent
    - No equity issue
    - Changing the system would benefit those who are more talented increasing growth and social welfare

Johannes Gutenberg-University Mainz
Bachelor of Science in Wirtschaftswissenschaften

Wealth distributions

Summer 2022

Niklas Scheuer (lecture) and (tutorials)

www.macro.economics.uni-mainz.de
April 21, 2022

Part II
A simple model of wealth distributions

4 Background: Expected utility maximization

4.1 Introduction

- Before we go deeper into models of wealth and wealth distributions, we need some background related to uncertainty
  - How does homo oeconomicus treat risk?
  - Answer: they maximize expected utility
- We take a standard example from business cycle analysis as motivation
  - This extends business cycle analysis from Makroökonomik I to allow for uncertainty
  - A textbook treatment is in Wälde (2012, ch. 8.1.4)

4.2 Budget constraints

- An individual lives for two periods
- We could allow for uncertainty in one-period setup or in many periods as well
- Budget constraint in the first period (period t)

$$w_t = c_t + s_t$$

  where w_t is the wage, c_t consumption and s_t stands for savings in t

- Budget constraint in the second period (period $t+1$, imagine individual is retired) reads

$$(1 + r_{t+1})\, s_t = c_{t+1}$$

  where left-hand side is (capital) income in period $t + 1$ (savings plus interest r_{t+1} on savings) and right-hand side is consumption expenditure

- Identical to setup in Makroökonomik I
- Now assume that the interest rate r_{t+1} is uncertain, i.e. its value is not known in the first period t

4.3 Uncertainty

- How do we model this uncertainty technically?
- $\Longrightarrow$ We make r_{t+1} a random variable
- We assume that r_{t+1} is a *discrete* random variable with realizations $r_{i,t+1}$ and probabilities p_i with $i = 1, ..., n$ such that $\sum_{i=1}^{n} p_i = 1$
- As a consequence, the individual in period t does not know the consumption level c_{t+1}
- A random variable in other words is an example of what decision theorists call a "lottery"
    - A collection of outcomes and their probabilities
    - For our example, the lottery (without the time index) is $(r_1, p_1, r_2, p_2, ..., r_n, p_n)$
    - Some authors define a lottery as the collection of probabilities only $(p_1, p_2, ..., p_n)$

4.4 Preferences

- Individual consumes in both periods
- Instead of the “Makroökonomik I” utility function

$$U_t = \gamma \ln c_t^y + (1-\gamma) \ln c_{t+1}^o \text{ with } \gamma \epsilon (0,1),$$

  we write

$$U_t = u(c_t) + \beta u(c_{t+1})$$

- Hence, we generalize the instantaneous utility function from ln to some concave $u(.)$, with $0 < \beta < 1$
- We also need to take uncertainty into account!

- The individual needs to form expectations as consumption c_{t+1} is uncertain

$$U_t = u(c_t) + \beta \mathbb{E}_t u(c_{t+1})$$

- $\mathbb{E}_t$ is the expectations operator:
  - Individuals form expectations in t
  - They take all knowledge up to and including t into account

- We talk about an *expected utility maximizer*, when the utility function is of the von Neumann-Morgenstern form, i.e. when we write

$$\mathbb{E}_t u(c_{t+1}) = \sum_{i=1}^{n} p_i u(c_{i,t+1})$$

- In words: there are utility levels $u(c_{i,t+1})$ for each outcome $c_{i,t+1}$ and p_i is the probability of this outcome

4.5 The maximization problem

- First, replace consumption levels by expressions from budget constraints
- This gives nice trade-off for choosing s_t

$$U_t = u(w_t - s_t) + \beta \mathbb{E}_t u((1 + r_{t+1}) s_t)$$

  and simplifies the analysis

- Next, let us be clear, what this expectations operator is
- Then, assuming that the individual forms *rational expectations*, we write

$$\mathbb{E}_t u((1 + r_{t+1}) s_t) = \sum_{i=1}^{n} p_i u((1 + r_{i,t+1}) s_t)$$

- Forming rational expectations means that the individual uses
  - (i) the correct model and
  - (ii) probability distributions to form expectations (e.g. Sargent, 2008)

- The individual knows how
  - (i) utility is related to savings and interest rates and
  - (ii) correctly applies probabilities to realizations of utility
- Finally, we compute the first-order condition

$$\begin{aligned}\frac{dU_t}{ds_t} &= \frac{d}{ds_t}\left[u\left(w_t - s_t\right) + \beta \sum_{i=1}^{n} p_i u\left(\left(1 + r_{i,t+1}\right) s_t\right)\right] \\ &= -u'\left(w_t - s_t\right) + \beta \sum_{i=1}^{n} p_i u'\left(\left(1 + r_{i,t+1}\right) s_t\right)\left(1 + r_{i,t+1}\right) = 0\end{aligned}$$

- We rewrite this as

$$\begin{aligned}u'\left(w_t - s_t\right) &= \beta \sum_{i=1}^{n} p_i u'\left(\left(1 + r_{i,t+1}\right) s_t\right)\left(1 + r_{i,t+1}\right) \\ &= \beta \mathbb{E}_t\left[u'\left(\left(1 + r_{t+1}\right) s_t\right)\left(1 + r_{t+1}\right)\right]\end{aligned}$$

- Using the budget constraints again, we get

$$u'\left(c_t\right) = \beta \mathbb{E}_t\left[u'\left(c_{t+1}\right)\left(1 + r_{t+1}\right)\right]$$

- Understanding the first-order condition
  - First-order condition again

$$u'(c_t) = \beta \mathbb{E}_t \left[u'(c_{t+1})(1 + r_{t+1}) \right]$$

  - Optimal behaviour from the perspective of t compares marginal utilities – just as in deterministic world where the optimality rule would read

$$u'(c_t) = \beta u'(c_{t+1})(1 + r_{t+1})$$

    (only expectations operator is missing)
  - There are many possible marginal utilities in $t+1$ depending on the value the interest rate will take (i.e. on its realization)
  - The individual therefore looks at some average marginal utility (taking the term $1 + r_{t+1}$ also into account) where the weights for each realization is the probability p_i
  - The role of the discount factor β and the interest rate $(1 + r_{t+1})$ is the same as in the deterministic world

4.6 An example

- Let us use a Cobb-Douglas utility function (like in Makroökonomik I)

$$U_t = \gamma \ln c_t + (1-\gamma)\,\mathbb{E}_t \ln c_{t+1}$$

- We get simple rules for optimal behaviour

$$\begin{aligned} c_t &= \gamma w_t \\ s_t &= (1-\gamma)\, w_t \\ c_{t+1} &= (1+r_{t+1})\,(1-\gamma)\, w_t \end{aligned}$$

- Is there any uncertainty left?
  - Yes, r_{t+1} is unknown in t and so is c_{t+1}
  - This implies that realized consumption in $t+1$ differs from expected consumption

5 The model

5.1 Background

- Questions and (some) answers
- How can we understand wealth distributions
  - theoretically and
  - empirically?
- What are *theoretical mechanisms* that allow us to understand that some have more wealth than others?
  - born rich (inheritance)?
  - saved a lot over time (preference)?
  - high labour income (intelligent and income-oriented)?
  - luck on the labour market (always had good paying jobs, never lost the job)?
  - wealthy because old (life-cycle considerations)?
- Can we construct economic models that explain wealth distributions (and their dynamics) in a *quantitatively satisfactory* way?

5.2 The setup

- The idea (see Bossmann, Kleiber and Wälde, 2007)
  - Individuals live in a 2-period OLG general equilibrium world
  - The economy evolves in a deterministic way at the aggregate level (as in 'Makroökonomik I')
  - There are no aggregate shocks (*no* TFP shocks)
  - Two novel features
    * Idiosyncratic shocks: labour income is uncertain - that means:
    * Ability when born and skills when entering the labour market are random
    * Bequests:
      · Individuals inherit wealth when born and
      · Leave bequests at the end of life

- The formal structure for an individual i
  - First-period budget constraint

$$w_t l_{it} + b_{it} + g_t = c_{it}^y + s_{it} \tag{5.1}$$

    - b_{it} denotes after tax inheritance received from the parent
    - $w_t l_{it}$ stochastic income depending on (deterministic)
    - wage w_t per efficiency unit and the
    - random ability of the individual l_{it}
    - g_t is the uniform lump-sum transfer received from the government in case it levies a tax on bequests
    - s_{it} savings
  - The distribution for individual ability

$$\mathbb{E}(l_{it}) = \bar{l} \equiv 1, \quad \mathrm{Var}(l_{it}) = \sigma^2, \quad \mathrm{Cov}(l_{ir}, l_{is}) = 0 \text{ for } r \neq s. \tag{5.2}$$

    - l_{it} are identically and independently distributed (i.i.d.)
    - Hence, mean and variance are the same for all t (identically distributed) and ...
    - Covariance is zero (independently distributed)
    - Without loss of generality, we set $\bar{l} = 1$

- Second-period constraint

$$s_{it}\left(1+r_{t+1}\right)=c_{it+1}^{o}+\left(1+\tau\right)b_{it+1}, \tag{5.3}$$

  * r_{t+1} is the second period certain (!) interest rate
  * c_{it+1}^{o} is second period consumption
  * τ is the proportional tax rate on
  * bequests b_{it+1}

- Preferences

  * Individuals enjoy consumption and bequests ("warm-glow" motive)

$$U_{it}=U\left(c_{it}^{y},c_{it+1}^{o},b_{it+1}\right) \tag{5.4}$$

  * They choose consumption c_{it}^{y} when young, c_{it+1}^{o} when old, and the bequest b_{it+1} passed on to the child
  * Utility depends on the amount b_{it+1} the child receives after tax
  * Joy-of-giving idea: "consumers leave bequests simply because they obtain utility directly from the bequest"
  * Next generation also has index i such that i is the "name" of a family/ dynasty

5.3 Equilibrium

- Optimal behaviour
  - After some (not complicated but time-consuming) steps,
  - employing a Cobb-Douglas utility function

$$U_{it} = \alpha \ln c_{it}^{y} + (1-\alpha) \left[\beta \ln c_{it+1}^{o} + (1-\beta) \ln b_{it+1} \right]$$

and defining wealth as $a_{it+1} \equiv s_{it}$, we get

$$a_{it+1} = (1-\alpha) w_t l_{it} + \underbrace{(1-\alpha) \frac{(1-\beta)(1+r_t)}{1+\tau} a_{it}}_{\equiv b_{it}} + \underbrace{(1-\alpha) \frac{\tau (1-\beta)(1+r_t)}{1+\tau} k_t}_{\equiv g_t} \quad (5.5)$$

which shows that wealth of dynasty i in period $t+1$ depends on

* wealth a_{it} of previous generation (via bequests b_{it})
* aggregate capital stock k_t per worker (via government transfers g_t)and
* individual skills l_{it} (via labour income)

- Macroeconomic equilibrium and microeconomic dynamics
  - Employ a simplifying assumption which is common to very many macroeconomic models
  - At the aggregate level, the economy is in a steady state, i.e.
    $$k_t = \bar{k},\ r_t = \bar{r},\ w_t = \bar{w} \tag{5.6}$$
    are constant over time
  - At the microeconomic level, there is still idiosyncratic risk via ability l_{it} of individual/ dynasty i
  - Some family i becomes richer over time
  - Some families become poorer
  - Some remain at more or less the same level

- The evolution of wealth
  - Fundamental wealth equation for family i

$$a_{it+1} = c_3 l_{it} + c_4 a_{it} + c_5 \tag{5.7}$$

  - c_3 to c_5 are abbreviations for parameters and constant variables (w, r, k) as shown in (5.5)
  - c_1 and c_2 were used earlier in paper
  - This is the reduced form of the model – no further simplification possible

6 The distribution of wealth

6.1 What we would like to understand

- What does this model tell us about the evolution of wealth of one family i?
  - Not very much
  - As individual skills are uncertain, so is individual wealth
  - Wealth evolves over time
    - it can rise
    - it can fall
    - almost anything can happen
  - Model makes hardly any prediction about the *realization* of wealth at some future point in time t
  - But, we
    - do know something about the *probability* that wealth is within a certain range
    - and we can compute *expected* wealth
- Simple but powerful principle

- A very simple example which has the same properties: playing dice
- Before someone throws a die, one cannot say a lot about the *realization* (apart from numbers between 1 and 6)
- But one can say something about the *probability* to throw between 3 and 5, or to throw 1 (or other)
- This is the case with all models containing some source of uncertainty
- They make predictions about probabilities or - more generally - distributional properties

• What does this model tell us about inequality?

  - This depends on how we define inequality
  - Various measures are available (as we know from above):
    * Variance and standard deviation
    * Coefficient of variation
    * Wealth held by richest $x\%$ (Ratios of percentiles and so on)
  - We start with a simple measure: variance
  - The coefficient of variation (standard deviation divided by the mean) would have the advantage of being scale-independent

- From individual probabilities to cross-sectional distributions
  - So far, we only discussed, for some future point in time t,
    * probability of an individual to have wealth within a certain range
    * expected wealth of a person
    * variance of wealth of a person
  - We also want to know what the expected wealth level is for a *group* of people
  - Imagine we look at many individuals that all start with the same initial wealth level $a_{i0} = a_{\text{low}}$ (we look at "the poor") or $a_{i0} = a^{\text{high}}$ (we look at "the rich")
  - Employing a law of large numbers, one can show that
    * the *probability of an individual* to have wealth within a certain range also gives the *share of individuals of this group* within this certain range
    * *expected wealth of a person* also gives *average wealth of this group*
    * *variance of wealth of a person* also gives *variance of wealth of this group*
    * we obtain a distribution of wealth for a cross-section of individuals for any point in time
  - To illustrate, think about playing dice

6.2 The mean and variance of the wealth distribution

- Let us now compute the variance and coefficient of variation (for which we need the mean) for the wealth distribution for one dynasty i
- We compute the wealth level a_{it} of an individual i at $t > 0$ by solving equation (5.7)
- We obtain (Wälde, 2012, ch. 2.5.3) wealth a_{it} as a function of parameters, time t, the initial wealth level a_{i0} and luck, i.e. the realization of skills l_{is} for family i at each point in time between 0 and $t-1$,

$$a_{it} = c_5 \sum_{s=0}^{t-1} c_4^s + c_4^t a_{i0} + c_3 \sum_{s=0}^{t-1} c_4^{t-1-s} l_{is} \tag{6.1}$$

- What does this tell us?
  - If we knew l_{is} already in 0, we could perfectly predict (no uncertainty) what the wealth level a_{it} is in t
  - As we do not know the l_{is}, a_{it} is unknown
  - Initial wealth a_{i0} matters and c_4 is a measure of social mobility: the lower c_4, the less social background ("wealth of parents") matters (see Charles and Hurst, 2003)
  - Apart from a_{i0}, why are some people rich and some poor? The rich were lucky, the poor were not: a_{it} is basically the sum of past luck l_{is}

- Is there "equality of chances"?
  - Same equation as (6.1) above

$$a_{it} = c_5 \sum_{s=0}^{t-1} c_4^s + c_4^t a_{i0} + c_3 \sum_{s=0}^{t-1} c_4^{t-1-s} l_{is} \tag{6.2}$$

  - If uncertain skills l_{is} come from the same distribution for all individuals, there is an "equality of chances" in this economy
  - If social background also affects luck, there is no equality of chances
  - Examples for absence of equality of chances
    * the share of students at university coming from parents with a university degree is larger than the share of parents with a university degree in society
    * the share of ethnic group A in government is larger than the share of this group in society

- Computing the mean

  - Define expected wealth as $\mu_{it} = E_0 a_{it}$
  - In words, μ_{it} is the expected wealth of dynasty i for some future point in time t when we form expectations at $t = 0$
  - Apply this to (6.1) and get

$$\mu_{it} = c_5 \sum_{s=0}^{t-1} c_4^s + c_4^t a_{i0} + c_3 \sum_{s=0}^{t-1} c_4^{t-1-s} \tag{6.3}$$

    where we use $E(l_{it}) = \bar{l} \equiv 1$ from (5.2)
  - Why does the expected wealth level still depend on the dynasty
  - That is, why is there an index i in μ_{it}?
  - Because of initial wealth a_{i0} of dynasty i

– Rewriting $c_5 \sum_{s=0}^{t-1} c_4^s$ and $c_3 \sum_{s=0}^{t-1} c_4^{t-1-s}$ using the formula for the geometric series, we obtain

$$\begin{aligned} \mu_{it} &= c_4^t a_{i0} + (c_3 + c_5) \left(\frac{1 - c_4^t}{1 - c_4} \right) \\ &= c_4^t a_{i0} + \left(\frac{c_3 + c_5}{1 - c_4} \right) - c_4^t \left(\frac{c_3 + c_5}{1 - c_4} \right) \end{aligned}$$

– We know from the paper $c_4 = \frac{\Delta}{1-\tau}$; $c_5 = \frac{\tau}{1+\tau} \Delta \bar{k}$; $c_3 = (1 - \Delta) \bar{k}$

– Then we are able to show that

$$\left(\frac{c_3 + c_5}{1 - c_4} \right) = \frac{(1 - \Delta) \bar{k} + \frac{\tau \Delta}{1+\tau} \bar{k}}{1 - \frac{\Delta}{1-\tau}}$$

Rewriting the fractions eventually leads to

$$\begin{aligned} \left(\frac{c_3 + c_5}{1 - c_4} \right) &= \frac{1 + \tau - \Delta - \Delta\tau + \tau\Delta}{1 + \tau - \Delta} \bar{k} \\ &= \bar{k} \end{aligned}$$

- Knowing $\left(\frac{c_3+c_5}{1-c_4}\right) = \bar{k}$, we can rewrite (6.3)

$$\begin{aligned}\mu_{i,t} &= c_4^t a_{i0} + \left(\frac{c_3 + c_5}{1 - c_4}\right) - c_4^t \left(\frac{c_3 + c_5}{1 - c_4}\right) \\ &= c_4^t a_{i0} + \bar{k} - c_4^t \bar{k}\end{aligned}$$

- Ultimately, we get a very intuitive result

$$\mu_{it} = \left(a_{i_0} - \bar{k}\right) c_4^t + \bar{k}$$

- Expected wealth in t depends on initial wealth a_{i_0}, wealth per capita, $\bar{k}$, in the economy and the social mobility parameter c_4
- In the presence of equality of chances
  - ∗ "family background" does not matter, $\mathbb{E}\left(l_{it}\right) = \bar{l} \equiv 1$
  - ∗ wealth regresses to the mean $\bar{k}$ from (5.6) given $c_4^t < 1$
  - ∗ initial wealth matters from generation to generation, but not in the long run
- This is a relatively "optimistic model" with respect to inequality as
  - ∗ Race, gender, country of origin do not play a role
  - ∗ Hard to believe!
  - ∗ Empirically hard to support!

- A numerical solution of the mean

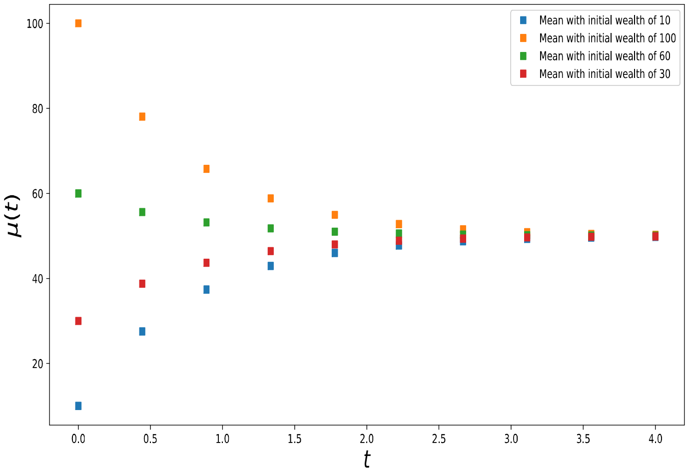

Figure 16 *Mean wealth for Bossmann et al. (2007)*

- What do we see?
- Depending on initial wealth (i.e. wealth at intial point in time $t = 0$) mean wealth or the expected level of wealth
  - increases over time (initial wealth is 10 or 30)
  - or decreases over time (initial wealth is 60 or 100)
- It appears to be converging to a long-run value
- This makes sense as w, r, k
  - are part of $a_{i,t}$ and $\mu_{i,t}$ and
  - obtain a constant value in the long-run, i.e. $k_t = \bar{k},\ r_t = \bar{r},\ w_t = \bar{w}$ (compare (5.6))
- Ultimately, the numerical solutions depend on assumptions made
- $\implies$ Careful interpreting

- Computing the variance
  - We are interested in the variance of wealth a_{it} as given in (6.2)

$$a_{it} = c_5 \sum_{s=0}^{t-1} c_4^s + c_4^t a_{i0} + c_3 \sum_{s=0}^{t-1} c_4^{t-1-s} l_{is}$$

  - Note that we can look at a_{it} as a standard random variable
    * It is true that a_{it} changes from one point in time to the next
    * When we are interested in the variance (or any other moment), we hold time t fixed and use standard rules for standard random variables
  - We therefore need to understand the variance of a sum of parameters and random variables

- Computing the variance (cont'd)
  - Starting from (6.2)
$$a_{it} = c_5 \sum_{s=0}^{t-1} c_4^s + c_4^t a_{i0} + c_3 \sum_{s=0}^{t-1} c_4^{t-1-s} l_{is},$$
    we get (using knowledge from Statistik I and II)
$$\begin{aligned} \mathrm{Var}\,(a_{it}) &= \mathrm{Var}\left(c_5 \sum_{s=0}^{t-1} c_4^s + c_4^t a_{i0} + c_3 \sum_{s=0}^{t-1} c_4^{t-1-s} l_{is} \right) \\ &= 0 + 0 + c_3^2 \mathrm{Var}\left(\sum_{s=0}^{t-1} c_4^{t-1-s} l_{is} \right) \\ &= c_3^2 \sum_{s=0}^{t-1} \left(c_4^{t-1-s} \right)^2 \mathrm{Var}\,(l_{is}) \end{aligned}$$
    where the second equality employed that the variance of a constant is zero and the second equality used (5.2), especially the covariance of zero
  - Using (5.2) further and Wälde (2012, ch. 2.5.1), we find
$$\mathrm{Var}\,(a_{it}) = c_3^2 \sigma^2 \sum_{s=0}^{t-1} \left(c_4^{t-1-s} \right)^2 = c_3^2 \sigma^2 \frac{1 - c_4^{2t}}{1 - c_4^2}$$
    which tells us that the variance increases over time (but approaches a constant)

- A numerical solution of the variance

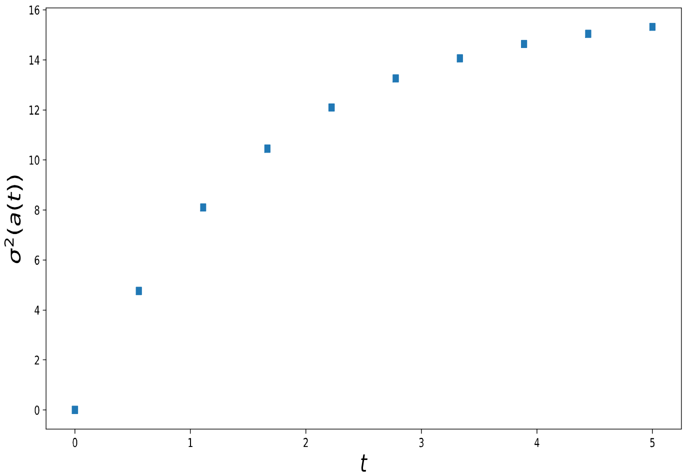

Figure 17 *Variance evolution over time*

- We see that the variance is concave over time
- This is over time the variance increases but the increase itself is decreasing

6.3 Plotting the evolution of the distribution

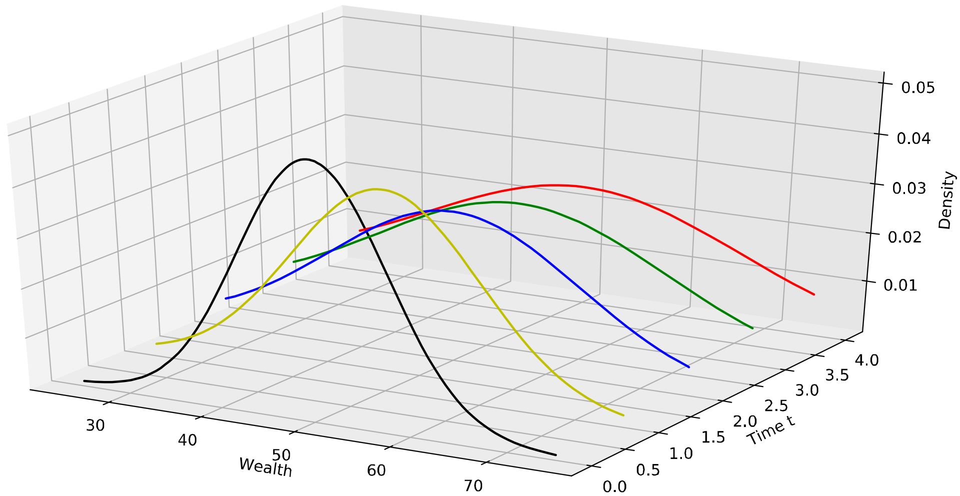

Figure 18 *Wealth evolution over time*

6.4 Wealth inequality

- Given variance and mean, what statements can be made with respect to inequality?
- Especially in the light of bequests

$\Longrightarrow$ In this section we investigate inequality and the role of bequests in this context

- Reminder:

$$CV(a_{it}) = \frac{\sqrt{\mathrm{Var}(a_{it})}}{E(a_{it})} \tag{6.4}$$

- Using this measure raises at least two questions:
    - 1. Why the coefficient of variation and
      2. why the coefficient of variation of wealth?
- Reasons:
    - popular
    - has a geometric interpretation just like the familiar Gini index

- Unlike the Gini coefficient, it is a simple function of the first and second moments of the underlying distribution $\Longrightarrow$ suitablity for our model with its emphasis on closed form solutions
- Regarding the second issue:
  * wealth inequality can be argued to be of interest per se
  * wealth makes also sense as we specifically look at is distribution and, thus, are interested in the degree of inquality
  * it can easily be shown that the determinants of our wealth distribution are identical to the determinants of the distribution of utility in our model

• On the many possible factors within the topic 'inequality', we concentrate on

  - Bequests
  - Taxing bequests

6.4.1 Bequests

- Inserting expressions of the expected value and the variance into the coefficient of variation
- Then comparing two economies A and B
  - in A, $\beta = 1$
  - in B, $\beta \in (0, 1)$
- Given both economies and their assumptions regarding β, we can compare the stationary (i.e. $t \to \infty$) CV of a representative individual i

$$CV\left(a_{i\infty}^{A}\right) = \sigma$$
$$CV\left(a_{i\infty}^{B}\right) = \sigma\sqrt{\frac{1-c_4}{1+c_4}}$$

- Comparing both we see that $\sqrt{\frac{1-c_4}{1+c_4}} \in (0, 1)$ and, thus,

$$CV\left(a_{i\infty}^{A}\right) > CV\left(a_{i\infty}^{B}\right)$$

- Hence, economy A inherits more inequality than B, which leads to bequests having a inequality decreasing effect

6.4.2 Taxing bequests

- We focus on taxing bequests at rate $\tau > 0$
- Affects decisions only on microeconomic level, as on macroeconomic level the average capital stock is constant
- Calculating the coefficient of variation for an economy with and without tax, we obtain

$$CV\left(a_{it}^{wt}\right) < CV\left(a_{it}^{nt}\right)$$

- The dispersion of family wealth decreases when government levies a tax on bequests and redistributes revenue among the young generation
- Hence, the redistributive policy of the government reduces intragenerational inequality

6.5 What do we learn from this?

- Imagine we have a real world distribution (reminder)

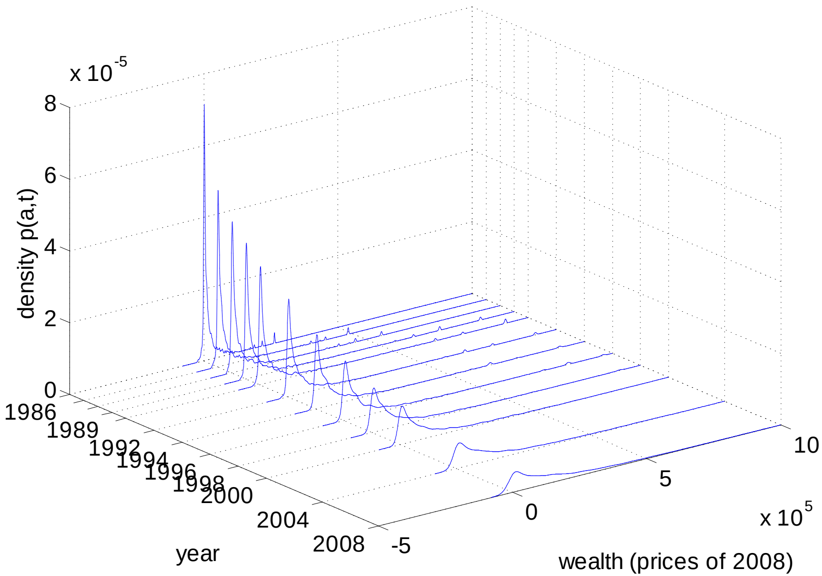

Figure 19 *Empirical wealth distribution of the 1979 cohort in the NLSY (1986 to 2008)*

- We can then ask the following question
  - Can we understand this increase in inequality to be consistent with 'equality of chances'?
  - Let us imagine we consider 'equality of chances' to be important – think of "all men are created equal" or "All human beings are born free and equal in dignity and rights"
  - More precisely speaking: if each generation has i.i.d. ability l_{it}
    * (a) can we replicate this empirical evolution of wealth in our model?
    * (b) can we do so with realistic parameter values?
  - If not, what is the source of large inequality and why is 'equality of chances' being violated?
- We can ask further questions
  - What would happen to the wealth distribution if we had a social security system or if we had a (progressive) tax system? Would the wealth distribution become more equal?
  - Is there a trade-off between average wealth (imagine society wants to become richer) and inequality? (Think about the Kuznets curve in economic development)
  - ... and much more ...

7 Conclusion

- Basic questions

  - Why are some people rich and some poor?
  - Why do some people even die with debt, i.e. with negative wealth?
  - What is the role of personality, family background, social background, education and work life?
  - Which role does the tax and redistribution system play?

- Framework of analysis

  - We got to know a simple but powerful analytical framework that allows us to think about these questions
  - With its two-period structure, it allows for many analytical findings
  - It seems a useful framework to answer questions *in principle*
  - For a *careful explanation of data*, a many-period structure (probably with life-cycle features) would be more promising

- Real world relevance?
  - Hard to deny
  - Think about discussions about rising inequality of all sorts in many OECD and G7 countries
  - Think about the outcome of (pre-) elections and a referendum in the US and the UK
  - For more background and a starting point, see OECD (2015, 2008) or Wälde (2016)

Johannes Gutenberg-University Mainz
Bachelor of Science in Wirtschaftswissenschaften

Wealth distributions

Summer 2022

Niklas Scheuer (lecture) and (tutorials)

www.macro.economics.uni-mainz.de
April 21, 2022

Part III
Principles of other approaches

8 Models on Pareto distributions

8.1 Motivation

- What do we mean by this expression
  - Models whose wealth distribution will converge to a Pareto distribution
  - Models which subsequently analyze wealth inequality
- How does a typical wealth distribution look like:

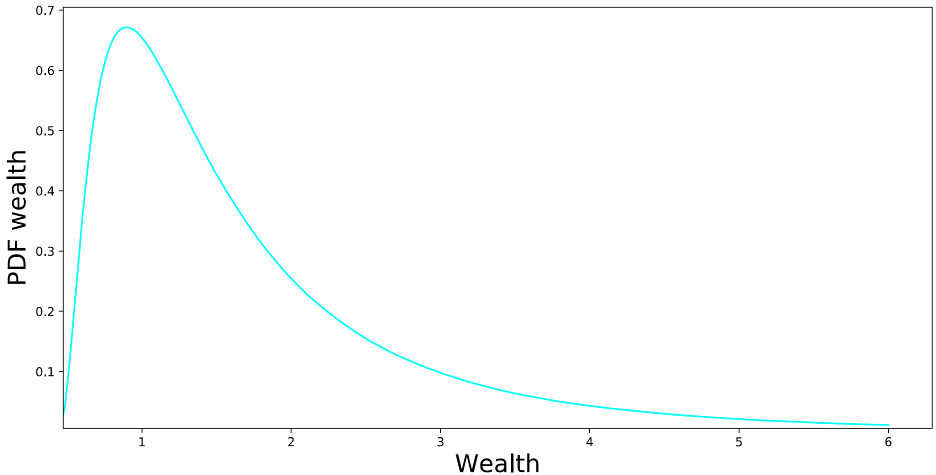

Figure 20 *Typical wealth distribution*

- What does this have to do with Pareto?
- If we focus on the right tail, we see?

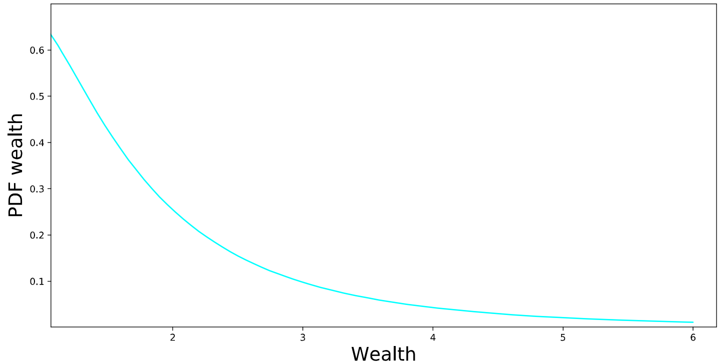

Figure 21 *Right tail of a typical wealth distribution*

- What do we see?
  - We observe the Pareto distribution for the right tail of the wealth distribution
  - This allows us to carefully characterize inequality
- Hence, we have to consider models leading to a Pareto distribution to properly apply data when intending to analyze top end inequality

- Literature for this section we want to focus on here:
  - Benhabib, Bisin and Zhu (2011) - The distribution of wealth and fiscal policy in economies with finitely lived agents
  - Benhabib and Bisin (2018) - Skewed Wealth Distributions: Theory and Empirics
  - Piketty and Zucman (2015) - Wealth and Inheritance in the Long Run
  - Castaneda, Gimenez and Rios-Rull (2003) - Accounting for the U.S. Earnings and Wealth Inequality

8.2 How to derive a Pareto distribution

8.2.1 Motivation

- This section intends to show a mechanism that leads to a Pareto distribution given a wealth process
- Used in many papers (!) but never properly shown $\implies$ This is why we want to show this
- There are many ways to do that
- However, the one presented here is rather trivial, yet elegant
- Some of it is outlined inter alia in Wackerly et al. (2008)
- Peu à peu we show which steps are necessary:
  - We begin with age being the variable that causes different wealth levels
  - So, heterogenity can be inferred from age only to keep it as simple as possible
  - Then we focus on wealth accumulation and
  - how this can be translated to represent a distribution

8.2.2 Age

- The idea of age, life and death is

$$dx = bdt - xdq_\delta \tag{8.1}$$

  where

  - age x increases over time with b (b is usually equal to 1)
  - Poisson process with arrival rate δ describes death
  - What happens in case a jump occurs?
  - Individual loses all her accumulated amount of years of living
  - hence, the individual dies

8.2.3 Wealth (the budget constraint)

- We consider a differential equation = typical budget constraint in an optimization problem

$$da_t = (r - \tau - \alpha)\, a_t dt - [a_t - \bar{a}] dq_\delta \tag{8.2}$$

  where:

  - da_t is the evolution of wealth at point in time t
  - r is the interest rate
  - τ is the tax rate
  - α is the share of wealth with which individuals consume
  - δ is the arrival rate of dq_δ mimicking death
  - No labor income

- where (continued):
  - Individuals who die, lose all their wealth, i.e. a_t
  - The state collects and accumulates all of that
  - $\bar{a}$ is equally redistributed among newborns
- Where does (8.2) come from
- Based on ideas of
  - Jones (2015)
  - Benhabib et al. (2016)
  - and many more
- STOP for an excursion

- Excursion: Poisson process[1]
  - Counting process $N(t), t \geq 0$ is said to be a Poisson process having arrival rate $\lambda, \lambda > 0$ if
    - (i) $N(0) = 0$
    - (ii) The process has stationary and independent increments
    - (iii) $\Pr(N(t) = 1) = \lambda t$
  - means
    - (i) At the initial point in time $t = 0$, the number of jumps is 0
    - (ii) Expected value, variance and other properties are the same for every increment (stationarity) and do not depend on said measures of other increments (independence)
    - (iii) Probability/Expectation of a jump taking place at t is the time period $(t-0)$ multiplied with the arrival rate of a jump λ
  - In the model above and also further below, *death* is a Poisson process with arrival rate δ
  - Examples for a Poisson process are Corona or Tipping points (climate change)

[1] See Ross (1993) for details.

– See Wälde (2012)

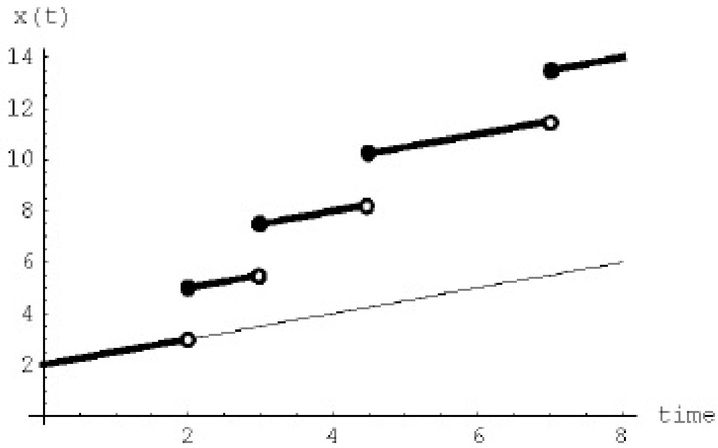

Figure 22 *Poisson process*

- Continue after excursion

8.2.4 Deriving the Pareto distribution

- Within this section we want to show the mechanism how to achieve a Pareto distribution coming from
- Steps we undertake:
  1. Set up the probability mass inflows and outflows for a certain time interval
  2. Derive the stationary distribution of age
  3. Use the method of distribution functions (Wackerly et al. (2008))
  4. Obtain the stationary distribution of wealth, i.e. the Pareto distributed wealth
- These steps essentially have two general requirements:
  1. The variable we want to obtain the Pareto distribution for, needs to grow exponentially with another variable
  2. This "other" variable needs to be exponentially distributed
- Et voilà - A Pareto distribution emerges

- What do we want to show theoretically?

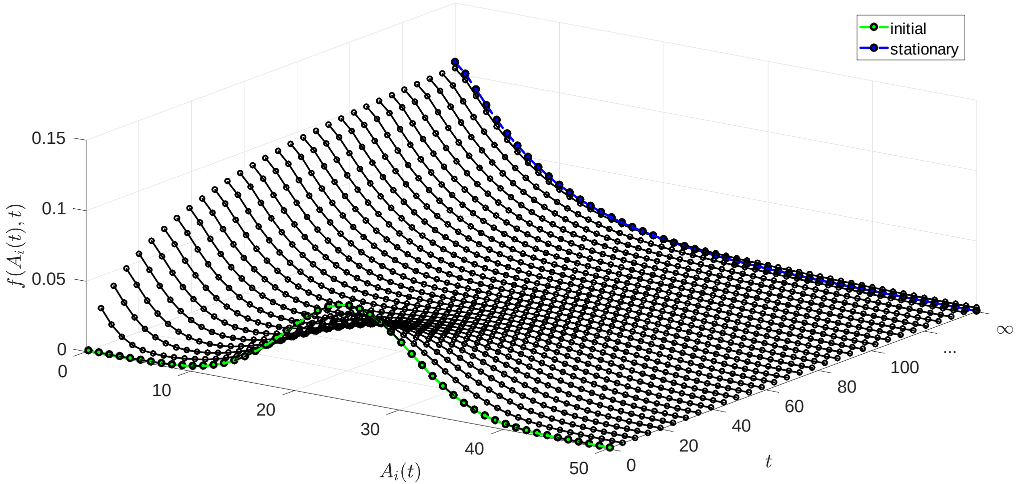

Figure 23 *Intuition of deriving a Pareto disrtibution for wealth*

- Looking at (8.1)
- We begin with

$$Pr(y,t) \equiv P(0 \leq x_t \leq y) \text{ for } y \geq 0 \tag{8.3}$$

- If we fix t and take a very small $h > 0$, we can write

$$\begin{aligned} P\left(y, t+h\right) = {} & e^{-\delta h} P\left(y - bh\right) \\ & + \delta h e^{-\delta h} \left[P\left(y - bh\right) + \left(1 - P\left(y, t\right)\right)\right] \end{aligned} \tag{8.4}$$

- For a visualization of (8.4) see

- We see

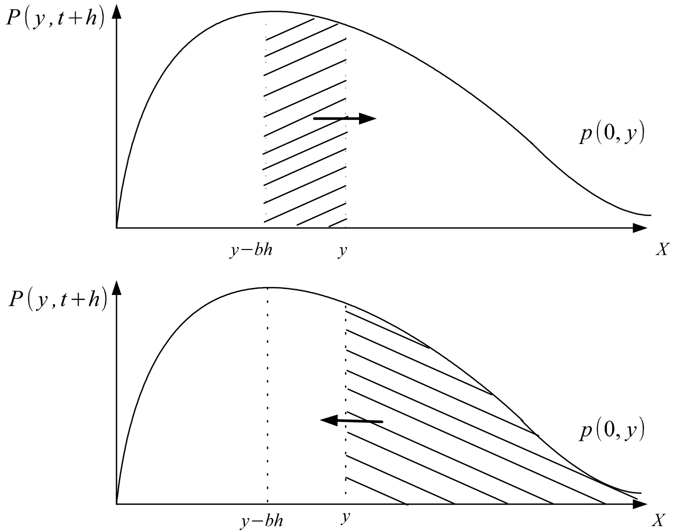

Figure 24 *Probability flows over the time interval h*

- Following a series of steps, we eventually obtain a partial differential equation (PDE) for change of the age distribution over time

$$\frac{\partial P\left(y,t\right)}{\partial t} = -b\frac{\partial P\left(y,t\right)}{\partial y} - \delta P\left(y,t\right) + \delta \tag{8.5}$$

- This partial differential equation (PDE) shows us the change of the probability mass according to our age process (8.1)

- As in the long run $t \implies \infty$, we know that $\frac{\partial P(y,t)}{\partial t} = 0$

- Thus, from (8.5) follows

$$\frac{\partial P\left(y,t\right)}{\partial y} = -\frac{\delta}{b}P\left(y\right) + \frac{\delta}{b}$$

- Given this ordinary differential equation (ODE), we are able to show the existence of a unique long-term stationary distribution of age

- Ultimately, we get

$$P\left(y\right) = 1 - e^{-\delta y} \tag{8.6}$$

- We can transform (Wackerly et al. (2008)) the stationary exponential age distribution into the stationary distribution of wealth
- Rewriting using $a_t = a_0 e^{(r-\tau-\alpha)x_t}$ (this will be shown in the tutorial) yields

$$P\left(a_t \leq z\right) = P\left(a_0 e^{(r-\tau-\alpha)x_t} \leq z\right) \tag{8.7}$$

- Reallocating with respect to x_t gives

$$P\left(a_t \leq z\right) = P\left(x_t \leq \ln\left(\frac{z}{a_0}\right)\frac{1}{r-\tau-\alpha}\right) \tag{8.8}$$

- Knowing (8.6), we can replace y with $\ln\left(\frac{z}{a_0}\right)\frac{1}{r-\tau-\alpha}$ and obtain the stationary distribution of wealth
- The stationary distribution of wealth reads

$$P\left(z\right) = 1 - \left(\frac{z}{a_0}\right)^{\frac{-\delta}{r-\tau-\alpha}} \tag{8.9}$$

- This expression is simultaneously the cumulative distribution function of a Pareto distribution with a Pareto inequality of $\frac{r-\tau-\alpha}{\delta}$

- At this point what usually follows is comparative statics:
  - How does inequality change with r
  - How does inequality change with τ
  - How does inequality change with δ
- and so on...
- Furthermore, this is not the only way to obtain a Pareto distribution
- Other mechanisms involve
  - Fokker-Planck equations (Kallenberg, 1997)
  - Laplace Theorem (Kase & Lei, 2018, or Gabaix et al. 2016)
- After we saw how to achieve a Pareto distribution for a certain differential equation/budget constraint let us consider *different differential equations* driven by *different factors*

8.3 An overview of Pareto models - Driving factors

- We intend to give a nice overview about the most prominent Pareto models $\Longrightarrow$ This section ought to characterize their commonalities and differences
- In recent years, we experienced a rise in publications of wealth papers[2] especially those featuring a Pareto distribution (sparked inter alia through Piketty (2014) - Capital in the Twenty-First century)
- This section covers theoretical studies and their underlying theoretical economic mechanisms to generate skewness and thick tails
- We also establish the link between empirical work and theoretical underpinnings
- Generally, given
  - a cross-section of countries and
  - time periods
- wealth distributions are skewed to the right displaying thick upper tails $\Longrightarrow$ Importance of Pareto models

[2] By that, we mean papers incorporating analysis of wealth distributions.

- Recall:

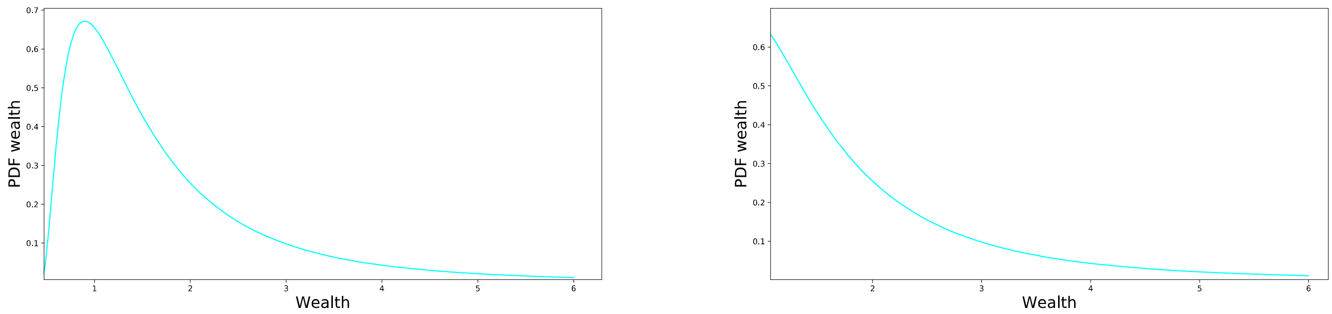

Figure 25 *Wealth distribution (left hand side) and its tail (right hand side)*

- What they all have in common is a micro foundation and the Pareto distribution

$$F(x) = 1 - \left(\frac{x_m}{x}\right)^{\alpha} \tag{8.10}$$

for $x \epsilon [x_m, \infty)$ and $x_m, \alpha > 0$

- x_m is the minimum value and α the parameter determining Pareto inequality

- Depicting (8.10) brings us back to the graph from before
- But what can be used to generates a Pareto distribution/a skewed distribution of wealth?
- In this section, we consider
  - Savings
  - Skewed earnings and stochastic returns to wealth
  - Explosive wealth accumulation
- in various forms and combinations applied to a general wealth accumulation equation:

$$w_{t+1} = r_{t+1} w_t + y_{t+1} - c_{t+1} \tag{8.11}$$

where

$$\begin{aligned} w_t &= \text{wealth at } t \\ w_{t+1} &= \text{wealth at } t+1 \\ r_{t+1} &= \text{interest rate at } t+1 \\ y_{t+1} &= \text{labour income in } t+1 \\ c_{t+1} &= \text{consumption in } t+1 \end{aligned}$$

8.3.1 Savings

- Assuming linear savings (=linear consumption) leads to

$$c_{t+1} = \varphi w_t + \chi_{t+1} \tag{8.12}$$

  where $\varphi, \chi_{t+1} \geq 0$

- As φ is the share of wealth used for consumption, χ_{t+1} marks the fraction of discounted sum of earnings consumed as well as precautionary savings (becomes evident when solving, not obvious at this stage)

- (8.11) becomes

$$w_{t+1} = (r_{t+1} - \varphi)\, w_t + y_{t+1} - \chi_{t+1} \tag{8.13}$$

- Normally models do not induce an exact linear consumption function

- However, useful benchmark to establish some basic properties of wealth accumulation processes

- Consider economies populated by agents with identical constant relative risk aversion (CRRA) preferences over consumption at any date t

$$u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma} \tag{8.14}$$

- who discount utility at a rate $\beta < 1$
- Wealth at any time can only be invested in an asset paying constant return r (for now)
- We distinguish
  - infinite horizon
  - OLG

- Excursion: Constant Relative Risk Aversion-CRRA (See also tutorial)
  - At this point it is useful to explicitly introduce CRRA (constant relative risk aversion) utility functions
  - Widely used in various applications
  - The measure of relative risk aversion is $\frac{-cu''(c)}{u'(c)}$
  - An individual with uncertain consumption at a small risk would be willing to give up a relative amount of consumption proportional to the measure of relative risk aversion
  - The CRRA utility function is the same function as the constant elasticity of substitution (CES) utility function
  - Inserting a CES utility function $\frac{\left(c^{1-\sigma}-1\right)}{(1-\sigma)}$ into this measure of risk aversion gives a measure of relative risk aversion σ (which is minus the inverse of the intertemporal elasticity of substitution)
  - This is why the CES utility function is also called CRRA utility function
  - It seems more appropriate to use the term CRRA in setups with uncertainty only
  - In a certain world without risk, risk-aversion plays no role

Infinite Horizon

- Consider an infinite horizon Bewley-Aiyagari-Huggett economy:
- They introduce heterogeneity (Heterogenous Agent Models) into the representative infinitely-living consumer
  - Wealth is
    * unproductive bonds (Huggett, 1993)
    * or capital (Aiyagari, 1994)
  - Idiosyncratic risk is
    * exogenous endowment shocks (Huggett, 1993)
    * labor income (Aiyagari, 1994)
- CRRA preferences, optimization subject to budget constraint
- The latter entails
  - a borrowing constraint
  - stochastic earnings
- $\Longrightarrow$ Precautionary motive for saving and accumulation

- If r_t is deterministic, i.e. $r_t = r$
  - Consumption is concave, and marginal consumption declines with wealth
  - Precautionary motive declines with higher wealth levels
  - The higher your wealth the less you have to worry, hence, the less you save in anticipation of future shocks
- If r_t is stochastic
  - The similar micro-foundations with CRRA preferences
  - Still a concave, asymptotically linear consumption function
  - Precautionary motive dies out for large wealth levels
  - Same as in the deterministic case

Overlapping generations

- See Benhabib, Bisin and Zhu (2011) for an OLG model
- The initial wealth of each dynasty maps into a bequest T periods later
- This is the initial condition for the next generation
- Equation (8.13) holds intergenerationally:

$$w_{n+1} = (r_n - \varphi)\, w_n + y_{n+1} - \chi_{n+1}$$

- Why is here an n and $n+1$ instead of t and $t+1$
  - In the case of overlapping generations we think in generations
  - So the wealth of generation $n+1$ is determined by
    * The rate of interest of the previous generation, r_n lowered by the share of wealth used for consumption φ multiplied by wealth of the previous generation
    * Additionally wealth of generation $n+1$ increases with the respective labor income of that generation, i.e. y_{n+1}
    * The wealth of generation $n+1$ decreases with precautionary savings, i.e. χ_{n+1}

8.3.2 Skewed earnings and/or Stochastic returns to wealth

- Discrete time
- We consider the deterministic case $r_t = r$, hence, the relevant budget constraint is (8.13)
- Assuming labor income y_t is described by a stationary distribution featuring a thick tail (implies an exponent of a cdf, say β between 0 and 2, see Crovella and Taqqu, 1998, p.55) with tail-index β
- Then (8.13) induces a stationary wealth distribution for wealth with tail $\alpha < \beta$
- The tail of wealth cannot be thicker than the tail of the earnings distribution
- Continuous time (wealth distribution is mostly developed in discrete time in the economics literature) and a Markov Process X change wealth accumulation:

$$dw = r\left(X\right) wdt + \sigma\left(X\right) dw \tag{8.15}$$

  where dw is a Brownian motion
- Through X the net rate of return $r\left(X\right)$ is stochastic

- Excursion: Brownian motion
- What is that?
- Looking at Ross (1993), a Brownian motion is defined accordingly:
  - A stochastic process $X(t), t \geq 0$ is said to be a Brownian motion process if
    * (i) $X(0) = 0$
    * (ii) $X(t), t \geq 0$ has stationary and independent increments
    * (iii) for every $t > 0$, $X(t)$ is normally distributed with mean 0 and variance $\sigma^2 t$
  - means:
    * (i) The value of X at the initial point in time $t = 0$ is zero
    * (ii) Between every point in time (increments) the properties (expected value and variance) do not change (stationarity) and do not relate to each other (independence)
    * (iii) Every value X at every point in time t is drawn (comes from) from a normal distribution with a certain mean and variance

- For instance:

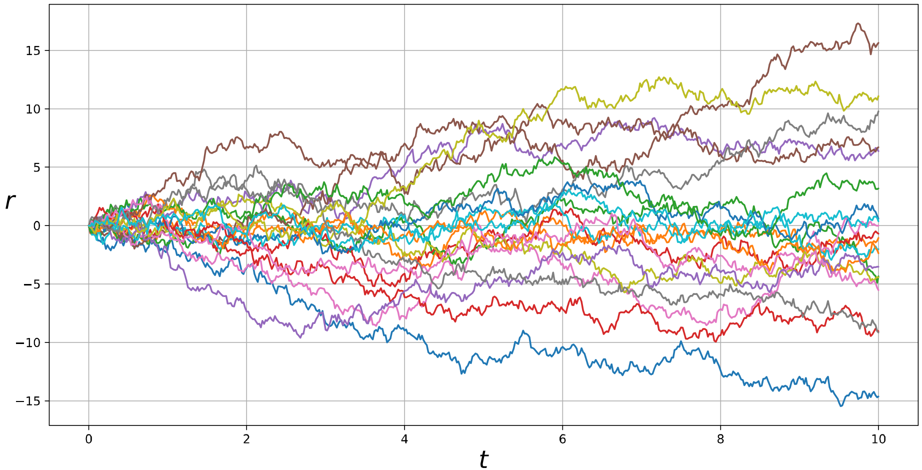

Figure 26 *A stochastic r over time t representing a Brownian Motion*

8.3.3 Explosive wealth accumulation and how to tame it

- "Explosiveness" triggered by many things
- "Explosiveness" means what?
- How can we then achieve a stationary wealth distribution skewed to the right?
- Solution: "births and deaths" (compare section 8.2.4):
- Blanchard (1985) introduces the perpetual youth model to tame explosiveness
- The model:
  - a deterministic explosive rate of return r
  - Perpetual youth: Constant mortality rate δ
  - The only stochastic variable generating wealth heterogeneity is the Poisson death rate
  - Constant earnings y
  - The discount rate is θ but agents discount future at rate $\theta + \delta$ incorporating the risk of death
  - Now - CUT

- Excursion: Constant death rate vs. Gompertz-Makeham
  - Independently of age (here X), individuals always face the same probability of dying
  - Whether you are 10 or 80 years old, probability is the same in this setup, i.e. δ
  - What do we see in reality?
  - In the beginning of life as an infant the risk of dying is high, but with increasing age the probability of dying decreases until with higher age it increases again
  - Gompertz-Makeham had the following idea:

$$\delta\left(X\right)=\gamma e^{\beta X(t)}+\rho \tag{8.16}$$

    * where γ is considered to be the initial mortality
    * β is considered to be the exponential mortality coefficient and
    * ρ, which was added by Makeham, is considered to be another constant including all other risks like accidents that may lead to death

– If we visualize this:

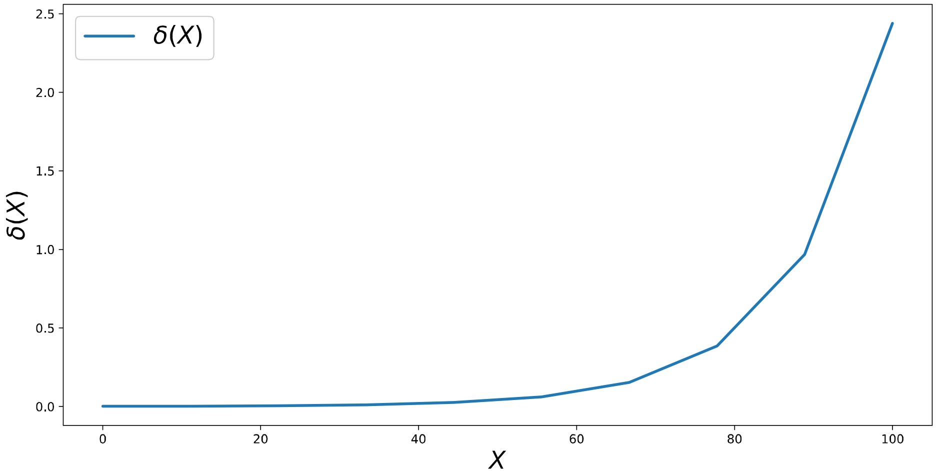

Figure 27 *Endogenous death rate δ, depending on age X*

- Now - CONTINUE
  - Individuals who die are 'replaced' with a new one
  - Newborns inherit/start with wealth w_0
  - $\implies$ Only driver of wealth is age
  - This leads to a Pareto distribution for large levels of wealth with the inverse of the exponent being again the measure of inequality $\alpha = \frac{r-\theta}{\delta}$
  - Interpretation:
    * If δ decreases, people live longer and longer and enjoy wealth growth for a longer time which increases α and, hence, inequality
    * If r increases, individual accumulate wealth faster which inceases inequality
    * If θ rises, individuals discount the future more and do not value the future, hence, less is saved contributing to less wealth inequality
- We now want to look at a specific model (out of the many we outlined above) generating a Pareto distribution

9 Precautionary saving

- What does this mean?
- People are afraid of what the future might bring$\Leftrightarrow$Individuals are risk-averse
- Individuals try to insure against future risk by displaying a precautionary motive
- We want to analyze such behavior using three models
  - At first, a standard utility maximization problem with two periods and no uncertainty
  - Then we introduce interest rate uncertainty to show what introducing uncertainty means in terms of the model and also its solutions
  - And finally we consider labor income uncertainty

9.1 A model without precautionary saving

9.1.1 The Individual

- One individual living two periods
- Labor income :
  - w_t in the first period (t)
  - w_{t+1} in the second period $(t+1)$
- Utility for the first and second period coming from consumption c_t and c_{t+1} with discount factor $0 < \beta < 1$

$$u = u(c_t) + \beta u(c_{t+1}) = \ln c_t + \beta \ln c_{t+1} \tag{9.1}$$

- Budget constraint in t and $t+1$ are

$$c_t + s_t = w_t \tag{9.2}$$

$$c_{t+1} = [1+r]\, s_t + w_{t+1} \tag{9.3}$$

  - r: interest rate
  - $s_t > 0$: savings from the first period (positiv)
  - $s_t < 0$: credit from the first period

9.1.2 Optimal behaviour

- Closed-form solution for logarithmic utility function
- Present Value of lifetime earnings

$$PV \equiv w_t + \frac{w_{t+1}}{1+r}$$

- Consumption levels

$$c_t = \frac{1}{1+\beta} PV, \quad c_{t+1} = \frac{\beta}{1+\beta} (1+r) PV \tag{9.4}$$

- In words: A fraction $\frac{1}{1+\beta}$ will be used for consumption in the first period
- The remaining fraction $\frac{\beta}{1+\beta}$ (plus capital income) for consumption in the second period
- Savings

$$s_t = w_t - c_t = w_t - \frac{w_t + \frac{w_{t+1}}{1+r}}{1+\beta} = \frac{\beta w_t - \frac{w_{t+1}}{1+r}}{1+\beta}$$

- No precautionary savings as no uncertainty present
- Individual is in a 'certain' world

9.2 A model with precautionary saving - labor income uncertainty

9.2.1 The individual

- One individual living two periods
- Labor income is uncertain in both periods:
  - $\tilde{w}_t$ in the first period (t)
  - $\tilde{w}_{t+1}$ in the second period ($t+1$)
- Utility for the first and second period coming from consumption c_t and c_{t+1} with discount factor $0 < \beta < 1$ (see above)
- Budget constraint in t and $t+1$ are

$$c_t + s_t = \tilde{w}_t \tag{9.5}$$

$$c_{t+1} = [1 + r]\, s_t + \tilde{w}_{t+1} \tag{9.6}$$

  - r: interest rate
  - $s_t > 0$: savings from the first period (positive)
  - $s_t < 0$: credit from the first period

- We consider a budget constraint with labour income uncertainty in period t:

$$\tilde{w}_t = \left\{ \begin{array}{c} w^{\text{high}} \\ w^{\text{low}} \end{array} \right\} \text{ with probability } \left\{ \begin{array}{c} p_t \\ 1 - p_t \end{array} \right\} \tag{9.7}$$

- In words: Individuals receive a wage w^{high} with probability p_t or wage w^{low} with probability $1 - p_t$ in period t

- Labour income uncertainty in period $t + 1$:

$$\tilde{w}_{t+1} = \left\{ \begin{array}{c} w^{\text{high}} \\ w^{\text{low}} \end{array} \right\} \text{ with probability } \left\{ \begin{array}{c} p_{t+1} \\ 1 - p_{t+1} \end{array} \right\} \tag{9.8}$$

- In words: Individuals receive a wage w^{high} with probability p_{t+1} or wage w^{low} with probability $1 - p_{t+1}$ in period $t + 1$

- This has to be incorporated throughout the analysis

9.2.2 Optimal behaviour

- We begin again with

$$\max_{\{c_t, c_{t+1}\}} U_t = \mathbb{E}_t \left[u\left(c_t\right) + \beta u\left(c_{t+1}\right) \right], \quad \beta \in (0, 1) \tag{9.9}$$

  where we can replace $u(c_t)$ and $u(c_{t+1})$ with $\ln c_t$ and $\ln c_{t+1}$, respectively

- We get

$$\max_{\{c_t, c_{t+1}\}} U_t = \mathbb{E}_t \left[\ln c_t + \beta \ln c_{t+1} \right] \tag{9.10}$$

- Replacing c_t and c_{t+1} with the budget constraints

$$\max_{\{s_t\}} U_t = \mathbb{E}_t \left[\ln\left(\tilde{w}_t - s_t\right) + \beta \ln\left(\left[1 + r\right] s_t + \tilde{w}_{t+1}\right) \right] \tag{9.11}$$

- Determining the first-order conditions

$$\frac{\partial U_t}{\partial s_t} = \mathbb{E}_t\left[-\frac{1}{\tilde{w}_t - s_t} + \beta\frac{1+r}{[1+r]\,s_t + \tilde{w}_{t+1}}\right] = 0$$

- Rewriting by using the additive separability of the expectation operator yields

$$\frac{\partial U_t}{\partial s_t} = \mathbb{E}_t\left[-\frac{1}{\tilde{w}_t - s_t}\right] + \beta\mathbb{E}_t\left[\frac{1+r}{[1+r]\,s_t + \tilde{w}_{t+1}}\right] = 0 \tag{9.12}$$

- We cannot go any further as the expected value of a fraction, such as $\mathbb{E}_t\left[-\frac{1}{\tilde{w}_t - s_t}\right]$ cannot be determined
  - Jensens inequality
  - Consider for instance $\mathbb{E}\left(\frac{1}{X}\right)$ for random variable X
  - Is this equal to $\frac{1}{\mathbb{E}(X)}$?
  - As $\frac{1}{X}$ will be convex for a strictly positive random variable X, we obtain $\mathbb{E}\left(\frac{1}{X}\right) \neq \frac{1}{\mathbb{E}(X)}$
- Hence, we have to turn to a numerical solution

9.3 Numerical solution

- How do we start?
- We have to find a value for s given assumptions regarding the other exogenous parameters, such as $\tilde{w}_t$, $\tilde{w}_{t+1}$, β and r, such that (9.12) is fulfilled
- We start with some function f depending on s, stemming from (9.12)

$$f(s) = \mathbb{E}_t\left[-\frac{1}{\tilde{w}_t - s_t}\right] + \beta\mathbb{E}_t\left[\frac{1+r}{[1+r]\, s_t + \tilde{w}_{t+1}}\right] \tag{9.13}$$

- We will obtain four cases

1) $w_t^{\text{high}}, w_{t+1}^{\text{high}}$	2) $w_t^{\text{high}}, w_{t+1}^{\text{low}}$
3) $w_t^{\text{low}}, w_t^{\text{low}}$	4) $w_t^{\text{low}}, w_{t+1}^{\text{high}}$

- A graphic visualization will illustrate what we intend to show and help us find a first initial value for s_0 we will begin solving with
- According to the case, we plot every possible $f(s)$

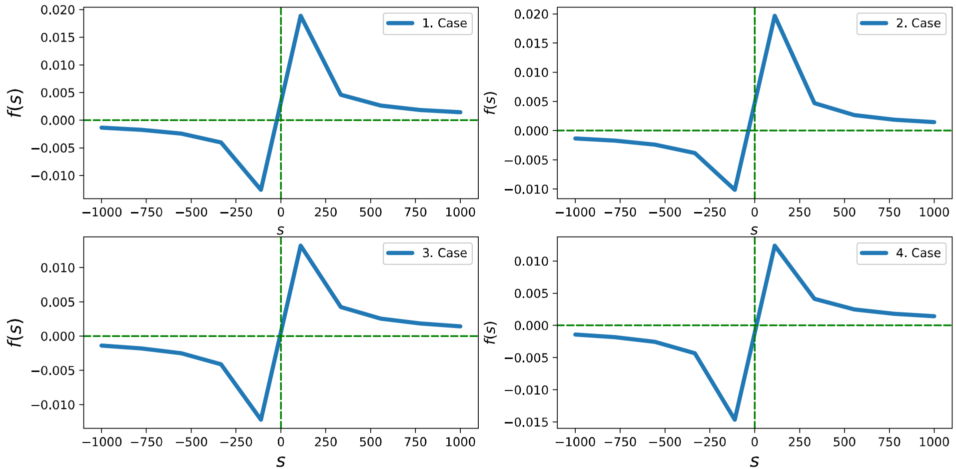

Figure 28 *All four possible cases of $f\left(s\right)$ for various combinations of $\tilde{w}_t$ and $\tilde{w}_{t+1}$*

- Keep in mind that this solution is only valid for assumptions we made regarding $\tilde{w}_t$, $\tilde{w}_{t+1}$, β and r
- Said parameters are defined as

$$\begin{aligned}
\text{high hourly wage, } w^{\text{high}} &= 50 \\
\text{low hourly wage, } w^{\text{low}} &= 10 \\
\text{discount rate, } \beta &= 0.4 \\
\text{deterministic interest rate, } r &= 0.03
\end{aligned}$$

- What do we see in the pictures?
  - Even though similar, the shape of every $f(s)$ changes with w
  - It appears that given our assumptions optimal saving, i.e. where $f(s)$ is 0 appears to be slightly negative or even 0 (fourth case)
  - Again, results highly susceptible to parameter

- We intend to find the s at which $f(s) = 0$ for every respective case
- In order to do so, we use build in routines for either MATLAB (fzero) or Python (fsolve)
- For instance for Case 1 we get $s = -20.38$
- This then also determines c_t and c_{t+1}
- for, again, given values of labor income in t and $t+1$
- What if had not assumed this structure regarding labor income uncertainty but just assumed labor income to be between 0 and 10000 Euros per hour?
- Same principle as before but solving numerically then takes a long time, as every possible labor income for both periods is incorporated $\Longrightarrow$ You'd achieve a vector of optimal saving levels
- This means that depending on case, the optimal solution for s is either
  - positive, meaning it is optimal to actually save in period t to consume more in $t+1$
  - or negative, meaning it is optimal to dissave and consume more than the budget constraint offers for period t
  - Within every case, there is no uncertainty

- What if *within case 1* the probability of having a high income in both periods is 30 %, respectively?
- This means

$$p_t = 0.3 \wedge p_{t+1} = 0.3$$

- Then the solution from $s = -20.38$ changes to

$$s = -0.04$$

- What does this mean?
  - If we look at one specific case and do introduce uncertainty for that very case(!) by explicitly considering the probabilities to be unequal to 1, *individuals dissave less or save more(!)*
  - Individuals precautionary savings are higher because of the uncertainty they are facing

Johannes Gutenberg-University Mainz
Bachelor of Science in Wirtschaftswissenschaften

Wealth distributions

Summer 2022

Niklas Scheuer (lecture) and (tutorials)

www.macro.economics.uni-mainz.de
April 21, 2022

Part IV
Programming

10 Python

10.1 Introduction

- Programming language such as C++, Java and so on
- Freeware
  - Language is free
  - Editors to edit programs are mostly free
  - Online Editors
- Find more here
- So please open up
  - PyCharm or a similar editor available via ZDV Apps
  - jupyterhub (Web based)

- Difference?
  - Former needs packages but no internet
  - Latter requires internet but no packages
- A typical first command is the *print* command
- print('Hello World') will yield a simple "Hello World"
- Basic operations like
  - plus, minus, multiply, divide
  - squared, root
- are always possible
- Before we go any further, we need to establish certain variable types
- Why?
- Well, because words cannot be multiplied!

10.2 Variables and types

- Beware of how you input variables

10.2.1 String

- Words
- Characters
- Letters

10.2.2 Numbers

- integer - A number without a fraction
- float - A number with a fraction
- For instance:
  - 7 is an integer
  - 2.5 is a float

10.3 if Conditions

10.3.1 if else

- if something is fulfilled
- then do something
- if something is not fulfilled
- then do something else

```
In [5]: x=1
        if x == 1:
            print('x is '+str(x))
        else:
            print('x is not '+str(x))

        x is 1
```

Figure 29 *Code example for If-else*

- First we set x equal to 1 (top line)
- We ask *if* x *is equal to* 1 (notice that for if conditions you use two equality signs)

  - then we print a sentence and add the value of x
  - Note that we have to change x from an integer to a string to avoid an error

- We *else* just print a sentence

10.3.2 elif

- if something is fulfilled
- then do something
- else if something is fulfilled
- then do something
- else
- but if the else condition contains another

In [2]:

```
x=2
if x ==1:
    print('x is '+str(x))
elif x ==2:
    print('x is '+str(x))
else:
    print('x is neither')
```

```
x is 2
```

Figure 30 *Code example for If-else-if*

- First we set x equal to 2 (top line)
- We ask *if* x *is equal to* 1 (notice that for if conditions you use two equality signs)
  - then we print a sentence and add the value of x
  - Note that we have to change x from an integer to a string to avoid an error
- We *else* ask *if* x *is equal to* 2 just print a sentence
  - then we print a sentence and add the value of x
  - Again we change x from an integer to a string to avoid errors
- We *else* just print a sentence

10.4 loops

10.4.1 for loops

- What do they do?
- *for* certain conditions particular actions will follow
- Imagine the following example:

```python
for i in range(10):
    x = 1+i
    print('i has the following value: '+str(i))
    print('x has the following value: '+str(x))
    print()
```

Figure 31 *Typical for-loop in python*

- What happens here?
  - The loop goes from $i = 0$ to $i = 9$ (this is because using the range command you need to subtract from the number you used to find the total amount of loops)
  - For every loop we calculate x, which is $1 + i$
  - What will be the value of i and x at the end of that loop?

10.4.2 while loops

- While some condition is fulfilled
- The while loop executes a program
- Difference to for-loop?
  - The for-loop specifies particularly how many times the commands are executed
  - e.g. *for i in range (10)* specifies clearly how many times the commands within the for-loop are exectued
  - Whereas the commands within a while-loop run until the condition is no longer fulfilled
  - This might be one turn but could also be a million turns, it is not particularly specified!

11 Assignments

11.1 Yathzee

- Popular game
- German term: "Kniffel"
- How does the game work?
  - Five dice within a dice cup
  - Then you roll the dice and decide which dice you want to keep outside and which you want to put back into the cup and roll again
  - Then you roll a second time and decide again
  - Then you roll a third time and decide again

- We do not look at the combinations of dice you can score
- It is just about simulating the game using Python
- Make use of:
  - *for*-loop
  - *if*-condition
  - *while*-loop

11.2 Dart Counter

- Dart is also a popular game
- You play a certain amount of points down to zero
- How does it work?
  - You throw three darts onto a board
  - The board has certain fields
    - 1 until 20
    - fields that triple your score for every number
    - fields that double your score for every number
- Make use of
  - *for*-loop
  - *if*-condition
  - *while*-loop

12 Economic Purpose behind Python

12.1 Plotting

- Remember the example for the cdf and pdf in the very beginning?
- Short reminder:

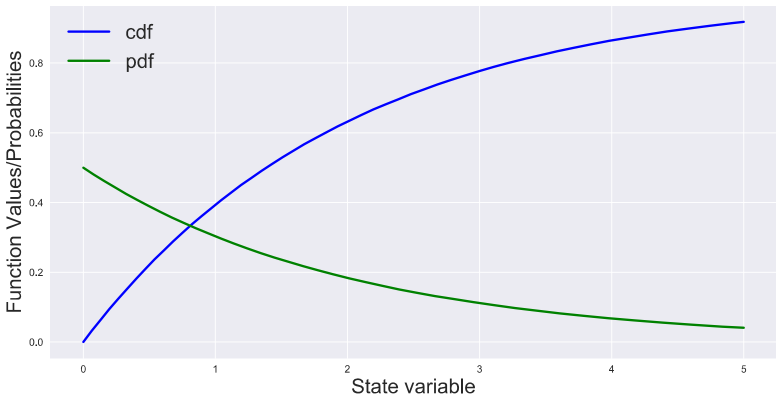

Figure 32 *cdf and pdf of an exponential distribution*

- How did we actually plot this using Python in PyCharm

12.1.1 Idea

- We begin with looking at the functional forms for the cdf and pdf of an exponential distribution
  - pdf reads
    $$f(x, \lambda) = \lambda e^{-\lambda x}$$
  - cdf reads
    $$F(x, \lambda) = 1 - \lambda e^{-\lambda x}$$
- How do we proceed in order to plot the last two equations

- In terms of programming, we begin with making sure that we are able to put in formulas and plot equations
  - We need the 'numpy' - package
  - We need the 'math' - package
  - We need the 'matplotlib' - package
- We know from the previous two equations, we have a rate parameter $\lambda > 0$ and a state variable x
- Hence:
  - Make assumptions regarding λ, e.g. set it to be 0.5
  - and state space x, e.g. set it from 0 to 5
- Write down functions and then use the
  - plt.plot, plt.legend, plt.xlabel, plt.ylabel
  - plt.show

12.1.2 Solution:

```
import math as m
import numpy as np
import matplotlib.pyplot as plt

x = np.linspace(0, 5, 5000)
l = 0.5

p = l*np.exp(-l*x)
c = 1-np.exp(-l*x)


plt.plot(x, c, color="blue", linewidth=2.5, linestyle="-", label='cdf')
plt.plot(x, p, color="green", linewidth=2.5, linestyle="-", label='pdf')
plt.legend(fontsize=20)
plt.xlabel('State variable', fontsize=20)
plt.ylabel('Function Values/Probabilities', fontsize=20)
plt.show()
```

Figure 33 *Code for PDF and CDF of the exponential distribution*

- Exercise: Explain line by line? What do you see and what does it mean?

12.1.3 Exercise

- Imagine the following equation:

$$\mu_i(a(t)) = a_0 e^{\left(r - \frac{\psi}{\sigma} - (1-\sigma)\delta\right)(t - t_B)},$$

  with $t_B = 0$, $a_0 = \chi a$ and the rest of parameters are between 0 and 1

- Plot μ over time t for a four values of a, i.e. $a_1 = 10$, $a_2 = 100$, $a_3 = 60$, $a_4 = 30$
- Ideas?
- Hints:
  - Do not forget packages to load
  - Is there an elegant way to plot this instead of writing the same code for *four different* a's
- If you got it, play around and interpret results

12.1.4 Solution

- Step 1:

```python
# ----------------------------------------
import math
import matplotlib.pyplot as plt
import numpy as np
```

Figure 34 *Solution Step 1*

- Step 2:

```python
# --------------------------
# define values
# --------------------------
psi = 0.5
sigma = 0.2
alpha = 0.5
delta = 0.7
r = 0.5
rho = 0.6
chi = 0.3
t_0= 1
# --------------------------
```

Figure 35 *Solution Step 2*

- Step 3:

```python
# ------------------------------
# define function
# ------------------------------

t = np.linspace(0, 10)
a0 = 50

```

Figure 36 *Solution Step 3*

- Step 4:

```python
def mu(t, a0):
    return a0*np.exp(((r-psi)/sigma -(1-chi)*delta)*(t-t_0))


plt.plot(t, mu(t, 10), linewidth=2.5, label= "$ \chi_a=10$")
plt.plot(t, mu(t, 100), linewidth=2.5, label="$ \chi_a=100$")
plt.plot(t, mu(t, 60), linewidth=2.5, label="$ \chi_a=60$")
plt.plot(t, mu(t, 30), linewidth=2.5, label="$ \chi_a=30$")
plt.xlabel("$t$", fontsize=20)
plt.ylabel("$\mu_i (a(t))$", fontsize=20)
plt.legend(fontsize=20)
plt.show()

```

Figure 37 *Solution Step 4*

12.2 Numerical Solving

- For most models and optimization problems we achieve a reduced form or, in other words, an analytical solution
- What does this mean?
  - A reduced form is a solution that entirely consists of exogenous parameters/variables
  - This goes hand-in-hand with an analytical solution
- What if we do not?
- Then we need to solve the optimization numerically
- What does this mean?
  - Roughly said: Use Software and let numbers run through your model until a solution is found
  - To be more specific, you approximate parameters and variables to solve the maximization function such that residuals are as small as possible

- Applications:
  - Imagine an optimization problem involving a Bellman equation in which you derive the Keynes-Ramsey rule (i.e. optimal consumption)
  - In order to not only have a general expession, you specify functional forms for utility
  - Guess the value function and the functional form of optimal consumption
  - This then needs to be verified that guesses are indeed a possible solution of a functional form
  - Often this is analytically impossible but needs to be done numerically by comparing residuals
- All this is programmed through MATLAB, Mathematica, Maple, Dynare or similar
- These are expensive programme outside the academic world $\Longrightarrow$ Use PyCharm or Jupyterhub
- Learning Python is not only exciting but also saves a lot of money

12.3 An example

- Imagine the following function

$$\frac{\partial U_t}{\partial s_t} = \mathbb{E}_t\left[-\frac{1}{\tilde{w}_t - s_t}\right] + \beta\mathbb{E}_t\left[\frac{1+r}{[1+r]\,s_t + \tilde{w}_{t+1}}\right] = 0$$

- How do you proceed?
  - What do you look for? $\Longrightarrow$ Element s such that this derivative is equal to 0 and lifetime utility is maximized
  - Plot the derivative to see where it crosses zero in order to find the root
  - Use a routine
    - fzero in MATLAB
    - fsolve in Python
  - To find a numerical solution (see section 9.3)

12.4 An excercise

- Given:

$$f(x) = \frac{2x^2 + 6x}{x^2 + 3}$$

- This is a very simple function
- The procedure is the same whether for trivial functions or more sophisticated ones
- Goal: Find the *roots* solving this function
- Plot the function to
  - See how the function looks like
  - To form a guess
- Solve $f(x)$ to find the roots based on your guess

- Step 1: Import necessary packages

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import fsolve
```

Figure 38 *Step 1*

- New here, the function to solve, i.e. *fsolve*
- This is the build-in solver of functions
- Step 2: Write the function and introduce the vector over which you want to find the roots
- $\Longrightarrow$ We find local roots whether these are coinciding with global roots is not always obvious

```python
def f(x):
    return (2*x**2+6*x)/(x**2+3)


x = np.linspace(-5,5)
```

Figure 39 *Step 2*

- Step 3: Define the vector over which you look for roots and plot the function

```
x = np.linspace(-5,5)
print(x)

plt.plot(x,f(x))
plt.plot(x,np.zeros(len(x)))
plt.show()
```

Figure 40 *Step 3*

- When we look at the function's plot, we form guesses

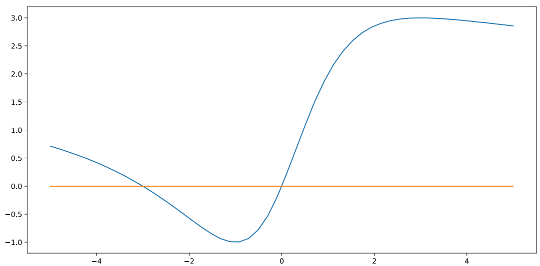

Figure 41 *Function plot of $f(x)$*

- What are our guesses?
  - We start with 0 as we see that the function should pass the x-axis at 0
  - The second root we see is somewhere between -4 and -2, so we guess -4 and let the solver do the rest
- Step 4: Use guesses and fsolve to find the roots

```
x = fsolve(f,0)
print(x)

x = fsolve(f, -4)
print(x)
```

Figure 42 *Step 4*

- After some seconds, we get the solution, which is

```
[0.]
[-3.]
```

- **Figure 43** *Solution*

References

Aiyagari, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," Quarterly Journal of Economics, 109, 659–84.

Beaumont, G. P. (2005): Probability and random variables. Elsevier.

Benhabib, J., and A. Bisin (2018): "Skewed Wealth Distributions: Theory and Empirics," Journal of Economic Literature, forthcoming, 1–47.

Benhabib, J., A. Bisin, and S. Zhu (2011): "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents," Econometrica, 79(1), 123–157.

——— (2016): "The Distribution of Wealth in The Blanchard-Yaari Model," Macroeconomic Dynamics, 20(2), 466–481.

Blanchard, O. J. (1985): "Debt, Deficits, and Finite Horizons," Journal of Political Economy, 93, 223–47.

Bossmann, M., C. Kleiber, and K. Wï£¡lde (2007): "Bequests, taxation and the distribution of wealth in a general equilibrium model," Journal of Public Economics, 91, 1247 – 1271.

Caputo, A., L. Fahrmeir, R. Künstler, S. Lang, I. Pigeot-Kübler, and G. Tutz (2009): Arbeitsbuch Statistik. Springer-Verlag.

Castaneda, A., J. Diaz-Gimenez, and J.-V. Rios-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," Journal of Political Economy, 111, 818–857.

Charles, K., and E. Hurst (2003): "The Correlation of Wealth across Generations," Journal of Political Economics, 111(6), 1155–1182.

Gabaix, X., J.-M. Lasry, P.-L. Lions, B. Moll, and Z. Qu (2016): "The Dynamics of Inequality," Econometrica, 84(6), 2071–2111.

Grossmann, V., and G. Kirsch (2016): "Erben ist ungerecht - und sollte mit 100 Prozenht besteuert werden," Sueddeutsche Zeitung.

Huggett, M. (1993): "The risk-free rate in heterogeneous-agent incomplete-insurance economies," Journal of Economic Dynamics and Control, 17, 953–969.

Jones, C. (2015): "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality," Journal of Economic Perspectives, 29(1), 29–46.

Kallenberg, O. (1997): "Foundations of Modern Probability," Springer-Verlag.

Kasa, K., and X. Lei (2018): "Risk, uncertainty, and the dynamics of inequality," Journal of Monetary Economics, 94, 60–78.

OECD (2008): Growing Unequal? Income Distribution And Poverty In OECD Countries. OECD, Paris.

——— (2015): In It Together - Why Less Inequality Benefits All. OECD, Paris.

Piketty, T. (2014): Capital in the 21st Century. Harvard University Press.

Piketty, T., and G. Zucman (2015): Wealth and Inheritance in the Long Run,pp. 1304–1366. Handbook of Income Distribution, Atkinson, A. and Bourgignon, F. (eds.), Volume 2B, 2015 Elsevier B.V.

Ross, S. M. (1993): Introduction to Probability Models, 5th edition. Academic Press, San Diego.

Ryglewski, S. (2019): Erweiterte Antwort auf kleine Bundestagsanfrage Nr. 177 November 2019. Bundesministerium der Finanzen, Berlin.

Saez, E., and G. Zucman (2016): “Wealth inequality in the United States since 1913: Evidence from capitalized income tax data,” The Quarterly Journal of Economics, 131(2), 519–578.

Sargent, T. (2008): Rational ExpectationsSteven n. Durlauf and Lawrencs E. Blume (Eds.) The New Palgrave Dictionary of Economics. Second Edition. Palgrave Macmillan.

Tiefensee, A., and M. M. Grabka (2017): "Das Erbvolumen in Deutschland dürfte um gut ein Viertel größer sein als bisher angenommen," DIW-Wochenbericht, 84(27), 565–571.

Wackerly, D., W. Mendenhall, and R. Scheaffer (2008): "Mathematical statistics," .

Wälde, K. (2012): Applied Intertemporal Optimization. Know Thyself - Academic Publishers, available at www.waelde.com/KTAP.

——— (2016): "Pareto-Improving Redistribution of Wealth - The Case of the NLSY 1979 Cohort," mimeo Johannes Gutenberg University Mainz.