RAND CORPORATION

Periods and cohorts

Ernesto F. L. Amaral

April 5, 2016

References:
Wachter KW. 2014. Essential Demographic Methods. Cambridge: Harvard University Press. Chapters 2 (pp. 30–47), 7 (pp. 153–173).
Fleurence RL, Hollenbeak CS. 2007. "Rates and probabilities in economic modelling: Transformation, translation and appropriate application." Pharmacoeconomics, 25(1): 3–6.

Periods and cohorts

(Wachter 2014, Chapter 2, pp. 30–47)
(Fleurence, Hollenbeak 2007)

- Lexis diagrams
- Period person-years lived
- Crude rate model
- Infant mortality rate
- Person-years and areas
- Cohort person-years lived
- Stationary population identity

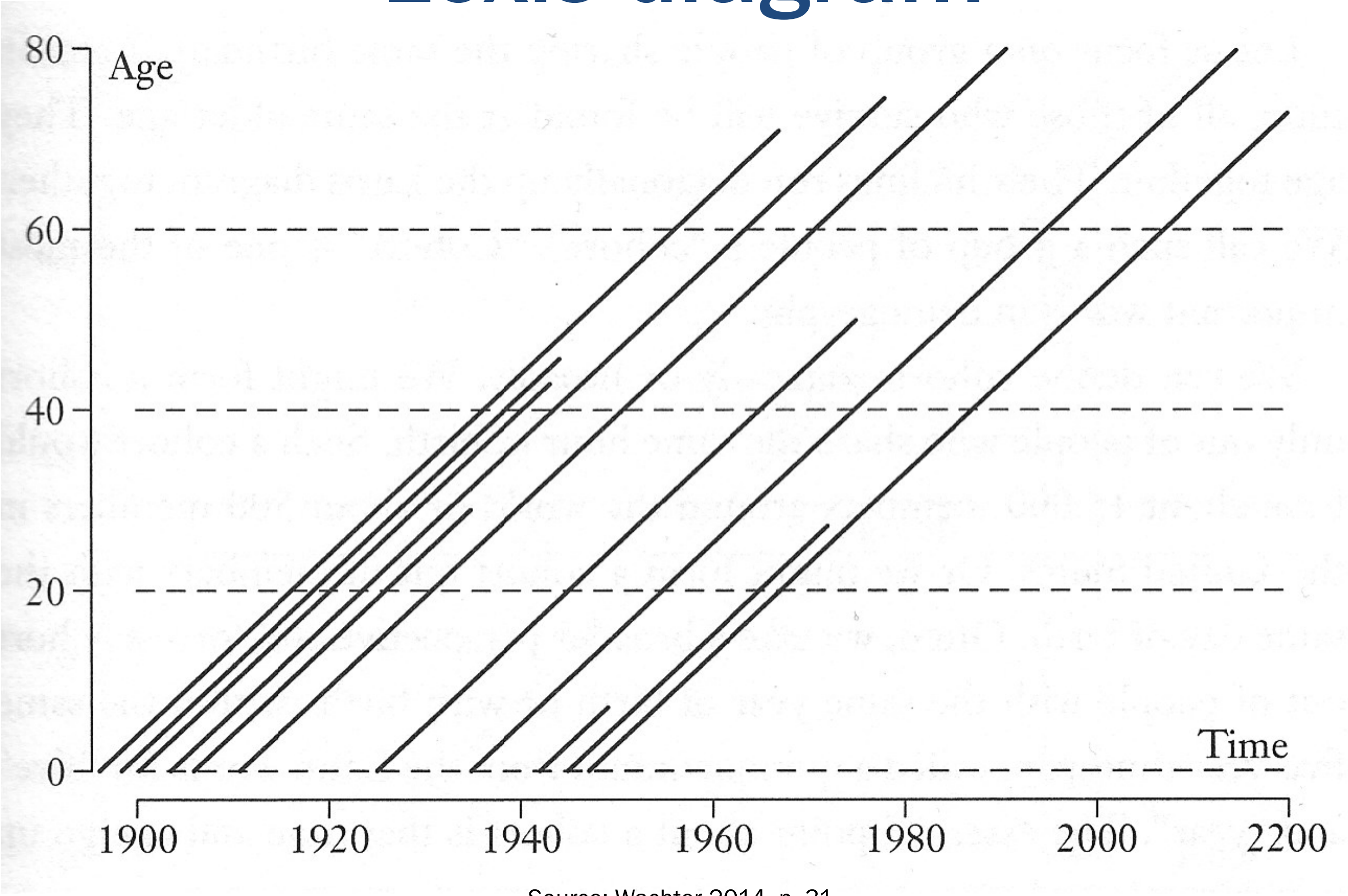
Exponential population growth model

- The exponential model treats all people as if they were alike
  - No mention to *age*
  - However, people are aging in the population
- Time enters demography in two ways
  - Chronological time: calendar dates, same for everyone
  - Personal time: age for each set of people who share same birthdate

Lexis diagram

- Lexis diagram provides relationships between chronological time *t* (horizontal) and age *x* (vertical)
- Each person has a lifeline on a Lexis diagram
  - Starting at (t_b, 0), where t_b is the person's birthdate and 0 is the person's age at birth
- Line goes up to the right with a slope equal to 1
  - People age one year in one year
- Lifeline goes up until time and age of the person's death

Lexis diagram



Source: Wachter 2014, p. 31.

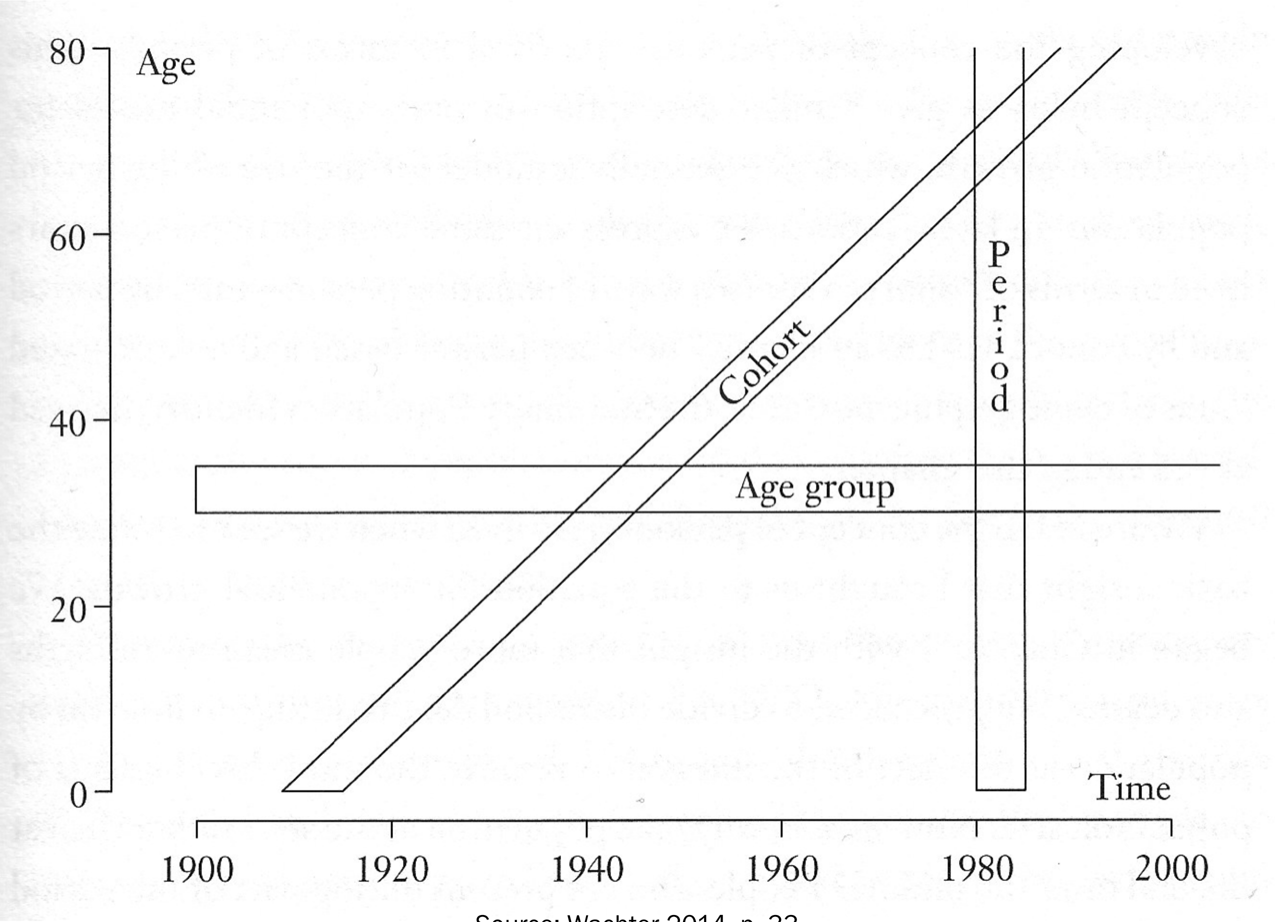
Exploring Lexis diagram

- To find population size
  - Draw vertical line upward from the time point
  - Count how many lifelines cross vertical line
- To find how many people survive to some age
  - Draw horizontal line across at the height corresponding to that age
  - Count how many lifelines cross that horizontal line
- Immigrants start at age and time of immigration

Cohort

- Group of people sharing the same birthdate
- Group of individuals followed simultaneously through time and age
- Their lifelines run diagonally up the Lexis diagram together
- In a cohort, time and age go up together
- A cohort shares experiences

Age, period, cohort



Source: Wachter 2014, p. 33.

Exponential growth

- For the equation for exponential growth
  - We divided births and deaths during an interval by population at the start of the interval

$$K(1) = K(0)\left(1 + \frac{B(0)}{K(0)} - \frac{D(0)}{K(0)}\right)$$

- Why not population at the end or in the middle?
  - People who are present during part of the period can also have babies or become corpses
  - More people present for more time in the denominator generate higher exposure (“risk”) to births and deaths

Period person-years lived (PPYL)

- Person-years is the sum of each individual's time at risk of experiencing an event (e.g. birth, death, migration)
  - For those who do not experience event, person-years is the sum of time until end of period
  - For those who experience event, it is the time until the event
- PPYL take into account that people are present during part of the period (fraction of years)
  - Each full year that a person is present in a period, he/she contributes one "person-year" to the total of PPYL
  - Each day a person is present in the population, he/she contributes 1 person-day, or 1/365 person-year, to PPYL

Calculating person-years

- Whenever we know the population sizes on each day over the period of a year
  - We can add up the person-years day by day
  - Take the number of people present on the first day times 1/365 of a person-year for each of them
  - Add up all contributions for following days
  - When our subintervals are small enough, our sum is virtually equal to the area under the curve of population as a function of time during the period

Approximation for PPYL

- When sequence of population sizes throughout a period are unknown
  - Take the population in the middle of the period and multiply by the length of the period
    - E.g., for 2005–2015, we take the mid-period count of 308.745 million in the U.S. from 2010 Census and multiply by 10 years to obtain 3,087 million person-years in the period
  - Or take the average of the starting and ending populations and multiply by the length of the period

Rates (Fleurence, Hollenbeak 2007)

- Rates are an instantaneous measure that range from zero to infinity
- Rates describe the number of occurrences of an event for a given number of individuals per unit of time
- Incidence rate describes the number of new cases of an event during a given time period over the total person-years of observation
- <u>**Numerator**</u>: number of events (e.g. births, deaths, migrations)
- <u>**Denominator**</u>: number of “person-years of exposure to risk” experienced by a population during a certain time period
- The denominator is the number of person-years
- Time is included directly in the denominator
- Rates take into account the time spent at risk

Probabilities (Fleurence, Hollenbeak 2007)

- Probabilities describe the likelihood that an event will occur for a single individual in a given time period and range from 0 to 1
- Does not include time in the denominator
- Divides the number of events by the total number of people at risk in the relevant time frame
- Conversion between rates and probabilities:

$$\text{probability: } p = 1 - e^{-rt}$$

$$\text{rate: } r = -1/t * \ln(1-p)$$

Ratios

- Describe a relationship between two numbers indicating how many times the first number contains the second
- Compares the size of one number to the size of another number
- Denominator is not at “risk” of moving to the numerator
- Examples
  – Total dependency ratio = population of children (0–14) plus elderly population (65+) divided by working-age population (15–64)
  – Sex ratio = population of males divided by population of females

CBR and CDR

- Crude Birth Rate (CBR or *b*)
  - Number of births to members of the population in the period divided by the total period person-years lived
- Crude Death Rate (CDR or *d*)
  - Number of deaths to members of the population in the period divided by the total period-years lived

Crude rate model

- Imagine a population
  - In which each person, each instant, is subject to constant independent risks of dying and having a baby
  - *b*: expected numbers of births per person per year
  - *d*: expected number of deaths per person per year
- Assumptions
  - Closed population
  - Homogeneous risks among people
  - No measurement of change over time inside the period

Growth rate

- Expected size of population has exponential growth
  - Growth rate = R = $b - d$
- Most actual populations are not closed and risks are not homogeneous over time
  - Need a measure of Crude Net Migration Rate (MIG)
  - Crude Growth Rate (CGR) = CBR – CDR + MIG

Most populous countries, 2012

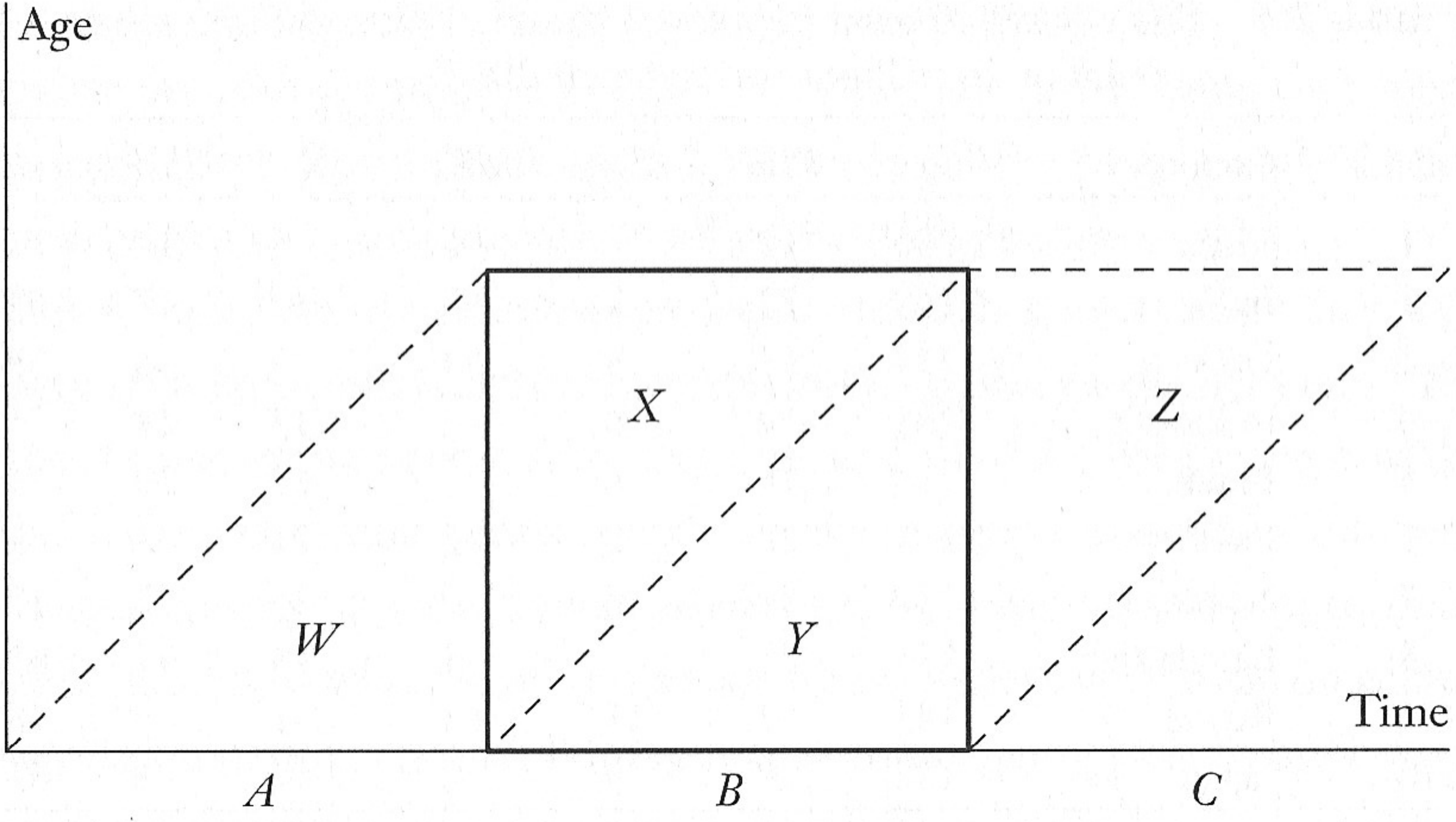
Rank	Country	Pop. (million)	CBR (‰)	CDR (‰)	MIG (‰)	R (‰)	IMR (‰)	e_0
1	China	1,350	12	7	–0	5	17	73
2	India	1,260	22	7	–0	16	47	65
3	USA	314	13	8	+3	9	6	78
4	Indonesia	245	19	6	–1	12	29	71
5	Brazil	194	16	6	–0	11	20	73
6	Pakistan	188	28	8	–2	21	64	63
7	Nigeria	170	40	14	0	24	77	47
8	Bangladesh	153	23	6	–3	14	43	65
9	Russia	143	12	15	+2	–1	8	68
10	Japan	128	9	9	0	0	3	83
	World	**7,017**	**20**	**8**	**0**	**12**	**46**	**69**

Infant mortality rate (IMR)

$$IMR = \frac{the\ number\ of\ deaths\ under\ age\ 1\ in\ the\ period}{the\ number\ of\ live\ births\ in\ the\ period}$$

- IMR is a period measure
- It uses current information from vital registration
- It can be computed for countries without reliable census or other source for a count of the population at risk by age
- Infants borne by teenagers and by older mothers are at higher risk

IMR contributions on a Lexis diagram



Source: Wachter 2014, p. 38.

Understanding previous figure

- Any lifeline which ends within the square
  - Contributes a death to the numerator of the IMR
- Any lifeline that starts on the base of the square
  - Contributes a birth to the denominator of the IMR

Still on previous figure

- Babies born outside the period in the preceding year (A) may die as infants during the period (X)
  - Counted in the numerator, but not in denominator
- Babies born during the period (B) may die after the end of the period (Z)
  - Counted in the denominator, but not in numerator
- Usually mismatched terms balance each other
  - IMR is close to the probability of dying before age 1

Period ≠ Cohort

- Period deaths and period person-years lived
  - Come from deaths and lifelines in the square (*X*, *Y*)
  - Dividing these deaths by person-years gives a period age-specific mortality rate (*M*)
- Cohort deaths and cohort person-years lived
  - Come from deaths and lifelines in parallelogram (*Y*, *Z*)
  - Dividing these deaths by person-years gives a cohort age-specific mortality rate (*m*)

Person-years and areas

- PPYL in the period between time 0 and time T is the area under the curve K(t) between 0 and T

$$PPYL = \int_0^T K(t)dt$$

- When growth is constant (exactly exponential)

$$PPYL = K(0)(e^{RT} - 1) / R = (K(T) - K(0)) / R$$

$$Growth\ Rate = R = CBR - CDR$$

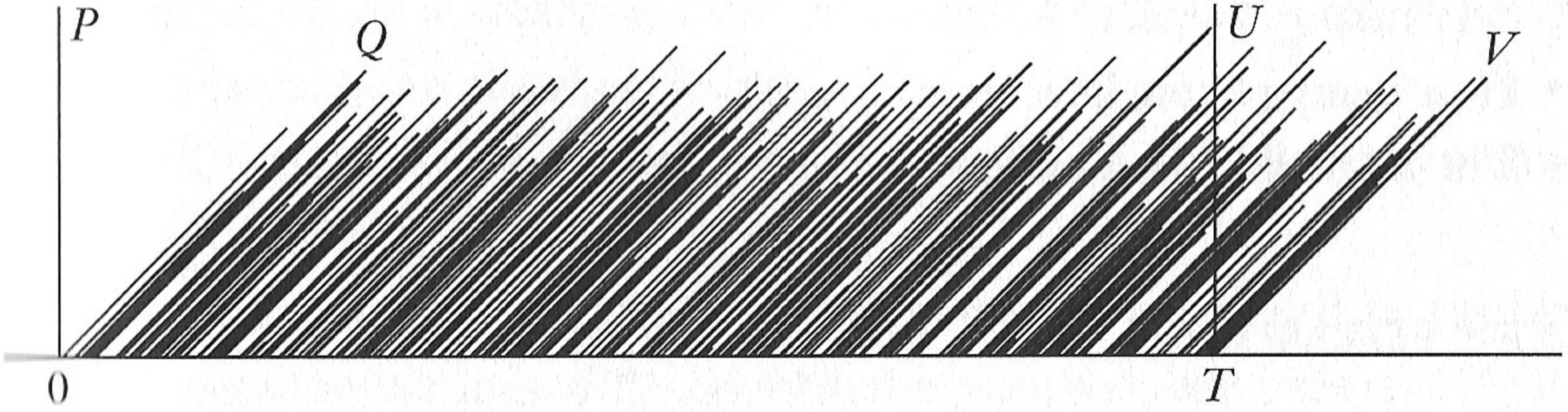
Cohort person-years lived (CPYL)

- We get CPYL when we add up all person-years lived by all members of the cohort
  - Instead of counting people from a rectangle of the Lexis diagram, we consider a parallelogram
- If we divide by the total number of members of the cohort (counted at birth)
  - We get expectation of life at birth (e_0)
  - Average number of person-years lived in their whole lifetimes by members of the cohort

Stationary population identity

- Stable population
  - Demographic rates are unchanging
  - Size might be growing, constant or declining
- Stationary population
  - Numbers are unchanging
  - Total population is the same from year to year (B=D)
  - # births is constant = B = Population * CBR = Kb
  - # deaths is constant = D = Population * CDR = Kd
- PPYL ≈ CPYL, so we have: $K\ T = K\ b\ e_0\ T$
- Stationary population identity: $1 = b\ e_0$ *when* $R=0$

Lexis diagram
for a stationary population



Source: Wachter 2014, p. 45.