Summer School on Capability and Multidimensional Poverty

28 August-9 September, 2008

New Delhi, India

HDCP-IRC

The Human Development, Capability and International Research Centre
Instituto Universitario di Studi Superiori
www.iusspavia.it

OPHI

Oxford Poverty & Human Development Initiative
University of Oxford
www.ophi.org.uk

Poverty Dynamics

Gaston Yalonezky (OPHI) and
Maria Emma Santos (OPHI)

30 August, 2008

Topics in poverty dynamics

- Contents:
  - Conceptual issues. Why studying poverty dynamics?
  - State of the art of the literature: measurement (indices) and models
  - Examples of indices
  - Examples of modelling techniques

Conceptual Issues

- **Motivation:** Some of the poor are persistently poor while others are not. The set of the poor is heterogeneous.

- Then, poverty is ***dynamic***.
  - The economic aspects of poverty (income) are particularly dynamic.

Typical Dataset

$$\begin{array}{ccc} & \text{Periods} & \\ y = & \begin{bmatrix} y_{11} & \cdots & y_{1T} \\ y_{21} & \cdots & y_{2T} \\ \cdots & & \\ & & \cdots \\ y_{n1} & \cdots & y_{nT} \end{bmatrix} & \text{People} \end{array}$$

Conceptual Issues

- **Objective:** Distinguish the ***chronic poor*** from the ***transient poor***, measure each type of poverty, and characterise it ("explanations" of poverty).

- **Ultimate Purpose:** To inform policy design to tackle more effectively each type of poverty.

- **Requirement:** Longitudinal data which tracks individuals over time. This makes the analysis different from static poverty analysis and poverty trends.

Links with other literatures

- Economic Mobility (Poverty dynamics focuses on mobility at the lower end of the distribution).

- Vulnerability.

- Intergenerational transmission (persistence) of poverty.

State of the Art of the Literature: Indices

- Sen's two-steps for poverty measurement also apply here.
  1. **Key Q/:** Who are the chronic poor and who are the transient poor? (Identification Step).

  2. How do we aggregate individual poverty?

Spells Approach: Identification

- The chronic poor are identified based on the number of periods they spend in poverty (persistence criterion).
- Two thresholds (Foster, 2007):
  - Z: poverty line
  - τ: duration line
- d_i: fraction of time in poverty.

Spells Approach: Identification

$$CP = \{i : d_i \geq \tau\}$$

$$TP = \{i : 0 < d_i < \tau\}$$

$$IP = \{i : d_i > 0\}$$

Foster (2007)-Spells Approach: Aggregation

- Foster (2007) proposes ***duration adjusted*** FGT class of chronic poverty measures.

$$F_{\alpha}^{C}(Y;z,\tau)=\mu(g^{\alpha}(\tau))$$

  - With α=0 satisfies time monotonicity.
  - With α=1 satisfies both time and income monotonicity.
  - With α=2 also satisfies distribution sensitivity.

Example:

z=10

$$y = \begin{bmatrix} 8 & 8 & 8 & 8 \\ 4 & 6 & 12 & 4 \\ 12 & 6 & 4 & 14 \\ 15 & 15 & 12 & 18 \end{bmatrix}$$

$$g_{it}^{\alpha} = \left(\frac{z - y_{it}}{z} \right)^{\alpha} \quad if \quad y_{it} < z$$

$$g_{it}^{\alpha} = 0 \quad if \quad y_{it} \geq z$$

If τ=0.75

$$g^{0} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathrm{d} = \begin{bmatrix} 1 \\ 0.75 \\ 0.50 \\ 0 \end{bmatrix} \quad g^{0}(\tau) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Example: Headcount Ratio

$$\mathrm{d}(\tau) = \begin{bmatrix} 1 \\ 0.75 \\ 0.50 \\ 0 \end{bmatrix} \quad g^0(\tau) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$H = q/n = 0.5$$

q: number of people in set CP.

Note that if individual 2 becomes poor in period 3, H does not change:
Violates time monotonicity.
Also violates income monotonicity.

$$\mathrm{d}'(\tau) = \begin{bmatrix} 1 \\ 1 \\ 0.50 \\ 0 \end{bmatrix}$$

$$H' = 0.5$$

Example: Adjusted Headcount Ratio

$$d(\tau)=\begin{bmatrix}1\\0.75\\0\\0\end{bmatrix}\quad g^0(\tau)=\begin{bmatrix}1&1&1&1\\1&1&0&1\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

$$F_0^C=\mu(g^0(\tau))=\frac{|g^0(\tau)|}{nT}=\frac{7}{16}=0.43$$

$$F_0^C=HD=\left(\frac{q}{n}\right)\left(\frac{|d(\tau)|}{q}\right)=\left(\frac{2}{4}\right)\left(\frac{1.75}{2}\right)=0.43$$

- Satisfies time mon.
- If any poverty income of the CP decreases, K0 does not change. **Violates income monotonicity.**

D: Average duration among the chronically poor.

Example: Adjusted Poverty Gap

$$y = \begin{bmatrix} 8 & 8 & 8 & 8 \\ 4 & 6 & 12 & 4 \\ 12 & 6 & 4 & 14 \\ 15 & 15 & 12 & 18 \end{bmatrix} \qquad g^1(\tau) = \begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 \\ 0.6 & 0.4 & 0 & 0.6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$F_1^C = \mu(g^1(\tau)) = \frac{|g^1(\tau)|}{nT} = \frac{2.4}{16} = 0.15$$

$$F_1^C = HDG = \left(\frac{q}{n}\right)\left(\frac{|d(\tau)|}{q}\right)\left(\frac{|g^1(\tau)|}{|g^0(\tau)|}\right) = \left(\frac{2}{4}\right)\left(\frac{1.75}{2}\right)\left(\frac{2.4}{7}\right) = 0.15$$

G: Average gap across all poverty spells of the chronically poor.

Satisfies income monotonicity. But a decrement in income in poorest spell has same impact as a decrement in income in least poor spell. **It is not distribution sensitive**.

Example: Adjusted Squared Poverty Gap

$$y = \begin{bmatrix} 8 & 8 & 8 & 8 \\ 4 & 6 & 12 & 4 \\ 12 & 6 & 4 & 14 \\ 15 & 15 & 12 & 18 \end{bmatrix} \qquad g^2(\tau) = \begin{bmatrix} 0.04 & 0.04 & 0.04 & 0.04 \\ 0.36 & 0.16 & 0 & 0.36 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$F_2^C = \mu(g^2(\tau)) = \frac{|g^2(\tau)|}{nT} = \frac{1.04}{16} = 0.065$$

$$F_2^C = HDS = \left(\frac{q}{n}\right)\left(\frac{|d(\tau)|}{q}\right)\left(\frac{|g^2(\tau)|}{|g^0(\tau)|}\right) = \left(\frac{2}{4}\right)\left(\frac{1.75}{2}\right)\left(\frac{1.04}{7}\right) = 0.065$$

S: Average squared gap across all poverty spells of the chronically poor.
A decrement in income in poorest spell will increase chronic poverty more than a decrement in income in least poor spell.
It is distribution sensitive.

Duration Adjusted FGT

- Transient Poverty is obtained as a residual of what might be called "Inter-temporal" Poverty and Chronic Poverty.

- Inter-temporal poverty: Average FGT across people and periods.

$$F_{\alpha}^{T}(Y;z,\tau)=\mu(g^{\alpha}(0))-\mu(g^{\alpha}(\tau))$$

Example:

$$y = \begin{bmatrix} 8 & 8 & 8 & 8 \\ 4 & 6 & 12 & 4 \\ 12 & 6 & 4 & 14 \\ 15 & 15 & 12 & 18 \end{bmatrix} \quad g^0 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad g^0(\tau) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$F_0^T(Y; z, \tau) = \mu(g^0(0)) - \mu(g^0(\tau)) = \frac{9}{16} - \frac{7}{16} = 0.125$$

Example:

$$y=\begin{bmatrix} 8 & 8 & 8 & 8 \\ 4 & 6 & 12 & 4 \\ 12 & 6 & 4 & 14 \\ 15 & 15 & 12 & 18 \end{bmatrix} \quad g^1=\begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 \\ 0.6 & 0.4 & 0 & 0.6 \\ 0 & 0.4 & 0.6 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad g^1(\tau)=\begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 \\ 0.6 & 0.4 & 0 & 0.6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$F_1^T(Y;z,\tau)=\mu(g^1(0))-\mu(g^1(\tau))=\frac{3.4}{16}-\frac{2.4}{16}=0.0625$$

Example:

$$y = \begin{bmatrix} 8 & 8 & 8 & 8 \\ 4 & 6 & 12 & 4 \\ 12 & 6 & 4 & 14 \\ 15 & 15 & 12 & 18 \end{bmatrix} \quad g^2 = \begin{bmatrix} 0.04 & 0.04 & 0.04 & 0.04 \\ 0.36 & 0.16 & 0 & 0.36 \\ 0 & 0.16 & 0.36 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad g^2(\tau) = \begin{bmatrix} 0.04 & 0.04 & 0.04 & 0.04 \\ 0.36 & 0.16 & 0 & 0.36 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$F_2^T(Y; z, \tau) = \mu(g^2(0)) - \mu(g^2(\tau)) = \frac{1.56}{16} - \frac{1.04}{16} = 0.0325$$

Spells Approach: Pros and Cons

- Advantages:
  - Intuitive.
  - Responds to persistence criterion.
- Disadvantages:
  - Does not consider the depth of poverty at the identification step.
  - Implicitly assumes that there is no possibility for income substitution across periods.

Components´s Approach Jalan y Ravallion (2000)

- Distinguish a ‘permanent’ income (or consumption) component from its transitory fluctuations.
- They distinguish a chronic and a transient component from an individual’s poverty.
- ***Identification:*** the chronic poor are those with a mean income over time below the poverty line.
- ***Aggregation:*** FGT_2

Jalan and Ravallion

- Following the same notation, chronic poverty is given by:

$$J^C(y;z) = \mu(g^2(\bar{y};z)) \quad \bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$$

$\bar{y}$ is the vector of inter-temporal means

Jalan & Ravallion

- Transient Poverty is obtained as a residual of what they call "Inter-temporal" Poverty and Chronic Poverty.

- Inter-temporal poverty: Average FGT across people and periods.

$$J^{T}(Y;z,\tau)=\mu(g^{2}(y))-\mu(g^{2}(\overline{y}))$$

Example: z=10

$$y = \begin{bmatrix} 8 & 8 & 8 & 8 \\ 4 & 6 & 12 & 4 \\ 12 & 6 & 4 & 14 \\ 15 & 15 & 12 & 18 \end{bmatrix} \quad \bar{y} = \begin{bmatrix} 8 \\ 6.5 \\ 9 \\ 15 \end{bmatrix}$$

$$g_i^2 = \left(\frac{z - \bar{y}_i}{z} \right)^2 \; if \quad \bar{y}_i < z$$

$$g_i^\alpha = 0 \qquad if \quad \bar{y}_i \geq z$$

$$g^2 = \begin{bmatrix} 0.04 & 0.04 & 0.04 & 0.04 \\ 0.36 & 0.16 & 0 & 0.36 \\ 0 & 0.16 & 0.36 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad g^2(\bar{y}, z) = \begin{bmatrix} 0.04 \\ 0.1225 \\ 0.01 \\ 0 \end{bmatrix}$$

$$J^C(y; z) = \mu(g^2(\bar{y}; z)) = \frac{0.1725}{4} = 0.043$$

$$J^T(y; z, \tau) = \mu(g^2(y)) - \mu(g^2(\bar{y})) = \frac{1.56}{16} - \frac{0.1725}{4} = 0.054$$

J&R´s Approach: Pros and Cons

- Differs from the spells approach in that a household identified as chronic poor can have a positive transient poverty value.

- Advantages:
  - Considers the depth of poverty at the identification step.

- Disadvantages:
  - Implicitly assumes perfect income substitutability across periods.

Other Approaches

- Foster & Santos (2006): Variant of Jalan and Ravallion, allowing imperfect substitutability.
- Duclos, Araar & Giles (2006): incorporates inequality in poverty gaps over time.
- Cruces & Wodon (2007): ‘Risk-adjusted’ poverty measure.
- Calvo & Dercon (2007): See discussion on discounting, direction of time and ‘bunching’ of poverty spells.

State of the art of the literature

- Micro-growth models
  - Similar to cross-country economic growth assessments
  - Conditional convergence/ divergence versus poverty traps
  - Studies of insurance against covariate and idiosyncratic risk
  - Household level; dependent variable is usually change in log consumption
  - Examples: Cochrane (1991); Mace (1991); Jalan and Ravallion (2002); Dercon (2004); Premand (2008)

State of the art of the literature

- Micro growth models
  - Examples regarding risk insurance:

$$\Delta \ln C_{it} = f\left(\alpha_i; C_{it-1}; C \vee aggregate\ risk; idiosyncratic\ events\right)$$

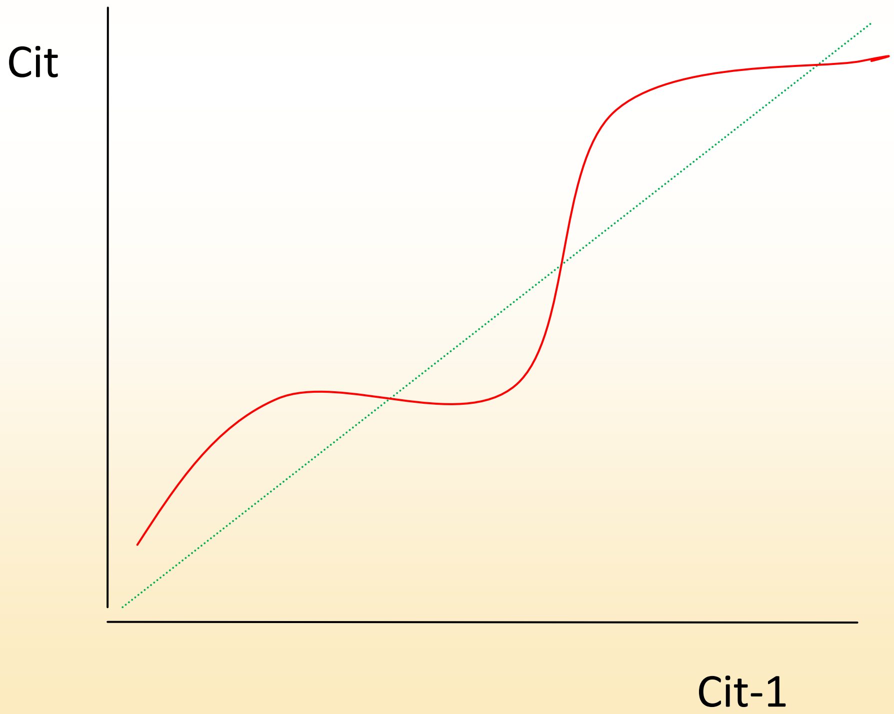
  - Full insurance would imply idiosyncratic events not statistically different from zero

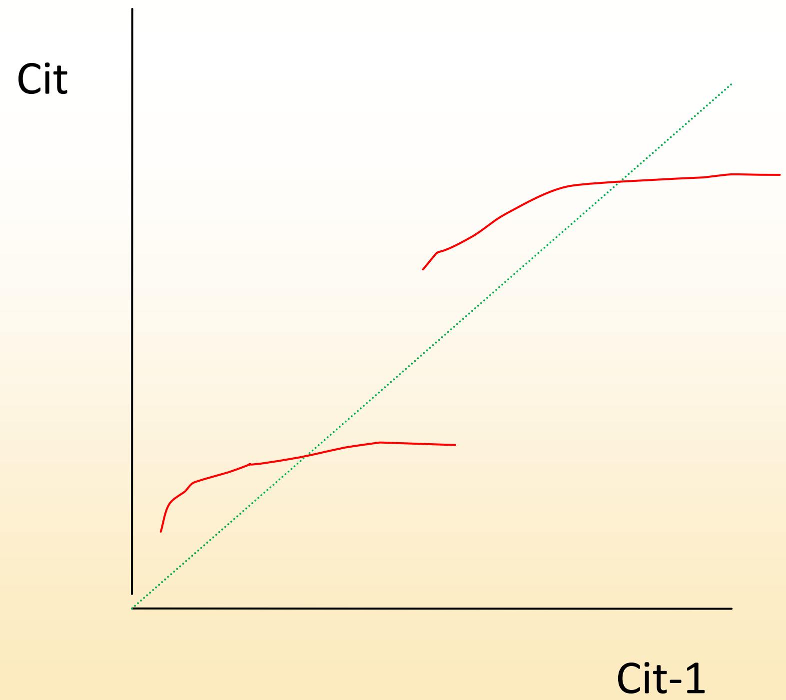
State of the art of the literature

- Micro-growth models:
  - Potential statistical challenges (Dercon and Shapiro, 2006):
    - Lagged income is endogenous in dynamic panel
    - Measurement error generates spurious transitions
    - Individual heterogeneity (e.g. intercept)
    - Non-random sample attrition may overstate persistence
    - Short panel duration prevents observing sufficient movements

State of the art of the literature

- Micro growth models and stochastic wealth dynamics
  - Conditional convergence or poverty traps?





State of the art of the literature

- Stochastic wealth dynamics
  - Similar to micro-growth models
  - Poverty traps versus conditional convergence
  - Roles of risk, ability, etc.
  - Related to the asset-based approach to poverty (Carter and Barret, 2007).
  - Several applications to rural populations. Example: Role of tragedy of the commons among rural populations
  - Examples: Lybbert et. al. (2004); Carter and Santons (2006); Dercon and Shapiro (2006); Antman and McKenzie (2007)

State of the art of the literature

- Stochastic wealth dynamics:
  - Examples of techniques:
    - Fitting polynomials (e.g. Dercon and Shapiro, 2006)

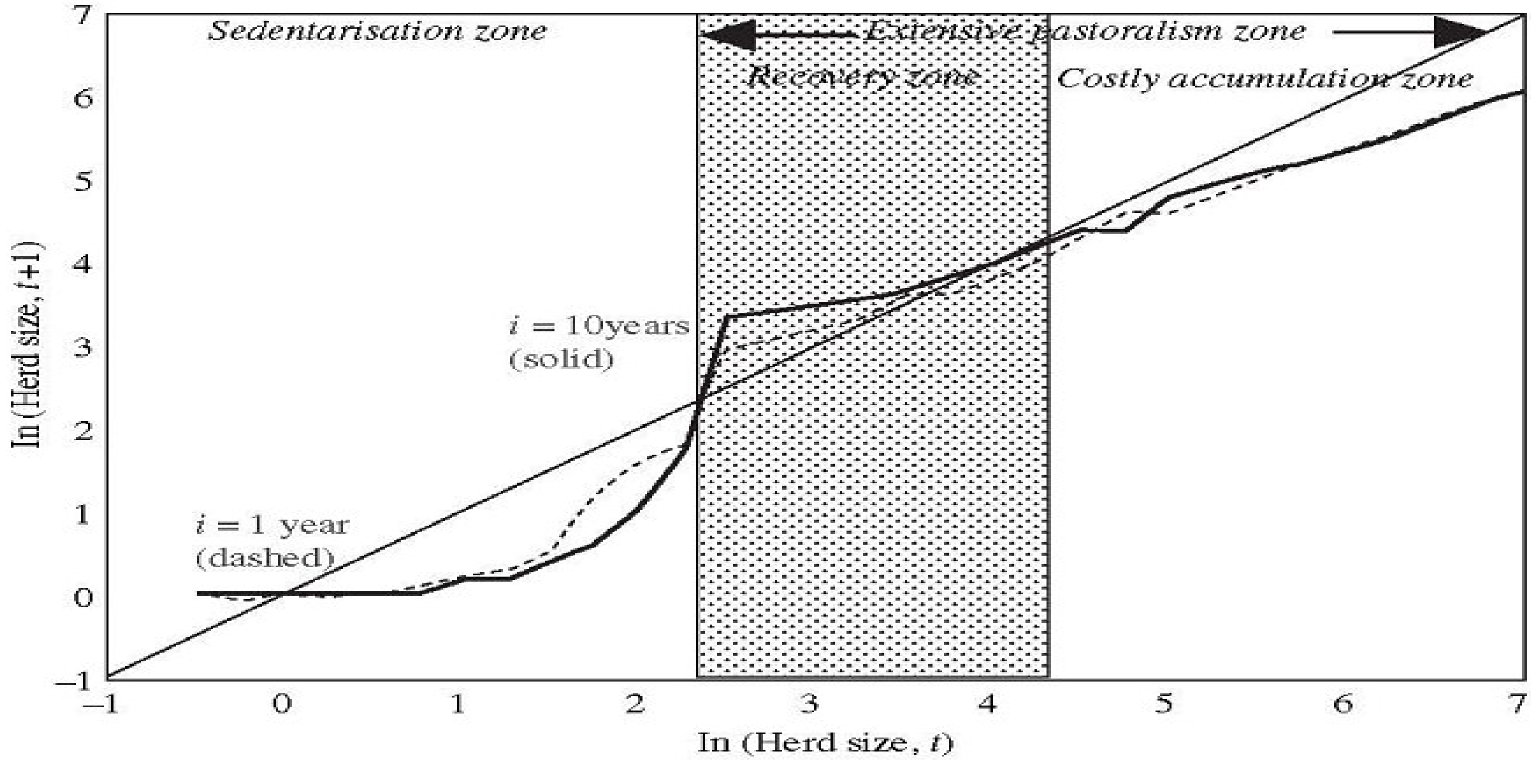
$$C_{it} = \alpha_i + \sum_{j=1}^{P} \beta_j C_{it-1}^{j} + \varepsilon_{it}$$

    - Non-parametric regression (e.g. Lybbert et. al. , 2004). For instance the Nadaraya-Watson kernel regression:

$$\widehat{m}_h(x) = \frac{n^{-1} \sum_{i=1}^{n} K_h(x - X_i) Y_i}{n^{-1} \sum_{i=1}^{n} K_h(x - X_i)}$$

State of the art of the literature

- Stochastic wealth dynamics:
  - Example from Lybbert et. al. (2004):
    - The sedentarization result does not necessarily imply a poverty trap although “there are few nonpastoral options available for stockless, pastoralists, the vast majority of whom are illiterate”



Nadaraya-Watson estimates using Epanechnikov kernel with bandwidth ($h = 1.5$)

State of the art of the literature

- Stochastic wealth dynamics:
  - Testing for a poverty trap with the polynomial model:
    - Step 1: Find the equilibrium points:

$$C^* = \alpha_i + \sum_{j=1}^{P} \beta_j C^{*j}$$

    - Step 2: test whether the derivatives evaluated at those points are higher or lower than 1

$$\frac{dC_{it}}{dC_{it-1}}(C^*) = \sum_{j=1}^{P} j\beta_j C^{*j-1} \leq\geq 1$$

    - A trap should have at least three equilibrium points with one of them being unstable (derivative higher than 1)

State of the art of the literature

- Markov and semi-Markov models
  - They model the probability of being in a welfare state as conditional on having transited a prior welfare trajectory
  - Enable estimation of equilibrium distributions
  - Absolute and relative mobility analysis can be conducted with them using indices (other mobility indices do not require this models though)
  - Discrete time versus continuous time (intensities of transition)

State of the art of the literature

- Markov and semi-markov processes
  - Example: a first-order discrete Markov model (a transition matrix) :
  - Absolute mobility:

p11	p12	p13
p21	p22	P23
1-p11-p21	1-p12-p22	1-p13-p23

  - Relative mobility:

p11	p12	1-p11-p12
p21	p22	1-p21-p22
1-p11-p21	1-p12-p22	p11+p21+p12+p22-1

State of the art of the literature

- Markov and semi-markov processes
  - Example of semi-Markov process: the mover stayer model:
    - A proportion of the population is estimated to be permanently in specific welfare states. The rest behaves according to a transition matrix.
    - Implies distributional traps.
  - Extensions:
    - More heterogeneity (the mover-stayer model implies more heterogeneity than the first-order Markov model)
    - More memory (higher order models)

State of the art of the literature

- Spells approach (discrete duration)
- Continuous duration models
  - Probabilities of entry into and exit from poverty as functions of duration in prior states of being (and other characteristics)
  - Both try to capture “cumulative inertia” or “duration dependence”
  - Attempt to classify individuals according to rate of movement /propensity to transfer to particular states, including duration dependence
  - For more details see Singer and Spilerman (1976)

State of the art of the literature

- Error components models:
  - Persistence modeled as a function of both observable characteristics and residual components (unobserved characteristics)

Examples of modelling techniques

- Spells approach
  - Bane and Ellwood (1986):
    - Interest in heterogeneity of the poor.
    - Most people entering poverty only to have short spells
    - Most people found in poverty experiencing long spells
    - Contribution I: estimation of duration-dependent (discrete) exit (from poverty) probabilities
    - Contribution II: estimation of poverty distributions:
      - Completed spells for people entering a spell of poverty
      - Completed spells for those poor at a given time
      - Uncompleted spells for those poor at a given time
    - Contribution III: Identification of beginning and ending (of poverty) events

Examples of modelling techniques

- Bane and Ellwood (1986):
  - Exit probabilities: they decline as time in the poverty spell increases
  - Why?:
    - Poverty itself may make things more difficult
    - They reflect heterogeneity, as the spell extends transiently poor are selected out leaving the persistently poor
    - Hard to disentangle which one of the two

Examples of modelling techniques

– Bane and Ellwood (1986):

TABLE 1
POVERTY SPELL EXIT PROBABILITIES, SAMPLE SIZES, AND STANDARD ERRORS BY LENGTH OF SPELL TO DATE

Spell Length to Date (years)	Exit Probability	Standard Error	Sample Size
1	0.445	0.010	5,872
2	0.285	0.012	3,220
3	0.246	0.013	2,145
4	0.208	0.016	1,504
5	0.197	0.018	1,096
6	0.145	0.017	759
7	0.128	0.019	516
8	0.074	0.016	334
9	0.083	0.024	223
9–29	0.100*	—	—
30	1.000*	—	—

Source: Probabilities and standard errors are derived from the Panel Study of Income Dynamics and are weighted.

* Value assumed

Examples of modelling techniques

– Bane and Ellwood (1986):

- Calculation of distributions
  - They are all based on exit probabilities (implicit assumption of no re-entry into poverty, later challenged by Stevens, 1999)
  - Exit probability after t years in poverty: p(t)
  - Fraction of people who have spells lasting exactly t years: D(t)
  - Distribution of completed spell duration for those entering poverty (T is maximum length of spell):

$$D(1) = p(1)$$

$$D(t) = p(t)\left[1 - p(t-1)\right]\ldots\left[1 - p(1)\right] = p(t)\left[1 - \sum_{j=1}^{t-1} D(j)\right]$$

$$D(T) = 1 - \sum_{j=1}^{T-1} D(j)$$

Examples of modelling techniques

- Bane and Ellwood (1986)
  - Calculation of distributions:
    - Distribution of completed spells at a point in time (assumes no-growth steady state of spells)
    - F(t) is the fraction of poor people who will have a spell lasting t years:

$$F(t) = \frac{tD(t)}{\sum_{j=1}^{T} jD(j)}$$

    - Distribution of uncompleted spells for poor people at a given time. G(t) is the fraction of people beginning spells t years earlier who would still be poor assuming a steady state:

$$G(t) = \frac{1 - \sum_{j=1}^{t-1} D(j)}{\sum_{j=1}^{T} \left[1 - \sum_{k=1}^{s-1} D(k) \right]}$$

Examples of modelling techniques

– Bane and Ellwood (1986)

TABLE 2
DISTRIBUTIONS OF COMPLETED AND UNCOMPLETED SPELLS OF POVERTY FOR NONELDERLY PERSONS

	Persons Beginning a Spell	Persons Poor at a Given Time*	
Spell Length (Years)	(1) Completed Spell Distribution	(2) Completed Spell Distribution	(3) Uncompleted Spell Distribution
1	44.5	10.6	23.9
2	15.8	7.6	13.3
3	9.8	7.0	9.5
4	6.2	5.9	7.1
5	4.7	5.6	5.7
6	2.8	4.0	4.5
7	2.1	3.5	3.9
8	1.0	2.0	3.4
9	1.1	2.3	3.1
Over 9	12.0	51.5	25.6
Totals	100.0	100.0	100.0
Average	4.2	12.3	6.2

Source: Table derived from exit probabilities reported on Table 1.
* Distributions derived assuming no growth steady state.

Examples of modelling techniques

– Bane and Ellwood (1986)

- Beginning and ending events

TABLE 4

DISTRIBUTION OF BEGINNING TYPES AND MEAN DURATION OF COMPLETED SPELLS FOR PERSONS BEGINNING A SPELL AND FOR PERSONS POOR AT A GIVEN TIME

Beginning Type: Primary Reason for Beginning	Persons Beginning a Spell of Poverty		Persons Poor at a Given Time**	
	Percent of Beginnings	Mean Duration of Completed Spell (years)	Percent of Beginnings	Mean Duration of Completed Spell (years)
Earnings of head fell	37.9	3.3	30.1	9.8
Earnings of wife fell	3.7	3.1	2.8	11.6
Earnings of others fell	7.7	6.5	12.1	15.0
Transfer income fell	8.0	5.2	10.0	14.3
Needs/poverty level rose	8.2	5.3	10.5	13.4
Child* became head or wife	14.7	2.4	8.5	6.9
Wife became female head	4.7	3.7	4.2	11.8
Child of male head became child of female head	6.4	4.0	6.2	10.8
Child was born into poverty	8.6	7.6	15.7	16.9
Total/Average	100.0	4.2	100.0	12.3

For all beginning types exit probabilities are assumed constant at 0.15 after the ninth year.
* Includes child and grandchild and other relative of head.
** Assuming no-growth steady state.

Examples of modelling techniques

– Bane and Ellwood (1986)

TABLE 5

ENDING TYPES BY FAMILY RELATIONSHIP IN LAST YEAR OF POVERTY SPELL

Ending Type: Primary Reason for Ending	All Persons	Members of Families with Children: Male Headed: Heads	Male Headed: Wives	Male Headed: Children	Female Headed: Heads	Female Headed: Children	Married Couples With No Children: Heads	Married Couples With No Children: Wives	Single Heads With No Children: Males	Single Heads With No Children: Females	Other Relative Of Head
Earnings of head rose	50.2	64.4	59.7	56.2	33.0	26.1	63.8	48.7	80.2	48.7	16.3
Earnings of wife rose	7.2	10.6	12.2	11.7	—	—	12.0	13.8	3.8	0.9	3.5
Earnings of others rose	15.8	12.8	12.1	18.6	18.4	26.2	3.1	5.1	2.7	8.5	37.3
Unearned income rose	13.8	7.7	8.6	8.0	19.0	20.1	16.8	24.4	9.0	20.8	25.4
Needs/poverty level fell	2.5	4.2	0.9	0.9	3.3	3.4	4.4	3.6	4.3	2.3	2.9
Female head became wife	4.7	—	5.9	—	26.4	—	—	3.6	—	18.9	—
Child of female head became child of male head	5.4	—	—	4.7	—	23.2	—	—	—	—	10.7
Child* became head or wife	0.4	0.3	0.5	0.0	—	1.1	0.0	0.8	0.0	0.0	3.9
Percent of all endings	100.0	9.5	10.3	28.2	7.9	16.3	4.5	4.8	6.1	9.6	2.7

* Includes child and grandchild and other relative of head.

Examples of modelling techniques

- Stevens (1999):
  - Bane and Ellwood (1986) did not consider re-entry probabilities and experiences of multiple spells of poverty. Extends Bane and Ellwood (1986) to account for re-entry into poverty
  - Develops a discrete-time hazard model with sophisticated handling of observed and unobserved heterogeneity
  - Unobserved heterogeneity is especially introduced to account for the correlation across spells in each individual
  - Deals with censoring of spells
  - Presents a novel version of the error components model
  - Compares predicted distributions of the hazard model with directs tabulations from the data and with error components model

Examples of modelling techniques

– Stevens (1999):

- Basic hazard rate model:
  - Probability of ending a spell of poverty after d years in time t for individual i:

$$\lambda_{idt}^{p} = \frac{\exp(y_{idt})}{1+\exp(y_{idt})}$$

$$y_{idt} = \alpha_{d}^{P} + \beta^{p} X_{it}$$

  - The alphas are the duration effects

Examples of modelling techniques

– Stevens (1999):

- The basic hazard rate model:
  - The probability of observing a complete poverty spell of d years

$$f(d)=\left[\prod_{s=1}^{d-1}\frac{1}{1+\exp(y_{ist})}\right]\left[\frac{\exp(y_{ist})}{1+\exp(y_{ist})}\right]$$

  - Probability of observing a right-censored spell of poverty for d years:

$$1-F(d)=\prod_{s=1}^{d}\frac{1}{1+\exp(y_{ist})}$$

Examples of modelling techniques

– Stevens (1999):

- The basic hazard rate model:
  - Probability of (re-)entering poverty after d years:

$$\lambda_{idt}^{n} = \frac{\exp\left(z_{idt}\right)}{1+\exp\left(z_{idt}\right)}$$

$$z_{idt} = \alpha_{d}^{n} + \beta^{n} X_{it}$$

  - Again, the alphas are duration effects
  - Probability of observing a spell out of poverty for d years:

$$g\left(d\right) = \left[\prod_{s=1}^{d-1} \frac{1}{1+\exp\left(z_{ist}\right)}\right]\left[\frac{\exp\left(z_{ist}\right)}{1+\exp\left(z_{ist}\right)}\right]$$

Examples of modelling techniques

– Stevens (1999)

- The basic hazard rate model:
  - Probability of observing a right-censored spell out of poverty for d years:

$$1 - G(d) = \prod_{s=1}^{d} \frac{1}{1 + \exp(z_{ist})}$$

- Treatment of unobserved heterogeneity:
  - Every individual has a pair of constant types such that:

$$\underline{\theta}_i = \begin{pmatrix} \theta_j^p \\ \theta_l^n \end{pmatrix}$$

$$y_{idt} = \theta_i^p + \alpha_d^p + \beta^p X_{it}$$

$$z_{idt} = \theta_i^n + \alpha_d^n + \beta^n X_{it}$$

Examples of modelling techniques

– Stevens (1999):

- Treatment of unobserved heterogeneity (continued)

	θ_1^p	θ_2^p
θ_1^n	R_1	R_2
θ_2^n	R_3	R_4

$$\sum_{i=1}^{4} R_i = 1$$

Examples of modelling techniques

– Stevens (1999):

- Likelihood function for every individual:

$$L_i\left(\underline{\theta}_k\right)=\left[\prod_{s=1}^{n1i} f_{is}\left(d;\underline{\theta}_k\right)\right]\left[\prod_{s=1}^{m1i} g_{is}\left(d;\underline{\theta}_k\right)\right]\left[1-F_i\left(d;\underline{\theta}_k\right)\right]^{\phi}\left[1-G_i\left(d;\underline{\theta}_k\right)\right]^{1-\phi}$$

$$L_i=\sum_{k=1}^{4} R_k L_i\left(\underline{\theta}_k\right)$$

- The likelihood function is maximized with respect to the alphas, the betas, the Rs and the thetas

Examples of modelling techniques

– Stevens (1999):

Table 5

Coefficients from Maximum Likelihood Estimation of Hazard Rates with Unobserved Heterogeneity

	Blacks		Whites	
	Spells Out of Poverty	Spells in Poverty	Spells Out of poverty	Spells in Poverty
Duration (years)				
1	−1.15	−0.33	1.80	−1.08
	(0.02)	(0.03)	(0.25)	(0.18)
2	−1.56	−0.64	1.17	−1.78
	(0.17)	(0.06)	(0.22)	(0.19)
3	−2.12	−1.17	0.97	−2.16
	(0.12)	(0.31)	(0.26)	(0.21)
4	−2.32	−1.36	0.77	−2.24
	(0.03)	(0.07)	(0.23)	(0.19)
5	−2.24	−1.49	0.56	−2.36
	(0.08)	(0.16)	(0.28)	(0.25)
6 or more	−2.75	−1.78	0.20	−2.62
	(0.05)	(0.06)	(0.23)	(0.22)
Age				
<6	0.72	−0.51	0.43	−0.36
	(0.03)	(0.02)	(0.13)	(0.10)
6–17	0.36	−0.27	0.36	−0.14
	(0.03)	(0.02)	(0.06)	(0.06)
18–24	−0.09	0.25	0.10	0.20
	(0.01)	(0.05)	(0.07)	(0.05)
>54	−0.27	−0.14	0.28	−0.40
	(0.14)	(0.01)	(0.08)	(0.08)
Female head	0.88	−0.85	0.68	−0.74
	(0.06)	(0.10)	(0.08)	(0.10)
High school or more	−1.03	0.27	−0.26	0.32
	(0.16)	(0.26)	(0.09)	(0.06)
θ^e_1		0.85		1.66
		(0.10)		(0.20)
θ^s_1	1.80		−3.36	
	(0.41)		(0.21)	
θ^e_2		0		0
θ^s_2	0		0	
$R1$ (θ^e_1, θ^s_1)	0[a]		0.96	
			(0.04)	
$R2$ (θ^e_1, θ^s_2)	0.44		0[a]	
	(0.38)			
$R3$ (θ^e_2, θ^s_1)	0.14		0[a]	
	(0.22)			
$R4$ (θ^e_2, θ^s_2)	0.42		0.04	
	(0.43)		(0.04)	

Note: Standard errors in parentheses.
a. Probabilities for four support points were initially included in the model. The indicated probabilities, however, were estimated to be zero. The estimation was then performed restricting the probabilities of these "types" to be equal to zero.

Examples of modelling techniques

- Error components models
  - Lillard and Willis (1978):
    - Modelled individual earnings dynamics as functions of observable characteristics, time effects and an error term made of individual random effect component and an autocorrelated component

$$Y_{it} = X_{it}\beta + \Gamma_t + \mu_{it}$$

$$\mu_{it} = \delta_i + \nu_{it}$$

$$\nu_{it} = \gamma\nu_{it-1} + \eta_{it}$$

$$\delta_i \sim \left(0, \sigma_\delta^2\right)$$

$$\eta_{it} \sim \left(0, \sigma_\eta^2\right)$$

$$E\left(\delta_i\eta_{it}\right) = E\left(\eta_{it} \middle| X_{it}, \Gamma_t\right) = E\left(\delta_i \middle| X_{it}, \Gamma_t\right) = 0$$

Examples of modelling techniques

– Lillard and Willis (1978):

- The coefficients of the observable characteristics and the time dummies are estimated via OLS
- The parameters of the error components are estimated using maximum likelihood and the residuals from OLS noting that:

$$E\left(\mu_{it},\mu_{j\tau}\right)=\begin{cases}\sigma_\delta^2+\sigma_\nu^2=\sigma_\mu^2 & i=j\wedge t=\tau\\ \sigma_\delta^2+\gamma^s\sigma_\nu^2=\sigma_\mu^2\left[\rho+\left(1-\rho\right)\gamma^s\right] & i=j\wedge\left|t-\tau\right|=s>0\\ 0 & i\neq j\end{cases}$$

$$\sigma_\nu^2=\frac{\sigma_\eta^2}{1-\gamma^2}\quad,\quad\rho=\frac{\sigma_\delta^2}{\sigma_\mu^2}$$

Examples of modelling techniques

– Lillard and Willis (1978):

TABLE I

COMPONENTS OF VARIANCE, AUTOCORRELATED INDIVIDUAL COMPONENT MODEL

	No. Obs.	$\hat{\sigma}^2_\mu$	$\hat{\sigma}^2_\delta$	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_\upsilon$	$\hat{\gamma}$	$\hat{\rho}$
Log earnings							
All	1144	.307	.224	.069	.083	.406	.731
Blacks	103	.369	.299	.059	.070	.395	.811
Whites	1041	.291	.207	.070	.084	.408	.711
Residual-Simple Eqn.[a]							
All	1144	.206	.125	.068	.081	.402	.606
Blacks	103	.219	.146	.060	.073	.419	.667
Whites	1041	.206	.124	.069	.082	.399	.602
Residual-Comp. Eqn.[b]							
All	1144	.153	.072	.071	.081	.350	.471
Blacks	103	.154	.081	.064	.073	.350	.526
Whites	1041	.153	.071	.072	.082	.350	.464

[a] Including only race, years of schooling, experience and experience squared, and time dummies. The regression is in Appendix Table A2.

[b] Including a comprehensive set of explanatory variables. The regression is presented in Appendix Table A2.

Examples of modelling techniques

– Lillard and Willis (1978):
  - With such estimates conditional probabilities of income state with different memories can be estimated assuming normality in the distribution of etha and delta.
  - For first-order models it means that transition matrices can be predicted
  - Individual and aggregate probability estimations can be performed

Examples of modelling techniques

– Lillard and Willis (1978):

- Example 1: Individual transition matrix of in-and-out-of poverty between year t and year tau:

$$Y_{it} \le Y_{it}^{*} \Leftrightarrow \frac{\nu_{it}}{\sigma_{\nu}} \le \frac{\left(Y_{it}^{*} - X_{it}\beta - \Gamma_{t} - \delta_{i}\right)}{\sigma_{\nu}} \equiv b_{it}^{*}$$

$$\phi_{it} = F\left(b_{it}^{*}\right)$$

$$\phi_{it,\tau} = F\left(b_{it}^{*}, b_{i\tau}^{*}; \gamma^{|t-\tau|}\right)$$

$$\phi_{it,-\tau} = F\left(b_{it}^{*}, -b_{i\tau}^{*}; -\gamma^{|t-\tau|}\right)$$

$$etc.$$

$$P_{i-t|-\tau} = \frac{\phi_{i-t,-\tau}}{\phi_{i-\tau}}, \quad P_{it|\tau} = \frac{\phi_{it,\tau}}{\phi_{i\tau}} \rightarrow P_{it|-\tau} = 1 - \frac{\phi_{i-t,-\tau}}{\phi_{i-\tau}}, \quad P_{i-t|\tau} = 1 - \frac{\phi_{it,\tau}}{\phi_{i\tau}}$$

Examples of modelling techniques

– Lillard and Willis (1978):

- Example 2: Aggregate transition matrix of in-and-out-of poverty between year t and year tau for a homogeneous population:

$$Y_{it} \le Y_{it}^{*} \Leftrightarrow \frac{\mu_{it}}{\sigma_{\mu}} \le \frac{\left(Y_{it}^{*} - X_{it}\beta - \Gamma_{t}\right)}{\sigma_{\mu}} \equiv b_{it}$$

$$\phi_{t} = F\left(b_{it}\right)$$

$$\phi_{t,\tau} = F\left(b_{it}, b_{i\tau}; \rho + (1-\rho)\gamma^{|t-\tau|}\right)$$

$$\phi_{t,-\tau} = F\left(b_{it}, -b_{i\tau}; -\left(\rho + (1-\rho)\gamma^{|t-\tau|}\right)\right)$$

$$etc.$$

$$P_{-t|-\tau} = \frac{\phi_{-t,-\tau}}{\phi_{-\tau}}, \quad P_{t|\tau} = \frac{\phi_{t,\tau}}{\phi_{\tau}} \rightarrow P_{t|-\tau} = 1 - \frac{\phi_{-t,-\tau}}{\phi_{-\tau}}, \quad P_{-t|\tau} = 1 - \frac{\phi_{t,\tau}}{\phi_{\tau}}$$

Examples of modelling techniques

– Lillard and Willis (1978):

TABLE IV

AGREGATE POVERTY PROBABILITIES IN 1969 CONDITIONAL ON ALTERNATIVE POVERTY HISTORIES BY RACE

Probability Concept[a]	Whites Predicted	Whites Actual	Blacks Predicted	Blacks Actual
(a) Unconditional				
$P(69)$	.024	.026	.113	.097
(b) Conditional on past year				
$P(69 \mid 68)$	.469	.370	.652	.600
$P(69 \mid \sim 68)$	.013	.017	.037	.043
(c) Conditional on past two years				
$P(69 \mid 67, 68)$	.589	.600	.737	.625
$P(69 \mid \sim 67, 68)$	.346	.235	.423	.500
$P(69 \mid 67, \sim 68)$	.179	.211	.250	.200
$P(69 \mid \sim 67, \sim 68)$	.010	.013	.025	.034

[a] ~ indicates not poverty. For example, $P(69 \mid 67, \sim 68)$ is the probability of poverty in 1969 given poverty in 1967 and not poverty in 1968.

Examples of modelling techniques

– Lillard and Willis (1978):

TABLE V

PREDICTED AND ACTUAL k-STEP POVERTY TRANSITION PROBABILITIES FOR WHITES AND BLACKS

	Year (τ)					
	1968	1969	1970	1971	1972	1973
A. $p_{\tau\|t}$ = Pr (poverty in τ|poverty in 1967)						
Whites						
Predicted	.458	.367	.331	.303	.290	.276
Actual	.345	.345	.276	.241	.241	.276
Blacks						
Predicted	.652	.568	.468	.503	.438	.436
Actual	.615	.462	.385	.231	.385	.462
B. $p_{\tau\|\sim t}$ = Pr (poverty in τ|not poverty in 1967)						
Whites						
Predicted	.013	.015	.015	.013	.013	.012
Actual	.017	.017	.022	.031	.023	.022
Blacks						
Predicted	.039	.041	.029	.036	.026	.026
Actual	.022	.044	.011	.044	.033	.056

Examples of modelling techniques

- Error components models:
  - Stevens (1999):
    - Estimates the error components with minimum distance estimators, not for earnings but for earnings-to-needs ratios
    - Estimates alternative models:
      - Allows for variation in variance of etha
      - Makes the individual random effect dependent on age:

$$\mu_{it} = \delta_i + \lambda_i age_{it} + \nu_{it}$$

      - The autocorrelated error component follows an ARMA (1,1) process:

$$\nu_{it} = \gamma \nu_{it-1} + \eta_{it} + \theta \eta_{it-1}$$

    - Compares the hazard rate model (spells approach) with the error components model (using the ARMA (1,1) specification with constant variance of etha and no dependence on age for simulations

Examples of modelling techniques

– Stevens (1999):

Table 8
Comparison of Hazard Rate Estimates, Variance Components Estimates, and Direct Tabulations

	Male Head of Household			Female Head of Household		
Years Poor	Hazard Rate	Variance Components	Direct Tabulations	Hazard Rate	Variance Components	Direct Tabulations[a]
			Blacks			
1	17.3	20.6	18.5	3.3	14.8	8.7
2	16.0	14.8	23.8	4.7	11.9	4.0
3	13.5	12.3	10.1	5.8	11.0	10.0
4	11.5	10.6	8.1	7.0	10.5	4.4
5	9.2	9.1	5.2	8.5	9.5	2.9
6	8.1	8.0	14.3	9.6	8.9	6.3
7	7.1	7.0	1.7	10.5	8.3	11.3
8	6.3	6.1	5.5	12.6	7.7	8.8
9	4.5	5.5	9.0	14.7	7.6	15.9
10	6.5	6.0	3.8	23.2	9.8	27.9
Mean	4.4 years	4.30 years	4.2 years	7.0 years	5.0 years	6.9 years
N			307			146

Examples of modelling techniques

– Stevens (1999):

	Whites					
1	32.9	24.1	39.5	13.3	17.7	20.4
2	21.4	16.2	16.2	11.6	13.7	18.5
3	13.8	12.9	9.3	10.6	12.1	14.3
4	9.8	10.7	14.0	10.4	10.5	4.6
5	6.6	8.9	1.5	9.8	9.5	7.8
6	4.4	7.3	6.6	8.5	8.4	4.6
7	3.2	6.1	2.9	7.9	7.7	5.6
8	2.5	5.3	2.5	7.8	6.8	9.1
9	2.5	4.2	5.4	7.8	6.4	12.2
10	2.9	4.5	2.1	12.3	7.1	2.8
Mean	3.1 years	4.0 years	3.1 years	5.2 years	4.6 years	4.3 years
N			217			71

Note: Distributions shown are simulated for individuals beginning a stay in poverty at age 30, and where the household head has less than a high school education.

a. For this category only, tabulations are based on individuals spending six or more of the ten years in a female-headed household. All other columns hold sex of the head constant over the full ten years.

Examples of modelling techniques

- Stevens (1999):
  - "The hazard model reproduces observed patterns of poverty persistence somewhat better than the variance components model"
  - Why? The error components model estimates parameters of the full distribution of the variable while the hazard rate model is based on poor individuals
  - In other words: there might be heterogeneity in income dynamics across different parts of the distribution!