The Econometrics of Income and Social Mobility

The Dynamics of Poverty

Michel Lubrano

January 2019

Contents

1 Introduction

In the previous chapter, we alluded to two polar cases for comparing distribution:

1. Using a Pareto criterion where possibly only the rich get an improvement,

2. Using a Rawlsian criterion which looks at the fate of the poor, irrespective to what happens to the others.

This opposition motivates a particular interest to consider the fate of the poor from a dynamic point of view. There is a vast literature on this topic, starting roughly with the work of Lillard and Willis (1978) which develops a dynamic econometric model explaining the income to needs ratio. Bane and Ellwood (1986) prefer to estimate a distribution free hazard for the length of poverty spells. Rodgers and Rodgers (1993) introduce the possibility of transferring money between two periods and thus introduce the concepts of chronic versus transitory poverty. However long panel are needed in these cases, at least ten years. The approach of Cappellari and Jenkins (2004) is more econometric in a way as it relies on a three equation model for a panel of only two periods, even if they use a longer panel. An equation describes the initial state. A second equation explains attrition between the two periods and the third equation models poverty dynamics, conditional of the initial state. As a result are estimated poverty exit and poverty persistence.

2 The early literature for poverty dynamics

2.1 The covariance structure model of Lillard and Willis (1978)

Household do not stay all the time in poverty. They have poverty spells, they enter into poverty and get out of it. Stevens (1999) got interest in explaining the duration of these poverty spells for the USA. In her paper, she proposes several models. I shall explain here the parametric model of Lillard and Willis (1978). The endogeneous variable is the log of the income-to-needs ratio that is going to be explained by exogeneous variables, but also by the dynamics of the error process. The model is then used to make judgement about the persistence of poverty spells in the USA in order to evaluate the economic situation of an household. The income-to-needs ratio is computed by considering the household income which does not include transfers and by dividing it by the official poverty rate corresponding to the household composition. The basic model is as follows:

$$\log\left(\frac{y_{it}}{z}\right) = x_{it}\beta + \delta_i + v_{it} \quad (1)$$

$$\delta_i \sim N(0, \sigma_\delta^2) \quad (2)$$

$$v_{it} = \gamma v_{it-1} + \eta_{it}. \quad (3)$$

The log of the income to needs ratio is explained by individual variables that are time independent as sex and education level, and by individual variables that are time varying. There is a random individual effect δ_i for unobserved heterogeneity. Parameter γ models a permanent effect common to all individuals. We can says that the individuals receive permanent shocks v_{it}. Under a normality assumption for δ_i and η_{it}, Stevens (1999) simulates this model for 20 years and compute the mean period spent in a poverty state. When estimating this model using the PSID data set, we find that the average period spent in a state of poverty is slightly longer if the head of the household is black or if it is a woman.

2.2 The hazard model of Bane and Ellwood 1986

With the use of longitudinal data, poverty appears to be very heterogeneous: most of the poor are poor for only a few years, while a minority is persistently poor. The difference between the two is interesting for targeting anti-poverty policies. So it is interesting to study the length of poverty spells and the determinants of poverty entry and exit.

Bane and Ellwood (1986) study the length of poverty spells in a PSID sample of 12 years from 1970 to 1982. A poverty spell is defined as the length of time in which the cash income of a household is below a defined poverty line after some adjustments to cope with the PSID specificities. For a given sample of finite length, we have completed spells for which we observe the starting point and the end point. But also uncompleted spells, which can be censored at the beginning or at the end. It is important to take into account of censoring for estimating the various quantities.

Estimation of *exit probabilities* using all the spells for which the *beginning is observed*. Sample sizes are quite large for estimating exit probabilities for the first years. Exit probabilities up to nine years are calculated.

For spells for which they observe a beginning or and end, they try to identify a particular family event.

The results obtained in this paper are not surprising: entry in poverty is due mainly to a fell in income of the head of the household and exit from poverty occurs when the latter recover his income. Moreover, the econometrics is not clearly explained. It relates to survival analysis. There is by the way a package `survival` in `R`. Finally, the definition of poverty is rather restrictive. The model recognizes only one type of characteristics, and not for instance the severity of poverty.

3 Chronic and transitory poverty in Europe

The drawback of the two previous papers is that they recognize only one dimension of poverty, using either a distribution-free approach or a parametric model. We shall first present three different dimension of poverty before distinguishing between chronic and transitory poverty. But there will be no explanatory variables in our model.

3.1 The three I's of poverty

Before considering dynamics, it is useful to go back to the FGT indices of Foster et al. (1984) because they allow to distinguish several aspects of poverty. The general formula is, supposing that the income variable $x_{[i]}$ is ordered and that z is the poverty line:

$$P_\alpha(x,z) = \frac{1}{n}\sum_{i=1}^{q}\left(\frac{z-x_{[i]}}{z}\right)^\alpha, \qquad P_\alpha(x,z) = \int_0^z \left(\frac{z-x_{[i]}}{z}\right)^\alpha f(x)\,dx$$

where q is the number of individuals below the poverty line. Depending on the value of α, we have several possible measures

- For $\alpha = 0$, we have poverty incidence, or in other terms the head-count ratio.
- For $\alpha = 1$, we have poverty intensity, which measures the average poverty gap, the average distance to the poverty line.
- For $\alpha = 2$, we have poverty inequality, which measures the distribution of the poor below the poverty line. Are individuals concentrated at certain locations below the poverty line or are they evenly distributed?

Reporting these different indices is very informative. For instance, Thuysbaert (2008) report those indices for Belgium using as a poverty line 50% of the mean contemporary income. We see that poverty incidence has dropped from

Table 1: FGT measures of poverty for Belgium

Indices	1976	1985	1988	1992	1997
P_0	0.0786 (0.0035)	0.0505 (0.0028)	0.0515 (0.0042)	0.0606 (0.0047)	0.0729 (0.0044)
$P_1 \times 10$	0.157 (0.009)	0.087 (0.007)	0.076 (0.008)	0.125 (0.018)	0.117 (0.009)
$P_2 \times 100$	0.54 (0.29)	0.29 (0.19)	0.19 (0.003)	0.62 (0.16)	0.34 (0.05)

8% in 1976 to 5% in 1985 and then progressively risen to reach 7% in 1997. The progression of poverty intensity does not follow the same pattern as well as poverty inequality.

How could we summarize these figures? Jenkins and Lambert (1997) propose to summarize these three aspects of poverty, namely incidence, intensity and inequality into a cumulative curve of poverty gaps which has a number of nice properties. This curve is named the TIP curve because it means the three I's of poverty. Let us consider an income distribution $F(x)$ and a poverty line z. The TIP curve is a function of p, a proportion of individuals below the poverty line when the income distribution is $F(.)$:

$$TIP(p,z) = \int_0^{F^{-1}(p)} \left(\frac{z-x}{z}\right) \mathbb{1}(x \leq z) dF(x).$$

where $\mathbb{1}(.)$ is the indicator function. The TIP curve is estimated by considering the ordered incomes, $x_{[i]}$ and

$$TIP(k/n,z) = \frac{1}{n}\sum_{i=1}^{k} \left(\frac{z-x_{[i]}}{z}\right) \mathbb{1}(x_{[i]} \leq z).$$

It becomes very easy to plot these graphs, plotting the sequence k/n against the corresponding values $TIP(k/n,z)$. The point on the horizontal axis where the curve becomes flat and horizontal is the head-count ratio $P_0(z)$ or incidence point. The corresponding ordinate on the vertical axis is equal to $P_1(z)$. Finally, the curvature of the TIP curve reflects inequality of income distribution among the poor, offering a symmetric view of a Lorenz curve.

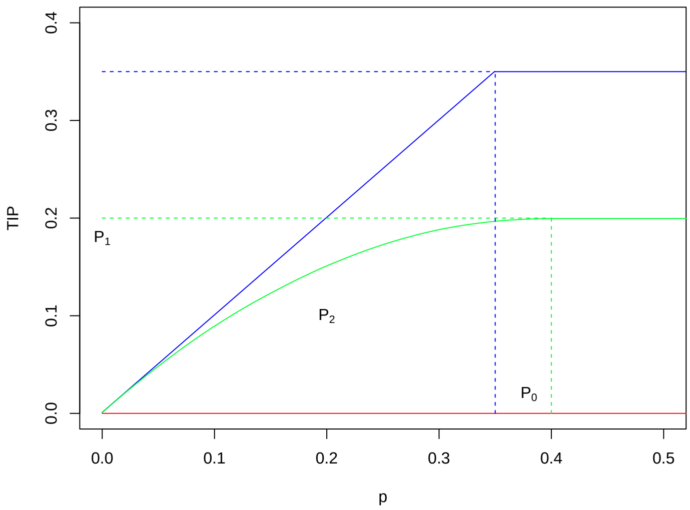

Figure 1: TIP curves from different income distributions

Figure 1 is interesting for understanding the functioning of TIP curves. The **blue curve** corresponds to a situation where **all the poor have a zero income**. This is a situation of maximum poverty. It corresponds to a straight line where poverty incidence is equal to poverty intensity ($P_0 = P_1$). The **green curve** is an intermediate situation. It was built using a uniform distribution of income. Note however that if poverty inequality and poverty intensity are lower, we have a higher poverty incidence. P_0 was chosen greater that in the previous case. So even if that green curve is lower than the blue curve, we have a better situation only for two poverty criterion, not for three. The **red curve** corresponds to a situation where **all the poor have an income equal to the poverty line** z. The TIP curve is horizontal.

TIP curves are closely related to the Generalised Lorenz curve. Decomposing the integral in two, we have, for the un-normalised version of the TIP curve:

$$TIP(p,z) = z\int_0^{F^{-1}(p)} 1\!\!\mathrm{I}(x \leq z) f(x)dx - \int_0^{F^{-1}(p)} x 1\!\!\mathrm{I}(x \leq z) f(x)dx$$

The first integral is just equal to p, so that its value is $z\,p$. The second integral is the definition of the Generalised Lorenz Curve (it is not divided by the mean) times the indicator function $1\!\!\mathrm{I}(x \leq z)$. So the final result is:

$$TIP(p,z) = zp - GLC(p).$$

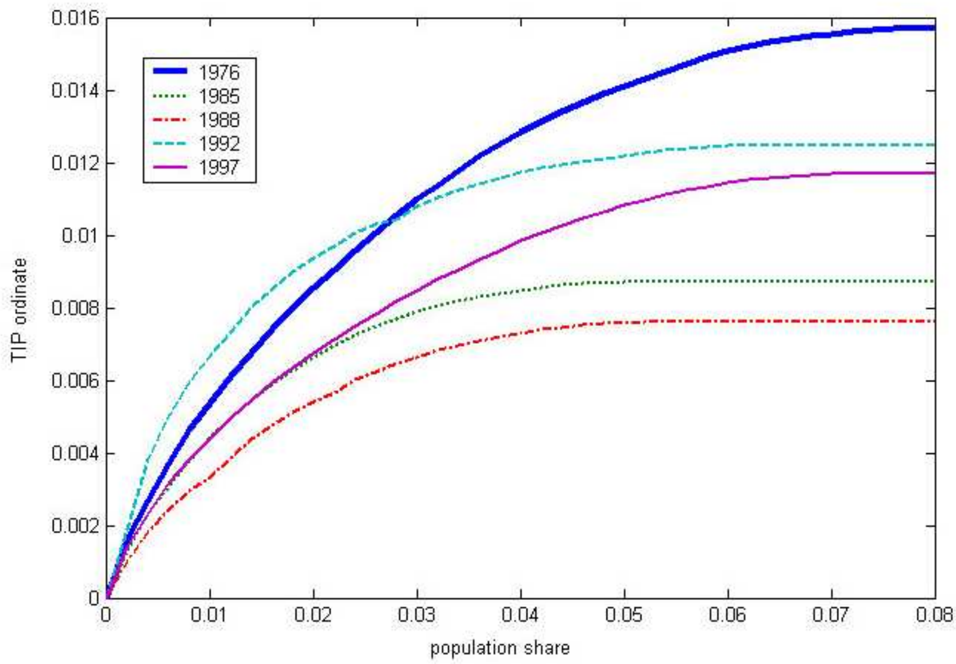

Figure 2: TIP curves for Belgium

This equivalence is valid only for the lower part of the Lorenz curve, till $p = q/n$ the headcount ratio. The normalised version is obtained by dividing by z. It follows from this partial equivalence that we can rank distributions using the TIP curve in a similar way as can be done using Generalised Lorenz curves, which means second order stochastic dominance. This open the way to a ranking distributions, with a focus on poverty. Jenkins and Lambert (1997) define TIP dominance as follows:

Definition 1 *Let us consider two income distributions A and B with a common poverty line z. Let us call TIP_A and TIP_B their associated TIP curves. Distribution A TIP dominates distribution B if*

$$TIP_A(p, z) \geq TIP_B(p, z), \quad \forall p \in [0, 1].$$

Which means here that there is more poverty in A than in B counter intuitive definition). These notions are further developed in Jenkins and Lambert (1998a,b). When the curves intersect, of course we cannot compare them. This is the case for 1976 and 1992. Poverty intensity is lower on average in 1992. However, poverty intensity is greater in 1992 for $p \in [0, 0.028]$. For the poorest of the poor, the situation has deteriorated.

3.2 Two approaches for poverty dynamics

Income is a random variable, the level of which is determined by the level of human capital, to say it short (Mincer's equation). Consequently, a cross-section vision of poverty can be false for a given individual, if it corresponds to a temporary state of unemployment. It is in general useful to distinguish between **permanent income** and **transitory income** to deliver a more realistic vision of poverty or to allow for income transfers between the years (savings).

Kuchler and Goebel (2003) distinguish two possible approach to characterize poverty and they try to combine them in their paper.

1. A first approach, they call the $N - Times - Poor$ or NTP, counts the number of times a person is in a state of poverty over the duration of the panel survey. If that person is only occasionally poor, he will belong to the transitory poor category. If he remains poor over the whole sample, he will belong to the persistently poor category. The poverty line is defined for every year. One can count the number of poverty spells. However, as poverty spells are censored (uncompleted spells), this counting method leads to biased results. A better way is to model the length of poverty spells and compute the probability of exit as in the very often quoted paper of Bane and Ellwood (1986). The status of a person is determined only by considering a static definition of poverty. This is just a kind of generalisation of the cross section approach. And no account is taken of the intensity of poverty.

2. A second approach supposes that the individual is able to smooth his income over the period of the panel survey in order to define a kind of permanent income. This approach was advocated by Rodgers and Rodgers (1993). This second approach is less frequent and is illustrated for instance in Hill and Jenkins (2001). This is the smooth income poverty approach (SIP). One of its advantages is that it is no longer on just a counting approach, but allows for poverty intensity and inequality.

3.3 Smoothed income poverty

Rodgers and Rodgers (1993) give the following example which is quite illuminating for understanding the difference between chronic and transient poverty.

Example 1 *Let us suppose that the poverty line is* $z = 100$.

Person	*Year 1*	*Year 2*	*Year 3*	*Year 4*	*Year 5*	*Year 6*
A	*300*	*300*	*300*	*99*	*99*	*300*
B	*101*	*101*	*101*	*10*	*10*	*101*
C	*300*	*300*	*99*	*99*	*99*	*300*

If we measure the duration of poverty spells, person C has the longest poverty spell compared to A and B. But if we suppose that a person can smooth his income over the years, spare money and transfer money to the next period, then clearly person B is in a state of chronic poverty.

We shall consider a linear decomposable poverty index and a period of length T. The most common choice is the FGT index. We first compute that index for every period and note it P_t for a given poverty line z_t. Then we have the following definitions of average, chronic and transitory poverty rates:

1. *Average annual poverty rate.* It is defined as a weighted sum of annual poverty measures. Most of the time the weights w_t are equal to $1/T$. This average is possible because the index P is linearly decomposable:

$$A_P(T) = \sum_{t=1}^{T} w_t P_t.$$

   There is no possible inter year income transfers. This is in a way the maximum poverty rate as said in Hill and Jenkins (2001) (verify).

2. *Chronic poverty.* We now assume that it is possible to transfer income between the years. We call Y_i^* the permanent income of person i. There are n individuals in the sample. So chronic poverty is defined as a poverty index applied to the series of the n smoothed incomes:

$$C_P(T) = P(Y_1^*, \cdots, Y_n^*).$$

   Rodgers and Rodgers (1993) have a complicated way of computing the permanent income. Hill and Jenkins (2001) and Kuchler and Goebel (2003) use a much simpler formula.

3. *Transitory poverty.* As the poverty index P is supposed to be linearly decomposable, transitory poverty can be found using a difference

$$T_P(T) = A_P(T) - C_P(T).$$

   A positive $T_P(T)$ represents the amount of poverty which is not chronic for an average year. Negative values are possible according to Rodgers and Rodgers (1993), depending on the chosen P and the way permanent income is computed.

When using the FGT index, we measure average or total poverty and chronic poverty:

$$A_{FGT}(T) = \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{q_t}(1-y_{it}/z)^{\alpha} \quad (4)$$

$$C_{FGT}(T) = \frac{1}{n}\sum_{i=1}^{q_{Y*}}(1-Y_i^*/z)^{\alpha} \quad (5)$$

Using the PSID over 1977-1986 (10 years) and P_0, Rodgers and Rodgers (1993) found that with the permanent income approach that:

1. Total poverty $A = 9.40\%$,
2. Chronic poverty $C = 6.25\%$,
3. Transitory poverty $T = 3.15\%$.

Using the tabulation approach, the usual definition of chronic poverty is that an individual must be in poverty for 8 years or more out of 10. In this case 3.8% of individuals were in this case when those who were at least once poor were 26.6% which means that chronic poverty represented 14.3% of the once ever poor. *With the permanent income approach, the proportion of chronic poverty is much higher with $C/A = 66\%$.* In order to get comparable figures between the two approaches, one has to define chronic poverty has being poor in at least 6 years out of 10.

The great advantage of the SIP approach is that it can be based on P_α for any value of α, when the tabulating or NTP approach relies only on P_0 (head-count). Using P_1, Rodgers and Rodgers (1993) found that the proportion of chronic poverty over total poverty converges to around 37% when the period for computing the smoothed income is increased up to 10 years.

3.4 Poverty dynamics in Europe

Kuchler and Goebel (2003) start from the relative income position of individual i in the sample of size n at time t which is

$$y_{it}^r = \frac{y_{it}}{\bar{y}_t}, \quad \bar{y}_t = \frac{1}{n}\sum_{i=1}^{n} y_{it}.$$

Dividing by the sample mean allows to avoid having to divide by a price index. Incomes are made comparable using the modified OECD scale. Using these data, it is possible to compute average annual poverty, the index

$A_{FGT}(T)$ of Rodgers and Rodgers (1993), using the data displayed in Table 1 of Kuchler and Goebel (2003). The poverty line is 50% of the average y_{it}^r. This is called **total poverty** in Hill and Jenkins (2001). So Table 2 repre-

Table 2: Average annual income poverty: 1994-1997
Total poverty

Country	Incidence P_0	Intensity $P_1 \times 10$	Inequality $P_2 \times 100$
Denmark	5.63	1.28	0.54
Netherlands	9.98	3.83	2.43
Germany	13.85	4.98	3.01
France	14.68	3.83	1.84
Italy	16.95	6.40	3.96
Belgium	16.45	4.85	2.55
UK	18.25	6.73	4.08
Spain	19.10	6.58	3.68
Ireland	20.10	3.95	1.47
Greece	21.38	7.63	4.04
Portugal	24.53	8.43	4.55

Source: Kuchler and Goebel (2003), Table 1 and own calculations.

sents total maximum poverty. Quite different pictures of total poverty are obtained when considering incidence or intensity. Denmark and the Netherlands have the smallest poverty incidence. At the other extreme, Spain, Ireland Greece and Portugal have the highest poverty incidence. However, when considering the intensity of poverty, Denmark and the Netherlands remain in the group where intensity is the smallest, but they are joined by France and by Ireland. Greece and Portugal remain in the group where poverty is highest. But they are joined by Italy, the UK and Spain.

Let us now turn to **chronic poverty** which aims at measuring poverty when we allow for inter-temporal income transfers. Smoothed or permanent income can be computed in different ways. Rodgers and Rodgers (1993) adopt a complicated mechanism based on borrowing and lending which might lead to apparent incoherencies (negative transitory poverty). The later literature adopted some kind of smoothing. We could imagine exponential smoothing, non-parametric smoothing following the time series literature where a topic is the decomposition of a time series in permanent and cyclical components. Kuchler and Goebel (2003) adopt the simplest way to define permanent income, using in fact just the mean income, resulting in a single value for each individual. So the time dimension is compressed. In a panel of size T, the smoothed relative income position of individual i is:

$$\bar{y}_i^r = \frac{1}{T} \sum_{t=1}^{T} y_{it}^r.$$

The poverty line will be defined now as 50% of the mean smoothed relative income position. It results the following picture of chronic poverty as depicted in Table 4. Chronic income poverty is a minor phenomenon in Denmark, and

Table 3: Smoothed or chronic income poverty: 1994-1997

Country	Incidence P_0	Intensity $P_1 \times 10$	Inequality $P_2 \times 100$
Denmark	2.4	0.2	0.04
Netherlands	6.1	1.0	0.38
Germany	8.2	2.0	0.91
France	13.8	2.2	0.55
Ireland	17.1	2.3	0.48
UK	13.5	2.8	1.03
Belgium	13.1	3.1	1.28
Italy	12.4	3.2	1.38
Spain	14.8	3.6	1.43
Greece	17.5	4.6	1.77
Portugal	21.6	6.7	3.11

Source: Kuchler and Goebel (2003), Table 2.

also in the Netherlands; while Portugal and Greece are at the other extreme. Between total and chronic poverty, the ranking does not change, except for France which has a higher chronic poverty as measured by P_0.

It is interesting to analyse and compare three countries which can look similar: France, Germany and the UK. For Germany, we are well before 2003, the time when Gerhard Schröder launched his cuts in the social welfare system. And for the UK, we are well after the Thatcher's period.

1. For total poverty, the UK is well above Germany and France while France has the lowest intensity and inequality.

2. For chronic poverty, France and the UK have very similar incidence, well above that of Germany.

3. For chronic poverty **severity and inequality**, the UK is in the least favourable position, while France and Germany become comparable.

Kuchler and Goebel (2003) have chosen to classify these 11 countries in reference to a welfare regime typology: a liberal welfare state (the UK, Ireland) together with the Mediterranean countries (Italy, Spain, Greece, Portugal) on one side and what they call the corporatist-conservative welfare regime (Germany, France, Belgium, the Netherlands), together with nordic countries, here Denmark . Using the TIP curves reported in Figures 3 and 4, these countries can be ranked. The vertical lines intersecting with the x axis represents chronic poverty incidence while the horizontal lines intersection with the y axis represents chronic poverty intensity. The curvature represents chronic poverty inequality. However, these curves cannot be ranked in term of poverty dominance because the poverty lines are not the same.

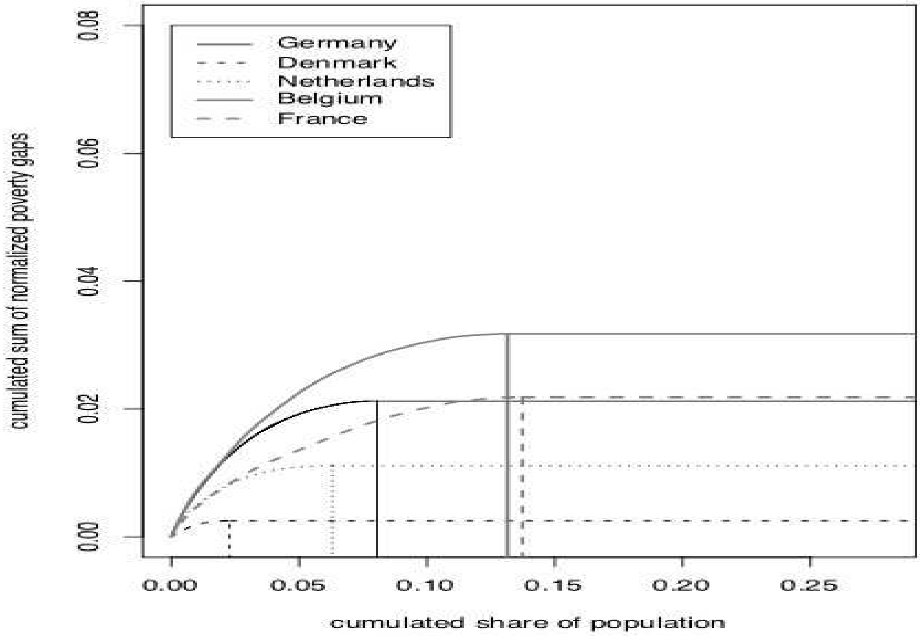

Figure 3: TIP curves for nordic European countries

Smoothed net equivalent disposable household income position, modified OECD equivalence scale, poverty line: 50% of mean of smoothed relative income position. Source: UDB ECHP 2001: Wave 1 (1994) - Wave 4 (1997), balanced panel, weighted. Figure extracted from Kuchler and Goebel (2003).

In Figure 3, Belgium has the highest poverty intensity. At the other extreme, Denmark and the Netherlands have the lowest intensity and incidence. The case of France might be similar to that of Germany for intensity, but is quite different for incidence and severity. The two curves are intersecting for these two countries.

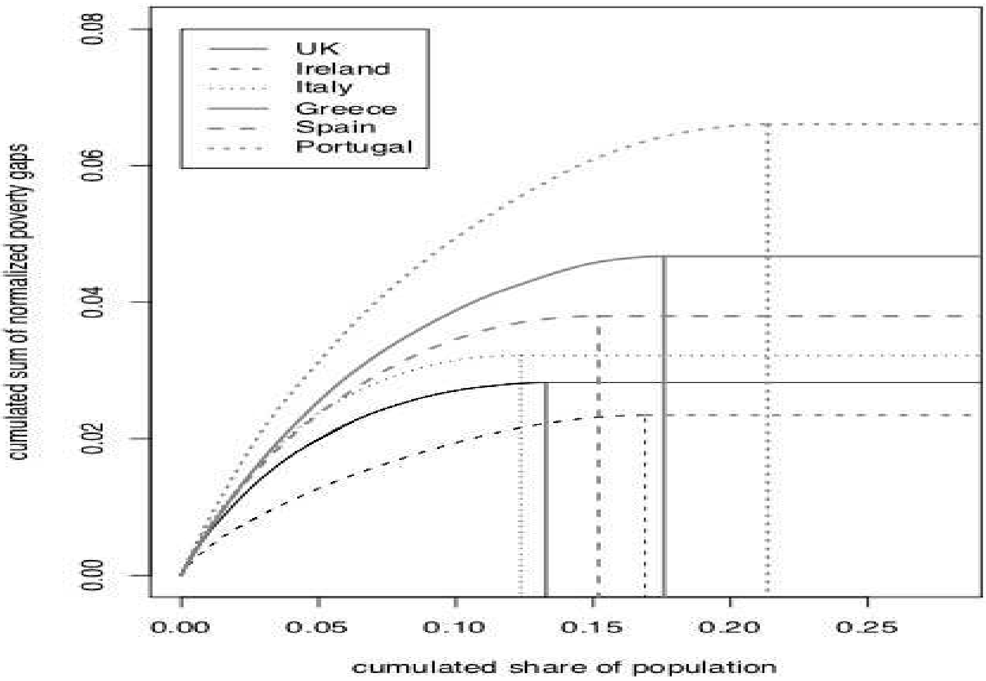

Figure 4: TIP curves for liberal and southern European countries

Smoothed net equivalent disposable household income position, modified OECD equivalence scale, poverty line: 50% of mean of smoothed relative income position. Source: UDB ECHP 2001: Wave 1 (1994) - Wave 4 (1997), balanced panel, weighted. Figure extracted from Kuchler and Goebel (2003).

In Figure 4, Portugal and Greece have the highest chronic poverty intensity and incidence. Even if the TIP curves do not intersect, the other countries can be ranked according to chronic poverty intensity, but have a different ranking for poverty incidence. Ireland has the lowest poverty intensity and inequality, but a higher poverty incidence than Italy, the UK and Spain. So these four countries are difficult to rank. Examining only one indicator is not enough.

By inspecting the proportion of chronic poverty over total poverty, we can have an idea of social mobility among the poor. Here again Denmark and the Netherland have the best position according to intensity. Portugal and Greece are at the bottom, but paradoxically with Belgium. Further analysis would be needed in order to study social mobility.

Table 4: Proportion of chronic poverty over total poverty: 1994-1997

Country	Incidence P_0	Intensity $P_1 \times 10$	Inequality $P_2 \times 100$
Denmark	43.67	15.69	07.37
Netherlands	61.16	26.14	15.62
Germany	59.21	40.20	30.28
UK	73.97	41.64	25.23
Italy	73.16	50.00	34.87
Spain	73.40	54.75	38.89
France	94.04	57.52	29.97
Ireland	85.08	58.23	32.77
Greece	81.87	60.33	43.81
Belgium	79.64	63.92	50.30
Portugal	88.07	79.53	63.32

Source: Kuchler and Goebel (2003), Table 1 and 2 and own calculations.

4 The dynamic model of Cappeliari and Jenkins

Except for the model of Lillard and Willis (1978), we have no explanatory variables. The model we want now to detail is exposed in two different papers: a technical paper with Cappellari and Jenkins (2004) published in the *Journal of Applied Econometrics* and a more applied paper with Cappellari and Jenkins (2002) published in *the Economic Journal*.

4.1 Selection bias and attrition

When considering panel data, there is always the question of attrition. For estimating income dynamics, one needs a balanced panel. However, household might get out of the panel, just by chance and this is not a problem. Or they can get out of the panel not at random, just because of the phenomenon we want to study. If we want to study the dynamics of poverty, a household might get out of the panel because they have lost their dwelling, because they are ashamed of getting poor. So our conclusion will be biased.

There is a similar problem when we study wage formation. Before estimating a wage equation, we must explain first the probability of participating to the labour market. In this case this is a pure selection bias problem, but still a very serious one. In their model of poverty dynamics,

Figure 5: The impact of attrition

Table I. Annual poverty inflow and outflow rates (row %), with and without missing income data

Poverty status, year $t-1$	Poverty status, year t		
	Not poor	Poor	Missing
(a) Sample with non-missing income at t			
Not poor	94.2	5.8	
Poor	41.5	58.5	
All	*87.9*	*12.1*	
(b) All individuals			
Not poor	84.3	5.2	10.6
Poor	36.3	50.8	13.3
All	*78.3*	*10.8*	*10.9*

Note: Pooled transitions from British Household Panel Survey, waves 1–9. Sample size (panel b) = 44, 772. Adults aged 20–59, excluding full-time students. The poverty line is 60% of median contemporary equivalized real net household income. Missing income data at t arise from either sample attrition or incomplete response within a respondent's household. See text for further details.

Source: Cappellari and Jenkins (2004)

Cappellari and Jenkins (2002, 2004) have an equation for modelling attrition.

4.2 What is measured?

The data comes from the BHPS, waves 1-9 which means 1991-1999. The sample was restricted to individuals aged 20 to 59 years who were not in full-time education. Focus is on poverty among adults of working age rather than child poverty or pensioner poverty. Disposable income (after taxes and transfers) scaled by a McClements equivalence scale, at 2000 prices. The poverty line is 60% of median income. Covariates refer to the household head or to the household.

The log of the income to needs ratio is taken as in the model of Lillard and Willis (1978). Essentially we have here an improvement over this model. Because this ratio

$$\log\left(\frac{y_{it}}{z}\right)$$

is positive if i is not poor at time t and negative otherwise, the explanatory model will be a probit and a multivariate probit because three equations will be considered. The interest is to explain

1. The initial condition in a Markov process of order one,
2. The attrition,
3. The probability of staying poor or exiting poverty, conditionally on the initial conditions.

So we have a model of poverty persistence and poverty entry.

4.3 The econometric model of Cappellari and Jenkins (2004)

The econometric model is difficult to understand in the first paper. We have to read the *Journal of Applied Econometrics* paper in order to understand clearly what is going on. A first-order Markov model of poverty transitions. Contrary to the model of Lillard and Willis (1978), their model, like hazard models, introduce non-linearities by distinguishing between rich and poor and thus assume different dynamics for rich and poor.

The first equation describe a latent propensity to be in a state of poverty at time $t-1$, modelling the initial condition of the two period process:

$$p^*_{i,t-1} = \beta' x_{i,t-1} + u_{i,t-1}, \qquad u_{i,t-1} \sim N(0,1).$$

The authors assume additional individual random effects μ_i which do not seem to lead to special treatment. So I discard them from the exposition. Then, the following variable is both an observation rule and a future explanatory variable:

$$P_{i,t-1} = 1, \quad \text{if } p^*_{i,t-1} > 0, \quad \text{zero otherwise.}$$

The third equation describes the conditional propensity of being poor in the second period:

$$p^*_{i,t} = [P_{i,t-1}\gamma_1 + (1-P_{i,t-1})\gamma_2] z_{i,t-1} + \epsilon_{i,t}, \quad \epsilon_{i,t} \sim N(0,1).$$

We note that the parameters of the explanatory variables can have a different value depending on the observed initial state.

The second equation models the propensity to remain in the sample when present at period $t-1$.

$$r^*_{it} = \psi' w_{i,t-1} + v_{it}, \quad v_{it} \sim N(0,1).$$

The correlation structure is as follows

$$\rho_1 = \text{corr}(u_{i,t-1}, v_{i,t}), \quad (6)$$

$$\rho_1 = \text{corr}(u_{i,t-1}, \epsilon_{i,t}), \quad (7)$$

$$\rho_1 = \text{corr}(v_{i,t}, \epsilon_{i,t}), \quad (8)$$

and this correlation structure is interpreted as covariances between individual random effects.

1. The correlation ρ_1 summarizes the association between unobservable individual-specific factors determining base year poverty status and income retention.

2. The correlation ρ_2 is the correlation between unobservable individual specific factors determining base year poverty status and poverty transitions

3. The correlation ρ_3 summarizes the association between unobservable individual-specific factors determining retention propensities and those determining conditional current poverty status.

The model can be identified thanks to its non-linearity (tri-variate probit). However, as soon as the correlations are different from zero, it is better to impose exclusion restrictions. Variables entering x or w should not enter z, the variables explaining conditional poverty status in period two in order to have a sufficient condition for identification.

4.4 Poverty transition probabilities

Poverty persistence

$$s_{it} = \Pr(P_{it} = 1 | P_{i,t-1} = 1) = \frac{\Phi_2(\gamma_1 z_{it-1}, \beta' x_{it-1}; \rho_2)}{\Phi(\beta' x_{it-1})}$$

Poverty entry

$$s_{it} = \Pr(P_{it} = 1 | P_{i,t-1} = 0) = \frac{\Phi_2(\gamma_1 z_{it-1}, -\beta' x_{it-1}; -\rho_2)}{\Phi(-\beta' x_{it-1})}$$

where Φ_2 is the bivariate cumulative of the Normal and Φ the CDF of the univariate standard normal. Under the assumption of stationarity the duration of a poverty spell is:

$$\frac{1}{1 - s_{it}}.$$

The mean duration of a spell of non-poverty is:

$$\frac{1}{e_{it}}.$$

4.5 State dependence

State dependence is a psychological notion designing the influence of memory on a learning process. Here, the notion dates back to the seminal paper of Heckman. Individual that have made an experience in the part have a stronger tendency to make the same experience that those who have not made that experience in the past. The present paper distinguishes two state dependence: aggregate and genuine. In the aggregate version, no account is taken of heterogeneity:

$$ASD = \frac{\sum_{i \in \{P_{it-1}=1\}} \Pr(P_{it-1} = 1 | P_{it-1} = 1)}{\sum_i P_{i,t-1}} - \frac{\sum_{i \in \{P_{it-1}=0\}} \Pr(P_{it-1} = 1 | P_{it-1} = 0)}{1 - \sum_i P_{i,t-1}}$$

GSD takes into account individual heterogeneity and corresponds to the following average across individuals

$$GSD = \frac{1}{N} \sum_{i=1}^{N} [\Pr(P_{it-1} = 1 | P_{it-1} = 1) - \Pr(P_{it-1} = 1 | P_{it-1} = 0)]$$

The absence of GSD can be tested by $H_0 : \gamma_1 = \gamma_2$.

4.6 Inference and testing model specification

Inference in the three-variate Probit can be relatively complicated because the evaluation of the likelihood required the evaluation of a tri-variate integral. Bivariate is still feasible, but tri-variate requires the help of simulation methods. Here the simulated MLE is used together with the GHK simulator for which we should know how to draw in a truncated normal.

Tests of the joint significance of the correlation parameters

1. The correlation between un-observables affecting initial poverty and income retention ρ_1 was negative and statistically significant, indicating a lower retention propensity among the initially poor compared to the non-poor

2. The correlation between unobservables affecting initial poverty and conditional current poverty ρ_2 was also negative and statistically significant. Since this measures the correlation between unobservables affecting initial poverty status and conditional current poverty status, poverty transition propensity, in other words, the negative sign can be interpreted as an example of Galtonian regression towards the mean.

3. Finally, the correlation between unobservables affecting income retention and poverty transition ρ_3 was not precisely estimated.

Exogeneity of initial conditions would imply that ρ_1 and ρ_2 were jointly zero, but such a hypothesis was strongly rejected. Exogeneity ρ_1 and ρ_3 were jointly zero strongly rejected at 1%, this result indicates that retention was endogenous for poverty transitions.

Overidentification restrictions tested.

Problem when testing for normality. Ok for current equation, but rejected for retention and initial state.

4.7 Marginal effects

Marginal effects in multivariate Probit models are always a problem. A change in a variable in one equation has an impact on other equations because of correlation. And s_{it} is computed as a ratio. The denominator (the initial state) has to be held constant.

1. First we computed the *predicted probability of poverty* in the base year for all respondents who were poor, and took the average of these values, call it c.

2. We then substituted $d = \Phi^{-1}(c)$ into (5), which means

$$s_{it} = \frac{\Phi_2(\gamma_1' z_{it-1}, d; \rho_2)}{d}$$

3. A marginal effect is the derivative of s_{it} with respect to the elements of z_{it-1}

4. Marginal effects for binary explanatory variables were calculated as the change in this expression implied by a unit change in each characteristic, ceteris paribus, relative to the characteristics of a reference person.

5. The reference person was defined by setting the continuous covariates (age and household head's age) equal to the sample median values (37 and 41), and all the remaining binary variables to zero.

4.8 Empirical results

The main conclusions of Cappellari and Jenkins (2002), also detailed in Cappellari and Jenkins (2004):

1. The correlation between the three equations of the probit are significant. So we have to take into account initial conditions. The estimated correlation implies that the poor are more likely to induce attrition from the sample. Not taking attrition into account would lead to underestimate poverty entry and persistence.

2. There are many more variables explaining poverty entry than explaining poverty persistence.

3. Poverty persistence is greater in households headed by a female and when there are several children between 3 and 11.

4. Entry entry has a large number of causes (heterogeneity), which reflect the standard results of the literature: single headed households, ethnic Chinese and Pakistani, all children up to 15 years. We could conclude that there is a deficit in family policy in the UK in term of poverty. 58% of poverty persistence and 6% of poverty entry. Full-time work and A-level are protecting factors for avoiding poverty entry.

5. There is a strong state dependence, comparable to what was measured in the US. The paper assumes a different coefficient for state dependence, depending on the initial status. The total state dependence is the difference of the two dependencies: poverty persistence minus poverty entry, 0.58-0.06 = 0.52.

5 Conclusion

An adapted version of the model of Cappellari and Jenkins (2004) was used in Sadeq and Lubrano (2018) in a Bayesian context for making inference on the consequences of the Wall on poverty dynamics in the occupied territories of the West bank.

Transition matrices were used in Chen and Cowell (2017): income mobility has risen after 2000 together with income inequality. But rank mobility has decreased after 2000.

References

Bane, M. J. and Ellwood, D. T. (1986). Slipping into and out of poverty: the dynamics of spells. *Journal of Human Ressources*, 21(1):1–23.

Cappellari, L. and Jenkins, S. P. (2002). Who stays poor? who becomes poor? evidence from the british household panel survey. *The Economic Journal*, 112(478):C60–C67.

Cappellari, L. and Jenkins, S. P. (2004). Modelling low income transitions. *Journal of Applied Econometrics*, 19(5):593–610.

Chen, Y. and Cowell, F. A. (2017). Mobility in China. *Review of Income and Wealth*, 63(2):203–218.

Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, 52(3):761–766.

Hill, M. S. and Jenkins, S. P. (2001). Poverty among British children: Chronic or transitory? In Bradbury, B., Mickelwright, J., and Jenkins, S., editors, *Falling in,climbing out: the dynamics of child poverty in industrialised countries*, chapter 7, pages 174–195. Cambridge University Press.

Jenkins, S. P. (2000). Modelling household income dynamics. *Journal of Population Economics*, 13:529–567.

Jenkins, S. P. and Lambert, P. J. (1997). Three I's of poverty curves, with an analysis of UK poverty trends. *Oxford Economic Papers*, 49(3):317–327.

Jenkins, S. P. and Lambert, P. J. (1998a). Ranking poverty gap distributions: Further TIPs for poverty analysis. In Slottje, D., editor, *Research on Economic Inequality*, volume 8, pages 31–38. JAI Press Inc.

Jenkins, S. P. and Lambert, P. J. (1998b). Three I's of poverty curves and poverty dominance: TIPs for poverty analysis. In Slottje, D., editor, *Research on Economic Inequality*, volume 8, pages 39–56. JAI Press Inc.

Kuchler, B. and Goebel, J. (2003). Incidence and intensity of smoothed income poverty in European countries. *Journal of European Social Policy*, 13(4):357–369.

Lillard, L. A. and Willis, R. J. (1978). Dynamic aspects of earning mobility. *Econometrica*, 46(5):985–1012.

Rodgers, J. R. and Rodgers, J. L. (1993). Chronic poverty in the United States. *The Journal of Human Resources*, 28(1):25–54.

Sadeq, T. and Lubrano, M. (2018). The wall's impact in the occupied west bank: A bayesian approach to poverty dynamics using repeated cross-sections. *Econometrics*, 6(2).

Stevens, A. H. (1999). Climbing out of poverty, falling back in. *The Journal of Human Ressources*, 34(3):557–588.

Thuysbaert, B. (2008). Inference for the measurement of poverty in the presence of a stochastic weighting variable. *The Journal of Economic Inequality*, 6(1):33–55.