Poverty Dynamics in Denmark, Germany and the USA

Mohammad Azhar Hussain

Welfare Distribution
Working Paper 29:2002

Working Paper
Socialforskningsinstituttet
The Danish National Institute of Social Research

Poverty Dynamics in Denmark, Germany and the USA

Mohammad Azhar Hussain

Welfare Distribution
Working Paper 29:2000

Poverty Dynamics in Denmark, Germany and the USA

by

Mohammad Azhar Hussain*

Abstract: *This study measures poverty persistence by extending the period of analysis and detects differences among different welfare states, Denmark, Germany and the USA. Poverty mobility patterns are explained by modelling the probability of poverty in a given year, conditional on poverty in a past initial year. Because the initial state could be endogenous, the initial conditions problem must be addressed. Applying a bivariate probit model with endogenous selection does that. The estimated parameters are used to simulate counter factuals in order to detect mobility differences among the three countries. The results suggest that the initial condition has some importance. The Danish and German welfare systems lower poverty persistence but have adverse effects on individual characteristics, and vice versa for the USA.*

JEL code: C53, C23, I32, J6

Key words: Poverty, non-poverty, state transition probability, initial state, poverty persistence, and welfare state

Acknowledgements: I am grateful to Nina Smith, Tor Eriksson, participants in the second (2000) workshop of European Political Economy Infrastructure Consortium (EPIC), Corfu, Greece, and participants at the seminar at the Department of Economics, Aarhus School of Business, for useful comments and suggestions. All remaining errors are mine.

* Department of Economics, The Aarhus School of Business, DK-8210 Aarhus V, Denmark, and CIM, CLS and SFI. E-mail: mh@sfi.dk

1. Introduction

That poverty and inequality measures based on annual income data are far from revealing the whole story about the distribution of income in a society, is now widely recognized. Considerable income mobility is observed even when the overall income distribution shows only little change from one year to the next. Thus stability at the macro level does not imply stability at the individual level. Panel data are necessary to complete the picture of low income or poverty, so that individual developments over time can be observed. This leads to the issue of mobility, which defines movements in and out of low income and the persistence in the low-income state. The question is whether low income is persistent, with the same people at the bottom of the income distribution or whether there is a considerable transitory income component. From a policy point of view this is important, because poverty, if temporary and therefore more equally distributed as in the latter case, may not result such a big problem. But if poverty is of a long run nature, as in the former case, it might call for targeted policies towards the poor. The dynamic perspective is also important in comparative studies because a country with a yearly high incidence of poverty compared to another country may have a more equal distribution of lifetime income as a consequence of higher income mobility. Finally, the extent to which a low-income person moves up or a high-income person moves down is a factor that probably affects motivation and future expectations of individuals: If it is the perception that escaping poverty is a difficult and hopeless task, then, that itself might discourage individuals from even trying to do something about their economic situation.

Here the aim is to identify how different factors affect the probability of transitions between low income and non-low income states. This is done by applying a bivariate probit model and by treating the initial state as endogenous, so that the initial conditions problem can be addressed in the context of poverty dynamics. Differences between welfare settings are analyzed through a comparison of three "representatives" (Esping-Andersen, 1994): Denmark, Germany and the USA. Denmark and the USA can be seen as the two extreme cases regarding government intervention to reduce poverty. It is assumed that the more market oriented approach in the USA implies a higher poverty rate, at the same time, though, the more flexible labour market generates greater mobility, so that over a longer period of time poverty might not be much higher in the USA compared to e.g. Denmark.

The next Section reviews earlier studies on economic mobility. In Section 3 there is a brief discussion about equivalence scales and poverty thresholds and the data is described. Section 4 presents the statistical model. Descriptive statistics and transition probabilities on relative poverty are presented in Section 5. In Section 6 the transition probabilities are modelled. Finally, conclusions are presented in the last Section.

2. Earlier studies

Different approaches have been applied in order to measure the extent of income mobility and its correlates. The studies can be categorized into non-parametric and parametric methods.

The basis for many income mobility analyses is to look at the transition matrix into which each person is categorized depending on the income class position at time t and then at time $t+? t$. The definition, length and number of income classes vary between studies. Estimation of transition probabilities between states over time is the central issue when analyzing income mobility. Contributions to these methods include Shorrocks (1978), which measures mobility by how much inequality is reduced when the accounting period is extended. Trede (1999) establishes asymptotic sampling distributions for this mobility measure and also for transition matrix mobility. His empirical findings suggest wide confidence intervals and therefore he favors to report confidence intervals when possible. Atkinson et al. (1992) surveys the economic mobility literature both concerning methodology and empirics, but particularly focusing on wages.

A significant proportion of mobility studies use descriptive statistics in order to estimate transition probabilities. There are examples of empirical applications from different countries with various definitions of income classes and length of accounting period. Adelman (1994) proposes a methodology for estimating income-mobility matrices from census data and applies this to Brazil, using 1970 and 1980 census data. Goodness-of-fit tests show that the applied income mobility measure is accurate and unbiased. The results indicate that workers at all income levels experienced growth, but that the growth was at the same time biased towards higher incomes. By using the estimated income-mobility matrices to calculate life-cycle incomes, it is concluded that upward mobility made life-cycle income distributions more equal than the distribution of starting incomes for workers. With German Socio-Economic Panel data Schluter (1997) finds that the intra-distributional mobility measured by transition probabilities is time varying. The observed mobility profiles are partly explained by focusing on the bottom of the income distribution instead of the entire transition matrix. Gittleman and Joyce (1999) examine the mobility of individuals in the United States by tracking movements across quintiles, and looks at the extent of income mobility change over time. The changes are found to be rather small from the 1970s to 1980s, despite major socioeconomic changes. Hauser and Fabig (1999), also using German Socio-Economic Panel data, compare individual income mobility after the German reunification, using different income measures. They see a strong convergence between former West and East Germany, particularly for net equivalent income mobility, which suggests that the social protection system has reduced mobility risks associated with the transformation process in the Eastern states of Germany.

The modelling of disposable income itself by income variance component or structural models is not an issue here, although this is another way to model poverty dynamics. Burgess and Propper's (1999) propose a framework for modelling household income dynamics. The role of marital partnership formation and dissolution, childbirth and labour supply decisions is emphasized. With their framework it is possible to address the issues of household, as distinct from individual, income and poverty dynamics. Their model seems to fit quite well for American Blacks and Whites, especially when bearing in mind that the poverty status is never used in the estimation process. Deding et al. (2000) uses a somewhat comparable strategy, although fewer behavioural decisions are estimated. Their simulations are good at replicating the Danish income distribution, e.g. concerning inequality and poverty measures.

Some parametric models of poverty include a variable for poverty status in last period as a regressor, whereby conditional probabilities of poverty can be inferred, see Pedersen and Smith (1998) for an application on Danish data.

Stewart and Swaffield (1999) take a step further and endogenizes the initial state in order to correctly model true state dependence on UK data. They use a bivariate probit model with endogenous selection and parental variables as instruments. Their study shows that the estimated parameters are generally considerably overstated, when the initial condition is not properly modelled. Deding (2000) applies a comparable set up, but her transitions are over longer periods of time, whereas Stewart and Swaffield only looks at one year transitions. In addition, Deding compares mobility in different welfare states, but does not find differences at the micro level in contrast to the aggregate level. Both studies' variable of analysis is low wage rather than low income, which is the topic here. Generally, the advanced econometric treatment of mobility is still primarily in the area of wage mobility; see Asplund et al. (1998) for more recent European studies on low wages. But it is also necessary to consider income mobility, because low wage is a very poor indicator of low income, see Marx and Nolan (1998), who shows that these two states need not and do not overlap much in the numerous Western countries examined in their study. This study fills the gap by using equivalent disposable income, which is equal to needs adjusted family income less taxes, as the basis for analysis. An earlier attempt to use equivalent income in international mobility comparisons is Aaberge et al (1996). With non-parametric methods they decompose and try to detect mobility differences. Their study does not support the hypotheses that more market-oriented economies like the US have higher mobility than the Scandinavian welfare states.

3. Data description

The Danish data used in the analysis are built from administrative registers in Statistics Denmark and supplied to the Centre for Labour Market and Social Research (CLS) at Aarhus Business School and University of Aarhus, Denmark. It is a longitudinal ½ percent representative unbalanced panel data set of the Danish population for the years 1976-1997. The data set contains information on a number of demographic, educational, income and labour market variables. The German and American data are from the Cross-National Equivalent File 1984-1997, compiled by Cornell University, USA, see Burkhauser and Daly (1995).

No standard way of proceeding with analysis of an equivalent income measure exists. There is thus some variation in defining needs, what income or tax components to include, which unit of calculation and which accounting period to apply. In this study the income concept is equivalent annual family income after transfers and taxes (EDI). Its calculated from EDI=DI/$S^?$, where $S^?$ is the equivalence scale and DI is family disposable income, defined as the sum of total family gross income minus total family taxes.

Gross income is the sum of labour earnings, asset flows, imputed value of owner occupied housing, private transfers and public transfers such as sickness benefits, unemployment insurance benefits, pensions and social assistance. Asset flows include income from rent, dividends and rent. For the USA public transfers include the face value of food stamps.

For Denmark, public non-taxable transfers such as housing benefits and individual supplementary means tested benefits for pensioners, are not included. Because these benefits are targeted at the bottom of the income distribution this might overstate the percentage in poverty. On the other hand the consequence is also a downward biased poverty line, so the total effect of excluding these benefits is not clear. Further, non-public transfers between separated parents of a child are not included.

$S^{?}$ is family size S raised to the power ?, which represents the elasticity of the equivalence scale with respect to family size. A number of recent international comparisons of poverty and inequality use this scale; see Förster (1995), Atkinson et al. (1995) and Buhmann et al. (1988). In this paper ? =½, which is most commonly used in international comparisons. In the calculation of equivalence scales, the German and American data restrict S to a maximum of six, so that the maximum equivalence scale is 2.45. This restriction is also imposed on Danish data, although it only affects less than about 1% of the observations in any country, because a family size exceeding six persons is rarely observed. These weights approximately produce the poverty rates seen in national studies using national weights, see e.g. Burkhauser et al (1995) and Pedersen and Smith (1998).

Table 1. Poverty lines, 1985-1997, 1,000 units of national currency

	Denmark DKK	Germany DEM	USA USD
1985	41.3	12.2	7.2
1986	43.0	12.4	7.5
1987	46.1	13.3	7.9
1988	48.5	13.6	8.4
1989	50.5	14.2	8.7
1990	54.5	14.7	9.2
1991	53.0	15.2	9.4
1992	54.5	14.4	9.8
1993	55.3	15.4	10.3
1994	59.0	16.2	9.4
1995	61.5	16.0	9.7
1996	64.0	16.9	10.1
1997	66.5	17.2	10.5
Changes			
Absolute, 1985-97	25.2	5.0	3.4
- Absolute yearly average	2.3	0.5	0.3
Relative, 1985-97, %	61.0	41.0	47.2
- Relative yearly average, %	4.4	3.2	3.6

There has not been obtained any consensus about the definition of an appropriate poverty line, which may cause implicit poverty lines in the different countries to differ. In Denmark the relative growth of public individual transfers (e.g. pensions, unemployment benefit and means tested transfers) follows the relative growth in hourly labour wages. This implies that Denmark implicitly has a relative poverty line, which does not only account for price increases but also for increases in productivity. In contrast the official American poverty line is absolute, so that changes only occur due to price changes and not due to productivity gains. Studies show that the (absolute) American poverty line is roughly equal to about 40% of median equivalent disposable income, see Burkhauser et al. (1995). In this study the 50% of median EDI is applied. The poverty line is in all cases calculated from incomes of persons aged 18 to 74 years, because this is the maximum age range simultaneously available in all three country-samples. In order to avoid extreme incomes to affect the poverty line, individuals with incomes above the 99th percentile (calculated with positive incomes only) and individuals below the 1st percentile were excluded from the analysis. Defined in this way the poverty lines have an average growth rate of about 4½% in Denmark (2,291 DKR), 3½% in USA (307 USD), and about 3¼% in Germany (455 DEM), see Table 1, which also contains the poverty line levels. The lower growth in Germany results from a high growth in 1985-1991, compared to 1991-1997. The decrease is probably caused by the costs of reunifying West and East Germany in 1991, but generally lower inflation could also be an explanation. Reunification does not affect which individuals are included in the sample, because all has to be present in 1984, seven years before unification, but the poverty line is affected by the unification, because that is based on the whole population as such.

Apart from the poverty line, all analysis is made on data restricted to people aged 26 to 59 years in order to exclude students and retiring persons.

The unit of analysis is the individual, although the household is the unit of calculation. The household consists of one or two adults, and children less than 18 years living with an adult. For Denmark, it has not been possible to pool the income of cohabiting partners, because income information only exists for both partners if they are married. Children over 18 years living with an adult are treated as singles. This will tend to overstate the estimated proportion living under the poverty threshold. But because the analysis is based upon persons aged 26-59 years, the single categorization is expected to be less of a problem.

The data are time wise limited to the period from 1984 to 1997, which is the maximum number of years available for all three countries.

Sample statistics for some of the sample years are presented in Table A1 in the Appendix.

4. Econometric model

Modelling of transition probabilities is complicated by the initial conditions problem, which arises when the initial state is not exogenous, cf. Heckman (1981). The endogeneity of relative poverty has been found in several studies. The incidence of poverty thus varies according to e.g. gender, age, education, number of children, marital status, labour market attachment, see e.g. Pedersen and Smith (1998) for a parametric treatment and Gallie and Paugam (2000) for descriptive statistics. Non-exogenous initial state can cause sample selection bias, implying that true representation and inferences may be distorted with standard modelling techniques. The initial conditions problem will be dealt with by using a model with selectivity and different variables to instrument the initial state. The degree of bias is detected by showing estimates on the assumption of endogenous as well as of exogenous sample selection.

The interest lies in the probability that a poor individual in year t is also in poverty in year $t+?\ t$. Assume that the following process generates the equivalent disposable income of an individual

$$f_1(y_{it}^*) \ ? \ x_{it} ?^* \ ? \ ?_{i1}$$

(1)

where y_{it}^* is disposable income for individual i in year t, x_{it} is a vector of income-determining characteristics, $?_{i1}$~N(0,1), and f_1 is an unspecified suitable monotonic transformation, ensuring the standard normal distribution of $?_{i1}$. The probability that income falls below the poverty threshold is given by

$$P(y_{it} \ ? \ 1) \ ? \ P(y_{it}^* \ ? \ ?_t) \ ? \ P(f_1(y_{it}^*) \ ? \ f_1(?_t)) \ ? \ ? \ (f_1(?_t) \ ? \ x_{it} ?^*) \ ? \ ? \ (x_{it} ?)$$

(2)

where y_{it} is an indicator variable equal to 1 if income falls below the poverty threshold $?_t$ and 0 otherwise. ? is the standard normal cumulative distribution function. The coefficients $-?^*$ are equal to the estimated coefficients $?$.

If the income in year $t+?\ t$ depends on the poverty status in year t, then for a poor person in year $t+?\ t$ income is generated by the process

$$f_2(y_{i,t??t}^*) \ ? \ z_{i,t??t} ?^* \ ? \ ?_{i2}$$

(3)

z now represents transition-determinants that are variables explaining income status in year *t+? t*, given poverty in year *t*. The monotonic transformation f_2 ensures the standard normal distribution of $?_{i2}$. The same process but with a different $?^*$-vector is assumed to apply for individuals who were not poor in year *t*,

Assume further that the error term from the income generating process and from the income transition process ($?_{i1}$, $?_{i2}$) have a bivariate standard normal distribution with the correlation coefficient ?, -1<?<1. Then the joint probability of being poor in both years is given by

$$P(y_{it} ? 1, y_{i,t??t} ? 1) ? P(y^*_{it} ? ?_t, y^*_{i,t??t} ? ?_{t??t}) ? ?_2(x_{it}?, z_{i,t??t}?, ?)$$

(4)

where ? is a transformation of the $?^*$-vector. Using the definition of conditional probability, the probability that an individual is poor in year *t+? t* given poverty in *t*, is

$$P(y_{i,t??t} ? 1 | y_{it} ? 1) ? \frac{P(y_{it} ? 1, y_{i,t??t} ? 1)}{P(y_{it} ? 1)} ? \frac{?_2(x_{it}?, z_{i,t??t}?, ?)}{?(x_{it}?)}$$

(5)

Exogenous initial conditions are represented by the ? being equal to 0. Zero correlation implies lack of dependence between poverty status in year *t* and poverty status in year *t+? t* and the probability of poverty in year *t+? t* reduces to the standard univariate probit model

$$P(y_{i,t??t} ? 1 | y_{it} ? 1) ? \frac{P(y_{it} ? 1) P(y_{i,t??t} ? 1)}{P(y_{it} ? 1)} ? ?(z_{i,t??t}?)$$

(6)

The probability of being non-poor in year *t+? t* and poor in year *t*, is defined analogously to (4)

$$P(y_{it} ? 1, y_{i,t??t} ? 0) ? ? _2(x_{it} ? ,? z_{i,t??t} ? ,? ?)$$
(7)

Because year $t+?t$ information is only used for those poor in year t, the model is a bivariate probit model with endogenous selection, where the likelihood contribution for individual i is

$$\ln L_i ? y_{it} y_{i,t??t} \ln ? _2(x_{it} ? , z_{i,t??t} ? , ?) ? y_{it}(1 ? y_{i,t??t}) \ln ? _2(x_{it} ? ,? z_{i,t??t} ? ,? ?) ? (1 ? y_{it}) \ln ? (? x_{it} ?) \quad (8)$$

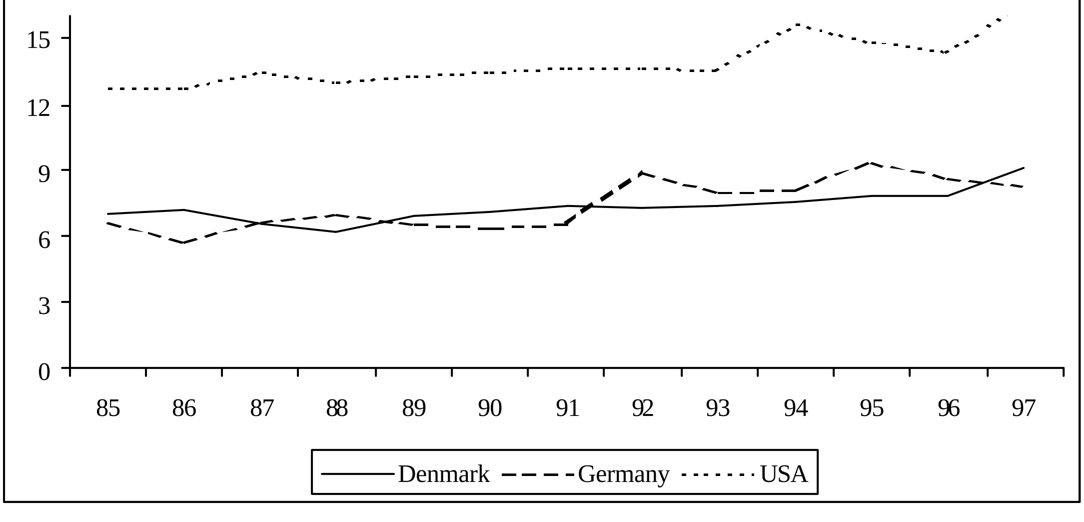
Estimation results from the empirical analysis for the three countries are presented in Section 6.

5. Aggregate poverty rates and transition probabilities

Poverty rates

Figure 1 shows the poverty rates for the population aged 26-59 years[1]. This age group primarily consists of persons on the labour market. Younger people under 26 years are excluded because they will often be attending education and older people over 59 years are excluded to avoid effects of retirement decisions. The fraction of the population under the poverty line is quite stable with only minor changes from one year to the next, cf. Figure 1. The average risk of poverty is highest in the USA, where the level is around 13 and 15%. The poverty rates in Germany and Denmark are lower and fluctuate around between 6 and 9%.

Figure 1. Poverty Rates, 26-59 years, 1985-1997, %



[1] In order to compare these poverty rates to other sources, Figure A1 in the Appendix show poverty rates for people in the age group 18-74 years, which is also the age group upon which the poverty lines are constructed. The percentage under the threshold in figure A1 are very much like in Pedersen and Smith (1998) and Burkhauser et al. (1995). Applying new OECD equivalence scales, Pedersen and Smith (1998) finds poverty rates between 10 and 12½ % in Denmark in the years 1980-1994. In Burkhauser et al. (1995) the poverty rate in 1989 in Germany is 8.8 % when using German equivalence scales and 9.2 % when using American equivalence scales. The latter study also contains American poverty rates for 1989, which are 18.1 % with American equivalence scales and 17.4 with German equivalence scales. Poverty rates in this study are thus close to comparable studies for the three countries. Compared 18-74 years old people, individuals aged 26-59 have lower poverty, because the latter group is on the labour market and thus have higher incomes than students and the elderly, which are included in the first group.

Transition probabilities

Data in Figure 1 are a cross-section, so that a person in e.g. 1985 is not necessarily in the sample in 1991 or 1997. When analysing transitions between states, a person has to be present in both the start year and the destination year, but do not have to be present in all years between the two of interest. Initially persons may also be in a state, which is not relevant for the analysis or enter such a state from the base year. The base year or initial year through out the paper is 1984 and the destination years are in this Section are 1985-1997.

Top- and bottom coding does not reduce the sample size much. For instance in the sample consisting of individuals present in 1984 and 1985, the top and bottom coding affected only 1.4 % of the sample persons in Denmark, see Table 1. Also zero, negative or missing income, which affected about 2% in each country, were excluded.

Table 2. The probability of being in a given state in 1985 (year *t+? t*), conditional on state in 1984 (year *t*), %

	Year *t+? t* =1985 state						
Year *t*=1984 state	Non-poor	Poor	More than 59 years	Top or bottom coded	Invalid income	All	Distribu-tion in 1984
Denmark							
Non-poor	95.1	1.8	2.2	0.4	0.5	100	90.1
Poor	24.8	68.1	3.9	1.9	1.2	100	6.5
Top or bottom coded[1)]	54.5	8.0	2.3	26.1	9.1	100	1.4
Invalid income[2)]	26.2	4.9	4.9	4.9	59.0	100	1.9
All	88.6	6.2	2.4	1.0	1.8	100	100
Germany							
Non-poor	92.7	3.7	2.1	0.0	1.5	100	92.6
Poor	41.0	50.8	4.5	0.5	3.2	100	5.3
Top or bottom coded[1)]	70.3	9.9	6.8	.	13.0	100	0.3
Invalid income[2)]	51.3	8.2	7.1	3.2	30.2	100	1.9
All	89.1	6.3	2.4	0.1	2.1	100	100
USA							
Non-poor	93.1	4.0	1.9	.	1.0	100	85.1
Poor	32.5	61.1	3.3	0.3	2.8	100	12.3
Top or bottom coded[1)]	36.1	25.3	.	.	38.6	100	0.1
Invalid income[2)]	32.6	12.6	2.4	2.2	50.2	100	2.6
All	84.0	11.3	2.1	0.1	2.5	100	100.0

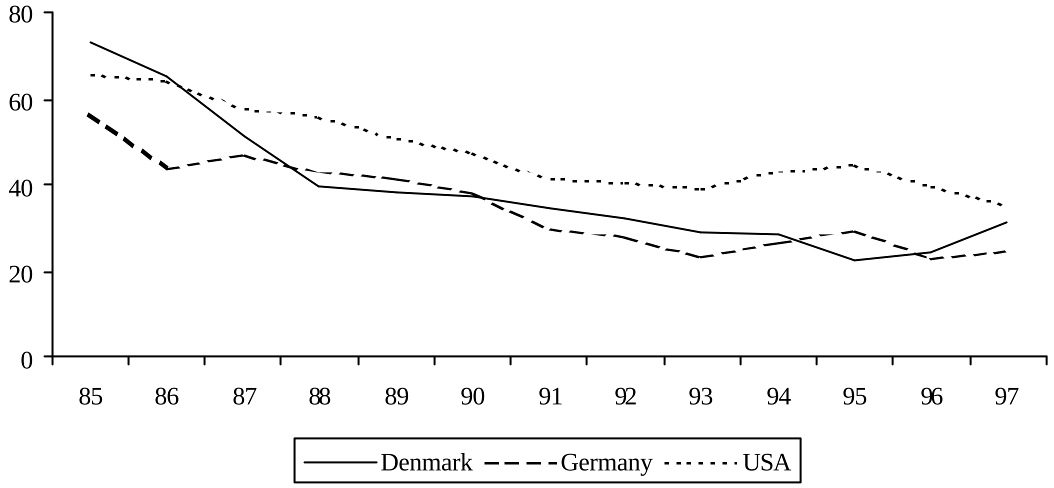
[1)] Bottom coded if income is below first percentile and top coded if above 99 th percentile. Individuals affected by this restriction are different from 2%, because the restriction is imposed on the broader age group 18-74 years, which is used to calculate the poverty line.
[2)] Zero, negative or missing income

The remaining sample is classified as either poor or as non-poor in 1984. Although these observations are valid in the initial year, about 3% of the poor Danes in 1984 had too high or too low or invalid incomes in 1985, and the same is true for about 1% of the non-poor. About 4% of the poor were above the upper age limit in 1985, whereas this was only the case for about 2% of the non-poor. The remaining cells represent transitions between poverty and non-poverty. About 68% of Danes poor in 1984 are also poor the year after, but this is only the case for about 2% of the non-poor in 1984. The risk of poverty thus very much seems to depend on whether an individual was poor the year before. When the sample is restricted to individuals who are classified into either of the two poverty categories, poverty persistence seems more serious because the conditional probability of poverty increases about 5%-points in Denmark to 73.3% (not shown in, but calculated from Table 2). The increased persistence is also present in Germany and the USA, when restricting the material to a 2x2 matrix. At the aggregate level poverty persistence is 56.1% in Germany and 65.8% in the USA. Germany thus has the lowest poverty persistence. For Sweden, which also belongs to the Scandinavian welfare model, poverty persistence is 61.2% (Gustafsson, 2000). This figure is lower than for Denmark, but comparability is not complete, because different equivalence scales are applied.

Conditional poverty risks are also calculated for later years than 1985, so the distance between the initial years of poverty, 1984, and some future year is increased, see Figure 2. But the same individuals are not followed from e.g. 1984 to 1986 as e.g. 1984 to 1985, because fewer individuals are present in both 1984 and 1986 compared to 1984 to 1985. This is due to attrition caused by people reaching the age limit of 59 years. Thus the longer apart we look at poverty persistence, the fewer observations we have.

Poverty persistence is reduced over time. Germany has huge drops in poverty persistence in the two first years after poverty in 1984, so that only 44% of the poor in 1984 are also poor in 1986, whereas 64-65% of Danes and Americans are still in poverty after 2 years. After four years less than half of the poor in 1984 are still poor in Denmark and Germany, while 56% are still poor in the USA. In the short run poor people in Germany seems to escape poverty fairly quickly, whereas the poor in the USA seems to have more difficulties in escaping poverty. From 1988 and onwards until 1996 Germany and Denmark follow each other closely on a lower level than the USA.

Figure 2. The Probability of Poverty in Year *t+dt* , Conditional on Poverty in Year *t* =1984, %



For 1989, five years after initial poverty, it is possible to compare poverty persistence in Germany with Hauser and Fabig (1999). Using net equivalent income, which is somewhat similar to the income concept used in this study, they find that 40.5% of the West German persons, who were below half mean income in 1990, were also below half mean income in 1995, five years later. In Figure 2, the German poverty persistence is 41.5% in 1989, which is close to Hauser and Fabig, although they use the mean instead of the median, and the periods of analysis are not overlapping. The two populations are not from different geographical areas, because conditioning on poverty in 1984 implies that only West Germans are actually included in this study.

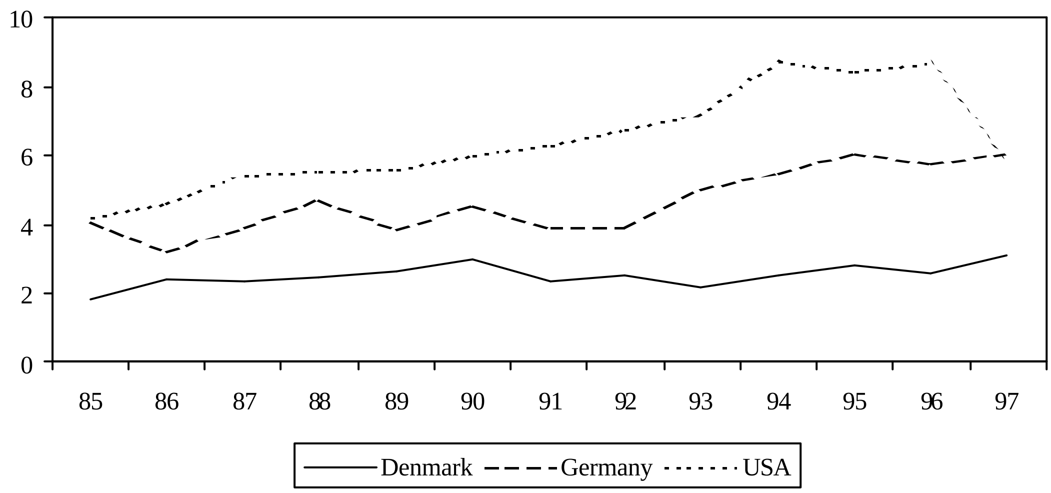
In 1991, 7 years after initial poverty in 1984, the conditional poverty is still considerable and reaches 41½% in the USA and around 30-35% in Denmark and Germany. Although the conditional poverty risk is high for the USA, when compared internationally, it is not high when controlled for the general poverty risk level of the country (figure 1 compared to figure 2). Initially poor Americans' poverty risk is about 3 times bigger than the unconditional average poverty risk in 1991, whereas the conditional risk is about 4½ times bigger than national unconditional risks in Germany and Denmark. When measured relative to national averages the USA actually only (marginally) exceeds Germany or Denmark in two years (1995 and 1996).

12-13 years after poverty starts in 1984, there are still between ¼ (Germany) and 1/3 (Denmark and the USA) under the poverty threshold in 1996 or 1997. These individuals are not necessarily poor in every year between 1985 and 1997, but they are poor in both 1984 and 1997. Poverty persistence is thus reduced markedly over the 13-year period, but the poverty risk for the initially poor is still high. Compared to national averages the 1984-poor have a 3½ times higher risk in Denmark, 3 times higher in Germany and 2¼ higher in the USA in 1997.

Poverty persistence show different patterns in the three countries, but in the long run all three countries end up with about the same poverty persistence. So over time the difference in poverty persistence between the three countries is reduced. This is somewhat surprising bearing in mind that the three countries each represents different welfare settings. A priori it would be difficult to say any thing about who should have greatest poverty persistence or mobility. The USA with a more flexible labour market and low replacement rates in the social benefit system could be expected to have lower poverty persistence, because people are dissatisfied with low income while receiving social benefits, and thus would be expected to escape poverty quickly. On the other hand poor people might have low qualifications and thus have problems escaping poverty, even when they find a job, because wages are low. In Europe one could expect higher poverty persistence because benefits are comparatively high relative to the wage rate. This might lower incentives to escape poverty.

In Deding (2000) the aggregate probability of low pay in 1984 and 1984+*? t*, *? t* ={2,4,6}, is about ¼ in Denmark and USA, and about 1/3 in Germany for *? t*=2. This indicates much higher low *income* (poverty) persistence than low *pay* persistence in both the short run (2 years) and also in the longer run (6 years). Further, low pay persistence falls over time, but not nearly as much as poverty persistence.

Figure 3. The Probability of Poverty in Year *t+dt* , Conditional on Non-Poverty in Year *t* =1984, %



It is seen from Table 1 that transitions also take place from non-poverty in 1984 to poverty in 1985. But this transition is much less frequent, so that the probability of poverty in a future year, conditional on non-poverty in 1984 is quite low. This is also the case when looking at transitions over a longer period, see Figure 3. This conditional probability shows an upward trend over time in the USA and Germany and ends in 1997 at about 6% in both countries. For Denmark it is more stable and fluctuates around 2½%. Compared to the overall poverty risk, the non-poor's risk is closest to the average in Germany (about 2/3 of the overall average) and lowest in Denmark (about 1/3 of the average). Generally, increasing conditional probabilities for poverty for non-poor and decreasing probabilities for the originally poor, means that the poor and non-poor's poverty risk converges over time. But even after 13 years there is quite a gap between the two groups.

Although the probability of poverty, given non-poverty, is very low compared to probability of poverty, given poverty, a major part of any year's group of poor is originally non-poor. The reason why low conditional poverty risks for the non-poor contribute largely to the group of poor in a given year is that, the non-poor group is much bigger than the poor group in any year.

From the aggregate measures it seems clear that the initial state of a person is very decisive regarding the future risk of poverty. The group of poor people in 1984 thus has much higher probability of poverty in later years compared to the average person. The non-poor on the other hand has much lower poverty risk then the average. There thus seems to be considerable state dependence in the probability of poverty in a given year. State dependence occurs when a given state, e.g. poverty, changes behaviour, because poverty could have adverse effects on motivation and social networks. This can come about because falling into poverty could force the family to shift residence from a socially well to a socially less well functioning neighbourhood with more unemployed people or people with drug abuse or alcoholism as a problem. People in these neighbourhoods are probably less aware of career or job opportunities and are therefore less likely to get out of the poverty state. Poverty could also change behaviour because the poor might get “used” to being poor and thereby lower expectations about escaping poverty. In both cases becoming poor might itself cause future poverty.

Some factors behind the observed poverty transition and poverty persistence pattern could be high-lighted by looking at which kind of individuals stays in poverty and which kind of individuals makes the transition from poverty to non-poverty. The characteristics of these two kinds of persons are presented in Table 3, where y_t is poverty status in year t, y={0,1}={non-poor, poor}.

Stayers are more often individuals without a partner compared to individuals making a transition. E.g. in Denmark, 59% of the individuals who were poor in 1984 and 1985 were singles, while 49% of individuals who jumped from poverty in 1984 to non-poverty in 1985 were singles. But there is not generally a lower or higher fraction of singles when the moving individual’s marital status is examined in the base year and the destination year (1984+? t). In Denmark, e.g. 54% among the stayers were singles in the destination year 1985, while 58% among the movers were singles.

Among the poor in 1984, men more frequently than women moves from poverty into non-poverty, because 66% of individuals remaining poor from 1984-1985 were women, while only 57% of the poverty escapers were women. An exception to this pattern is Germany in 1997. The size of this gender gap varies over time and countries.
Apart from Germany, age does not determine whether a person will stay in poverty or move in to non-poverty. Children’s effect is also not clear, and both stayers and movers

have fewer children in the destination year than in the initial year, but individuals leaving poverty generally have a higher decrease in the number of children than stayers.

Table 3. Average of background variables in the initial year *t* (*t*=1984) and in the destination year *t*+? *t* (? *t* ={1,7,13})

	Year $t+?t$ =1985		Year $t+?t$ =1991		Year $t+?t$ =1997	
	y_t=1,$y_{t+?t}$ =1	y_t=1,$y_{t+?t}$ =0	y_t=1,$y_{t+?t}$ =1	y_t=1,$y_{t+?t}$ =0	y_t=1,$y_{t+?t}$ =1	y_t=1,$y_{t+?t}$ =0
Denmark						
Average of variables in initial year t=1984, for the sample present in both t and t+?t						
Woman	0.66	0.57	0.73	0.60	0.66	0.63
Single	0.59	0.49	0.48	0.57	0.62	0.52
Age	38.99	36.31	35.61	36.01	34.18	34.42
No of children	0.94	1.03	1.23	1.10	1.25	1.23
Education, years	10.40	10.96	9.67	10.89	9.62	10.95
Annual hours employed/100	3.50	5.05	4.59	4.33	4.86	5.02
Average of variables in destination year t+?t =1984+?t, for the sample present in both t and t+?t						
Woman	0.66	0.57	0.73	0.60	0.66	0.63
Single	0.54	0.58	0.39	0.67	0.64	0.62
Age	39.99	37.31	42.61	43.01	47.18	47.42
No of children	0.93	0.85	0.97	0.53	0.48	0.36
Education, years	10.53	11.37	9.74	11.25	9.80	11.48
Annual hours employed/100	4.04	7.63	4.58	7.34	1.98	9.59
Germany						
Average of variables in initial year t=1984, for the sample present in both t and t+?t						
Woman	0.55	0.55	0.64	0.50	0.43	0.53
Single	0.50	0.30	0.39	0.44	0.59	0.42
Age	41.02	38.18	38.48	34.45	36.00	34.42
No of children	0.85	1.28	1.38	1.11	1.00	1.21
Education, years	10.95	11.52	10.43	11.71	9.70	12.33
Annual hours employed/100	4.29	7.12	5.45	6.72	6.45	5.68
Average of variables in destination year t+?t =1984+?t, for the sample present in both t and t+?t						
Woman	0.55	0.55	0.64	0.50	0.43	0.53
Single	0.49	0.32	0.50	0.37	0.41	0.31
Age	42.02	39.18	45.48	41.45	49.00	47.42
No of children	0.80	1.22	0.88	1.05	0.66	0.73
Education, years	11.00	11.52	10.43	11.82	9.70	12.48
Annual hours employed/100	5.00	9.91	6.34	14.37	1.55	13.54
USA						
Average of variables in initial year t=1984, for the sample present in both t and t+?t						
Woman	0.68	0.63	0.67	0.64	0.69	0.67
Single	0.55	0.39	0.55	0.45	0.58	0.38
Age	37.30	37.61	34.63	34.59	33.28	32.69
No of children	1.71	1.50	1.80	1.93	1.69	1.99
Education, years	11.01	11.82	10.73	12.09	11.16	11.99
Annual hours employed/100	8.34	10.60	6.45	10.69	8.53	11.04
Average of variables in destination year t+?t =1984+?t, for the sample present in both t and t+?t						
Woman	0.68	0.63	0.67	0.64	0.69	0.67
Single	0.56	0.37	0.61	0.41	0.70	0.41
Age	38.30	38.61	41.63	41.59	46.28	45.69
No of children	1.66	1.37	1.38	1.31	0.94	0.85
Education, years	11.23	12.11	10.98	12.47	11.31	12.40
Annual hours employed/100	9.57	16.29	8.78	17.43	8.63	17.17

Stayers are in all countries less educated than people who leave poverty, and individuals making the transition have in all cases, but one (1984-1985 transition in Germany), a higher increase in educational level, from 1984 to the destination year, than stayers. Working hours are almost always higher for people in 1984 if they later leave poverty. Further it's worth noting, that the transiting individuals always increase their time on the job and that this increase is on average always greater than for people staying poor, who some times actually decrease their time on the job. From the definition of disposable income it can be seen that the central components are wages and working hours. According to human capital theory, the wage rate is primarily determined by the level of education and number of years of job experience, which is sometimes approximated by age. A reduction in either of these variables can thus cause poverty.

These effects of different background variables on the probability of poverty is rather well documented empirically, for instance concerning gender, race, marital status, unemployment and number of children, cf. Burkhauser et al. (1988), Pedersen and Smith (1998). The latter study finds a significantly increased risk for singles, individuals with children or low education, and the unemployed.

6. Results

The probability of poverty dependent on poverty experience in 1984 is calculated for the beginning of the sample period (1985), the middle (1991) and the last year (1997).

The probability of poverty in the initial year is estimated with the same variables, some of whom are time varying, as the conditional probability of poverty in year $t+?t$, but for identification purposes two other variables are included. The two extra variables included in year t are instruments for the endogenous selection into the initial state. The first is the aggregate unemployment rate in the year of graduation divided by numbers of years since graduation. The assumption is that workers who entered the labour market in years of high unemployment are disadvantaged compared to other workers in terms of unemployment and wages, see De Vreyer et al. (2000). Dividing with the years since graduation implies, that the effect of conditions at the year of labour market entry is reduced over time. The other variable is public transfers as a percentage of disposable income.

Estimated parameters from (8) are presented in Table 4. Generally, many parameter estimates are significant and have the expected signs. A positive (negative) parameter implies higher (lower) probability of poverty in year $t+?\,t$, conditional on poverty in year t. A more detailed interpretation of the parameters is made in the comments to Table 5. The correlation coefficient ? is negative in all estimations. This corresponds well with the fact (not shown here) that individuals have lower income *growth* from year t to year $t+?\,t$, the higher their income *level* in the initial year t. Negative ? is also found in Stewart and Swaffield's (1999) study of low pay in Britain. ? becomes more negative, the higher the $?\,t$, but it is only significant in all pair of years for Germany. The negative trend implies that the linear relation between initial income level and income growth is stronger as time passes. The dependence between level and change could also have been great in earlier years as well, but that would not have been in a linear way. ?'s significance in some year or another for each country suggests that the initial state seems to be endogenous, and that the assumption of exogenous initial state must be rejected, except for Denmark in 1984-1985 and 1984-1991 and for USA in 1984-1997. Addressing the initial conditions problem is thus important.

The formula for calculating the pseudo coefficient of determination, R^2, is as in the ordinary least square model, but only for the individuals, who were poor in year t. R^2 is above 0.4 in all estimations, which indicates a reasonable model fit, especially in the case of micro data. In all estimations the model fit is better in the year right after initial poverty compared to later years.

Because the statistical model is non-linear it is more instructive to look at marginal effects of the explanatory variables, conditional on some values of the explanatory variables, as in the following manner, see Greene (2000). For the initial year, let x_j be the dummy variable, whose parameter $?_j$ is being evaluated, and X_{-j} are all the explanatory variables excluding x_j, and $?_{-j}$ is the parameter vector excluding $?_j$. In the destination year, the dummy's parameter is $?_j$, and Z_{-j} are all the explanatory variables excluding x_j, and $?_{-j}$ is the parameter vector excluding $?_j$. The individual probability of poverty, conditional on poverty in year t and conditional on x being 0, is

$$P(Y_{it}=1 | Y_{i,t+?t}=1 ? (x_{ji,t},x_{ji,t+?t})=(0,0))$$

and the conditional probability, given x being 1, is

$$P(Y_{it}=1 | Y_{i,t+?t}=1 ? (x_{ii,t},x_{ij,t+?t})=(1,1))$$

The marginal change in the conditional probability of poverty when x changes from 0 to 1 is then given by

$$P[Y_{i,t+?tt}=1 | Y_{i,t}=1 ? (x_{ji,t},x_{ji,t+?t})=(1,1)]- P[Y_{i,t+?t}=1 | Y_{i,t}=1 ? (x_{ji,t},x_{ji,t+?t})=(0,0)]$$

$$= \frac{?_2(X_{i,?jt}'?_{?j} ? ?_j, Z_{i,?j,t??t}'?_{?j} ? ?_j, ?)}{?(X_{i,?jt}'?_{?j} ? ?_j)} ? \frac{?_2(X_{i,?jt}'?_{?j}, Z_{i,?j,t??t}'?_{?j}, ?)}{?(X_{i,?jt}'?_{?j})}$$

The marginal effects are averaged over the initially poor sample. The value of other explanatory variables is the actual value of the individual. For the continuous variables the same procedure is applied. So age is increased with one year for each individual, the number of children is increased with 1, and the same applies for education and the number of hours employed. Because of non-linearities, number of employed hours and education are also changed with 5 each.

Table 4. ML-estimates from the bivariate probit model, modelling of the probability of poverty in year* t+? t*, given poverty in year* t*,* t*=1984, ?* t*={1,7,13}. (Explanatory variables are from the destination year* t+? t*)

	From 1984 to *t+? t* =1985		From 1984 to *t+? t* =1991		From 1984 to *t+? t* =1997	
	Estimate	St. error	Estimate	St. error	Estimate	St. error
Denmark						
Intercept	1.4156	1.4438	4.9012	3.7219	-2.4447	7.9331
Woman	0.1504	0.1415	0.1662	0.1964	-0.0850	0.2126
Age	-0.0166	0.0688	-0.1238	0.1636	0.2229	0.3294
Age2/100	0.0284	0.0822	0.1290	0.1807	-0.2736	0.3370
Single	0.0727	0.1605	-0.3001	0.1811	0.0056	0.2767
No of children	0.1177	0.0704	0.4356	0.1088	0.7293	0.1860
Education, years	-0.0397	0.0282	-0.2147	0.0538	-0.1532	0.0565
Annual hours employed/100	-0.0426	0.0124	-0.0353	0.0168	-0.1166	0.0230
Correlation coefficient, *?*	-0.1405	0.1300	-0.2160	0.1439	-0.3051	0.1555
Log likelihood	-1,086.1		-712.2		-519.5	
Pseudo R^2 (*)	0.7260		0.4346		0.5196	
No of observations	5,856		4,179		3,117	
Germany						
Intercept	2.0631	1.4657	8.6813	4.4902	32.2437	33.6240
Woman	-0.1191	0.1392	0.0889	0.2286	-0.5387	0.5344
Age	-0.0786	0.0704	-0.2375	0.1905	-0.9907	1.2718
Age2/100	0.1103	0.0843	0.2571	0.2096	0.8945	1.2523
Single	0.2796	0.1723	-0.3819	0.2577	-0.8187	0.6959
No of children	0.0172	0.0713	0.3898	0.1375	0.2137	0.5510
Education, years	-0.0095	0.0237	-0.3401	0.1026	-0.4544	0.2583
Annual hours employed/100	-0.0175	0.0085	-0.0021	0.0211	-0.0033	0.0592
Correlation coefficient, *?*	-0.3640	0.0926	-0.4697	0.1893	-0.7344	0.2940
Log likelihood	-9,251.2		-459.1		-123.6	
Pseudo R^2 (*)	0.5549		0.4108		0.4320	
No of observations	5,807		3,603		2,240	
USA						
Intercept	1.9604	1.1425	2.2780	1.9297	8.5314	9.1269
Woman	-0.1648	0.0855	-0.2476	0.1055	-0.0421	0.1723
Age	-0.0150	0.0568	-0.0012	0.0796	-0.3287	0.3833
Age2/100	0.0139	0.0678	-0.0062	0.0860	0.3535	0.4006
Single	0.4243	0.1058	0.1648	0.1174	0.5349	0.1842
No of children	0.1001	0.0344	0.0825	0.0407	0.1493	0.0657
Education, years	-0.0625	0.0180	-0.1033	0.0257	-0.0774	0.0370
Annual hours employed/100	-0.0333	0.0050	-0.0494	0.0058	-0.0549	0.0082
Correlation coefficient, *?*	-0.1652	0.0898	-0.3137	0.0806	-0.1043	0.1573
Log likelihood	-2,164.4		-842.5		102.1	
Pseudo R^2 (*)	0.6296		0.4663		0.4223	
No of observations	5,940		4,155		2,263	

* The pseudo R^2 is calculated as 1 minus the sum of squared residuals (poverty state 0/1 minus the fitted probability) divided by the sum of squared observed poverty state difference (actual poverty state minus average probability of poverty). R^2 is only calculated for the sample, which is poor in the initial as well as in the destination year.

Table 5. Estimated marginal effects of the conditional probability of poverty in year *t+? t*, given poverty in year *t*, *t*=1984, *? t* ={1,7,13}, evaluated at individual values of explanatory variables.

	1984 to 1985		1984 to 1991		1984 to 1997	
	Conditional probability, %	Effect, %-points	Conditional probability, %	Effect, %-points	Conditional probability, %	Effect, %-points
Denmark						
Representative person	73.2	.	34.6	.	31.9	.
Rep.person, but all "are" men[1)]	69.8	.	30.6	.	32.3	.
- All "are" women[1)]	75.2	5.4	36.9	6.3	31.6	-0.7
Age	73.2	0.1	34.2	-0.5	30.7	-1.2
Rep.person, but all "are" women[2)]	71.2	.	38.4	.	30.5	.
- All "are" singles[2)]	74.9	3.7	31.5	-6.9	33.1	2.6
Children	77.3	4.2	49.8	15.1	52.3	20.4
Education, 1 year more	71.9	-1.3	28.2	-6.4	27.9	-3.9
Education, 5 years more	66.4	-6.8	9.5	-25.2	14.7	-17.2
Annual hours emp.,100 more	71.7	-1.5	33.4	-1.3	28.8	-3.1
Annual hours emp., 500 more	65.4	-7.8	28.5	-6.2	17.9	-13.9
Germany						
Representative person	55.4	.	26.4	.	38.2	.
Rep.person, but all "are" men[1)]	58.7	.	24.5	.	43.0	.
- All "are" women[1)]	52.7	-6.0	27.7	3.2	33.4	-9.6
Age	55.8	0.4	25.9	-0.5	35.8	-2.4
Rep.person, but all "are" women[2)]	49.9	.	27.4	.	43.2	.
- All "are" singles[2)]	63.6	13.7	24.0	-3.4	29.3	-13.9
Children	56.9	1.6	38.7	12.3	43.2	5.0
Education, 1 year more	54.8	-0.5	18.7	-7.7	29.7	-8.5
Education, 5 years more	52.8	-2.5	2.9	-23.5	5.8	-32.3
Annual hours emp., 100 more	54.5	-0.9	25.9	-0.5	37.9	-0.2
Annual hours emp., 500 more	51.0	-4.4	24.1	-2.2	37.1	-1.1
USA						
Representative person	69.1	.	46.7	.	34.7	.
Rep.person, but all "are" men[1)]	72.4	.	52.3	.	35.6	.
- All "are" women[1)]	67.1	-5.3	43.7	-8.6	34.4	-1.1
Age	68.9	-0.2	46.4	-0.3	34.7	-0.1
Rep.person, but all "are" women[2)]	61.3	.	41.8	.	26.9	.
- All "are" singles[2)]	77.2	15.9	52.4	10.7	44.7	17.8
Children	72.9	3.7	50.8	4.1	39.7	4.9
Education, 1 year more	66.7	-2.5	42.5	-4.2	32.2	-2.5
Education, 5 years more	56.1	-13.0	27.0	-19.7	23.0	-11.7
Annual hours emp., 100 more	68.0	-1.1	44.9	-1.8	33.1	-1.6
Annual hours emp., 500 more	63.3	-5.9	38.1	-8.6	26.9	-7.8

Note: The marginal effects are evaluated at every individual observation and afterwards the sample averages of individual marginal effects are used.

1) All persons are first assumed to be men and with other characteristics unchanged. Then only gender status is changed to female and marginal effects are calculated.

2) All persons are first assumed to be living in a couple and with other characteristics unchanged. Then everybody are assumed to be singles and marginal effects are calculated.

Evaluating parameter estimates in this fashion give the figures in Table 5. The first column represents the estimated probability in year 1984+? *t*, given poverty in 1984. Column two represents the difference between the conditional probability of poverty for the representative person and for a person with one characteristic (the one in the text column) changed.

For the representative person the probability of poverty in 1985, given poverty in 1984 is estimated to be 73% in Denmark, 69% in the USA and 55% in Germany. These conditional estimated probabilities are almost identical to the aggregate numbers in Table 2. The calculated probabilities deviate somewhat from the aggregate figures for the USA in 1985 and 1991. The estimated conditional probability shows some decline over time, but only with a clear pattern for Denmark. Generally, the main part of the reduction takes place in the first about 6 years from 1985 to 1991.

This decline over time is also clearly present for low wages, except for the USA, see Deding (2000). Because poverty dynamics have not been analysed in this fashion, the results are here compared to Deding's (2000) low pay mobility study.

Being a woman has only an insignificant effect in all countries, except the USA in 1991. This is not necessarily surprising, because individual equivalent income depends on income at the family level, and therefore differences among men and women will only result if *single* men and *single* women have different risks of conditional poverty. The results for low pay are mixed with Danish men having higher risk, and with German men having a much lower risk than women.

Not living with a partner is usually insignificant, except in the USA in 1985 and 1997, where poverty risks are about 16-18 %-points higher for singles. This is contrary to low pay dynamics, where singles have a lower conditional risk of low pay. Marital status is likely to affect equivalent income for the following reason. If two persons with equal disposable incomes in the initial year are singles but the next year form a couple, they might, all else equal, change status from being poor to being non-poor. That's because the disposable income is doubled whereas the equivalence scale is only increased by less than 100 per cent because of economies of scale, when more people live together. The equivalent income is thus increased without changes in income, but with changes in marital status.

Age is included both in a linear and in a cubic fashion. The correlation coefficient between the two parameters is high, and this could be a reason for their insignificance in all countries. But including e.g. only the age linearly (not shown in Table), makes the parameter very significant in most estimations, therefore age is commented upon as if it were significant. The impact of age on the conditional poverty risk is small and has different signs in the short run in Denmark and Germany. But in 1997 a one year older Danish person has a 1.2 %-points lower poverty risk, while the effect is two times higher in the same direction in Germany. Age thus seems to have a rather high effect in the longer run in these two countries. Age has a negative effect in all years in the USA, but with a small magnitude: 0.1 to 0.3 %-points.

The number of children significantly increases the risk of repeated poverty with about 4-20 %-points in Denmark, 4-5 %-points in the USA, and 12 %-points in Germany (only 1991). This increase in poverty is not surprising, because with initially two adults and a child in a household, the disposable income must increase 15%[2] in order to maintain the equivalent disposable income prior to child-birth. The high effect in Denmark is surprising at first sight, but could be explained with the fact that Danish women have a high labour market participation rate (almost equivalent to men) compared to both Germany and the USA. Because maternity leave has not always, or actually rarely, had a 100 % replacement rate in Denmark, and because the replacement rate is 100% when the woman is not on the labour market, women's absence from the labour market could have a higher impact on family income in Denmark than in the other two countries. Danish men thus not only have to earn the 15% extra but also have to compensate for the foregone earnings of the wife, whereas a man with a *house*wife only needs to increase income with 15%. Deding's study does not show a similar positive effect across countries.

Education has predictable negative signs in all estimations, but the parameters are not always significant. An extra year of education reduces the Danish poverty risk with 6 %-points in 1991 and 4 %-points in 1997, which is large relative to the conditional poverty *level*. In Germany education reduces the poverty risk in 1991 (-8 %-points) and for the USA the parameter is significant in all estimations, and is between -2.5 and –4.2 %-

[2] When the number of persons is increased from 3 to 4, the equivalence scale, which is also the denominator, when calculating equivalent income, increases with 15%. This increase can be offset by an equivalent increase in disposable income, which is the nominator.

points. In the low pay analysis, Deding's estimated education parameters are all strongly significant.

The number of working hours also has a negative impact in all countries, but it is not significant in Germany.

The above results clearly indicate the importance of modelling the endogenous initial state. To illustrate the effect on parameters of wrongly assuming an exogenous initial state, estimations in Table A2 in the Appendix are made with ?=0. Generally, the parameters under the assumption of exogenous initial poverty are more often significant. For instance gender and age are generally insignificant in Table 4, but are often strongly significant in Table A2. Further, the parameters under the exogeneity assumption are also generally numerically greater than in the endogenous case. This study thus shows that modelling the endogenous initial state is also important for low income, the same conclusion applies for low wage, as seen in Stewart and Swaffield (1999) and Deding (2000).

Differences in poverty mobility or persistence between welfare states do not seem to have much relation to the regime choice in the three countries. Denmark belongs to the Scandinavian welfare model with a comprehensive social safety net, a publicly financed health care system and a universal right to pension with no regard to labour market status in the non-pension years. The USA has a more market-based system with lower benefits and quicker exhaustion. Germany is the middle case and belongs to the continental European model. The implication of this should be that social events like a child birth in the family, divorce, or labour market events such as more or less working hours should have less of an impact in Denmark and more impact in the USA, because the tax system and benefit structure in Denmark reduces inequality more compared to the USA. Looking across the effects in Table 5, there is not much indication of this.

An alternative way to detect welfare state differences is to use the estimated parameters to calculate counter factuals, where the "system" as expressed by the estimated parameters is imposed on another country. In Table 6 the diagonal element contain the average conditional probability in a country, when that country's own population and own estimated parameters from Table 4 are used. The off-diagonal elements are a combination of one country's own population and another country's parameters. The entries are the averages over initially poor individuals estimated conditional probability of poverty in 1985, given poverty in 1984.

Table 6. Estimated probability of poverty in 1985, given poverty in 1984, using 1984/1985-parameter estimates for other countries, %.

	Parameter estimate origin		
	Denmark	Germany	USA
Population origin:			
Denmark	73.2	60.3	79.1
Germany	67.2	55.4	73.3
USA	61.2	47.5	69.2

One first result is that the American structure or combined effects of e.g. having a child, being a woman, etc., would clearly increase poverty persistence among Danes, from 73.2% to 79.1%. If the population had the characteristics of German citizens and the system of the USA, then poverty persistence would increase by almost 20%-points. On the other hand, imposing the Danish or German structure on the American population would markedly reduce poverty persistence in the USA. With German estimates, the conditional probability of poverty would fall from 69% to 48%, and it would fall to 61% if Danish effects were imposed.

Imposing Danish rules on Germans increase poverty persistence in Germany, and German rules on Danish characteristics produce lower poverty persistence. This seems to favour the German system compared to the Danish system. This ordering of Germany and Denmark is also found for transitions between 1984 and 1991, but not for 1984/1997, see Tables A3 and A4 in the Appendix. Nevertheless it is clear that there is some indication that European welfare states cause poverty to have less effect over time.

On the other hand the European populations have characteristics that promote poverty persistence. If Europeans were in a more market based American system, then poverty persistence would be greater than in the USA. The opposite is the case with Americans, who would have lower persistence than the European countries, if they had the European system. These characteristics, e.g. number of hours of work and time spent in education, could be an adaptation to the welfare regime. Americans thus have favourable characteristics, because public intervention to sustain living standard is less than in

Europe, so that the Europeans have less incentive to e.g. work longer or join education for a longer time, than the Americans. That this is actually what seems to happen is apparent from Table A1 in the Appendix. On average the Americans have higher education than Danes or Germans in any year, and also works more during the year, but the contribution of public transfers to disposable income is in all years lower. There is also a dynamic dimension to this point, because initially poor Americans seem to “change“ characteristics over time in order to reduce poverty persistence. An indication of this is Table 7, where national samples have national parameter estimates from the 1984/1985-transitions imposed upon them.

Table 7. Estimated probability of poverty in 1985, given poverty in 1984, using parameter estimates from 1984/1985-transitions, but samples for all years, %

	Sample origin		
	1985	1991	1997
Parameter estimate origin:			
Denmark	73.2	70.9	69.7
Germany	55.4	52.9	58.0
USA	69.2	64.7	61.2

Using future populations[3] decreases poverty persistence over time, except in Germany in 1997, which might be due to the reunification. But the change in characteristics over time reduces poverty persistence to a higher degree in the USA compared to Denmark or Germany.

It is of course debatable whether the estimates represent a whole welfare system, because the behaviour of individuals is not only affected by tax rules, public transfers and other public interference, but is also affected by the development on the labour market, macro economic impulses and effects from other countries. But even the labour market is regulated with relatively high minimum wages and high unemployment benefits in

[3] The future populations are older, but that will be the case in every country, so there has been no adjustment for that.

Europe, which e.g. avoids the working poor phenomenon existing in the USA. And this benefit structure actually also helps in avoiding some negative effects of globalisation, that implies export of low productivity jobs and import of cheaper goods, causing higher domestic unemployment. It can thus be argued that the parameters do resemble or at least are affected by the national choice of welfare regime.

6. Conclusion

The probability of poverty in year *t+? t*, conditional on poverty in year *t*, is much higher than the probability in year *t+? t*, given non-poverty in year *t*. The initial state therefore seems to affect future risks of poverty, indicating some persistence in poverty. Over time (higher *? t*) poverty persistence is reduced.

The observed poverty persistence pattern is partly explained by a number of explanatory variables. Age, the number of children, education and the number of working hours have some impact on poverty persistence, whereas gender and whether being single or not, is generally less important.

The study shows that not taking the endogeneity of the initial state into account can seriously distort estimated parameters. This seems clear, when comparing results with and without assuming endogenous initial poverty state. Specifically, the parameters under the exogeneity assumption are generally inflated and more often very significant.

At the aggregate level, mobility seems to be higher in Germany and Denmark than in the USA, but with some convergence over time. Looking at parameters or estimated effects, there is no clear pattern of differences among welfare state choices. But combining populations and estimated parameters of different countries to calculate counter factual conditional probabilities of poverty, lends some support to the view that the European system (represented by Denmark and Germany) promotes poverty persistence, whereas the system in the USA changes individual behaviour in order to protect themselves from poverty persistence.

References

Aaberge, R., A. Björklund, M. Jäntti, M. Palme, P. Pedersen, N. Smith and T. Wennemo (1996): „Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States", Statistics Norway, Research Department, Discussion Paper, no. 168

Adelman, I. et al (1994): "Estimating Income Mobility from Census Data". Journal of Policy Modeling. vol. 16. no. 2

Atkinson, A.B., F. Bourguinon and C. Morrisson (1992): "Empirical Studies of Earnings Mobility", Harwood Academic Publishers, Chur CH

Atkinson, A.B., L. Rainwater, T. Smeeding (1995): "Income Distribution in European Countries" in A.B. Atkinson's "Incomes and the welfare state: Essays on Britain and Europe", Cambridge

Buhmann, B. (1988): "Equivalence Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates across Ten Countries Using the Luxembourg Income Study (LIS) Database", Review of Income and Wealth, vol. 34, no. 2

Burgess, S. and C. Propper (1999): "An Economic Model of Household Income Dynamics, With an Application to Poverty Dynamics Among American Women", University of Bristol, Centre for Economic Policy Research, Discussion Paper 1830

Burkhauser, R.V., B.A. Butrica and M.C. Daly (1995): "The Syracuse University PSID-GSOEP Equivalent Data File: A Product of Cross-National Research", All-University Gerontology Center, Maxwell School of Citizenship and Public Affairs, Syracuse University

De Vreyer, P., R. Layte, M. Wolbers and M.A. Hussain (2000): "The Permanent Effects of Labour Market Entry in Times of High Unemployment" in D. Gallie and S. Paugam's "Welfare Regimes and the Experience of Unemployment in Europe", Oxford University Press, London, UK

Deding, M., P.J. Pedersen and N. Smith (1999): “Modelling the Determinants of Poverty“, Department of Economics. Aarhus School of Business. Denmark. Draft

Deding, M. (2000): “Mobility of Low-Wage Earners in Denmark, Germany and the United States”, Ph.D. dissertation, Institute of Economics, University of Aarhus

Feenberg, D. and E. Coutts (1993): “An Introduction to the TAXSIM Model”, Journal of Policy Analysis and Management, vol. 12, no. 1

Förster, M.F. (1994): “The Effects of Net Transfers on Low Incomes among Non-elderly Families”, OECD-Economic-Studies, no. 22

Gallie, D. and S. Paugam (2000): “Welfare Regimes and the Experience of Unemployment in Europe”, Oxford University Press, London, UK

Gittleman, M. and M. Joyce (1999): “Have Family Income Mobility Patterns Changed?”, Demography, vol. 36, no. 3

Greene, W.H. (2000): “Econometric Analysis”, fourth edition, Prentice Hall International, Inc., New Jersey

Gustafsson, B. (1994): “The Degree and Pattern of Income Mobility in Sweden”, Review of Income and Wealth, vol. 40, no. 1

Gustafsson, B. (2000): “Poverty in Sweden: Changes 1975-1995, Profile and Dynamics” in B. Gustafsson and P. Pedersen (eds.): “Poverty and Low Income in the Nordic Countries”, Ashgate Publishing Ltd., England

Hauser, R. and H. Fabig (1999): “Labor Earnings and Household Income Mobility in Reunified Germany: A Comparison of Eastern and Western States”. Review of Income and Wealth, vol. 45, no. 3

Heckman, J.J. (1981): “The incidental parameters problem of initial conditions in estimating a discrete time-discrete data stochastic process” in C.F. Manski and D.

McFadden's (eds.) "Structural Analysis of Discrete Data with Econometric Applications, Cambridge, MIT Press

Hill, M.D. Hill and R. Walker (1998): "Intergenerational dynamics in the USA: poverty processes in young adulthood" in L. Leisinger and R. Walker: "The Dynamics of Modern Society". The Policy Press

Marx, I. and B. Nolan (1998): "Low Pay and Household Poverty". paper at the LoWER-conference in Groningen

Pedersen, P. and N. Smith (1998): "Low Incomes in Denmark. 1980-1995". Centre for Labour Market and Social Research. Working Paper no. 19

Schluter, C. (1997): "On the Non-Stationarity of German Income Mobility", STICERD, London School of Economics Discussion Papers Series DARP/30

Stewart, M.B. and J.K. Swaffield (1999): "Low Pay Dynamics and Transition Probabilities", Economica, no. 66

Trede, M. (1999): "Statistical Inference for Measures of Income Mobility". Jahrbucher für Nationalökonomie und Statistik, vol. 218, no. 3-4

Appendix

Table A1. Sample statistics for data used in estimations in Tables 5 and A2

	Year *t*+? *t* =1985				Year *t*+? *t* =1991				Year *t*+? *t* =1997			
	Mean	St.dev.	Min.	Max.	Mean	St.dev.	Min.	Max.	Mean	St.dev.	Min.	Max.
Denmark												
Poor in 1984 indicator	0.07	0.25	0	1	0.06	0.24	0	1	0.06	0.24	0	1
Poor in 1984+? *t* indicator	0.06	0.25	0	1	0.04	0.20	0	1	0.05	0.21	0	1
Woman	0.50	0.50	0	1	0.50	0.50	0	1	0.50	0.50	0	1
Single	0.25	0.43	0	1	0.23	0.42	0	1	0.25	0.43	0	1
Age	41.45	9.34	27	59	45.24	7.45	33	59	48.93	5.75	39	59
No of children	0.88	1.00	0	6	0.76	0.98	0	5	0.49	0.82	0	5
Education, years	11.39	2.50	9.0	18.0	11.55	2.54	9.0	18.0	11.81	2.55	9.0	18.0
Hours employed/100	11.76	7.35	0.0	54.2	11.07	6.89	0.0	36.1	11.66	7.26	0.0	38.8
Unempl. rate at labour market entry	3.31	2.01	1.0	11.0	2.91	1.68	1.0	10.7	2.58	1.65	1.0	10.7
Transfers as % of disp. Income	8.11	23.74	0	100	7.16	22.39	0	100	7.04	22.78	0	100
Germany												
Poor in 1984 indicator	0.05	14.10	0	1	0.05	15.52	0	1	0.05	16.29	0	1
Poor in 1984+? *t* indicator	0.07	15.89	0	1	0.05	15.79	0	1	0.08	20.72	0	1
Woman	0.52	31.46	0	1	0.52	35.31	0	1	0.52	37.22	0	1
Single	0.22	26.12	0	1	0.24	29.96	0	1	0.27	33.24	0	1
Age	42.93	591.28	27	59	45.69	544.10	33	59	49.33	459.43	39	59
No of children	0.81	62.07	0	7	0.77	73.03	0	8	0.58	70.87	0	6
Education, years	11.59	174.67	7.0	19.0	11.75	206.59	7.0	19.5	11.83	218.47	7.0	19.0
Hours employed/100	14.51	693.37	0.0	50.9	15.50	730.52	0.0	52.0	14.32	773.94	0.0	52.0
Unempl. rate at labour market entry	2.77	127.90	0.3	6.9	2.32	142.23	0.3	6.9	1.56	99.09	0.3	6.9
Transfers as % of disp. Income	6.46	997.94	0	100	7.02	1092.08	0	100	7.79	1198.79	0	100
USA												
Poor in 1984 indicator	0.12	1.33	0	1	0.12	1.42	0	1	0.11	1.31	0	1
Poor in 1984+? *t* indicator	0.12	1.31	0	1	0.11	1.33	0	1	0.09	1.20	0	1
Woman	0.53	2.05	0	1	0.54	2.16	0	1	0.53	2.08	0	1
Single	0.26	1.79	0	1	0.28	1.94	0	1	0.28	1.87	0	1
Age	40.46	39.44	27	59	43.81	32.41	33	59	47.71	23.44	39	59
No of children	1.09	4.92	0	8	0.99	5.11	0	8	0.71	4.27	0	6
Education, years	13.05	10.39	2.0	17.0	13.24	10.54	3.0	17.0	13.51	9.76	3.0	17.0
Hours employed/100	16.55	38.63	0.0	58.1	17.09	40.85	0.0	57.1	17.08	39.32	0.0	52.5
Unempl. rate at labour market entry	5.02	7.72	1.2	20.1	5.08	5.50	1.2	9.9	5.23	5.18	2.9	8.3
Transfers as % of disp. Income	4.59	63.23	0	100	4.60	67.47	0	100	3.69	54.23	0	100

Table A2. Parameter estimates from probit model, modelling the probability of poverty in year *t+? t*, only with the sample who was initially poor in year *t*, *t*=1984, *? t* ={1,7,13}

	Year *t+? t* =1985		Year *t+? t* =1991		Year *t+? t* =1997	
	Estimate	St. error	Estimate	St. error	Estimate	St. error
Denmark						
Intercept	1.6486	1.5526	6.3875	3.6918	-1.1226	8.3141
Woman	0.1895	0.1527	0.2318	0.2010	-0.0724	0.2363
Age	-0.0179	0.0737	-0.1673	0.1628	0.1735	0.3423
$Age^2/100$	0.0278	0.0880	0.1618	0.1808	-0.2310	0.3506
Single	-0.2422	0.1534	-0.7832	0.1886	-0.1969	0.2499
No of children	0.0738	0.0762	0.3504	0.1200	0.6437	0.2023
Education, years	-0.0490	0.0302	-0.2264	0.0568	-0.1595	0.0631
Hours employed/100	-0.0534	0.0119	-0.0502	0.0164	-0.1328	0.0234
Log likelihood	-204.5		-126.9		-83.6	
No of observations	382		254		193	
Germany						
Intercept	0.9042	0.0283	-2.4569	0.0619	17.5634	0.2094
Woman	-0.3370	0.0027	-0.0594	0.0034	-0.9327	0.0052
Age	-0.0050	0.0013	0.0970	0.0027	-0.6177	0.0085
$Age^2/100$	0.0197	0.0016	-0.0732	0.0030	0.6646	0.0085
Single	0.4734	0.0027	0.3244	0.0033	0.4854	0.0056
No of children	-0.1242	0.0013	0.0295	0.0015	0.1243	0.0025
Education, years	-0.0461	0.0005	-0.0530	0.0006	-0.2936	0.0015
Hours employed/100	-0.0333	0.0001	-0.0377	0.0002	-0.1382	0.0005
Log likelihood	-753,636.1		-478,551.0		-182,741.3	
No of observations	342		204		128	
USA						
Intercept	2.0623	0.2717	6.8251	0.5475	14.0121	1.5464
Woman	-0.3913	0.0299	-0.4357	0.0325	-0.5086	0.0509
Age	-0.0398	0.0134	-0.2127	0.0250	-0.5421	0.0647
$Age^2/100$	0.0484	0.0161	0.2297	0.0281	0.5665	0.0674
Single	0.7684	0.0295	0.6870	0.0313	1.1560	0.0529
No of children	0.1954	0.0115	0.0988	0.0129	0.3071	0.0234
Education, years	-0.0637	0.0052	-0.1492	0.0062	-0.1189	0.0091
Hours employed/100	-0.0424	0.0014	-0.0551	0.0015	-0.0601	0.0024
Log likelihood	-6,617.5		-4,999.2		-2,112.6	
No of observations	1,102		729		292	

Table A3. Estimated probability of poverty in 1991, given poverty in 1984, using 1984/1991-parameter estimates for other countries, %.

	Parameter estimate origin		
	Denmark	Germany	USA
Population origin:			
Denmark	34.6	28.6	66.1
Germany	31.2	26.7	53.9
USA	28.6	23.3	46.7

Table A4. Estimated probability of poverty in 1997, given poverty in 1984, using 1984/1997-parameter estimates for other countries, %.

	Parameter estimate origin		
	Denmark	Germany	USA
Population origin:			
Denmark	31.9	40.9	51.6
Germany	28.1	38.2	43.7
USA	20.6	33.8	34.8

Figure A1. Poverty Rates, 18-74 years, 1985-1997, %