Economic incentives and individuals' choice between welfare programmes and work in Denmark

Martin Rasmussen

Welfare systems and policies
Working Paper 7:2004

Working Paper

Socialforskningsinstituttet
The Danish National Institute of Social Research

Economic incentives and individuals' choice between welfare programmes and work in Denmark

Martin Rasmussen

The Danish National Institute of Social Research, Herluf Trolles Gade 11, DK-1052 Copenhagen K, Denmark [1]

Abstract

We estimate the effect of welfare benefits and wages on individuals' choice between working or collecting one of three welfare programmes. We compare the magnitude of transitions between various welfare programmes with transitions between, say, work and disability benefit. We use simulation methods to estimate random parameters. Estimation results show significant effects of economic incentives and significant variations of estimated parameters. Experiments with the estimated model show that transitions within welfare programmes are important relative to transitions between such programmes and work.

Keywords: Disability benefit, Simulation methods, Random parameters, Labour supply

JEL classification: I38, J22

[1]Email address: mar@sfi.dk.. Phone: +45 33 48 09 10. Fax.: +45 33 48 08 33.

1. Introduction

We estimate empirically the effect of individuals' economic incentives of disability benefit or other types of welfare benefits in Denmark on individuals' 'choice' to participate in welfare programmes, to work or not to participate in any of the these 'states'. We use a multi state model to be able to compare transitions between various welfare programmes with transitions between work and (any kind of) social programmes. We use a random parameters model to describe unobserved heterogeneity. This gives a richer and policy-relevant description of transitions between states and we argue that it helps to take account of potential bias of the parameter of interest related to economic incentives. As a third feature of the model, we tried (with little success) to model individuals' eligibility for various states. The focus of the paper is especially on the disability benefit programme. This was however initially more pronounced because I tried to estimate eligibility for disability benefit.

We use panel data information about individuals in Denmark from 1992 to 1998. In the remainder of the introduction, we motivate in turn the three characterizing features of the study (many states, random parameters and eligibility constraints) and review related existing literature.

The motivation for using many states is as follows: The main incentive problem related to public benefits is the possible distortion of the labour supply-decision so that mutually gainful trades are not carried out (some unemployed would like to work at the wage rate they can obtain had it not been for the loss of benefits). Some empirical studies estimate participation in a public benefit programme using the benefit rate for this programme as an explanatory variable. In this case, a mutual increase of the benefit rate and participation might arise due to people moving to the programme studied from another public programme rather than from work. These kinds of problems are avoided using our model.[2]

We use simulation methods to estimate random parameters. The random parameters describe some of the unobserved variation in preferences (unobserved heterogeneity) otherwise not captured. One purpose of describing unobserved heterogeneity is as follows: suppose, for example, that people who might apply for disability benefit also have preference for early retirement programmes, and suppose further that the early retirement programme becomes less favourable. In that case, many potential applicants for early retirement might alternatively apply for disability benefit, so a policy reform of the early retirement scheme will have spill over effects on participation in the disability benefit programme. Estimates of unobserved correlations of preferences for these states therefore put light on such spill over effects. Another purpose with estimating random effects is that it might help to

[2] Other studies consider participation in the labour force (employed and unemployed) as explained by e.g. the benefit rate of the disability benefit scheme. Again, an increase of disability beneficiaries might be mirrored in fewer unemployed. Of course, a simpler solution to the problems described is to consider employment as a function of the disability benefit rate. In this case, it is reasonable to attribute a mutual increase of the benefit rate and a decrease of the employment rate to transitions between the two states. Our five-state model however gives a richer description of transition between various states.

Furthermore, for Denmark, the immediate link between employment and disability benefit is weak – the typical disability benefit is awarded years after loss of employment. This makes the estimation of economic incentives difficult in a two-state model (employment and disability benefit). With multiple states, we may estimate the choice of employment versus any welfare programme.

correct potential bias of the parameter related to economic incentives. However, we cannot with certainty claim that the estimated 'incentive-parameter' is unbiased. There are no large discrete changes in benefit rates or wages rates in the period we study and it is therefore difficult to take account of possible unobserved fixed effects with traditional fixed effects methods. Expected wage rates are calculated as averages or with OLS. Almost – but not exactly – the same variables that are used to estimate wages are used as covariates in the latent variables.

We use simulation methods to estimate random parameters because they are supposed to be computationally easier to use than numerical integration methods.[3] Another alternative is to use mass-points distributions to estimate random parameters.

Modelling of 'eligibility' is potentially important. For example, to be able to obtain disability benefit an individual has to be eligible, i.e. to be assessed to have a permanently reduced working capacity, and many policy initiatives are concerned about changing the eligibility criteria for e.g. disability benefit. It is also interesting to study "eligibility" to the labour market, since it offers a route to disentangle the effect of individuals' incentives and "rationing" via a low number of job offers. To anticipate results, I did not have success with estimating eligibility. I present a single estimation with deterministic modelling of eligibility criteria for early retirement benefit and disability benefit.

[3] Nevertheless, the estimations in the paper are computationally very burdensome – it takes about a month to for the estimations to converge even with stating values close to optimum.

The number of studies about disability benefit and incentive effects is large. We now survey some papers on the subject with special emphasis on the problems of identifying the effect of economic incentives.

In a number of papers from the 1980s, economic incentives in relation to disability benefit are estimated and the problems related to using cross-individual information of wages are discussed. Parson (1980, 1982) finds significant effects of replacement rates on non-labour-force-participation for individuals with poor health. Haveman and Wolfe (1984) note that the method may give biased estimates due to unobserved heterogeneity. Using selection methods, they seek to avoid such bias. They still find significant economic effects but the selection problems are found to affect parameters dramatically. Haveman, Wolfe and Warlick (1988) extend the model to a three-state model (disability benefit, other early retirement and labour force participation), use a careful description of income, and now find insignificant economic parameters. Parson (1982), however, notes that the use of explanatory variables did not affect 'economic' parameters a lot and therefore he doubts that unobserved heterogeneity is very important. Also, Parson (1980) tests the model estimated for 1969 by successfully tracking the actual aggregate time series.

A few recent papers use combinations of aggregate data and individual data that make it easier to take account of unobserved heterogeneity. Furthermore, some of these studies make use of large, discrete, politically controlled changes of disability benefits rates (in Canada) or large changes in earnings in certain parts of the US labour market (a coal boom in certain states) to estimate incentive-parameters. Gruber (2000) and Campolieti (2001a,b) find significant effects of discrete changes of benefit rates in Canada on the labour force participation (Campolieti) or the employment rate (Gruber). Black et al. (2002) estimate the effect of (potential) labour market income on disability benefit expenditures. Their data set is rather ingenious: they use coal prices as indicator of potential labour market earnings. They use a panel of US counties with important cross-section information (coal-counties versus non-coal-counties) and within-information arising from time-series changes of the coal prices. They find significant economic effects. Autor and Duggan (2003) use panel data for US states. They take good care to calculate benefits rates and finds significant effects from replacement rates. They also consider the effects of general labour market shocks and distinguish between effects for unemployed and employed.

Few studies for Denmark focus on disability benefit. Christiansen (2000) use data aggregated to national levels to study effects of policies relating to the behaviour of awarding authorities (i.e. in particular relating to municipalities in Denmark). Weatherall (2002) finds disability awards to correlate with annual wage income.

Simulation methods are described in Hajivassiliou and Ruud (1994). That paper describes simulations methods in general as well as the particular method we use in this paper (maximum simulated likelihood). Examples of use of simulation methods are Hyslop (1999) and McFadden and Train (2000).

The paper is organized as follows. The next section is about data. Section 3 is about the estimation method and section 4 contains results. Section 5 concludes.

2. Data

Data used in the paper are from registers from Statistics Denmark. These registers cover the whole population. We use a sample that covers the years 1992-1998. The estimations are made with a random sample of 9,289 individuals followed through the years 1990 to 1998 and aged 41-65 in 1998. The sample therefore covers the age groups with the highest propensity to be awarded disability benefit or early retirement benefit. Appendix 3 and Rasmussen (2004) contains a description of the policy context in general.

Observations of people who obtained disability benefit or early retirement benefit the previous year are excluded from the sample used in estimations. Hence, a person who is awarded obtains disability in 1995 is included in the estimations only until that year. The reason for this is pragmatic: very few people leave these states for other reasons than transition to old age pension (or death). Therefore, the interesting question is whether or not people enter these states.[4]

[4] Of course, it is also a very interesting question why people never leave e.g. disability benefit, but since there are very few observations on this in the sample, surveys question are more appropriate for this issue.

Data is described in detail in the appendix 1. Below, we will primarily and briefly describe the construction of 'states' and the 'income rates', i.e. the wage rate and the welfare benefit rates.

We classify individuals into 'states'. 'Disability beneficiaries' are those who collect the public disability benefit for more than half of the year. 'Early retired' are people who collect the public early retirement benefit for more than half of the year. 'Unemployed' are people who collect a significant amount of certain benefits intended for temporary income support (e.g. unemployment insurance benefit or social assistance but not e.g. benefit during maternal leave.) 'working' individuals are employed at least a quarter of the year or earn above a certain amount from self-employment. Of course, according to these definitions, some individuals are classified in two states. To classify each individual (each year) in exactly one state, the states have been given priority according to the sequence they are mentioned above. Those individuals not put in any of the four states are classified in the residual state labelled 'home'.

Benefit rates are calculated as averages across recipients. The wage rate is estimated with OLS (see section 3). We use these 'income rates' as expected values. For example, for a person not collecting disability benefit, the calculated average is the income she expects if awarded disability benefit. See appendix 1 for details about the calculation of these measures.

Other explanatory variables used are: age, gender, cohabitation, education, health indicators, and time trends.

3. The econometric model

The base model is a multinomial logit-model. Equation (1) displays the latent variable I_{sit} for individual i in state s in period t. This is a function of the economic variable, y (see below), observed individual characteristics, x, unobserved characteristics, η, which is assumed constant over time, and error term, e. The β's are related parameters to be estimated.

$$I_{sit} = \beta_y y_{sit} + \beta_{sx} x_{it} + \eta_{si} + e_{sit} \tag{1}$$

The state 'work' is the comparison state. In (1), the variable of interest, y, is 'relevant benefit rate minus wage rate'. As mentioned in the introduction, the x-variables used in (1) are almost – but not exactly – the same as used in the wage equations. The parameter β_y is therefore identified via the small differences in the set of variables and the minor changes of benefit rates that occur over time. Variables used in the wage equation are: an annual general growth variable, difference between general and regional growth, four education groups, a time trend, five age groups (<40, 40-44, 45-49, 50-54, 55-59, 60-65), gender, two variables for employment status the previous year (one dummy for being employed and one for no observation – this relates to estimation for 1989). Variables in the x-vector are: growth in each region (the sum of the two growth variables mentioned above), four educational groups, four variables combining gender and cohabitation, a time trend, three age groups (<50, 50-59, 60-65) and the age (age groups are not in the early retirement latent variable, because the programme is not open for young people), work status the previous year.

The variation of random parameters, η, across individuals is estimated using simulation methods. We decompose the vector of unobserved random parameters η into another vector of unobserved independent standard normal random variables, v, and a lower triangular matrix ρ, so that $\eta = \rho v$. The covariance matrix of η is $\mathrm{cov}(\eta) = \rho\rho'$. For each individual, we make R draws, $r = 1,...,R$ of v. If the value of v for individual i, v_i, were known, the true likelihood contribution from the individual would be $L(v_i)$. The approximate likelihood contribution used in estimations is (see appendix 4 for details about the method)

(2) $$\tilde{L}_i = \frac{1}{R}\sum_{r=1}^{R} L(v_{i,r})$$

4. Results

First we present the model with deterministic parameters with and without modelling eligibility. We then include eligibility equations, and finally present estimations with random parameters (without modelling of eligibility).

Deterministic parameters

Table 1 shows a single result from a static model with deterministic parameters, namely the estimate of β_y.

Table 1. Economic incentives and social states. Result from a model without modelling of eligibility.

Variable	Latent variable	Parameter	Standard deviation
Income rates	All latent variables	0.953	0.124

Source/notes: Registers from Statistics Denmark. Own calculations. We use a simple multinomial logit-model as in equation (1) with $\eta = 0$.

The income-parameter is strongly significant.

As mentioned above, we cannot be certain that income measures are not correlated with the error term and hence the parameter β_y might be biased. A way to avoid or at least reduce such bias is to include values of y for other periods, for example initial values of y. With this method, β_y is identified from the dynamic variation of wages, while the effect from cross-individual variation of wages is captured by the initial value of y. I have tried this method with no success. The parameter related to initial-period-y have the expected sign, but the parameter related to current-period-y has the unexpected sign. I presume the poor results were due to little dynamic variation of wages. Part of the motivation for estimating random parameters (below) is to take account of some of the potential bias of β_y. This motivation should be seen in light of these bad results with the most direct method to take account of variation between wages and error terms.

As mentioned in the introduction, some effort has been done to model and estimate 'eligibility' to various states in order to disentangle individuals' preferences from their options. Appendix 2 describes in more details the estimations I have tried. We end up with simple, deterministic 0-1-indicators that describe eligibility for each individual for each state. The description is as follows: all individuals can enter 'work', 'unemployment' and 'home'. Only individuals who are 'sick' can enter 'disability benefit'. A person is 'sick' if the yearly number of visits to the family doctor is greater than five. A person is eligible for 'early retirement' is the person has reached 60 years of age or is between 50 and 59 years of age in the years 1994-1996 and was unemployed the year prior to early retirement: the early retirement scheme was open for this group of people as a temporary labour market policy (see appendix 3).

Table 2 describes the result of the estimation with these simple eligibility rules.

Table 2. Economic incentives and social states. Result from a model with modelling of eligibility.

Variable	Latent variable	Parameter	Standard deviation
Income rates	All latent variables	1.046	0.125

Source/notes: See table 1.

The parameter is only slightly different from that in table 1. (The parameters related to health and age in utility equations (1) are very different in the models presented in table 1 and 2.)

A number of experiments are done with the model presented in table 2 in order to interpret the results. For each person in the sample we calculate probability of being in each of the five states. Table 3a-b shows the effect of the experiments on the average across individuals of the probabilities for being in each of the five states.

Table 3a. Experiments with the model in table 2, results on absolute proportions.

	Effect on probability for state					
Effect of:	'Disability'	'Early retirement'	'Unemployment'	'Work'	'Home'	Sum of all states
A 10% increase of[1] the ..						
.. disability benefit rate'	0.000436	-0.000041	-0.000142	-0.000080	-0.000174	0.00000
.. early retirement benefit rate	-0.000044	0.000883	-0.000213	-0.000203	-0.000423	0.00000
.. benefit in unemployment	-0.000173	-0.000238	0.003642	-0.001279	-0.001952	0.00000
.. wage rate	-0.000291	-0.000701	-0.003889	0.012177	-0.007296	0.00000
All four above	-0.000081	-0.000113	-0.000692	0.010616	-0.009728	0.00000
Removal of special opening of early retirement[2]	0.000386	-0.009880	0.002905	0.001578	0.005011	0.00000
5% more people become ill[3]	0.000607	0.000044	0.000305	-0.000882	-0.000075	0.00000
Initial distribution[4]	0.013241	0.029082	0.121007	0.648128	0.188541	1.00000

[1] In the estimated model, the used explanatory 'income rates' are measured as 'relevant benefit rate minus wage rate'. The experiments increase the particular benefit rate or the wage rate with 10%. (Hence the changes of the 'income rates' are not 10%. In the experiment with all benefit rates and the wage rates increased by 10% 'benefit rates minus wage rates' increases numerically.)

[2] The dummy for eligibility to early retirement becomes equal to a dummy for having reached 60 years of age. In the base-model, the dummy is also 1 for people aged 50-59 years during 1994-1996, who were unemployed the previous year. Calculations only use observations for the year 1996.

[3] The sickness-indicator is changed from 0 to 1 for 5 percent of the population for whom the sickness-indicator is observed equal to 0.

[4] The distribution differs from that in table A3 in the appendix 1. This is because we here exclude observations of people who obtained disability benefit or early retirement the previous year (see section 2).

Table 3b. ..., results as proportions of initial distribution.

	Effect on probability for state					
Effect of:	'Disability'	'Early retirement'	'Unemployment'	'Work'	'Home'	Sum of all states
A 10% increase of ...						
The disability benefit rate'	0.0329	-0.0014	-0.0012	-0.0001	-0.0009	0.0000
Early retirement	-0.0033	0.0304	-0.0018	-0.0003	-0.0022	0.0000
Benefit in unemployment	-0.0131	-0.0082	0.0301	-0.0020	-0.0104	0.0000
Wage rate	-0.0220	-0.0241	-0.0321	0.0188	-0.0387	0.0000
All three above	-0.0061	-0.0039	-0.0057	0.0164	-0.0516	0.0000
Removal of special opening of early retirement	0.0292	-0.3397	0.0240	0.0024	0.0266	0.0000
5% more people becomes ill	0.0458	0.0015	0.0025	-0.0014	-0.0004	0.0000

The effects of changes of income rates are modest. For example, the 10 percent increase in the disability benefit rate increases the proportion of people awarded disability benefit with 0.000463. This is equivalent to an increase of 3.29 percent of 'initial' disability beneficiaries, i.e. an elasticity of approximately 0.3. Perhaps more interesting is that only a small fraction of new disability beneficiaries would otherwise have been working according to the model (0.000080 out of 0.000436 or 18 percent).

By the face value, these experiments therefore suggest small incentive problems related to an increase of the generosity of social security programmes. Two reasons for being cautious about this immediate conclusion are (assuming the model and the estimated parameters are true): As concerns the effect of increase in disability beneficiaries, the experiment shows a kind of first-year-effect. This is so, because as explained above, the estimation sample consists of people not collecting disability benefit or early retirement benefit in the previous year (see section 2). Hence, the proportion 0.000436 is the people entering disability benefit the first year after an increase of the benefit rates. However, neither is the interpretation that the effect continues forever correct. A second modification is that the effect of finance the increased public expenditures via increases of taxes is not included in the experiment. This would presumably decrease incentives to supply labour.

Removal of the special opening of the early retirement programme (for long term unemployed aged 50-59 years in 1994-1996) reduces the number of people who enter the programme by 0.009880, which is a substantial proportion (a third) of the 'initial' entrants. Surprisingly few would alternatively 'choose' disability benefit (only 386/9880=3.9%), and many would alternatively choose 'home' or work.

If 5 percent of the people who are not 'sick' (empirically measured as an indicator for more than five visits a year to the family doctor) become so, a proportion of those working choose disability benefit. The number of awards of disability benefits increases by 4.6 percent.

Random parameters

We now turn to models with random parameters.

In table 4 we use the same deterministic parameters as in table 1. Random parameters capture the variance of the constant terms of the utility equations. Latent variables are

$$
(3) \qquad
\begin{aligned}
I_{sit} &= \beta_y y_{sit} + \beta_{sx} x_{it} + c_s + \eta_{si} + e_{sit}, \quad \text{where } s = D, R, U, H \\
\eta_{Di} &= \rho_D v_{Di} \\
\eta_{Ri} &= \rho_R v_{Ri} + \rho_{RD} v_{Di} \\
\eta_{Ui} &= \rho_U v_{Ui} + \rho_{UD} v_{Di} + \rho_{UR} v_{Ri} \\
\eta_{Hi} &= \rho_H v_{Hi} + \rho_{HD} v_{Di} + \rho_{HR} v_{Ri} + \rho_{HU} v_{Ui}
\end{aligned}
$$

Terms capturing pair wise covariance of the random constant terms are included.

Table 4 Economic incentives and social states. Result from a model with random constants, and without modelling of eligibility.

Variable	Parameter	Standard deviation
Income rates	0.322	0.147
ρ_D	0.314	0.106
ρ_R	0.753	0.198
ρ_U	-0.844	0.058
ρ_H	-0.090	0.085
ρ_{RD}	-1.609	0.095
ρ_{UD}	-0.737	0.085
ρ_{UR}	0.413	0.162
ρ_{HD}	0.452	0.063
ρ_H	-0.391	0.087
ρ_{HU}	0.250	0.068

Source/notes: See table 1. We use $R = 50$ draws in the simulations.

Except one, all ρ's are significant. The inclusion of random parameters leaves the income parameter slightly lower. Tables 5 and 6 below present standard deviations and correlation coefficients of the random constants η from table 4 using $\mathrm{COV}(\eta) = \rho\rho'$.

Perhaps surprisingly, the estimate of the incentive-parameter decreases significantly compared to table 1, and it is now on the margin of being significant.

Table 5. Standard deviations of random constant terms as implied by ρ.

State	Standard deviation
Disability	0.31
Early retirement	1.78
Unemployment	1.19
Home	0.65

Table 6. Correlation coefficients of random constant terms as implied by ρ.

State	Disability	Early retirement	Unemployment	Home
Disability	1	.	.	.
Early retirement	-0.91	1	.	.
Unemployment	-0.62	0.71	1	.
Home	0.69	-0.88	-0.90	1

Table 6 shows that individuals with high preference for 'disability' also tend to have preference for 'home' but preference against 'early retirement'. Hence, to return to the example mentioned in the introduction, if the early retirement programme is made less attractive (a current debate in Denmark), the *unobserved* tastes will not lead many people to collect disability benefit as an alternative to early retirement. To evaluate the importance of these correlations, table 7 presents effects on odds ratios of an increase by one-standard-deviation of constant terms.

Table 7. Importance of variation of constant terms. Effect on odds ratios of increase by one-standard-deviation of random constant terms.

	Effect from one-standard-deviation increase of constant term in state:			
Effect on OR for state:	'Disability'	Early retirement	'Unemployment'	'Home'
'Disability'	1.36	0.20	0.48	0.64
'Early retirement'	0.75	5.93	2.33	0.56
'Unemployment'	0.83	3.54	3.29	0.56
'Home'	0.81	0.21	0.34	1.92

Note: For example, $0.75 = \exp\left[\text{corr}(\eta_R, \eta_D) \cdot \text{std.dev}(\eta_d)\right] = \exp(-0.91 \cdot 0.31)$.

Hence, the table shows very important impacts from the variation of the constant terms. This is true for the impact of a random constant term in a particular state to that particular state (the diagonal of the table), but also between states.

Results of experiment of changes in the benefit rates and wage rates are shown in table 8.

Table 8. Experiments with the model in table 4, results on absolute proportions.

	Effect on probability for state					
Effect of:	'Disability'	'Early retirement'	'Unemployment'	'Work'	'Home'	Sum of all states
A 10% increase of the ..						
.. disability benefit rate'	0.000173	-0.000015	-0.000041	-0.000037	-0.000079	0.000000
.. early retirement	-0.000017	0.000266	-0.000084	-0.000061	-0.000104	0.000000
.. 'unemployment' benefit rate	-0.000051	-0.000093	0.000841	-0.000374	-0.000322	0.000000
.. wage rate	-0.000136	-0.000212	-0.001150	0.003795	-0.002296	0.000000
All four above	-0.000033	-0.000056	-0.000441	0.003323	-0.002794	0.000000

Comparing table 8 with 3a, the effects of increases in income rates are generally low – measured e.g. at the diagonal of the tables. This is explained by the lower estimate of the parameter to the income rates. The co-variation of unobserved constant terms has an impact on the experiments. Consider for example an increase of the early retirement benefit. The effect on the proportion of early retired is smaller than in table 3a (0.000266 vs. 0.000883, or 30 percent) but the effect on 'home' is disproportionately lower (-0.000104 vs. –0.000423, or 25 percent) due to estimation of a negative correlation between unobserved tastes of the two states.[5]

The estimations above take account of unobserved heterogeneity and hence we might reduce bias of the income rate-parameter due to possible correlation between income rate and unobserved tastes. Below, we more directly try to reduce possible bias. We estimate a model with a random constant term in each of the four latent variables and a random parameter to the income rate, and – especially – an intersection term for income rates and the unobserved constant term. For simplicity we do not – in contrast to table 4 – allow constant terms to correlate. The latent variables can be written

(1')
$$I_{sit} = \left(\beta_y + \rho_y v_{yi}\right) y_{sit} + \beta_{sx} x_{it} + \rho_{ycs} v_{csi} y_{sit} + \rho_{cs} v_{csi} + e_{sit}$$

5 Admittedly, the comparison is incorrect because in table 2 we model eligibility. A comparison with experiments made with the model in table 1 however gives similar results.

where v's are random variables and ρ's related parameters with ρ_{ycs} capturing intersection effects from the observed economic variable, y, and the unobserved constant term, v_{csi}.

Table 9 presents the estimation of the parameters of interest.

Table 9 Economic incentives and social states. Result from a model with random constants and income parameter, and without modelling of eligibility.

Parameter	Latent variable	Parameter estimate	Standard deviation
Mean income parameter (β_y)	All latent variables	0.831	0.160
Variance of income parameter (ρ_y)	All latent variables	-0.051	0.045
Variance of constant (ρ_{cs})	Disability benefit	-0.950	0.454
	Early retirement	1.131	0.502
	Unemployment	1.079	0.134
	Home	-2.180	0.216
Variance of intersection (ρ_{ycs})	Disability benefit	-1.073	0.484
	Early retirement	0.551	0.592
	Unemployment	-0.171	0.148
	Home	-0.882	0.172

Source/notes: See table 1.

The (mean) parameter to the income rate has the expected sign and the magnitude is about the same as in table 1, but greater than the estimate in table 2. Again, most random parameters are significant. The random component related only to the income parameter (ρ_y) is however practically equal to zero.

In the rest of the paper, we discuss interpretations of random parameters.

One way to interpret the intersection terms $\rho_{ycs} v_{csi} y_{sit}$ is as an effect from economic incentives (i.e. from y_{sit}). Using this interpretation, the term describing economic incentives is $(\beta_y + \rho_y v_{yi} + \rho_{ycs} v_{csi}) y_{sit}$. Ignore for the moment ρ_y, and consider disability benefit for a person with $v_{csi} = \frac{0.831}{1.079}$, namely $(0.831 - 1.071 \frac{0.831}{1.071}) y_{sit} = 0 \cdot y_{sit}$, that is, there are no effects from economic effects for a person with this value of v_{csi}. Interpreting the intersection terms as effects from economic incentives, the estimations shows that about 20% $(= \Pr(v_{csi} < 0.831/1.079))$ of the population have no (or unexpectedly signed) effects of economic incentives on disability benefit. The percentage is 15 for ‘home’, and lower for ‘unemployment’ and ‘early retirement’. Perhaps these incidences of low values of incentive-parameters in this estimation are the counterpart to the low average value of the estimate of the incentive-parameter found in table 4.

However, that interpretation of intersection terms presumes that ρ_{ycs} is unbiased. An alternative interpretation is that the intersection term $\rho_{ycs} v_{csi} y_{sit}$ is used to reduce bias of β_y due to correlation of y and error terms e_{sit}. With this interpretation, we write the random individual constant terms as

(4) $$\eta_{csi} = \rho_{cs} v_{csi} + \rho_{ycs} v_{csi} y_{sit}$$

We therefore directly allow for covariation of error terms and y (but remember as we wrote above, we have not been able to credibly estimate the model with y affecting the mean of the random constant term, i.e. a model with $\eta_{csi} = \rho_{cs} v_{csi} + \rho_{ycs} v_{csi} y_{sit} + \gamma_y y_{\text{initial value}}$).

As a final exercise, we will describe how wages and unobserved characteristics, η, are related: for example, do people with relatively high wage have relatively high or relatively low preference for disability benefit. We use the interpretation of η represented in equation (4). The details of the discussion are rather confusing even though the idea is simple, and therefore put in appendix 5. The conclusions are

- As concerns 'disability benefit' and 'unemployment': if a person is a type with high idiosyncratic preference for the state, this idiosyncrasy is increased if the person has a wage rate above average.
- As concerns 'early retirement' and 'home', the case is the reverse, i.e. a high wage reduces the person's idiosyncrasy.

5. Conclusion

With a multi state model with random parameters we found that: 1) labour market distortions from economic incentives exist, but transitions between various welfare programmes are quantitatively at least as important as transitions between work and some type of welfare programmes. 2) Unobserved variations of tastes are important.

This is relevant for predictions of the effect of policy changes. For example, individuals with preference for the early retirement programme do not appear to have unobserved preferences for the disability benefit programme. 3) We seek to use a random parameters approach to reduce potential bias in the parameter to the income rate by allowing for an intersection term between observed income rate and unobserved constant terms. The results are mixed. In one estimation (table 9), we try directly to take account of heterogeneity correlated with the income variable, and the income-parameter appears robust. In another estimation (table 4), we model random constants flexibly, and the income-parameter decreases significantly.

We have less success with estimating equations for eligibility to various states. This is unfortunate, since many policy initiatives are concerned with changing eligibility rules for e.g. disability benefit. As a poor substitute, we modelled eligibility deterministically and the estimation result in table 2 fortunately indicates that account of eligibility does not have significant impact on non-eligibility-parameters.

Appendix 1. Data

States

Each individual is each year characterized as being in one of the five states. The classification is based on the duration of the periods the various types of income are obtained. With respect to the social benefit system, Statistics Denmark directly calculates the fraction of a year each type of social benefit is received. With respect to work, Statistics Denmark publishes total wage income and an hourly wage rate. With these variables, the period of employment can be calculated. Also, income from self-employment is published.

Let T_s denote the fraction of a year an individual collects benefit s, where $s = d\,\text{(disability)},\ r\ \text{(early retirement)},\ u\ \text{(unemployment)}$. Disability benefit is a distinguished benefit. The same is true for early retirement benefit, but the classification covers a general early retirement programme open for most people reached 60 years (in Danish: 'efterløn') and a special opening of this programme during some years for unemployed people in their 50s (in Danish: ‘overgangsordningen'). People classified as ‘unemployed’ collects one of various benefits (unemployment insurance benefit, social assistance, or benefit during certain labour market programmes, but not people on various leave schemes even though many people on such schemes previously collected e.g. social assistance). A 'primary' classification for the states disability benefit, and early retirement is now possible (D_s is dummy for state s).

$$\begin{aligned} &D_d = 1 \Leftrightarrow T_d > 0.5 \\ &D_r = 1 \Leftrightarrow T_r > 0.5 \\ &D_u = 1 \Leftrightarrow \sum_{u=\text{"unemployment" programmes}} T_u > 0.5 \end{aligned} \tag{A1}$$

The classification of work makes use of information on yearly wage income and the individual's hourly wage rate. Also, profit from self owned company is used. The proportion of a year spend on earning wage income is estimated as

(A2) $$T_e = \left(\frac{\text{Yearly wage income}}{\text{Hourly wage rate } \times 8 \text{ (hours/day)}} \right) \frac{1}{360 \text{ (days/year)}}$$

The definition of work is

(A3) $$D_e = 1 \Leftrightarrow T_e > 0.25 \text{ or profit} > 200{,}000 \text{ DKK}$$ [6]

There condition $T_e > 0.25$ rather than '> 0.5' allows individuals with regular part time employment to be classified as working.[7]

According to this primary classification some individuals will for some years be put in more than one state. Table A1 informs about the number of observations put in more than one state.

---

[6] 200,000 DKK approximates 26,900 EUR.

[7] The construction will exclude individuals with a high wage rate and a low yearly income. In practice however, it is not a subject to be concerned with, since simple statistics show that very few have a high wage rate and a low yearly wage income.

Table A1. Number of people primarily classified in more than one state

	D_d	D_r	D_u	D_e
D_r	314	.	.	.
D_u	1,447	136	.	.
D_e	5,441	329	8,962	.
All	249,058	124,093	211,910	1,547,333

Source: Registers from Statistics Denmark. Own calculations.

About 4 percent of unemployed individuals are also classified as working. Overall, the number of double-classifications seems acceptable.

The state 'home' is defined as those not put in any other state. However, we would like to think of 'home' as the state for people not working and not obtaining social benefit, but e.g. living of spouse's income or rental income. To evaluate this interpretation, the states are compared in table A2 with respect to various measures.

Table A2. State specific averages (and standard deviations) of miscellaneous variables

Variable	State 'Disability'	 'Early retirement'	 'Unemploy-ment'	 'Work'	 'Home'
Yearly wage income	3,985 (19,278)	6,847 (21,738)	23,781 (38,920)	220,543 (116,909)	44,034 (80,039)
Yearly total income	108,273 (60,350)	136,830 (64,030)	127,269 (61,365)	265,241 (213,174)	124,898 (142,496)
Wealth	102,695 (375,206)	290,520 (494,476)	40,255 (309,858)	136,700 (1,277,702)	170,943 (1,650,676)
Number of visits to family doctor	11.9 (17.9)	5.6 (7.2)	6.6 (9.8)	4.3 (5.6)	5.8 (7.9)
Age	54.8 (9.3)	62.8 (3.3)	43.3 (9.6)	44.0 (8.8)	48.3 (11.5)
Women in couples	0.33	0.36	0.35	0.35	0.50
Women	0.61	0.49	0.53	0.44	0.61

Source: Registers from Statistics Denmark. Own calculations.

On average, individuals classified as ‘home’ have wage income greater than that for other people not classified as working. Total yearly income is comparable to that of other people who do not work. Wealth is relatively large and health is approximately the same as other individuals who do not obtain disability benefit. Average age for ‘home’ is in the middle compared to other states. Many women are ‘home’, in particularly women who live in couples.

Summing up, the state ‘home’ has some of the characteristics that one would expect is true for people working at their home (using everyday language), but it is certainly also a heterogeneous group (the standard deviations are large), which might have been guessed from the statistical classification as a residual group.

To construct the final classification, states are given priority according to the sequence they are mentioned in table A1, i.e. disability benefit first, early retirement second, unemployment third, and work fourth. A person is finally classified in a state if she is classified primarily in that state and not in any state with a lower number.

Table A3 gives the final distribution on states

Table A3. Final distribution of observations on states

State	Disability benefit	Early retirement	Unemployment	Work	‘Home’
Proportion	0.098	0.049	0.083	0.606	0.164
Number	249,058	123,779	210,328	1,532,617	414,989

Source: Registers from Statistics Denmark. Own calculations.

Surprisingly many are “home”.

To conclude, the definitions of states rarely put the same individual in more than one state. On the other hand, this is obtained by making the residual state, ‘home’, large. An alternative might be to lower the number of days on unemployment benefit (or social assistance etc.) required to be classified as unemployed. As table A2 shows, those called unemployed have very poor attachment to the labour market, and lowering the 'unemployment criteria' will presumably move a considerable share of those in the state 'home' to the state 'unemployment'.

Income rates

‘Income rates’ are measures of hourly wage rates for work and hourly benefit rates for people on each of the three states with public benefits. The income rate is 0 for the state ‘Home’. First, we calculate such rates for those actually in each state. Second, on the basis of these income rates, we calculate expected rates to be used in the estimations.

Actual income rates

Registers from Statistics Denmark contain statistics on hourly wage rates for individuals who have been employed. Information on rates of social benefit is calculated according to the following equation (numerator and denominator are from Statistics Denmark)

(A4) $$w_s = \frac{\text{Total income obtained from social benefit } s}{\text{Number of days obtaining benefit } s}$$

The measures w_s vary across years and individual.

Expected income rates

Expected benefit rates for 'disability', 'early retirement' and 'unemployment' is calculated as averages across participants. The rates are therefore identical across individuals in each year.

The expected wage rate is a simple OLS regression using growth, education, time-trend, age, gender and a dummy for work the previous year – see section 3.

Nominal income rates are deflated using a consumer price index in order to obtain real measures. In 1994, a number of public benefits were made taxable. This implied an increase of gross rates. (The political intentions were not to change net-value of benefits.) To correct for this in a simple way, the disability benefit rates and unemployment benefit rates are spliced (levels in 1992 and 1993 are increased proportionally so the 1993-level equals the 1994-level). In general, income rates measure income before tax. No tax rules are modelled.

Appendix 2. Eligibility and observed states

In this appendix, I describe the approach that was used to try to model eligibility.

In general terms, the probability of observing a certain state is derived from probabilities relating to eligibility and to individuals preferences. Denote a subset of states by m. If a person is eligible for the subset of states, m, she chooses the best for her among these. Hence, given the set of eligible states, m, the probability to observe a particular state s among alternatives $s' \in m$ is (assuming a logit-model)

(A4) $$p_{sit,m} = \frac{\exp(I_{sit})}{\sum_{s' \in m} \exp(I_{s'it})}$$

A simple – and not very realistic – way to calculate the probability of eligibility for a certain subset of states is to assume independence of unobserved eligibility characteristics. If q_{sit} is the probability of eligibility for state s, the probability of eligibility for exactly the subset of states, m, is

(A5) $$q_{mit} = \prod_{s \in m} (q_{sit}) \times \prod_{s \notin m} (1 - q_{sit})$$

Let M denote the set of subsets. The likelihood contribution from individual i in period t is (d_{sit} is 1 if state s is observed, 0 otherwise)

(A6) $$L_{sit} = \prod_{m \in M} \prod_{s \in m} \left(p_{sit,m} \times q_{mit} \right)^{d_{sit}}$$

Using the approach above, I tried to model eligibility for disability benefit as a random parameter depending on health, for example as

(A7) $$q_{\text{disability},it} = \frac{\exp(\gamma_d + \gamma_h h_{it})}{1 + \exp(\gamma_d + \gamma_h h_{it})}$$

where h_{it} is a health indicator for person i in period t , which is equal to 1 if the person has bad health and 0 otherwise. In turned out however that γ_h became very large so the value of the function in (A7) in practice was as described in section 4 in the text above table 2.

We experimented in a similar manner with ‘eligibility’ for employment. This was calculated as a function of a measure of the general growth in Denmark. The estimated parameters became insignificant or came out with unexpected signs. This is unfortunate, since the modelling potentially allows for a disentangling of supply side and demand side effects on unemployment. Similarly, modelling of eligibility for disability benefit potentially allows for entering variables that might affect behaviour of awarding authorities (municipalities) – in particular the sharing of the cost for disability between state and municipalities has been a policy instrument used (though not very frequently in the period we study).

Appendix 3. Policy context

The Danish disability benefit system was not changed in any important way during the period we study, 1992-1998.

However, some conditions outside the disability benefit system changed and may have affected the number of disability benefit awards. Furthermore, after the period we study, some changes of the disability benefit policy were made. These changes are worth noting, since the approach used in this paper potentially may predict the responses to such changes. A further description of the policy context is made in Rasmussen (2004).

In 1984 several benefits schemes were merged into the Danish disability benefit programme “førtidspension” (for example widow’s pension). During 1984 to 2002, the scheme had four levels of the benefit rate and the entitlement to these rates depended on health and age. In 2003 the scheme was simplified with only one rate of benefit approximately equal to the previous maximum rate. The scheme covers the whole population (i.e. no contributions) aged 18 until the lower age limit for old age pension (which is 67 years for the time period studied). For the major part, benefit levels are not means tested and do not depend on previous earnings.

The relationship between the early retirement programme and the disability benefit programme is important. The early retirement benefit scheme was open for most people who reached 60 years of age during the period we study. Hence, relatively few of that age were awarded disability benefit. During some years in the 1990s, the early retirement scheme was open for long-term unemployed aged 50 to 59 years. Hence, for this group we might suspect fewer to be awarded disability benefit.

The general business cycle changed during the 1990s, with low growth in the first part of the period and high growth and decreasing unemployment from the middle of the period.

In 1999, the early retirement programme was changed. Eligibility now depends on contributions to the specific scheme (rather than contributions to the unemployment insurance scheme). At the same time the age limit for old age pension was lowered from 67 to 65 years. In terms of the logic of the model in this paper (substitution between alternative programmes), the former could imply more applicants for disability benefit, whereas the latter means the opposite. As mentioned, from 2003, the disability benefit programme has been changed and the expected rate of benefit increases for many potential beneficiaries. On the other hand, more emphasis is put on exploitation of remaining work capacity through a 'flex job' scheme (a job scheme with subsidized wage and special working conditions offered prior to possible award of disability benefit). Hence it may be that the screening will become harder. In terms of the logic in the model in this paper, the introduction of the reform therefore may either increase (due to a higher benefit rate) or decrease (due to more flex job openings) the number of disability benefit awards.[8]

Appendix 4. Random parameters

In this appendix we are a little move detailed about the simulation methods used.

As mentioned in section 3, for each individual, we make R draws, $r = 1,...,R$ of v which is assumed to be standard normal. If the value of v for individual i, v_i, were known, the true likelihood contribution from the individual would be as follows. Let D_{sit}, d_{sit} denote the random state-dummy variable and its realisation (i.e. 1 or 0), and $P(D_{sit} = d_{sit} \mid v_i)$ the probability that state s is realised (i.e. $d_{sit} = 1$) or not ($d_{sit} = 0$) given v_i. Suppress notation for observed explanatory variables and parameters to be estimated. The true likelihood contribution given v_i is

(A8) $$L(v_i) = \prod_t \prod_s P(D_{sit} = d_{sit} \mid v_i)^{d_{sit}}$$

The true (but unknown) likelihood contribution is (f is the density of v_i)

(A9) $$L_i = \int L(v_i) f(\mathrm{d}v_i)$$

[8] It appears logical to guess that the sum of awards of disability benefits and flex jobs will increase due to the reform.

To approximate the likelihood with the simulated values of v_i, namely the R draws $v_{i,1},...,v_{i,r},...,v_{i,R}$, we use

(A10) $$\tilde{L}_i = \frac{1}{R}\sum_{r=1}^{R} L(v_{i,r})$$

which is identical to equation (2).

Appendix 5. Wages and the variation of random terms

In general, the greater the value of the random term v_{csi}, the greater is the numerical value of η_{csi}, i.e. the farther away from 'average preferences' is individual i. We like to know whether a high wage increases or decreases the individual's idiosyncrasy. Even though the idea below is as simple as that, the following is admittedly somewhat confusing.

First rewrite η_{csi} using average level of y (benefit minus wage), $\bar{y}_s$, as (drop subscript t for simplicity)

$$\begin{aligned}\eta_{csi} &= (\rho_{ycs} y_{si} + \rho_{cs}) v_{csi} \\ &= \rho_{ycs}(y_{si} - \bar{y}_s) v_{csi} + [\rho_{cs} + \rho_{ycs}\bar{y}_s] v_{csi}\end{aligned}$$

If the term in the square bracket is positive, we say that a person with average y_s have idiosyncratic preference for state s when $v_s > 0$, and preference against the state when $v_s < 0$. This idiosyncrasy may be increased or decreased depending in the sign and magnitude of $\rho_{ycs}(y_{si} - \overline{y}_s)$. Remember that $y_{si} = b_s - w_i$, where b_s is benefit and w is wage, and remember that benefits are measured as simple averages so that they are identical across individuals. If $\overline{w}$ denotes average wage, we therefore have $y_{si} - \overline{y}_s = \overline{w} - w_i$. Furthermore, for states other than 'home', $\overline{y}_s$ is approximately –1. For 'home' it is a bit greater numerically. Table 10 below indicates signs of relevant measures and figures 1-3 describes the relationship between unobserved variables v_{csi} and η_{csi}.

Table A4. Signs of relevant magnitudes describing the variation of random parameters

	sign$(\rho_{cs} + \rho_{ycs}\overline{y})$ using the approximation $\overline{y}_s = -1, s = D, R, U,$ and $\overline{y}_h = -1.3$	sign$(\rho_{ycs}(\overline{w} - w))$ when $w > \overline{w}$
Disability benefit	+	+
Early retirement	+	-
Unemployment	+	+
Home	-	+

Figure 1 **Sketch of variation of random constant terms, for 'disability benefit' and 'unemployment'.**

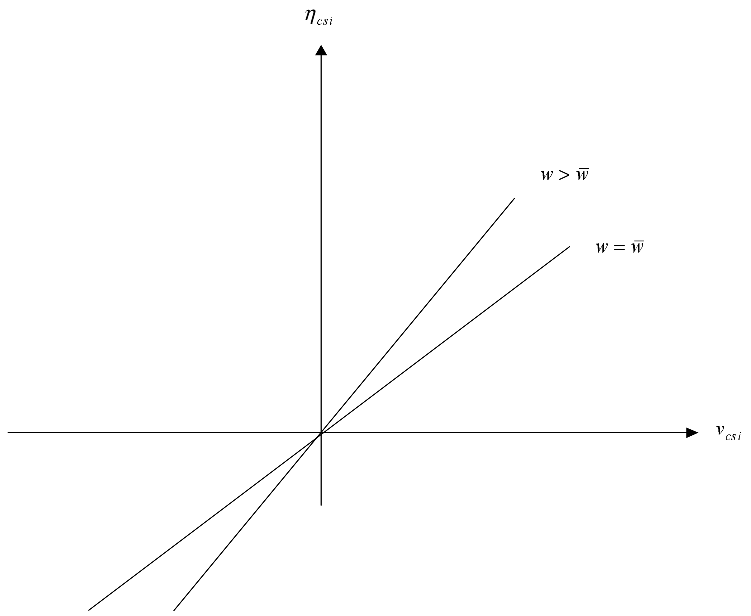


Figure 2 **Sketch of variation of random constant terms, for 'early retirement'.**

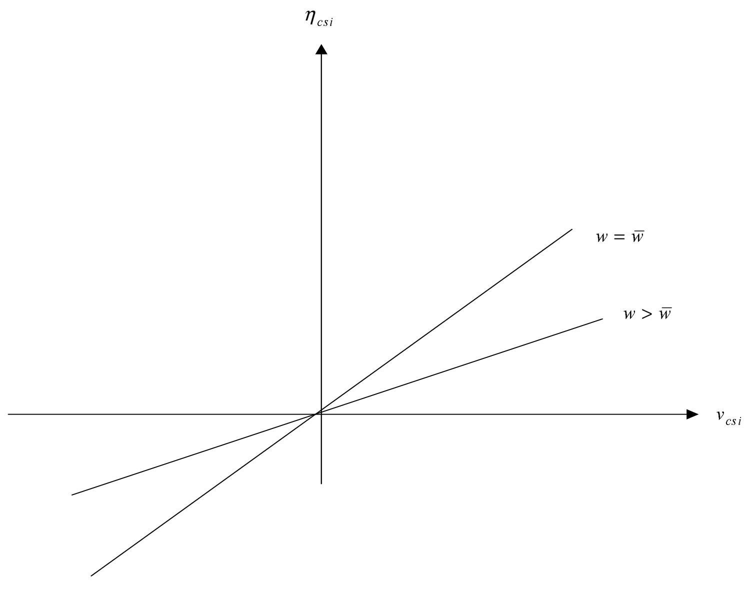
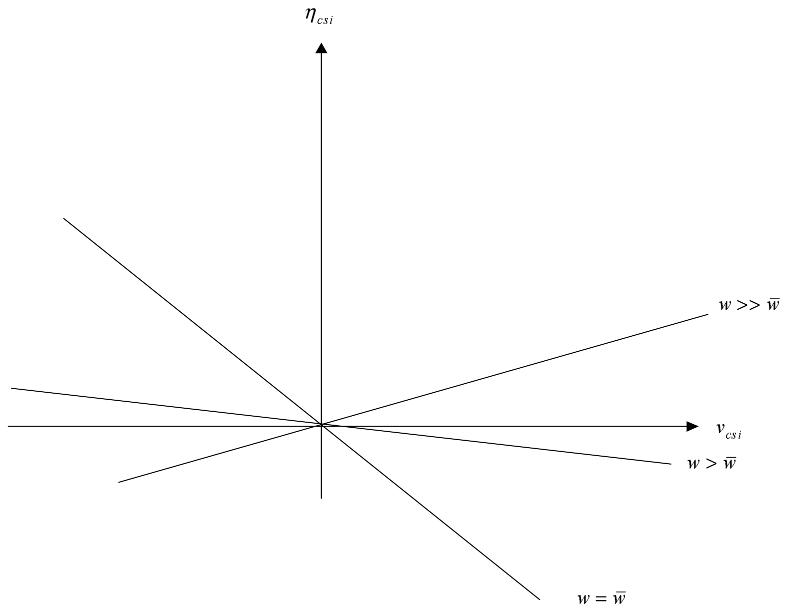


Figure 3 Sketch of variation of random constant terms, for ‘home’.



For states ‘disability benefit’ and ‘unemployment’, a wage above average increases the dispersion of random variables η_{csi}. The reverse is true for ‘early retirement’ and ‘home’.

References

Autor, D.H., Duggan, M.G., 2003. The Rise in Disability Rolls and the Decline in Unemployment. Quarterly Journal of Economics, vol. 118, 157-205.

Black, D., Daniel, K., Sanders S., 2002. The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust. American Economic Review, vol. 92, 27-50.

Campolieti, M., 2001a. The Canada/Quebec pension plan disability program and the labor force participation of older men. Economics Letters, 70, 421-426.

Campolieti, M., 2001b. Disability insurance and the labour force participation of older men and women in Canada. Canadian Public Policy, vol. xxvii.

Christiansen, F., 2000. Økonomiske incitamenter og tilkendelser. Samfundsøkonomen, 19-25.

Gruber, J., 2000. Disability Insurance Benefits and Labor Supply. Journal of Political Economy, vol. 108, 1162-1183.

Hajivassiliou, V. A., Ruud P.A., 1994. Classical Estimation Methods for LDV Models Using Simulation. Chapter 40 in Engle, R.F., McFadden, D.L, Handbook of Econometrics, volume IV, Elsevier Science B.V.

Haveman, R.H., Wolfe, B.L., 1984. Disability Transfers and Early Retirement: A Causal Relationship? Journal of Public Economics, 24, 47-66.

Haveman, R.H., Wolfe, B.L., Warlick J., 1988. Labor Market Behavior of Older Men: Estimates from a Trichotomous Choice Model. Journal of Public Economics, 36, 153-175.

Hyslop, D. R., 1999. State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women. Econometrica Vol. 67, 1255-1294.

McFadden, D., Train K., 2000. Mixed MNL Models for Discrete Response. Journal of Applied Econometrics Vol. 15, 447-470.

Parson, D.O., 1980. The Decline in Male Labor Force Participation. The Journal of political Economy, Vol. 88, 117-134.

Parson, D.O., 1982. The Male Labor Force Participation Decision: Health, Reported Health and Economic Incentives. Economica, 49, 81-91.

Rasmussen, M., 2004. Trends in disability benefit rates in post-industrial societies. Danish National Institute of Social Research, Working paper 01:2004,

Weatherall, J.H., 2002. "Vejen til førtidspension", Danish National Institute of Social Research, Rapport 02:9.