Dismissal through disability

Ridder, G.; Hassink, W.H.J.; van Ours, J.C.

published in
Economist
1997

DOI (link to publisher)
10.1023/A:1002942818866

document version
Publisher's PDF, also known as Version of record

DE ECONOMIST 145, NO. 1, 1997

DISMISSAL THROUGH DISABILITY

BY

WOLTER H.J. HASSINK, JAN C. VAN OURS AND GEERT RIDDER*

Key words: dismissals, disability, social security, employment

1 INTRODUCTION

A firm that wants to reduce the size of its workforce or change its composition, may choose between three methods to get rid of the redundant employees. It may induce employees to quit, it may dismiss employees, or it make employees eligible for disability benefits. In case of an induced quit there are no separation costs involved. The other two methods have costs that may differ substantially. In the following we will not attempt to give a comprehensive analysis of the choice and the use of the three methods but we restrict ourselves to the two methods that involve costs. In particular we focus on the use of enrolment into disability as an alternative to dismissal.

The rules that determine the eligibility for disability benefits restrict these benefits to persons who are no longer able to work due to health problems. Because of the imperfect verifiability of these health problems and the lower costs for both employer and employees associated with a transition to disability, it has been and still is attractive to use disability as an alternative to dismissal.

In The Netherlands the number of claimants of disability benefits is substantially larger than in other European countries. For example, in 1990 there were 139 claimants for every 1000 workers in The Netherlands. For Sweden this number was 78, for Germany 43 (Aarts, Dercksen, and De Jong (1993)). This raises the suspicion that a substantial fraction of the claimants was made eligible for enrolment because of redundancy, and not because of poor health. This suspicion is confirmed in a number of empirical studies which find that up to 50% of dis-

* During the period of research the first author was affiliated to the Department of Economics, Vrije Universiteit Amsterdam, The Netherlands. The second author is with the Erasmus University Rotterdam, The Netherlands. The third author is with the Department of Econometrics, Vrije Universiteit Amsterdam, The Netherlands. All three authors are also affiliated with the Tinbergen Institute, The Netherlands. Correspondence to: Jan C. van Ours, Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands; E-mail: ours@tir.few.eur.nl. The authors thank the Organisation for Labour Market Research (OSA) in The Hague, The Netherlands for permission to use their data and for financial support. They are grateful to Philip de Jong for his comments on an earlier version of the paper.

De Economist **145**, 29–46, 1997.

ability enrolment is related to redundancy. However, these studies are based on data collected before the 1987 major restructuring of the disability benefit system. The main objective of this restructuring was a reduction of the inflow into disability. Sofar, there has not been an empirical analysis of the relationship between disability enrolment and redundancies after the restructuring of the disability benefit system.

In this paper we use post-1987 data to investigate empirically to what extent separations into disability are still used as an alternative to dismissals. We use data on the dismissal and disability rates of Dutch firms to estimate the fraction of the desired dismissals that was channeled into disability. We find that about 10% of the new disability claimants became eligible because their job was redundant. Since this is substantially less than the numbers found in previous studies based on pre-1987 data we conclude that the social security reform of 1987 reduced the use of disability enrolment to avoid dismissals.

The paper is organised as follows. Section 2 gives a brief history of disability insurance and describes the arrangements with respect to dismissal and eligibility for disability benefits and discusses the incentives that result from these arrangements. Section 3 discusses previous empirical work on disability and unemployment. Section 4 describes the data. Section 5 presents the statistical model and the estimation results. Section 6 concludes.

2 DISMISSALS AND DISABILITY: INSTITUTIONAL ARRANGEMENTS

In 1967 the comprehensive disability insurance for employees (WAO) was introduced in The Netherlands.[1] Under the terms of this law workers were insured against wage loss due to long-term disability. From then on if a worker became ill, he was allowed to claim a benefit under the illness scheme for a maximum period of one year. After that he could claim a disability benefit. Workers were entitled to disability benefits after a so called disability examination, which consisted not only of a medical examination but also of an investigation of the labour market position of the worker. A worker could be considered disabled if there was no suitable job for him at his own educational level in his previous occupation. Furthermore, unemployment was 'internalized,' which means that those workers who were considered to be partially (more than 15%) disabled, could collect full disability benefits because it was assumed that partially disabled were doomed to remain unemployed. The benefit had a maximum of 80% of the wage in the last job. Disability benefits could be collected until age 65.

Since the introduction of the comprehensive disability insurance the number of workers collecting disability benefits has increased from 150,000 to about

1 In The Netherlands disability rules are given in the 'Wet op de Arbeidsongeschiktheid' (WAO) and the 'Algemene Arbeidsongeschiktheidswet' (AAW). Privately employed workers get benefits through the WAO-programme, whereas the AAW covers all residents, aged 18–64.

850,000 in the beginning of the 1990s. This huge increase in the number of disability benefits induced the government to change the structure of the disability benefit system.[2] In 1981 the rules to calculate the wage in the last job changed: tips and overtime were no longer included in the calculations. In 1985–1987 the maximum percentage of 80% of the previous wage was reduced to 70%. In 1987 there was a major restructuring of the disability benefit system of which the main objective was to reduce the inflow into disability. The most important change was the abolition of the 'internalization of unemployment' rules. Partially disabled workers were considered as such and were expected to find a job or claim unemployment benefits for their remaining work capacity. In the early 1990s there were some further changes. For employers the disability insurance premium was experience-rated, the disability examination no longer took the availability of suitable jobs with respect to education and previous occupation into consideration, the duration of the benefit was limited to five years after which a re-examination had to take place and, all disabled workers younger than 50 years had to be re-examined.

Our data relate to the end of the 1980s, a period in which the institutional arrangements with respect to dismissal and eligibility for disability benefits still provided incentives for employers and employees to use disability as an alternative to dismissal.[3]

A dismissal occurs if a labour contract is dissolved unilaterally by a firm. The Netherlands has, like many other European countries, laws that protect workers from unjustified dismissals. The employer must give a reason for the dismissal and not all reasons are acceptable. We can distinguish between three types of reasons.

First, an employer may dismiss a worker immediately, if the worker fails to perform his tasks. The worker may challenge this decision in court. Second, an employee may be dismissed, if the relationship between the individual employee and the employer has become untenable. Third, a dismissal may be the consequence of the redundancy of the employee's job. In order to dismiss a worker for one of these reasons, an employer must either ask permission from the regional employment office or obtain the approval of a district court.[4] Usually a permission from the employment office is needed in case of the dismissal of a group of workers. Approval by the court may be preferred by the employee, because court settlements, in contrast to permissions by the employment office, usually grant redundancy pay to the dismissed employees. In 1990 62,000 permits were given by the employment offices and 10,000 dismissals were approved by the courts.[5]

2 We do not discuss all the changes but restrict ourselves to some important examples.

3 See also Aarts and De Jong (1992, chapters 2, 5 and 8).

4 See article 1639w of the Dutch civil code.

5 The actual number of dismissals is higher because these numbers include dismissals of groups of workers, though it is likely that the court mainly has cases of individual workers.

Dismissing workers by either one of these methods is costly. First, the dismissal procedure takes time. The firm has to wait until the employment office or the court has reached a decision, which takes four to six weeks.[6] If the dismissal is approved, there is a period of advance notice with a maximum of thirteen weeks. This waiting period is expensive for the firm, because it has to pay the wage of the redundant employee. The employment office does not award redundancy pay. In case of a group dismissal, i.e. more than 20 workers within three months, redundancy pay may be claimed by unions. Redundancy pay may also be imposed in a court settlement. Finally, some, mainly large, firms supplement the unemployment benefits, which pay 70% of the last wage, to 100% for some period.[7] After dismissal a firm is not allowed to replace a worker within 3 months (Aarts and De Jong (1992)), and this may cause additional costs to the employer.

Dismissals are also costly for workers. The unemployed worker has to search for a job.[8] The period that the unemployed worker receives unemployment benefits depends on his work history and age. After exhaustion of the unemployment benefits the worker may apply for means-tested public assistance.

A transition into disability is costly for employers for several reasons. A worker only qualifies for disability benefits if he has been on sick leave for a year. Sick pay starts after two days of sick leave, and workers receive 70% of gross wage. Most firms supplement sick pay to the level of the net wage and also pay for the first two days (Ministry of Social Affairs and Employment (1989)). Because illness insurance is only partially experience-rated,[9] firms bear only a fraction of the direct costs of a sick leave. A sick leave has no effect on the worker's income. After one year a sick worker may apply for disability benefits. As indicated before, workers get both a medical examination and an investigation of their labour market position. On the basis of these examinations an expert determines what type of work the worker should be able to do. There is strong evidence that not all disability entrants were carefully examined. In 1987 37% of the applicants got a so-called 'reduced procedure' which did not involve a medical examination (Aarts and De Jong (1992)).

Disability insurance is only partially experience-rated since the insurance premiums are only industry-specific and do not depend on the disability record of the individual firm. Disability benefits are higher than benefits of long-term unemployed workers. Moreover, during the first (two) year(s) of disability 62% of the firms paid the difference between the disability benefit and the previous net wage of their former employees (Ministry of Social Affairs and Employment (1989)).

6 Cases in which there is *pro forma* defence on behalf of the worker last on average 2 weeks.

7 A survey of the Ministry of Social Affairs and Employment (1989) shows that 7.5 per cent of the firms supplemented the income of dismissed workers.

8 Except for persons older than 57.5 years, who will receive unemployment benefit until their retirement.

9 The premium only depends on the industry of the firm.

All in all, the choices of employers and employees with respect to the disability enrolment depend on the balance of costs and benefits. Employers who want to get rid of workers and have two options in doing so may choose the cheapest option. Employees will consider the expected residual lifetime income of the two options and may choose the option with the highest expected income. Due to incomplete experience rating employers may prefer disability over dismissal as a method to get rid of redundant employees. Because of the terms of the benefit systems redundant employees may prefer disability over long-term unemployment (Aarts and De Jong (1992), chapter 8). For both employers and employees disability may be the optimal choice.

3 EMPIRICAL STUDIES

Sofar, all empirical studies on the relationship between disability and dismissals refer mainly to the period before 1987, the year of the major restructuring of the disability benefit system. Some of the studies investigate the relationship between disability enrolment and economic variables. If a relationship is found this may indicate an improper use of the disability insurance. Other studies try to estimate the unemployment component of disability more precisely. Table 1 presents an overview of these empirical studies.

TABLE 1 – EMPIRICAL STUDIES ON DUTCH DISABILITY

Study	Data set	Unemployment component	Economic variables
Disability inflow			
Van den Bosch and Petersen (1983)	16 branches of industry 1968–1979		share of profits in added value, number of workers employed
Roodenburg and Wong Meeuw Hing (1985)	disaggregate[a)] 1971–1982	2%–35%	percentage of regional unemployment
Aarts and De Jong (1992, chapter 5)	2808 workers	33%–51%	expected future work capacity
Aarts and De Jong (1992, chapter 8)	1311 firms[b)]		firm size, employment change
Disability stock			
Westerhout (1995)	aggregate 1973–1992	50%	replacement rates disability and unemployment, number of inactive workers

a) Disaggregate to region, age, gender and economic sector.
b) The firms have at least 50 workers.

For each study the table presents the main features of the data set and an estimate of the unemployment component, that is the percentage of workers collecting disability benefits who are in fact unemployed. Furthermore, the table presents the main explanatory variables. These studies are classified according to whether they use stock or inflow data.

Van den Bosch and Petersen (1983) analyze time-series data on Dutch industries that cover a large part of the 1970s. They show that the growth of disability-incidence rates in the 1970s can to a large degree be explained by a deterioration of the economic situation in The Netherlands as indicated by the development of profits and employment. Their study also shows that the change in the perception of health is also an important determinant of the growth in disability transfers.

Roodenburg and Wong Meeuw Hing (1985) use time series of mainly 1970s data disaggregated with respect to region, age, gender, and industry to analyze the relationship between inflow into disability and unemployment. They conclude that in 1981 about 35 per cent of the annual inflow was related to unemployment.

Aarts and De Jong (1992, chapter 5) investigate to what extent workers who enrol in disability insurance are in fact able to work. Using microdata on workers of the 1980 inflow into disability they demonstrate that about 30–50 per cent of the workers collecting disability benefits were in fact 'hidden' unemployed workers because they were 'capable of performing a suitable job.'

Aarts and De Jong (1992, chapter 8) use data on disability enrolment in 1311 firms with at least 50 workers, observed in the years 1978–1979. They estimate a disability equation and conclude that there are underlying firm-specific economic reasons for disability enrolment. Aarts and De Jong also find that inter-firm differences in disability incidence are related to costs and benefits of retraining and replacing marginally efficient employees.

Westerhout (1995) uses time-series data of the period 1973–1992. So, his data include information before and after the 1987 reform of the disability benefit system. Unfortunately these are stock data so that the after-1987 data also contain information on workers who enrolled into the disability programme before 1987. Therefore, it is impossible to isolate the impact on the new disability enrolment after 1987. Westerhout shows that about half of the participation in disability schemes can be characterised as hidden unemployment.

We conclude that the various empirical studies show that before 1987 up to 50 per cent of the disability enrolment was due to non-health reasons. Hence, there is empirical evidence that before the restructuring of the disability benefit system disability enrolment has been used to a substantial degree to get rid of redundant employees.

4 THE DATA

Our data are from a panel survey of Dutch firms, conducted by the Organisation for Labour Market Research (OSA). We use the first two waves of this survey, 1988 and 1990.[10] The sampling units are organisations, which we refer to as firms. These firms are from all economic sectors including government and education, and have at least 10 employees. The sample is stratified according to economic activity and size of the firm (three classes: 10–49, 50–99, 100+ workers).

In each wave two questionnaires are used. The first questionnaire used by enumerators gathers qualitative characteristics and financial data of the firm. Firms that responded to this questionnaire received a second postal questionnaire to collect additional quantitative information about the firm. The response to the second questionnaire is 75–80% of the response to the first one.

In 1988 the sample consisted of 2041 firms, in 1990 of 2017 firms. A large number of the 1988 firms did not respond in 1990, had a substantial change in activities or experienced a merger. Restricting ourselves to the remaining firms, we retain a gross sample of 1168 firms. If we remove those firms of which we do not know dismissal rates (1988: 359 firms, 1990: 435 firms) and/or disability rates (1990: 433 firms), the variables of interest in this study, we obtain an intermediate sample of 559 firms.[11] After removing firms that did not answer questions on the explanatory variables to be used in the empirical analysis, a net sample of 225 firms remains.[12]

Our net sample contains substantially less firms than the original sample. Table 2 presents the distribution of firms over economic sectors in the three samples and gives a first impression of the selectivity of our dataset.

Because our statistical model contains lagged variables and we only have information on two years, we use a cross-section of firms. This means that the information used for the estimates is mainly from 1990. Definitions of the variables in our analysis are given in the Appendix. Employment is measured as the number of employees in December of each year irrespective of the number of hours they worked. Workers with temporary contracts shorter than one year are excluded.[13] In our empirical analysis both dismissals and transitions to disability are expressed as rates, defined as fractions of employment at the start of the year.

10 The surveys were carried out from March 1989 until March 1990, and May 1991 until December 1991, respectively.

11 The sole purpose of this intermediate sample is to investigate whether the selection introduces a bias.

12 The main reason for the loss of cases in going from the intermediate to the net sample is the lack of response to the questions 'output' (response of 49%) and 'wage' (response of 38%). 5 firms are not included in the net sample because they contain an outlier in one of the variables (a outlier for L (9 standard deviations from the mean), 1 for f_{88} (10 standard deviations from the mean), 3 for $Y_{90}/L_{90} - Y_{88}/L_{88}$ (at least 7.5 standard deviations from the mean). In the net sample f_{90} is 0.173 maximum and d_{90} is 0.100 maximum.

13 On average, the fraction of workers with a temporary contract shorter than one year is 9% of the total number of employees.

TABLE 2 – FIRMS BY ECONOMIC SECTOR IN SHARES: NET, INTERMEDIATE AND GROSS SAMPLE [a)]

	Net sample	Intermediate sample	Gross sample
SBI 0–3	0.29	0.25	0.31
SBI 5	0.05	0.08	0.08
SBI 6	0.05	0.09	0.10
SBI 7, 8	0.24	0.21	0.22
SBI 4 and 90	0.17	0.14	0.11
SBI 92	0.03	0.05	0.04
SBI 9, except 90, 92	0.17	0.18	0.14
	1.00	1.00	1.00
Number of firms	225	559	1168

a) The SBI classes are according to the Standard Industrial Classification of The Netherlands, Netherlands Central Bureau of Statistics, 1974. The seven strata are: agriculture, mining and manufacturing (0–3); construction (5); trade, hotels, cafés, restaurants and repair of consumer goods (6); transport, storage, communication, banking, insurance (7, 8); government and public utilities (4 and 90); education (92); other services (9, except 90 and 92).

Table 3 shows that there are many firms in both our net and our intermediate sample without dismissals or separations into disability in one or both years. In the net sample only 20% of the 225 firms have positive dismissal and disability rates in both years, while in the intermediate sample it is 16%. The differences between the net sample and intermediate sample are not very large.

TABLE 3 – DISMISSALS (F) AND TRANSITIONS INTO DISABILITY (D), PERCENTAGE OF FIRMS, NET AND INTERMEDIATE SAMPLE [a)]

	Net sample	Intermediate sample
$F = 0, D = 0$	0.45	0.50
$F = 0, D > 0$	0.24	0.21
$F > 0, D = 0$	0.11	0.13
$F > 0, D > 0$	0.20	0.16
	1.00	1.00
N	225	559

a) N is the total number of observations used.

Table 4 shows to what extent the dependent variables, the dismissal and disability rates, are affected by the selection process. The average dismissal rates are similar in the gross and net samples. The average disability rates are slightly different, but the difference can be attributed to sampling variability. Table 4 also shows that there are no substantial differences between the averages of other variables in both samples.

We conclude that the observed characteristics of the net sample do not substantially differ from those of less selective samples.

TABLE 4 – MEANS OF THE VARIABLES, NET AND GROSS SAMPLE [a)]

Variable	Net sample		Gross sample		Number of firms in gross sample [b)]
Dependent variables					
f_{90}	0.0081	(0.001)	0.0078	(0.001)	733
d_{90}	0.0074	(0.001)	0.0078	(0.001)	735
Firm characteristics					
$w_{90} - w_{88}$	0.190	(0.04)	0.216	(0.03)	439
$Y_{90}/L_{90} - Y_{88}/L_{88}$	0.210	(0.07)	0.161	(0.04)	572
q_{90}	0.065	(0.004)	0.076	(0.003)	737
f_{88}	0.0080	(0.001)	0.010	(0.001)	809
L_{89}	1.29	(0.10)	1.17	(0.05)	835
Adjustment costs					
Start-up $\geq$ 300 days	0.222	(0.03)	0.261	(0.01)	1168
Personnel characteristics					
Age $\geq$ 50	0.129	(0.005)	0.126	(0.003)	844
Part-time	0.196	(0.02)	0.182	(0.01)	830
Working conditions					
Bad conditions	0.787	(0.03)	0.738	(0.01)	1168
Illness in 1988	0.072	(0.003)	0.071	(0.002)	768

a) Standard deviations of the means are in parentheses. Units of the variables *f*, *d* and *q* are fractions of employment at the beginning of the year; age, part-time and illness in 1988 are fractions of employment at the end of the year; bad conditions and start-up are dummy variables; L is in hundreds of workers; $w_{90} - w_{88}$ is the change of the yearly gross real wage, measured in then thousands of guilders (average w_{88} is 4.19); $(Y_{90}/L_{90} - Y_{88}/L_{88})$ is in hundred thousands of guilders (average Y_{88}/L_{88} equals 3.61).

b) For every variable the number of firms in the gross sample is equal to 1168 minus the number of missing observations for that specific variable; the number of firms in the net sample is 225.

5 EMPIRICAL ANALYSIS

5.1 *The Statistical Model*

A firm has a desired dismissal rate (f^*) and a 'true' disability rate (d^*). Both rates are unobserved. f^* consists of the fraction of workers that should be dismissed because of economis reasons. d^* is the fraction of workers that enrol into disability, because they are unable to work for health reasons. d^* cannot be influenced by the firm in the short run, although over a longer period of time it may be affected by variables that are under control of the firm, for instance working conditions. f^* and d^* are both affected by a vector of exogenous variables $\mathbf{x}_1$. In addition, f^* is influenced by a vector of exogenous variables $\mathbf{x}_2$ and d^* depends on a vector of exogenous variables $\mathbf{x}_3$. We specify a linear regression model for f^* and d^*:

$$f_i^* = \beta_1' \mathbf{x}_{1i} + \beta_2' \mathbf{x}_{2i} + u_{1i}\,, \tag{1a}$$

$$d_i^* = \beta_3' \mathbf{x}_{1i} + \beta_4' \mathbf{x}_{3i} + u_{2i}\,, \tag{1b}$$

where

$$\mathrm{E}u_{1i}u_{2i} = 0$$

$$\mathrm{E}u_{ki} = 0$$

$$\mathrm{E}u_{ki}^2 = \tau_k^2$$

$$k = 1, 2$$

$$i = 1, \ldots, N$$

Index i refer to firm i. N is the number of firms, which in our case is 225. The assumption that the disturbances are uncorrelated after controlling for firm characteristics reflects the fact that the disability rate is given to the firm.

Instead of f^* and d^* we observe realized dismissal and disability rates (f and d). If a firm directs a fraction λ of its redundancies into disability, f is equal to $(1-\lambda)f^*$ and λf^* enrols into disability in the current year. Of course, these workers must be on sick leave for 12 months and hence, this strategy can only be used if there is a sufficiently large number of workers on sick leave. d is equal to $\lambda f^* + d^*$. Substituting in (1a) and (1b) gives:

$$f_i = (1-\lambda) f_i^* = \gamma_1' \mathbf{x}_{1i} + \gamma_2' \mathbf{x}_{2i} + v_{1i} \tag{2a}$$

$$d_i = \lambda f_i^* + d_i^* = \gamma_3' \mathbf{x}_{1i} + \gamma_4' \mathbf{x}_{2i} + \gamma_5' \mathbf{x}_{3i} + v_{2i} \tag{2b}$$

with

$$\mathrm{E}v_{ki} = 0\,, \qquad k = 1, 2$$

$$\mathrm{E}v_{1i}^2 = \sigma_1^2 = (1-\lambda)^2\tau_1^2$$

$$\mathrm{E}v_{2i}^2 = \sigma_2^2 = \lambda^2\tau_2^1 + \tau_2^2$$

$$\mathrm{E}v_{1i}v_{2i} = \sigma_{12} = \lambda(1-\lambda)\tau_1^2$$

$$\rho_v = \sigma_{12}/(\sigma_1\sigma_2)$$

$$\boldsymbol{\gamma}_1 = (1-\lambda)\boldsymbol{\beta}_1$$

$$\boldsymbol{\gamma}_2 = (1-\lambda)\boldsymbol{\beta}_2$$

$$\boldsymbol{\gamma}_3 = \lambda\boldsymbol{\beta}_1 + \beta_3$$

$$\boldsymbol{\gamma}_4 = \lambda\boldsymbol{\beta}_2$$

$$\boldsymbol{\gamma}_5 = \boldsymbol{\beta}_4$$

We complete the specification by choosing the variables that are included in $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$. The variables in $\mathbf{x}_1$, which affect both the autonomous disability rate and the latent dismissal rate are the growth (from 1988 to 1990) of the average real wage and the growth (from 1988 to 1990) of the average real productivity. Wage growth and productivity growth are both related to labour costs and therefore may affect the decision of the employer to reduce or expand their workforce in whatever way. The variables in $\mathbf{x}_2$, which are supposed to influence the desired dismissal rate are the quit rate, the lagged dismissal rate, the number of workers, the training period on the job, and the share of workers older than 50 years. These variables are quite traditional, all possibly influencing decisions of employers with respect to the optimal size of the workforce. The quit rate is complementary to the dismissal rate: the more voluntary separations there are the less involuntary separations are needed and *vice versa*. If the lagged dismissal rate has an influence this may reflect partial adjustment of the workforce. Furthermore, large firms may want to get rid of their workers more easily. Finally, the start-up period may be important because this reflects to what extent the firm has invested in its workforce and the age of the workforce may be important because the firing costs for older workers are higher.[14] The variables in $\mathbf{x}_3$ which influence the autonomous disability rate are the share of part-time workers, the

14 Hassink (1996) provides additional motivation for the choice of the variables in $\mathbf{x}_1$ and $\mathbf{x}_2$.

working conditions and the share of workers on sick leave. The first variable indicates to what extent workers are at risk of becoming disabled presupposing that longer working hours lead to more physical and psychical stress. The latter two variables are direct indications of the risks of becoming disabled. Table 3 presents the averages of the variables in our sample.

The statistical model must allow for the zeros in the dependent variables, and for the correlation of the disturbances. Hence, we estimate the model (2a)–(2b) as a bivariate Tobit. For ease of exposition we rewrite (2a, b) in obvious notation:

$$f_i = \boldsymbol{\omega}_1' \mathbf{z}_{1i} + v_{1i}$$

$$d_i = \boldsymbol{\omega}_2' \mathbf{z}_{2i} + v_{2i}\,. \tag{3}$$

Define the indicator functions as:

$$I_{ki} = I(k_i > 0)\,, \qquad k = f, d\,.$$

The likelihood function of (3) is:

$$\prod_{i=1}^{N} \left[\phi_2 \left(\frac{f_i - \boldsymbol{\omega}_1' \mathbf{z}_{1i}}{\sigma_1}, \frac{d_i - \boldsymbol{\omega}_2' \mathbf{z}_{2i}}{\sigma_2};\ \rho_v \right) \right]^{I_{fi} I_{di}}$$

$$* \left[\Phi_1 \left(\frac{-\boldsymbol{\omega}_1' \mathbf{z}_{1i} - \rho_v \dfrac{\sigma_1}{\sigma_2} (d_i - \boldsymbol{\omega}_2' \mathbf{z}_{2i})}{\sigma_1 (1 - \rho_v^2)^{1/2}} \right) * \phi_1 \left(\frac{d_i - \boldsymbol{\omega}_2' \mathbf{z}_{2i}}{\sigma_2} \right) \right]^{I_{di}(1 - I_{fi})}$$

$$* \left[\Phi_1 \left(\frac{-\boldsymbol{\omega}_2' \mathbf{z}_{2i} - \rho_v \dfrac{\sigma_2}{\sigma_1} (f_i - \boldsymbol{\omega}_1' \mathbf{z}_{1i})}{\sigma_2 (1 - \rho_v^2)^{1/2}} \right) * \phi_1 \left(\frac{f_i - \boldsymbol{\omega}_1' \mathbf{z}_{1i}}{\sigma_1} \right) \right]^{I_{fi}(1 - I_{di})}$$

$$* \left[\Phi_2 \left(\frac{-\boldsymbol{\omega}_1' \mathbf{z}_{1i}}{\sigma_1}, \frac{-\boldsymbol{\omega}_2' \mathbf{z}_{2i}}{\sigma_2};\ \rho_v \right) \right]^{(1 - I_{fi})(1 - I_{di})} \tag{4}$$

where $\phi_1(.)$, $\Phi_1(.)$ ($\phi_2(.,.;\rho_v)$, $\Phi_2(.,.;\rho_v)$) the univariate (bivariate) normal p.d.f and c.d.f., respectively. Maximum likelihood estimates of the parameters are obtained by maximizing (4) with respect to the parameters.

Estimates of the parameters of the structural model (1) can be obtained from the reduced form parameters by means of the minimum distance method (Chamberlain (1984)). There are 22 parameters of the reduced form model (2a, b) which

can be arranged in a 22-dimensional vector:

$$\boldsymbol{\theta} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \boldsymbol{\gamma}_4, \boldsymbol{\gamma}_5, \sigma_1, \sigma_2, \rho_v,) \; .$$

The structural model follows from the reduced form model by imposing the restriction

$$g(\boldsymbol{\pi}) = \boldsymbol{\theta} \tag{5}$$

where $\boldsymbol{\pi}$ is a 17-dimensional vector of structural parameters[15]:

$$\boldsymbol{\pi} = (\lambda, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4, \tau_1, \tau_2) \; .$$

Let $\boldsymbol{\theta}_N$ be the maximum likelihood estimator of $\boldsymbol{\theta}$. An efficient minimum distance estimate of $\boldsymbol{\pi}$ is obtained by minimizing the quadratic form:

$$Q_N(\boldsymbol{\pi}) = (\boldsymbol{\theta}_N - g(\boldsymbol{\pi}))' (\text{var}\, \boldsymbol{\theta}_N)^{-1} (\boldsymbol{\theta}_N - g(\boldsymbol{\pi})) \; .$$

The structural model is exactly identified if the dimensions of $\boldsymbol{\beta}$ and $\boldsymbol{\pi}$ are equal. The dimension of our $\boldsymbol{\theta}$ is larger than that of $\boldsymbol{\pi}$, which means that our structural model is overidentified. It is possible to test the overidentifying restrictions. Under the null hypothesis that the restrictions are satisfied Q_N has a Chi-square distribution and the number of degrees of freedom is equal to the number of overidentifying restrictions.

5.2 *Estimation Results*

The maximum likelihood estimates of the parameters of the reduced form model (3) are reported in the first column of Table 5. The other columns give the structural form parameters of model (2). These were obtained by the minimum distance method. The second column gives the structural estimates for the specification of section 5.1. The exclusion of 5 variables that appear in the dismissal equation from the disability equation allows us to identify the structural parameters. With this specification the structural model is overidentified. The key parameter λ is equal to 0.095, which means that 9.5% of the desired dismissals is through enrolment in disability insurance, or expressed as a fraction in the inflow into disability, that 12% of that inflow consists of employees that would have been dismissed otherwise. The test of overidentifying restrictions, which provides a check of the specification, rejects the specification in the second column. There-

15 $(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \boldsymbol{\gamma}_4, \boldsymbol{\gamma}_5, \sigma_1, \sigma_2, \rho_v)$ is related to $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{ß}_4, \tau_1, \tau_2, \lambda)$ by means of the following restrictions: $\boldsymbol{\gamma}_1 = (1-\lambda)\boldsymbol{\beta}_1$, $\boldsymbol{\gamma}_2 = (1-\lambda)\boldsymbol{\beta}_2$, $\boldsymbol{\gamma}_3 = \lambda\boldsymbol{\beta}_1 + \boldsymbol{\beta}_3$, $\boldsymbol{\gamma}_4 = \lambda\boldsymbol{\beta}_2$, $\boldsymbol{\gamma}_5 = \boldsymbol{\beta}_4$, $\sigma_1 = (1-\lambda)\tau_1$, $\sigma_2 = \lambda^2\tau_1^2 + \tau_2^2)^{0.5}$, $\rho_v = \lambda * \tau_i/(\lambda^2\tau_1^2 + \tau_2^2)^{0.5}$.

TABLE 5 – ESTIMATION RESULTS BIVARIATE TOBIT REGRESSION, REDUCED AND STRUCTURAL FORMS [a)]

	Reduced form		Structural form I		Structural form II	
Dismissal rate						
$w_{90} - w_{88}$	0.01	(1.66)*	0.013	(2.36)**	0.012	(2.33)**
$Y_{90}/L_{90} - Y_{88}/L_{88}$	−0.0006	(0.13)	−0.0009	(0.25)	−0.0007	(0.20)
q_{90}	−0.039	(0.52)	−0.049	(0.85)	−0.044	(0.80)
f_{88}	0.402	(2.05)**	0.472	(3.06)**	0.440	(2.93)**
L_{89}	0.005	(2.08)**	0.007	(3.30)**	0.005	(2.95)**
Start-up $\geq$ 300 days	−0.019	(1.75)*	−0.023	(2.72)**	−0.021	(2.56)**
Age $\geq$ 50	−0.062	(1.15)	−0.076	(1.80)*	−0.070	(1.70)*
Constant	−0.026	(2.13)**	−0.030	(3.16)**	−0.027	(2.97)**
σ_1	0.048	(10.32)***	–		–	
τ_1	–		0.053	(27.03)***	0.051	(26.67)**
Disability rate						
$w_{90} - w_{88}$	−0.007	(1.79)*	−0.008	(3.01)**	−0.008	(2.87)**
$Y_{90}/L_{90} - Y_{88}/L_{88}$	0.002	(0.98)	0.002	(1.20)	0.002	(1.22)
q_{90}	−0.025	(0.67)	–		–	
f_{88}	0.108	(0.99)	–		–	
L_{89}	0.004	(3.45)**	–		0.004	(4.38)**
Start-up $\geq$ 300 days	−0.006	(1.22)	–		–	
Age $\geq$ 50	−0.029	(0.99)	–		–	
Part-time	−0.012	(1.29)	−0.010	(1.60)	−0.012	(1.89)*
Bad conditions	0.046	(0.88)	0.060	(1.63)	0.050	(1.36)
Illness 1988	0.192	(3.46)**	0.176	(4.52)***	0.189	(4.86)**
Constant	−0.022	(2.74)**	−0.017	(3.66)***	−0.025	(5.03)**
σ_2	0.026	(12.07)***	–		–	
τ_2	–		−0.025	(65.79)***	−0.025	(95.24)**
ρ_v	0.085	(0.82)	–		–	
λ	–		0.095	(2.96)**	0.063	(1.86)*
$\chi^2_{(5)}$	–		11.95**		2.46	
N	225		225		225	

* Statistically significant from zero at the 10% level.
** Statistically significant from zero at the 5% level.
*** Statistically significant from zero at the 1% level.
a) Absolute *t*-values are in parentheses; N is the number of observations used to estimate the model; σ_k, $k = 1, 2$, is the standard error of the regression of the reduced form estimates; τ_k, $k = 1, 2$, is the standard error of the regression of the structural form; ρ_v is the correlation coefficient in the error structure of the reduced form; $\chi^2_{(5)}(\chi^2_{(4)})$ is a Chi-square test on the overidentifying restrictions with 5 (4) degrees of freedom.

fore we experimented with other specifications by including some additional variables in the disability equation. The smallest estimate of λ was obtained when we included employment in 1989 in the disability equation. Moreover, the 4 overidentifying restrictions are not rejected for that specification, which strongly supports this specification. The estimation results in the third column of Table 4 show that the other structural parameters do not change much. The structural parameter estimates in columns 2 and 3 are as expected (see also Hassink (1996)). In the dismissal equation wage growth is significantly positive, as are the lagged firing rate and lagged employment. Training costs have a significantly negative effect on dismissals. Wage growth has a significantly positive effect and absence due to illness has a significantly negative effect on the disability rate.

The estimates of λ in the second and third columns give bounds on our estimate of the fraction of dismissals that is channeled into disability. That fraction is between 0.063 and 0.095, which implies that between 8 and 12% of the inflow into disability consists of dismissals.[16]

If we compare our results with those from previous studies it is obvious that we find a much smaller part of the disability enrolment is due to dismissals. Instead of a share up to 50% we find a share of approximately 10%. The main difference between the previous studies and ours is that the previous studies use data from before 1987 major change in the disability benefit system while we use data from after that change. This comparison is indirect and of course we would have preferred a direct comparison of the situation before and after the restructuring of the benefit system. However, since there are no data for a direct comparison we stick to the indirect one. We conclude that there has been a decline in the use of disability enrolment for purposes which have nothing to do with health. And, we conclude that this decline is probably due to the restructuring of the disability benefit system.

6 CONCLUSIONS

If a firm wants to reduce its workforce, it may dismiss some of its workers. Alternatively, it may make some workers eligible for disability benefits. Upon examination these workers formally satisfy the conditions for disability enrolment. Because these conditions allow for a rather liberal interpretation of disability, these workers could have stayed in their job had they not become redundant. Note that this behaviour may not be in conflict with social security laws, because the workers satisfy the conditions for disability enrolment. Previous empirical studies find that before the 1987 reform of the disability benefit system up to 50% of the disability enrolment was related to redundancy of workers. We have

16 A plausible alternative is to include the fraction of workers over 50 in the disability equation. However, the coefficient is not significantly different from 0, the estimate of $\lambda = 0.094$ and the overidentifying restrictions are rejected.

demonstrated that at the end of the 1980s employers in The Netherlands still used disability enrolment as an alternative to dismissals. Our estimates imply that about 10% of the transitions into disability are due to redundancy of the worker. An implication of our result is that also after the social security reform of 1987 disability enrolment was used by some employers and employees to avoid dismissal. However, comparing our results with those from previous analyses we conclude that this type of behaviour is much less common than before the reform.

APPENDIX

DEFINITION OF THE VARIABLES

The variables are defined as follows. The quotes are translations of the relevant questions in the questionnaire.

- Employment, L:
  'How many workers were employed in your organization in December 1988 (1990) (excluding temporary workers). This concerns the number of employees irrespective of the number of hours worked.' The numbers of employees in December 1987 and December 1989 are constructed by means of the hires (H) and the separations (S) of employees during 1988 (1990), which are measured in the survey: $L_{t-1} = L_t - H_t + S_t$.
- Dismissals, F, and transitions into distability, D:
  'How many employees left your organization in 1988 (1990), excluding employees with a temporary contract shorter than one year.' The questionnaire distinguishes 5 reasons for leaving the firm:
  - pension, early retirement, death;
  - separations because of disability (D);
  - dismissal (F);
  - quit (Q);
  - end of temporary contract with a duration $>$ one year.
- Wage, w:
  The annual gross wage in the organization at the time of the survey. The employer is asked to distinguish the salary levels of the employees in the three groups where all have an equal number of employees. For each group, the maximum and the minimum wage are registered. The average wage is constructed as $\Sigma_i (L_i/L) * (w_i^{min} + w_i^{max})/2$; where w^{min} and w^{max} are the minimum and the maximum wage level, respectively. L_i is the number of employees in each group. The nominal wage is defined by 1988: 100, 1990: 103.
- Sales, Y:
  'Sales of the firm before taxes in thousands of guilders in 1988 (1990).' Unfortunately the questionnaire is not more specific about the tax. Rather we

would have used the value added of the firm, but this information is not available. Nominal sales are defined by 1988: 100, 1990: 103.

– Training period:
  Dummy variable which is one in case of a start-up period longer than 300 days. (This variable is from the first questionnaire gathering qualitative information.)
– Age $\geq$ 50:
  Share of employees older than 50 years.
– Part-time:
  Share of part-time employees in December.
– Bad working conditions:
  Dummy variable which is one in case of physically (or mentally) strenuous working conditions in the firm. (This variable is from the first questionnaire gathering qualitative information.)
– Illness:
  Average share of employees absent due to illness in 1988.

REFERENCES

Aarts, L.J.M., W.J. Dercksen, and Ph. R. de Jong (1993), 'Arbeidsongeschiktheid, een internationale Vergelijking' (Disability, an International Comparison), *Sociaal Maandblad Arbeid*, pp. 755–769.

Aarts, L.J.M. and Ph. R. de Jong (1992), *Economic Aspects of Disability Behavior*, Amsterdam, North Holland.

Bosch, F.A.J. van den and C. Petersen (1983), 'An Explanation of the Growrh of Social Security Disability Transfers,' *De Economist*, 131, pp. 65–79.

Chamberlain, G. (1984), 'Panel Data,' in: Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics*, 2, New York, North Holland.

Hassink W.H.J. (1996), *Worker Flows and the Employment Adjustment of Firms: An Empirical Investigation*. Dissertation Vrije Universiteit Amsterdam, Tinbergen Institute Research Series, No. 105.

Ministerie van Sociale Zaken en Werkgelegenheid (1989), *Bovenwettelijke uitkeringen na de stelselwijziging sociale zekerheid* (Extrajudicial Benefits After the Social Security Reform), Research Memorandum, The Hague.

Roodenburg, H.J. and W.J.M.L. Wong Meeuw Hing (1985), *De arbeidsmarktcomponent in de WAO* (The Labour Market Component of the Disability Insurance), Occasional Paper No. 34, The Hague, Centraal Planbureau.

Westerhout, E.W.M.T. (1995), *Arbeidsongeschikt of werkloos?* (Disabled or Unemployed?), Occasional Paper No. 120, The Hague, Centraal Planbureau.

Summary

DISMISSAL THROUGH DISSABILITY

If a firm wants to reduce its workforce, it may dismiss some of its workers. Alternatively, it may make workers eligible for disability benefits. Upon examination these workers formally satisfy the conditions for disability enrolment. Because these conditions allow for a rather liberal interpretation of disability, these workers could have stayed in their job had they not become redundant. We use data on Dutch firms to show that at the end of the 1980s about 10 per cent of the observed inflow into disability were in fact dismissals. Comparing our results with those from previous studies we conclude that due to the social security reform of 1987 the use of disability enrolment to avoid dismissals has been substantially reduced.