

Risto Herrala – Karlo Kauko

Household loan loss risk in Finland – estimations and simulations with micro data



Bank of Finland Research
Discussion Papers
5 • 2007

**Suomen Pankki
Bank of Finland
P.O.Box 160
FI-00101 HELSINKI
Finland
📞 + 358 10 8311**

<http://www.bof.fi>

Bank of Finland Research
Discussion Papers
5 • 2007

Risto Herrala – Karlo Kauko

**Household loan loss risk in
Finland – estimations and
simulations with micro data**

The views expressed are those of the authors and do not necessarily reflect the views of the Bank of Finland.

<http://www.bof.fi>

ISBN 978-952-462-354-4
ISSN 0785-3572
(print)

ISBN 978-952-462-355-1
ISSN 1456-6184
(online)

Helsinki 2007

Household loan loss risk in Finland – estimations and simulations with micro data

Bank of Finland Research
Discussion Papers 5/2007

Risto Herrala – Karlo Kauko
Monetary Policy and Research Department

Abstract

This discussion paper presents a microsimulation model of household distress. We use logit analysis to estimate the extent to which a household's risk of being financially distressed depends on net income after tax and loan servicing costs. The impact of assumed macroeconomic shocks on this net income concept is calculated at the household level. The microsimulation model is used to simulate both the number of distressed households and their aggregate debt in various macroeconomic scenarios. The simulations indicate that household credit risks to banks are relatively well contained.

Key words: financial stability, indebtedness, micro simulations, households

JEL classification numbers: D14, G21, E47, R29

Kotitalouksista johtuvat luottotappioriskit Suomessa – estimointeja ja simulointeja mikroaineistolla

Suomen Pankin keskustelualoitteita 5/2007

Risto Herrala – Karlo Kauko
Rahapolitiikka- ja tutkimusosasto

Tiivistelmä

Tässä keskustelualitteessa esitetään mikrosimulaatiomalli kotitalouksien riskistä joutua taloudellisiin vaikeuksiin. Logit-malleilla estimoidaan, kuinka kotitalouden riski joutua taloudellisiin vaikeuksiin riippuu verojen ja lainanhoidokulujen jälkeisistä nettotuloista. Oletettujen makrotaloudellisten sokkien vaikutus tähän nettotulokäsittelyseen on laskettu kotitalouksittain. Mikrosimulaatiomallilla simuloidaan sekä vaikeuksissa olevien kotitalouksien lukumäärää että niiden yhteenlaskettua velkamäärää erilaisissa makrotaloudellisissa skenaarioissa. Kotitaloudet eivät vaikuta pankkien kannalta merkittävältä luottotappioriskin lähteeltä.

Avainsanat: rahoitusmarkkinoiden vakaus, velkaantuminen, mirkosimulaatiot, kotitaloudet

JEL-luokittelu: D14, G21, E47, R29

Contents

Abstract	3
Tiivistelmä (abstract in Finnish)	4
1 Introduction	7
2 Previous literature.....	8
2.1 General observations	8
2.2 Indebtedness as a problem	9
2.3 The quantity of debt.....	11
3 The model.....	12
4 The data.....	14
5 Selection and estimation of an empirical model	17
6 The household distress simulator	21
7 Simulations.....	23
7.1 Macro scenarios	23
7.1.1 The basic scenario	23
7.1.2 The extreme scenario	23
7.1.3 Standardised shocks	24
7.2 Results	25
7.3 Robustness tests	27
7.3.1 The unemployment benefit system	27
7.3.2 Ability to postpone amortizations	28
8 Conclusions	29
References.....	31
Appendix 1 Decomposition of forecast error in model 2a	34
Appendix 2 Updating the data	37
Appendix 3 Proxies for dwelling prices	38
Appendix 4 Controlling for non-linear marginal tax rates	40
Appendix 5 Unemployment benefits.....	43

1 Introduction

We construct a micro simulation model which uses as an input the macroeconomic forecast of the Bank of Finland and a micro data set of households, and produces as an output a forecast of distress in the household sector. The study contributes to the growing empirical literature on the determinants of household distress. The analysis is also relevant for understanding the issue of whether bank loan losses and, in the extreme case, banking crises, can be predicted.

Traditionally, these issues have been analysed from aggregated (macro level) data. While such studies yield relevant insight, they are incomplete because aggregation hides distributions. In many fields of study, the representative agent point of view inherent in aggregated data suffices for analytical purposes, but in the field of financial fragility this may not be the case. One indication of this is that in theoretical models of financial fragility, equilibria usually depend on the joint distribution of liquidity and shocks in the underlying population. The benefits of working with micro level data readily present themselves to the empirical economist, who can match budget constraints, debts, and collateral at household level to study financial fragility.

The following section reviews the previous literature on household debt formation and distress. It is observed that there is a young and growing empirical literature on the use of micro data sets to model the link between macroeconomic developments and household distress. Much of the work, like the present paper, is driven by the interests of central banks to study and promote financial stability. Especially the work of May and Tudela (2005), and Del-Rio and Young (2005b) are important starting points for our efforts.

In section 3, we outline a theoretical model of household distress based on the household budget constraint: distressed households are those whose consumable income is ‘too low’. Our approach can be seen as a variant of Del-Rio and Young (2005b). We reformulate their model by taking into account loan instalments and the ability of households to sustain consumption by running down assets and incurring debt. Our model leads to the conclusion that the probability of household distress depends on a set of macroeconomic factors such as interest rates and income (as in Del-Rio and Young), and also the collateral value of wealth. Our formulation in particular illustrates that not all relevant variables are observable, and we discuss a set of assumptions under which we can formulate an estimable econometric equation on the macroeconomic determinants of household distress.

Section 4 reviews the micro data set, which consists of a sequence of annual household surveys ('the service data on income distribution') conducted by Statistics Finland. The surveys give information about some ten thousand

households. For the purposes of this study, we can use the surveys from 1999 onwards. A preliminary look at the data gives support for the idea that the relationship between household budgets and household distress accords qualitatively with the theoretical model outlined in section 3.

Section 5 presents the econometric model used for quantifying the link between household budgets and household distress. We analyse the parameter estimates, in sample fit and forecast performance of a number of empirical models that are consistent with our theoretical approach. In the end, the model with the best forecast performance is chosen for simulations.

Section 6 introduces the simulator. In the simulator, macroeconomic developments affect the budgets of individual households in a mechanical way. The econometric model with the best forecast performance is used to calculate the forecasts for the proportion of distressed households, debt at risk, and uncovered credit risk in different macroeconomic scenarios. The simulator can be calibrated to produce a rough forecast on loan losses.

Section 7 reports a number of simulations. In the first simulation, a scenario made by the Bank of Finland macroeconomic model is utilised to produce a projection of household distress under a relatively mild depression. In the second simulation, macroeconomic data from the period of the Finnish banking crisis in the early 1990's is utilised to simulate distress in an extreme depression. Finally, the section reports simulations on the direct, '*ceteris paribus*' effect of shocks in the different macroeconomic variables (keeping other variables constant). Interestingly, it is found that household distress is, *ceteris paribus*, particularly sensitive to changes in interest rates.

Section 8 concludes with a brief review of the results and suggestions for further analysis.

2 Previous literature¹

2.1 General observations

Micro level data on household debt seem to be available in many countries, including at least Finland, UK, Japan, Spain, Portugal, Italy, France, Norway, Sweden, US, Australia, Germany, Belgium, Thailand and Philippines. However, relatively few articles on household debt have been published in refereed journals in the last few years. Instead, most reports fall within the category of 'grey literature'. This applies especially well to contributions that consider indebtedness

¹ We thank Katja Taipalus from the Bank of Finland for her contribution to this section of the paper.

as a problem. It seems to be particularly difficult to find simulation based analyses on household debt. Instead, descriptive analyses seem to be more commonplace.

There are a few theoretical contributions on household debt. The permanent-income hypothesis, presented by Milton Friedman in 1957, states that people base consumption on what they consider their ‘normal’ income even though their incomes may vary considerably in the short term. This may create occasional needs to incur debt. The life cycle approach is somewhat similar to the permanent income hypothesis, but it emphasises the impact of changes in wealth on consumption. Neither of these approaches focuses on problems related to indebtedness.

This paper focuses on potential problems caused by household indebtedness. Hence, it is of particular interest to take a look at previous analyses on difficulties in debt servicing (section 2.2). However, section 2.3 contains also a brief survey of recent work on the related issue of the determinants of the quantity of debt.

2.2 Indebtedness as a problem

Unfortunately, there seem to be hardly any theorising on household indebtedness as a problem. Instead, there are numerous empirical analyses on households’ debt servicing difficulties.

May and Tudela (2005) used British data to study which factors determined the likelihood of having debt servicing difficulties. Unemployment, high levels of indebtedness and a high proportion of non-collateralised debt increased the likelihood of problems. Difficulties seemed to be of permanent nature; past problems were a good predictor of future problems. There was almost no evidence of housing wealth preventing difficulties. This might be due to the possibility that many interviewees considered themselves to be in difficulties if the time schedule of repayments had been renegotiated with the bank. A typical household is certainly unwilling to sell the dwelling in order to service loans. Hence, the availability of collateral is of limited use in preventing reported difficulties, even though collateral certainly reduces banks’ credit risks.

May and Tudela introduce a new concept, namely ‘debt-at-risk’, an indicator of credit risk. It is simply the sum of household level housing debt multiplied by the household level likelihood of financial distress. If it is possible to identify the determinants of the likelihood of distress, and if micro level data are available, it is possible to calculate the aggregate amount of ‘debt at risk’ and how it would react to changes in the parameters, such as interest rates. The sections 6 and 7 of this paper apply this concept.

Quercia, McCarthy and Stegman (1995) studied mortgage default rates in rural areas and among low-income workers in the US. Some of their results were

not consistent with previous findings; interest rate subsidies prevented problems, but there was no evidence of the relevance of the loan-to-value ratio as a determinant of mortgage default.

La Cava and Simon (2003) analysed indebtedness among Australian households and the risks of rapidly increasing debt. The study was based on 1998 and 1999 household data. Several factors contributed to a low default risk, including owner-occupied housing, advanced age and high income. It seemed that couples were less likely to have problems than single persons. If, instead, there were several unemployed family members, the household received income subsidies or if it had incurred a substantial amount of debt relative to its income, or if interest was paid on a credit card balance, problems were more likely to occur.

Liu and Lee (1997) used various methods to study the determinants of housing loan defaults in Taiwan. Several factors, including the loan-to-value ratio and the level of education, seemed to contribute to the occurrence of problems. Problems were particularly commonplace among the youngest.

Bowie-Cairns and Pryce (2005) found with British data that household debt servicing ability was an increasing function in the level of education, age and marital status. Families with many children were more likely to have problems. Geographic factors seemed to play some role. As to loans with no collateral, Del-Rio and Young (2005b) concluded that high levels of debt and factors related to marital status and ethnicity were relevant to the occurrence of problems among British households.

De Doncker (2006) analysed person related factors correlating with credit defaults. The data were extremely comprehensive; in principle, all the mortgage and consumer loans granted to individuals in Belgium were included. The credit risk of consumer loans appeared to be higher than the credit risk of housing loans. The risk of loans granted to a male person seemed to be particularly high. Loans with two debtors of different sexes seem to bear a particularly low risk.

There are some macro level studies on the determinants of mortgage defaults. Brookes, Dicks and Pradhan (1994) analysed British aggregate data. Surprisingly, household income seemed to be of limited importance in explaining difficulties in debt servicing. Elmer and Seelig (1998) studied the incentives of households to default on mortgages in the US. The ‘strategic option’ seemed to be the wrong way to explain defaults. Instead, defaults were typically caused by inability to service debt. The default risk was higher if there were negative shocks in real estate prices and household income. Instead, perhaps surprisingly, interest rate variations seemed to be of almost no importance.

Barrel & Davis (2004) have presented estimation results on the impact of banking and currency crises on consumption. The macroeconomic costs of crises tend to be higher if households are more dependent on debt.

2.3 The quantity of debt

There are a lot of descriptive and exploratory studies on the determinants of household debt in different countries. These papers do not form clearly identifiable schools that would be based on established theoretical approaches. Nevertheless, there seem to be some empirical regularities that have been corroborated by observations made in many different countries.

Del-Rio and Young (2005a) concluded that non-collateralised loans in Britain were typically taken by people who were in their twenties, had no children, were relatively well educated, were employed and had optimistic expectations. Cox, Whitley and Brierley (2002) studied the development of collateralised and non-collateralised debt at the household level. Brown, Garino, Taylor and Price (2003) found that optimism seemed to strengthen the demand for non-collateralised loans, even though after a certain level the degree of optimism had no impact. Factors such as spouse income and savings seemed to have no impact.

Martins and Villanueva (2003) studied the impact of interest-rate subsidies on long-term household debt in Portugal. Riiser and Vatne (2006) presented observations on Norwegian data from 1986–2003. Household debt had been on increase, especially among the youngest households and those with low income. The growth of financial wealth had taken place above all in households with no debt. Magri (2002) studied the occurrence of debt among Italian households. Higher income strengthen both the demand for loans and the availability of credit, whereas being an entrepreneur strengthened the demand for credit but made it more difficult to obtain loans.

Brown and Taylor (2005) studied the determination of household financial wealth and debt in the UK, Germany and the US. There was a clear correlation between financial wealth and debt. If households with no debt were excluded, the correlation vanished.

Tudela and Young (2005) analysed the determinants of household balance sheets in the UK, applying the hypothesis that households optimise their income and consumption over the life cycle. Crook (2001) analysed the 1995 US Survey of Consumer Finances data; the demand for household loans seemed to be an increasing function in income and family size.

Davydoff and Naacke (2005) presented a purely descriptive report on the distribution of housing and consumer loans in France, Britain, Germany and Italy. Housing loans were particularly commonplace in the UK but remarkably exceptional in Italy, even though owner-occupied housing is particularly commonplace in Italy. In all the countries the determinants of indebtedness were rather similar.

The availability of loans can also be a major determinant of household debt. Not every household can always get a loan, and problems related to credit

availability have been analysed by numerous studies. This literature has a close connection with an established theoretical discourse, namely life cycle studies and the Euler equation. If credit is not available, households may need to temporarily spend less than what optimal consumption smoothing would require. Credit constraints may be one of the main reasons why private consumption reacts to transitory income stronger than what the Euler equation approach would suggest. Most of this literature was published more than ten years ago; in many cases the data are from the US of the 1980s. (See Jappelli, 1990; and Cox and Jappelli, 1993). Ethnicity seems to be related to credit constraints (Duca and Rosenthal, 1993). Credit availability is particularly problematic if financial markets are regulated; credit constraints had a significant effect on the demand for motor vehicles before financial liberalisation (Alessie, Devereux and Weber, 1997).

3 The model

To model the relationship between macroeconomic developments and household distress, we use a variant of the approach by Del-Rio and Young (2005b). Their model of household distress is based on a household budget constraint: distressed household are those, whose surplus (which in their case is income diluted by interest payments on debt) falls below a certain threshold, the ‘comfortable’ level of consumption of the household. Del-Rio and Young show that, from this definition, it follows that household distress depends on the level of interest rate, the level of debt, the level of income, and certain other characteristics which affect the ‘comfortable’ level of consumption of the household.

In our formulation, we use t to index time, i to index household, Y to denote household disposable income, D to denote debt, r to denote the rate of interest on loans, and s to denote the loan instalment that the household is committed to. Household surplus SRPLS is defined

$$\text{SRPLS}_{t,i} \equiv Y_{t,i} - (r_t + s_t)D_{t-1,i} \quad (3.1)$$

Denote by $\overline{MC}_{i,t}$ the minimum level of consumption that a household is ‘comfortable with’ at period t and by $\overline{D}_{t,i}$ the pledgeable amount of wealth.² Household distress, DSTRS, is then defined

² Pledgeable wealth includes both assets that can be used directly to finance consumption (deposits etc.), and tangible or non-tangible wealth that can be used as collateral for a loan.

$$DSTRS_{t,i} \equiv \begin{cases} 1 & \text{if } SRPLS_{t,i} + \bar{D}_{t,i} < \bar{MC}_{t,i} \\ 0 & \text{if } SRPLS_{t,i} + \bar{D}_{t,i} \geq \bar{MC}_{t,i} \end{cases} \quad (3.2)$$

Under (3.2), a household is distressed if its surplus (income diluted by debt service payments) incremented by the possibility to incur new debt, is ‘small enough’. Our definition of distress deviates from the definition used by Del-Rio and Young in two respects. Firstly, our formulation allows the households’ prior commitment to make loan instalments to affect surplus and thus be a potential cause for distress. Secondly, our formulation underlines the fact that distress may depend on pledgeable wealth $\bar{D}_{t,i}$. Both empirically and theoretically, it is difficult to justify ruling out the possibility that households can sustain consumption temporarily by taking more debt or running down their stock of liquid assets.

Below, we use a micro data set to construct proxies for DSTRS and SRPLS at household level. The variables $\bar{MC}_{t,i}$ and $\bar{D}_{t,i}$, however, are not observable, and one has to make further assumptions to arrive at estimable relationship

- 1) As regards $\bar{MC}_{t,i}$, the level of consumption that the household is comfortable with, this is bound to depend on a number of idiosyncratic factors such as tastes, the local price level, health, family type, etc. Assume, for now, that $\bar{MC}_{t,i}$ is independent, and log-normally distributed across households with mean \bar{mc} . (The mean \bar{mc} becomes the ‘constant term’ in estimations, while the assumption about its logistic distribution enables the use of logit analysis for subsequent estimations).
- 2) As regards $\bar{D}_{t,i}$, a standard assumption in economic models is that households can borrow against some fixed proportion of their net wealth. Denote by $W_{t,i}$ the expected value of household tangible and nontangible net wealth at t , and by \bar{d} the pledgeable proportion of this value. Define

$$\bar{D}_{t,i} \equiv \bar{d}W_{t,i} \quad (3.3)$$

Under assumptions given in 1) and 2), the probability of distress of a household is

$$\Pr(DSTR_{t,i} = 1) = \Pr(\bar{mc} - SRPLS_{t,i} - \bar{d}W_{t,i}) \quad (3.4)$$

If surplus and wealth were observable at household level, the unobservable parameters in (3.4) could be estimated by logit. If forecasts of household level surplus and wealth were available, one could then utilise the estimated model to make forecasts about the number of distressed households in the economy.

However, estimates of household wealth are not included in the data set used in this study. In practice, estimation of wealth is problematic at best, especially when it comes to the net present value of both tangible and human wealth, defined as the discounted value of future wage income, which is the relevant wealth concept under an inter-temporal budget constraint. In practice, one has to work with equations of the type

$$\Pr(DSTR_{t,i} = 1) = \text{Pl}(\beta_0 + \beta_1 * SRPLS_{t,i} + \beta_2 * \text{other}) \quad (3.5)$$

where β are parameters and ‘other’ refers to a vector of potentially useful control variables that are included in the data set (to be introduced below). Theoretical considerations given lead one to expect that the parameter β_1 should be negative.³

4 The data

The data, Tulonjaon palveluaineisto (the service data on income distribution), is from Statistics Finland. The sample, which is collected annually by survey and appended from various registers, covers information about the socioeconomic status, income and debt of 27 000–30 000 persons (depending on the year), or 11 000–13 000 families, of which roughly a half have debt.

As we are here interested in the mapping between macroeconomic developments and loan defaults, the indebted family is the relevant unit of study.⁴ For modelling purposes, it is useful to have information about the family for both the current and the previous year. Each family participates in two consecutive surveys so that there are about 3 000 such families in the sample of each observation year.

Since the late 1990’s, the sample variables have included various indicators of economic distress

- a) a subjective opinion about whether the household debt level has risen above sustainability (four classes: yes; no; does not want to say; cannot say),

³ The estimated parameter of SRPLS does not need to be negative unity because scaling of the variables is arbitrary in (3.4): as long as the estimated parameter is negative, rescaling of SRPLS and the other variables by the absolute value of β_1 gives (3.4).

⁴ Use of the family as the relevant unit is necessitated by the data (certain key variables only refer to families). From a legal perspective, a debt contract usually binds individuals, not households, and wealth and income is by law individual in a marriage. However, as many loans (such as housing loans) are typically taken jointly by family members, family assets are typically used as collateral for loans, and as spouses typically do choose to assist each other in debt repayment and have a joint family budgets, the family may for most purposes be the relevant unit of study.

- b) is the household's property being collected by distress (yes; no; does not want to say; cannot say),
- c) a subjective opinion about whether the household has had difficulty with paying bills or making payments during the ongoing year (very often; often; sometimes; once; never; does not want to say; cannot say).
- d) a subjective opinion about whether the household has had difficulties in servicing loans (Once, more often than once, never, unable to say, unwilling to say).

After various trials, the approach chosen was to derive from the variables a), b) c) and d) one binary indicator of economic distress (DSTRS). The binary indicator variable DSTRS gets value 1 if the household either signalled an unsustainable debt level, or the household property was under collection by distress, or the person had very often or often difficulty with paying bills or making payments or has at least once had difficulties in servicing loans during the survey year.

The variable SRPLS (surplus) is constructed by diluting from household disposable money income (after tax) their estimated annual debt service payments: interest payments and instalments of housing and student loans plus estimated interest payments on other loans. The variable SRPLS/C is surplus per the number of consumption units in the family.⁵

Table 1 shows the means of the variables DSTRS, SRPLS and SRPLS/C in five consecutive samples 2000–2004. It is observed that the average share of distressed households varied between 13 %–19 % in the samples. To interpret, this is the proportion of indebted households that show even mild distress in the survey. Experiments with narrower measures of distress indicate that the simulation results appear to be qualitatively very robust to alternative definitions of distress.

The table also indicates that the average amount of surplus varied between 25 000 € and 30 000 € per household, and 15 000 € and 19 000 € per household consumption unit. It appears that the average proportion of distressed households was lower in the more recent samples compared to the earlier ones. At the same time the amount of surplus income available for each household for each household consumption unit increased.

⁵ Income and loan service payments are available in the data except that the average interest rate of other loans has been estimated by applying average interest rate of other household loans (from the Bank of Finland interest rate statistics) to the stock of other credit of the household during the previous year. The number of consumption units in a family has been measured by the modified OECD scale. An alternative measure of surplus, the debt service ratio, has at best a very weak explanatory power on distress in our data.

Table 1.

Sample means (weighted) of the distress variable and the surplus variables

Surplus variables = SRPLS and SRPLS/C; C = consumption units.⁶

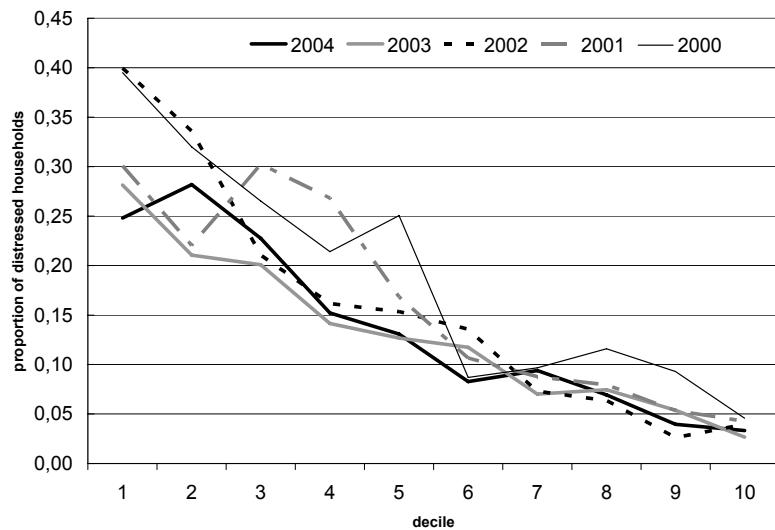
sample	DSTRS (%)	SRPLS (€)	SRPLS/C (€)
2000	18.8	25766	15349
2001	16.3	27526	16573
2002	16.0	28393	17105
2003	13.0	29335	18016
2004	13.6	29849	17946

Chart 1 shows the proportion of DSTRS in each decile of SRPLS/C in the five samples (households with smallest surplus are in class 1). In all samples, there appears to be a general tendency for distress to increase as the surplus decreases.⁷ The fact that a negative relationship exists between distress and surplus in all samples already gives some indication that surplus may be a useful indicator of household distress. Surplus correlates well with household vulnerability.

Chart 1.

The average of DSTRS in deciles of SRPLS/C in five data samples

(1=lowest decile and 10= highest decile of SPRLS).



⁶ The means are weighted by probability weights given in the original data.

⁷ With the debt service ratio, no such relationship exists in the data.

5 Selection and estimation of an empirical model

In addition to the absence of wealth, a major complicating factor in the econometric analysis is that the underlying model (3.4) is likely to be time dependent. For one thing, the level of minimum consumption \bar{m}_C is likely to depend on such factors as the general price level and demographics, which are not constant in time. Furthermore, the pledgeable proportion of household wealth \bar{d} varies with the credit policy of banks.

In principle, such temporal variation can, to some extent, be controlled in panels with a sufficiently long time dimension. However, the present data set does not allow for a construction of a long panel, as each household only participates in two consecutive surveys.⁸

Promising results have been obtained by working with the ‘mini panels’, each consisting of the about 3 000 households that participate in two consecutive surveys. From the available surveys, five such mini panels can be constructed at present time, the first one consisting of households that participated in the surveys of 1999 and 2000, and the latest mini-panel consisting of households that participated in the surveys of 2003 and 2004.

Table 2 shows the estimated logit -coefficients of four alternative variants of the model (3.5) in the five mini-panels. In models 1a and 1b, the surplus variable is SRPLS (surplus at household level), and in models 2a and 2b the SRPLS/C (surplus at household consumption unit level). 1a and 2a include no other explanatory variables, while models 1b and 2b include lagged distress (DSTRS(-1)) as an additional explanatory variable. Various other variants of equation (3.5) have been estimated but are not discussed here.

⁸ One possible approach is to group the data in repeated cross sections to estimate the parameters of model (3.5). This approach has not been successful: the estimated relationships have been weak, and the estimated coefficients typically had signs that are inconsistent with the theoretical model outlined above. Such problems may be due to the loss of information related to grouping, and the fact that the time dimension of the data is still relatively short.

Table 2.

**Logit coefficients in four models estimated
from the five ‘mini panels’.**

Probability of DSTRS=1 is the endogenous variable.

		2000	2001	2002	2003	2004
Model 1a	constant	-0.41	-0.58	-0.32	-0.70	-0.94
	SRPLS	-0.00005	-0.00004	-0.00006	-0.00005	-0.00004
Model 1b	constant	-2.0	-1.7	-1.5	-1.9	-2.3
	SRPLS	-0.00002	-0.00003	-0.00004	-0.00003	-0.00002
	DSTRS(-1)	3.0	2.5	2.6	2.7	2.8
Model 2a	constant	-0.31	-0.10	0.45	-0.35	-0.67
	SRPLS/C	-0.00008	-0.00011	-0.00015	-0.00010	-0.00007
Model 2b	constant	-2.1	-1.3	-0.8	-1.6	-2.2
	SRPLS/C	-0.00003	-0.00008	-0.00012	-0.00007	-0.00004
	DSTRS(-1)	3.0	2.4	2.5	2.7	2.7

In all estimations reported in table 2, the signs of the estimated parameters conform with prior expectations: surplus affects the probability of distress negatively, and past distress affects the probability of future distress positively. The large estimated coefficient of past distress suggests significant persistence in distress, and inclusion of this variable in the model diminishes the estimated effect of surplus on distress.

The marginal effect of a change in surplus on the probability of distress is in general significantly larger in the simpler models 1a and 2a than in the more complicated models 1b and 2b (typically by a factor of about 2).⁹ In 2004 (as also during the previous years), model 2a had the largest marginal effect. By a mechanical (ie linear) readjustment of scale we get that, in model 2a with 2004 coefficients, a 10% increase in the annual surplus per consumption unit from 17946 € to 19741 € leads to a decrease in the probability of distress of the household from 11.8% to 10.4%.

We have made some attempt to benchmark our estimated marginal effects with previous results obtained with at least roughly comparable Finnish data. In particular, Hyttinen et al (2006) study the probability of payment difficulties by logit with the 1998 wealth survey. They find that a 10 percentage point increase in the ratio of debt service payments to income leads to a 4.3 percentage point increase in the probability that a household had difficulties in making payments, and a 2.7 percentage point increase in the probability that a household has

⁹ In 2004, the marginal effect of surplus on output was -7.7E-6 model 1a and -3.0E-6 for model 1b, and the marginal effect of surplus per consumption unit was -3.7E-6 and -1.5E-6 for models 2a and 2b respectively. Marginal effects in models 1a and 1b are larger than in models 2a and 2b, because the exogenous variable in the latter models is scaled down by the number of consumption units in each household, which was on average 1.8 in 2004. After making a mechanical correction of consumption units, the marginal effects of the simpler models 1a and 2a are typically close to each other, as are the marginal effects of the more complicated models 1b and 2b.

difficulties in servicing debts. According to calculations based on model 2a (with 2004 coefficients), a 10 percentage point increase in loan service payments leads to a 1.8 percentage point increase in the probability of distress (from 11.8% to 13.6%). While differences in approach significantly complicate the comparisons, it appears tentatively that in our estimated models the probability of distress may respond more mutedly to changes in household budgets than in the models of Hyttinen et al. Further research is needed to explain these differences.

The in sample properties of the models given in table 2 have been studied by standard statistical tests. All in all, the more complicated models 1b and 2b significantly outperform the simpler models 1a and 2a in in-sample tests. The pseudo R² measure varies between 5%–10% for the two variable models 1a and 2a, and between 20% and 30% for models 1b and 2b. By the ROC curve analysis, models 1a and 2a have moderate and the models 1b and 2b good ability to distinguish between the different types. Test results indicate that all models may suffer from some form of misspecification, which could be the result of omitting some relevant variables eg wealth from the model specifications. While the apparent problems in model specification are worrying, they do not necessarily imply that the models are useless in forecasting. The forecast ability of the models remains an empirical issue which can be tested.

As the models are used for forecasting and simulation purposes, one could argue that the main emphasis in model selection should be placed on the estimated model's forecasting ability. To this end, the ability of the models to forecast the average share of distressed households in the sample of next year was tested in each mini-panel.¹⁰ This is a joint test of both the statistical model, estimated from previous year's panel, and the procedure for updating the exogenous variables used in the statistical model for next year by using aggregate statistics. The test was run under a number of updating routines, but differences in the updating of the exogenous variables did not affect the ranking of the models in terms of forecasting ability. In all cases, the updating procedure relied on available information about the development of aggregate income, interest rates and other macroeconomic variables. In this sense, the updating procedure used mimics the real life simulation environment in which the paths of the exogenous variables are updated by using scenarios generated by the Bank of Finland macroeconomic model 'Aino'.

The relative ability of the models to forecast average distress is reported in table 3.¹¹ The second column reports the mean square error of the forecast on

¹⁰ As data accumulates, one may wish to consider also many-step-ahead forecasts. See Kauppi and Saikkonen (2006) for a discussion of many-step-ahead forecast tests in binary response models.

¹¹ The updates of exogenous variables are made by a simple updating routine in which the growth rates of both household income and debt are assumed to be equal to the growth rate of aggregate household disposable income and the aggregate household debt stock)

average distress probability as an absolute amount, and the third column as a percentage of that of the best forecast (marked by 100%).

Table 3.

Mean square forecast error (MSFE) for the alternative models in absolute amount and as % of the best model

	MSFE	MSFE (as % of the best model)
model 1a	0.000199	106.9%
model 1b	0.000269	144.4%
model 2a	0.000187	100.0%
model 2b	0.000280	149.9%

The table shows that models 1a and 2a (the simpler models) perform much better than the more complicated models 1b and 2b in forecasting average distress. In our view, the problems model 1b and 2b may reflect the effect of the missing wealth variable: lagged distress functions as a proxy for wealth ‘in sample’, but it does not develop like wealth out of sample, during the forecast period.

Model 2a is the best forecaster and model 1a the second best forecaster of average distress probability. It is observed that the more complicated models 1b and 2b, which include the lagged distress as an additional regressor, underperform the best forecast by a considerable margin (their respective mean square forecast errors are 144.9% and 149.9% of the best model). The sources of forecast error of the best forecaster (model 2a) are discussed in appendix 1.

As a further benchmark of the forecast accuracy, the forecasting error of the best model 2a was compared with the forecast error of a random walk forecast, that next year’s average distress equals the previous year’s distress. The average mean square forecast error of the best model 2a was about 48% of the mechanical, random walk, forecast during the period 2000–2004. This exercise shows that the econometrically estimated model 2a outperforms the random walk forecast even during this period of relatively ‘average’ economic developments during which changes in the underlying exogenous variables are relatively small. The advantage of model 2a over the random walk forecast is likely to be significantly amplified during periods of more dramatic changes in the economy.

To summarise, the larger models 1b and 2b have superior in sample properties, but the simple models 1a and 2a outperform them in forecasting. Model 2a is the best forecaster over all. In analysis of financial stability, one on balance rather wants to use a model that exaggerates rather than dampens the risks. From this point of view, model 2a which has the largest marginal effects not only has the best forecasting record but it is also the safest choice.

6 The household distress simulator

The aim is to simulate the average probability of distress, debt at risk, uncovered credit risk, and loan losses in various macroeconomic scenarios. At present, the simulations are based on the latest survey of 2004. Consequently, as aggregate household debt and incomes have continued to increase fast, this implies that 2004 data do not correspond well to the current situation. In order to alleviate this problem the data were adjusted to better correspond to the situation in December 2005. These transformations are described in the appendix 2. The transformations give us a data set characterised by 2005 levels and 2004 cross-sectional distributions of variables.

As a first step in each simulation, the exogenous household level variables (income, loan rates etc.) are updated in line with the macroeconomic scenario used. Then, each household is assigned a probability of being in financial distress by using the latest coefficients of the logit equation 2a. The debt at risk of a household is the product of the probability of distress and the amount of debt of the household. Debt at risk is calculated separately for two types of debts, namely housing loans and other debt. To be more precise, the aggregate debt at risk for loan type k is calculated according to the following formula.

$$DR_k = \sum_{j=1}^N \frac{M_j D_{k,j} \text{Exp}(Z_j)}{1 + \text{Exp}(Z_j)} \quad (6.1)$$

where DR_k is the aggregate debt at risk of type k debt in the economy, j is the household number, N is the total number of households in the sample, M_j is the sample weight of the household j (total number of households in the economy represented by the sample household j), $D_{k,j}$ is the amount of debt of type k by household j , and Z_j is determined by the variables and coefficients of the logit model.

To calculate a proxy for the uncovered credit risk, the negative net housing equity of distressed households was calculated. Housing equity is said to be negative if the housing loan exceeds the value of the dwelling in the scenario. The formula for negative housing equity of distressed households is

$$NHE = \sum_{j=1}^N M_j \text{Max}(0, D_{j,housing} - H_j) \text{Exp}(Z_j) / \{1 + \text{Exp}(Z_j)\} \quad (6.2)$$

where NHE = aggregate negative housing equity, $D_{j,housing}$ is the stock of housing loans of household j and H_j is a proxy of the sales value of the dwelling of the household j . If there is no owner-occupied dwelling, this variable equals zero.

Aggregate uncovered credit risk is now defined as the sum of the negative housing equity of distressed households and the aggregate debt at risk of loans other than housing loans. The impact of different shocks on the aggregate amount of uncovered credit risk in the economy was calculated by adjusting the values of different income and interest rate expenditure variables at the household level. The sample weight, the number of consumption units and the amount of debt remain unchanged irrespective of the scenario in all the calculations.

Unfortunately there is no information on interest rate pegs at the household level in the data. Hence, an assumption on the interest rate peg was made at random for each household. It was assumed that all the loans of a household are either at variable or at fixed rates. It was assumed that 6% of household loans bear fixed rates whereas 94% of them bear variable rates. This assumption is based on aggregate level data. An interest rate shock is assumed to affect the debt servicing burden proportionately, provided the household has been assigned variable rate loans in the calculations. This simplistic approach ignores two factors that might be of importance. First, in some cases the debtor can purchase an insurance against negative income shocks, such as unemployment. Secondly, the interest rate structure of housing loans may be more complicated. For instance, variable rate loan contracts may include an interest rate ceiling clause; the rate of interest will not exceed a predefined level even if the reference rate does.

Interest paid on housing and student loans is tax deductible. If possible, the interest expenditure on these loans is deducted from capital income. If there is no capital income, the resulting deficit in capital income is tax deductible, implying that taxes decline by 28% of the interest expenditure. The upper limit for this tax reduction is EUR 1400 a year per adult. This ceiling is EUR 400 higher if the debtor has one child, and EUR 800 higher if there are more than one child. The simulator takes into account these tax deductions when it calculates the disposable income after tax and loan servicing costs (SPRLS).

Income taxation is a key factor affecting household disposable income. Income tax legislation is complicated, and it would be extremely time consuming to enter equations that would calculate the income taxes of each household precisely in all the scenarios. A naïve, exploratory method was used to approximate the marginal tax rate and to calculate a rough proxy for the change in disposable income at the household level in the various scenarios. This method is described in the appendix 4.

Unfortunately, the likely sales value of households' dwellings is not included in the original data. Hence, a proxy for this highly important variable was calculated (see appendix 3). In the various scenarios it is assumed that a price shock affects the prices of all the dwellings by the same percentage.

Each individual in the sample is assigned an unemployment risk. This risk is determined by a random variable drawn from a uniform distribution between 0 and 1. If an increase in unemployment is assumed, the employees with the lowest

variable values are assumed to lose their jobs. If, for instance, unemployment increases from 9 to 11 per cent, every employee whose unemployment risk variable is less than $(0.11-0.09)/(1-0.09) \approx 0.022$ is assumed to become unemployed. In this case, the income of the household is reduced by the difference of the reported individual wage income and the unemployment benefit. This approach does not take into account the fact that unemployment risk may depend on employee and employer specific factors. There are three different kinds of unemployment benefit systems in Finland. These benefits and their inclusion in the simulator are described in appendix 5.

7 Simulations

7.1 Macro scenarios

7.1.1 The basic scenario

First, the impact of a mild macroeconomic recession is simulated. This basic scenario was created with the help of the Bank of Finland macroeconomic model Aino. It manifests itself in several ways. The real GDP deviates by 3.9% from the BoF baseline forecast published in early 2006, implying a moderate 0.4 per cent decline in real GDP. Unemployment is assumed to surge by 3.3 percentage points and housing prices to decline by 12%. No interest rate shock is assumed; Euribor rates are assumed to evolve according to prevailing market expectations in spring 2006. Entrepreneur income is assumed to decline by 3.9% in the basic scenario; because of the lack of a better estimate, the elasticity of entrepreneur income with respect to deviations from the baseline GDP forecast is assumed to be equal to one, and to affect all the entrepreneur households by the same percentage.

7.1.2 The extreme scenario

Moreover, the impact of a major depression is simulated. In the early 1990s, the Finnish economy was hit by an exceptionally deep recession. No other developed country has experienced such a collapse in economic activity in the post World War II era. The impact of this crisis on the banking sector has been described by eg Koskenkylä and Vesala (1994).

A depression of this kind is highly unlikely, but it may be of interest to make some calculations on the impact of a comparable extreme scenario. Historical data is used to identify the most extreme four quarter changes in the three key variables. The most extreme four quarter drop in housing prices was 19%

(Q3/1991-Q3/1992). The highest increase of unemployment was 4.39 percentage points. The highest annual increase of the money market rate was 402 basis points. Entrepreneur income is assumed to decline by 11.4%, as happened in 1990–1991. This decline is assumed to affect all the entrepreneur households by the same percentage. These shock components are entered simultaneously in the simulator. The shocks would be even more extreme if they were based on observation periods longer than four quarters.

As can be seen in table 2, the coefficients of the logit equation vary over time. We tentatively interpret these changes as indicators of banks' credit policies (see appendix 1). During extreme conditions banks' credit policies might tighten, at least in the case of new loans, even though some banks might allow distressed customers to postpone amortizations. The extreme scenario was carried out using the most extreme coefficient, from the 2002 data, for the variable SRPLS/C.

7.1.3 Standardised shocks

The scenario analysis does not tell us much about the sensitivities of household loan servicing capability to various macroeconomic factors. These sensitivities, however, can be essential to our understanding of the potential stability concerns related to the rapid growth of household debt during the last few years. The three potential shock factors to be analysed are, again, unemployment, housing prices and interest rates.

Ideally the individual shocks should be somewhat comparable, but what kind of an interest rate shock would be comparable to an unemployment shock of, say, one percentage point? In order to calibrate these factors some basic descriptive statistics are needed. The standard deviations of these variables' differences were calculated using quarterly Finnish data for a 20 years period. As can be seen in table 4, in terms of being either extreme or moderate, a 0.6 percentage point increase in the unemployment rate in one quarter is comparable to a 3.4% decline in house prices or a 98 basis point increase in interest rates. In the following, these one standard deviation shocks are called standardised. By calculating the impact of a standardised shock we can obtain a rough measure of the sensitivity of uncovered credit risk with respect to adverse macroeconomic disturbances.

Table 4.

Some descriptive statistics on macroeconomic phenomena

Q1/1986-Q4/2005	Dwelling prices $(P_t/P_{t-1}-1)*100$	Unemployment rate, in % $U_t - U_{t-1}$	3 months money market rate, in % $R_t - R_{t-1}$
Mean	1.19	0.04	-0.11
Std Dev	3.42	0.59	0.98
	<i>Minimum</i>	<i>Maximum</i>	<i>Maximum</i>
Largest adverse 4 quarter change	-19.1 %	4.39	4.02

7.2 Results

The simulation results are reported in table 5. In the basic scenario, the total amount of problem loans does not change much in the simulation compared to the actually observed amount. In fact, the uncovered credit risk levels are in the basic scenario surprisingly insensitive to the underlying macroeconomic disturbance. In relative terms, the increase of negative housing equity of distressed households perhaps stands out as the only interesting exception. If the depression lasts for many years, dwelling price declines could cumulate and the total change would be stronger. However, the estimated level of negative housing equity in the initial situation is probably biased downwards because of the assumption that no household living in owner-occupied housing can have a negative housing equity before the shock. This makes the percentage increase appear high. Even after this shock housing loans would pose no serious threat to the financial system.

The extreme scenario gives stronger results. In this case the uncovered credit risk would increase by almost EUR 5 billion, when compared to the model-based estimate in the absence of shocks. In 2005, operating profits of the Finnish banking sector totalled EUR 1.9 billion (FSA 2006). The final sum of banks' loan losses would probably be significantly lower. However, we cannot rule out the possibility that under extreme conditions credit risks of household loans alone could pose major stability problems, even though this situation is highly unlikely.

Table 5.

Simulation results in various scenarios

mill €	% of indebted households distressed	Loans (excl housing loans) of distressed households	Housing loans of distressed households	Negative housing equity of distressed households	Uncovered credit risk¹⁾	Add impact on negative housing equity of distressed households	Add impact on other loans	Increase in uncovered credit risk
Observed 2005 situation - model based	13,9 %	2 572	5 964	426	2 997			
Basic scenario	14,1 %	2 616	6 075	635	3 251			
Extreme scenario - 2002 coefficients	20,9 %	5 725	12 410	2 001	7 725			
Add. unemployment (+0.59 %)	14,2 %	2 620	6 091	636	3 256	1	4	5
Additional dwelling price shock (-3.4%)	14,1 %	2 616	6 075	708	3 323	72	0	72
Interest rate shock (+98 basis points)	14,3 %	2 727	6 238	658	3 386	23	112	135

¹⁾ Negative housing equity of distressed households + non-housing loans of distressed households

The next step is to analyse the sensitivity of uncovered credit risk to each of the three standardised macroeconomic shocks (see table 4). These standardised shocks are added to the basic scenario. As can be seen in the table 5, the standardised interest rate shock affects the amount of uncovered credit risk more than any of the other shocks. Hence, when one evaluates credit risks of household loans, particular attention should be paid to the development of interest rates. On the other hand, this result is probably partly due to the observation period. In the early 1990s, interest rates were substantially higher and more volatile than what they have been in the monetary union. If the standard deviation were estimated with more recent data, we would probably have a weaker standardised interest rate shock, but the number of observations on the EMU era may be too limited for meaningful statistical calculations.

The impact of an additional housing price shock would also be of some importance, even though its impact is limited to the negative housing equity. If housing equity affected the probability of being financially distressed, house price shocks would have other effects as well.

Perhaps the most interesting finding is the surprisingly weak estimated impact of unemployment shocks. This may seem surprising, since one would expect that for many households loss of (wage) income due to unemployment presents perhaps the most important source of income risk. The result may be due to assumptions on unemployment benefits. New housing loans are typically granted to persons in full time jobs. In the above calculations, it has been assumed that such persons are entitled to the earnings-related allowance; in most cases this allowance helps to maintain a satisfactory income level, unless the duration of unemployment is prolonged.

7.3 Robustness tests

7.3.1 The unemployment benefit system

In the above simulations the unemployment shock seems to have only a very marginal contribution to the household sector credit risk. We think this result is counterintuitive and needs further analysis. More specifically, we think we need to check for the role of the unemployment benefit system in explaining this outcome. The simulations above assume that all the wage earners who have been employed for at least 10 months on a full time basis are entitled to earnings-related unemployment benefits in the case of unemployment. Is the result robust if this assumption is relaxed? Not every employee is member of an unemployment insurance fund. The duration of the earnings-related allowance is limited to 500 working days. Hence, if unemployment worsened permanently, this source of household income would gradually diminish.

In the following, a more pessimistic scenario is applied. It is assumed that none of the employees who lose their jobs in the scenario get more than the labour market subsidy, which is much less than the earnings-related benefit (see appendix 5). Furthermore this subsidy is conditional on spouse income, making it possible for an unemployed worker to end up in a situation where the system declares him or her not entitled to any unemployment benefit. This particular feature of the unemployment system has also been taken into account in the calculations. However, as we can see in table 6, even in this case the impact of an additional unemployment shock on the amount of uncovered credit risk remains surprisingly weak. It is moreover worth mentioning that these calculations do not take into account the likely increase of other transfers from the government, such as subsistence grants.

Table 6.

Results if no earnings-related unemployment benefits

Assumption - all the newly unemployed will get the labour market subsidy only; mill €	% of indebted households distressed	Loans (excl housing loans) of distressed households			Housing loans of distressed households with negative housing equity	Uncovered credit risk	Add impact on housing loans of distressed households with negative housing equity	Add impact on other loans	Increase in uncovered credit risk
		2 640	6 165	647	3 287	2	6	8	
Basic scenario	14,3 %								
Add unemployment shock (+0.59 %)	14,4 %	2 647	6 190	648	3 295	2	6	8	

Basic scenario described in section 7.1.1.

Even though this result may seem surprising, it may simply reflect the inadequacy of the postulated size of the unemployment shock in the type of exercise we are

interested in tables 4 and 5; to generate a sizable adverse income shock to household finances through unemployment we probably need a much larger increase in the unemployment rate. The results above do suggest that household income is quite well insured against relatively small shocks to unemployment.

7.3.2 Ability to postpone amortizations

The simulations assume that households cannot affect the financial burden of loan servicing. Both interest payments and amortizations of housing and student loans are exogenously given throughout the simulation exercise. Interest rate expenditure is likely to be exogenous even in reality. However, it is plausible that in practice households are, at least to some extent, able to postpone amortizations. In some loans this is an explicit part of the original contract. The monthly loan servicing cost may be fixed for the whole loan period, but the quantity of amortization depends on the sum of money needed for interest payments.

The disposable income variable (SRPLS) was recalculated. It was assumed to equal the difference of net income (after tax) and interest rate payments. But unlike in the basic model, no amortizations are deducted from SRPLS. As can be seen in the table 7, the uncovered credit risk is relatively insensitive to the assumption on the possibility to postpone amortizations, as the estimates are only slightly lower than in the basic model of section 7.2.

Table 7. **Results if households can postpone amortizations**

No compulsory amortizations; mill €	% of households distressed	Loans (excl housing loans) of distressed households		Housing loans of distressed households	Negative housing equity of distressed households	Uncovered credit risk
		households	distressed	households	households	
Observed situation - model based						
	12,4 %		2 352	4 723	334	2 686
	12,6 %		2 393	4 812	495	2 888

Basic scenario described in section 7.1.1.

8 Conclusions

Our analysis contributes to the literature on household indebtedness, in which studies on the determinants of household debt seem to be particularly commonplace. In contrast to many of these studies, we focus on loan servicing problems and out of sample forecasting and simulation, rather than purely on in-sample goodness-of-fit of an estimated model for household indebtedness. We use the conceptual framework of Del-Rio and Young (2005b). A household is financially distressed if its income after tax and loan servicing costs is ‘too low’.

Our proxy for financial distress was defined as a simple function of the information collected in interviews. The definition for distress used in these simulations is relatively broad: a household is classified financially distressed if the interviewee reports any problems related to loan servicing or paying bills. The number of distressed households declined in 2000–2004, but even in the most recent data almost 14% of households still experienced some distress. A number of logit models for the probability of a household being financially distressed were estimated. These estimations tried to measure the dependence of financial distress on a relatively limited number of key variables. These models were estimated separately using household level data for five consecutive years.

We tested whether the estimated model could be useful in forecasting the average number of distressed households in the next year’s sample under the assumption that an accurate macroeconomic forecast is available. The macroeconomic situation facing households at the micro level is generated under the assumption that the determinants of distress, such as interest rates, income and debt, evolve similarly for all the households. In terms of forecasting performance the best model was relatively simple, as it incorporates only one explanatory variable, income net of taxes and loan servicing costs per consumption unit. In particular, and somewhat surprisingly, our results suggest that past financial problems should not prove useful in forecasting financial problems in the household sector. Our simple model is shown to significantly outperform the competing alternatives, including a mechanical one where the current average distress is extrapolated into the next period, during the estimation period.

The model with the best forecast performance and the estimated logit coefficients were used in a number of simulation exercises concerning changes in households’ financial distress to macroeconomic shocks. The 2004 data was adjusted to correspond to the situation in 2005. We simulated the effects of shocks to unemployment, interest rate and housing prices on the degree of financial distress of the household sector. Debt at risk, defined as the total amount of debt of distressed households, seems to be surprisingly insensitive to these shocks. Shocks to unemployment seem to have only a marginal contribution to debtor households’ financial problems. Interest rate shocks have a somewhat stronger

impact because most of the household loan stock bears variable interest rates, but even this effect is moderate. The main findings corroborate conclusions drawn on past experience. In most states of the economy household loans bear a relatively low credit risk to banks. However, under extreme conditions with a coincidence of large and persistent adverse shocks to unemployment, interest rates and housing prices, even household loans could become a threat to financial stability.

Our approach is obviously quite simple and straightforward and does not take into account all the relevant factors affecting the fragility of household finances. For instance, under certain conditions housing loans of young households can be offered a partial government guarantee, which reduces banks' credit risks. Moreover, housing loans may be covered by an insurance against unemployment, health problems and other negative income shocks, which may also be an efficient risk mitigation technique. A useful extension of our approach would thus be to control for these in the estimation procedure.

Future work aims at developing the model further to increase its forecast accuracy. Other explanatory variables for distress could conceivably be found. The updating process can be improved. Moreover, the approach is clearly partial. It does not take into account indirect effects of household debt on banking crises. For instance, domestic industries are likely to suffer if consumers have to cut down expenditure, which might lead to bankruptcies and credit losses. Moreover the analysis does not fully incorporate the possibility of the coincidence of very large shocks to interest rates, unemployment and housing prices, ie of very high interest rates, very high unemployment rate and very low housing prices.

From the point of view of financial stability, it would be important to carry out a companion simulation study of corporate distress. Furthermore, the issue of feedback effects of financial instability on the macro economy should be placed high on the agenda for future research.

References

- Alessie, R – Devereux, M P – Weber, G (1997) **Intertemporal Consumption, Durables and Liquidity Constraints: A Cohort Analysis.** European Economic Review 41, 37–59
- Barrell, R – Davis, E P (2004) **Costs of Financial Instability, Household-Sector Balance Sheets and Consumption.** National Institute of Economic and Social Research discussion paper 243.
- Bowie-Cairns, H-A – Pryce, G (2005) **Trends in Mortgage Borrowers' Repayment Difficulties.** CML Housing Finance Issue 11/2005.
- Brookes, M – Dicks, M – Pradhan, M (1994) **An empirical model of mortgage arrears and repossession.** Economic Modelling 11, 134–144.
- Brown, S – Taylor, K (2005) **Household Debt and Financial Assets: Evidence from Great Britain, Germany and the United States.** University of Leicester Working Paper 05/2005.
- Brown, S – Garino – Taylor, K – Price, S (2003) **Debt and Financial Expectations: An Individual- and Household-level Analysis.** Economic Inquiry 43, 100–120.
- Cox, D – Jappelli, T (1993) **The Effect of Borrowing Constraints on Consumer Liabilities.** Journal of Money, Credit and Banking 25, 197–213.
- Cox, P – Whitley, J – Brierley, P (2002) **Financial Pressures in the UK household sector: evidence from the British Household Panel Survey.** Bank of England Quarterly Bulletin, Winter.
- Crook, J (2001) **The Demand for Household Debt in the USA: Evidence From the 1995 Survey of Consumer Finance.** Applied Financial Economics, Vol. 11, 83–91.
- Davydoff, D – Naacke, G (2005) **L'endettement des ménages européens.** Observatoire de l'épargne européenne, august 2005.
- De Doncker, H (2006) **Crédits aux particuliers – Analyse des données de la Centrale des Crédits aux Particuliers.** Banque Nationale de Belgique working paper 78.

Del-Rio, A – Young, G (2005a) **The Determinants of Unsecured Borrowing: Evidence from the British Household Panel Survey.** Bank of England 2005, Working Paper 263.

Del-Rio, A – Young, G (2005b) **The Impact of Unsecured Debt on Financial Distress Among British Households.** Bank of England 2005, Working Paper 262.

Duca, J V – Rosenthal, S S (1993) **Borrowing Constraints, Household Debt and Racial Discrimination in Loan Markets.** Journal of Financial Intermediation 7, 77–103.

Elmer, P J – Seelig, S A (1998) **Insolvency, Trigger Events, and Consumer Risk Posture in the Theory of Single-Family Mortgage Default.** FDIC Working Paper 1998–3.

Friedman, M (1957) **A Theory of the Consumption Function.** NBER Princeton.

FSA (2006) **Financial Supervision Authority, Rich profits for banks in 2005.** FSA Newsline 2/2006 – economic issue.

Hyytinens, A – Määttänen, N – Johansson, E (2006) **Omistusasunnon ja asunto-lainan rooli kodin taloudessa.** Kansantaloudellinen aikakauskirja. 102 vsk. 1/2006, 48–66 (In Finnish).

Jappelli, T (1990) **Who is Credit Constrained in the U.S. Economy?** The Quarterly Journal of Economics 105, 219–234.

Kauppi, H – Saikkonen, P (2006) **Predicting U.S. Recessions with Dynamic Binary Response Models.** Manuscript available at <http://www.valt.helsinki.fi/blogs/saikkonen/post65.htm>.

Koskenkylä, H – Vesala, J (1994) **Suomen talletuspankit 1980–1993: kasvun ja kriisin vuodet.** Kansantaloudellinen aikakauskirja 90, (2/1994), 171–196 (In Finnish).

La Cava, G – Simon, J (2003) **A Tale of Two Surveys: Household Debt and Financial Constraints in Australia.** Research Discussion Paper, No. 08–2003. Reserve Bank of Australia.

Liu, D-Y – Lee, S-P (1997) **An Analysis of Risk Classifications for Residential Mortgage Loans.** Journal of Property Finance, Vol. 8, No. 3, 207–225.

Magri, S (2002) **Italian households' debt: determinants of demand and supply.** Bance d'Italia, Temi di discussione 454, october 2002.

Martins, N – Villanueva, E (2003) **The Impact of Interest-Rate Subsidies on Long-Term Household Debt: Evidence from a Large Program.** Lokakuu 2003, UPF Working Paper No. 713.

May, O – Tudela, M (2005) **When is Mortgage Indebtedness a Financial Burden to British Households? A Dynamic Probit Approach.** Bank of England Working Paper 277, october 2005.

Riiser, M D – Vatne, B H (2006) **Utviklingen I husholdningenes gjeld. En analyse på mikrodata for perioden 1986–2003.** Norges Bank, Penger og Kreditt 1/2006.

Tudela, M – Young, G (2005) **The Determinants of Household Debt and Balance Sheets in the United Kingdom.** Bank of England, Working Paper No. 266.

Quercia, R G – McCarthy, G W – Stegman, M A (1995) **Mortgage Default Among Rural, Low-Income Borrowers.** Journal of Housing Research, Vol. 6, No. 2, 349–369.

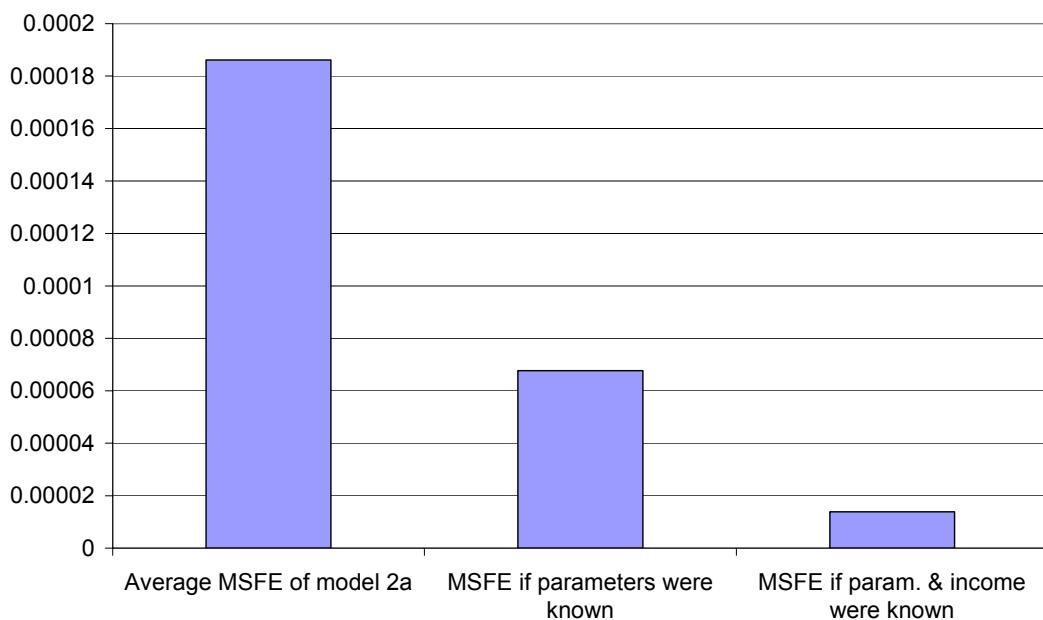
Appendix 1

Decomposition of forecast error in model 2a

Chart 2 illustrates the main sources of forecast error for model 2a, which is the best forecaster. The left side pillar shows the actual mean forecast error of model 2a in the whole sample. The centre pillar shows the ‘forecast’ error of this model, if the estimated coefficients had corresponded with the true coefficients. This component of the error is, therefore, solely caused by the updating procedure (used for updating the exogenous variables in the model). The rightmost pillar measures the mean forecast error in the hypothetical case when the estimated coefficients are correct and the estimates of income developments at the household level conform exactly with reality. The mean square error in this case is caused solely by the update routine for debt repayments.

Chart 2.

Forecast mean square error of model 2a with different update procedures



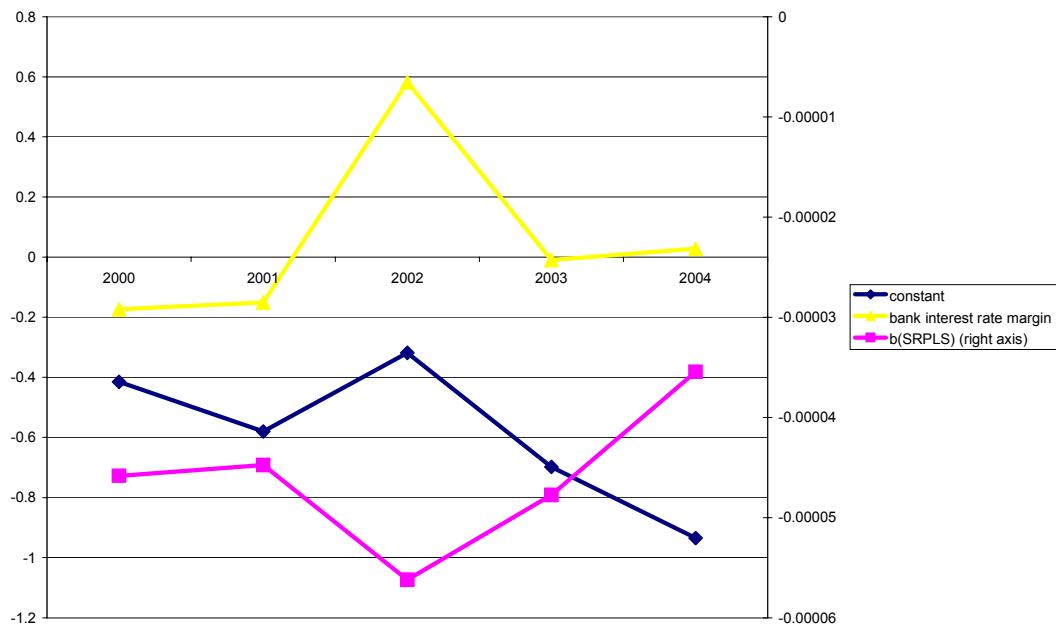
It is observed from the chart, that a lion’s share of the forecast error of model 2a is caused by changes in the model parameters during the forecast years. If the correct model parameters were known at the time when the forecast is made, this would reduce the mean square forecast error to close to one third of what it currently is. The remaining forecast error is caused by the fact that the forecaster does not know the development of income and debt at household level but, rather,

uses the aggregate growth rates of household income and debt to update household income and debt during the forecast year. A number of attempts were made to append the models with additional explanatory variables to improve their forecast performance. So far, these attempts have not been successful. It seems that, for forecasting purposes, the simplest models work best.

It is interesting that parameters of model 2a in fact move in time counter to each other. This is illustrated graphically by chart 3, which shows the evolution estimates of the constant term and the parameter for surplus in time. In 2002, the constant term was relatively small, and the effect of surplus large: during these years distress probability appears to have been relatively closely related to the economic surplus of the household. In years 2000, 2001, 2003 and 2004, on the other hand, the default probability is less dependent on current surplus. The fact that the parameter estimates are ‘mirror images’ of each other may be an indicator that there is a single underlying factor that is driving the evolution of the model parameters in time.

Chart 3.

Parameters of the model 2a and bank interest rate margin on consumer loans



The third line in the chart is the margin applied by Finnish banks when granting consumption loans. The evolution of this margin is, again, a mirror image of the effect of surplus on default probability. It could be, then, that one factor driving the instability of the model parameters is the credit policy of banks. During periods when banks apply tight margins for consumption credit (years 2000, 2001,

2003 and 2004), household distress is relatively less related to household surplus that during other periods, because household can balance their budget by increasing their borrowing. During periods when bank credit policy tight (2002), households found it harder and more expensive to borrow. The probability of distress is, subsequently, more related to household surplus income that during periods of slack credit policy.

Appendix 2

Updating the data

The net monetary income of each household was multiplied by

$$[1.034 \cdot \text{Wage income} + 1.0037 \cdot \text{Income transfers} + 1.048 \cdot \text{Entrepreneur income} + 0.956 \cdot \text{Capital Income}] / [\text{Wage income} + \text{Income transfers} + \text{Entrepreneur income} + \text{Capital income}].$$

Capital income does not include the imputed rent on dwellings. The multiplier 1.034 corresponds to the increase of the overall wage level in 2004–2005. The multipliers of entrepreneur and capital income correspond to the aggregate growth of these income components in 2004–2005. The multiplier 1.0037 corresponds to the increase of the KELA index, which should normally correspond to the increase in the cost of living; several public sector income transfers are indexed on it. The resulting augmented net monetary income is probably a satisfactory proxy for the net monetary income in 2005. This approach suffers from the problem that it does not take into account the impact of taxation on net income. Tax laws are subject to frequent changes, and not every income item is treated equally by tax laws. Moreover, income transfers are far from being homogenous, and not all of them are indexed. Transfers in the data include voluntary transfers between households.

On average, these adjustments increased the nominal income by 1.9%.

The amount of housing loans was multiplied by 1.167 for each household, which corresponds to the growth rate of households' mortgage loans at the aggregate level in 2004–2005.¹² Interest rates paid on housing loans were multiplied by 1.115, which corresponds to the growth of the loan stock and the decline in the average interest rate level.

The amount of student loans was held constant; there has been no clear trend in the amount of outstanding student loans.

Loans other than student and housing loans were multiplied by 1.113. This corresponds to the growth rate of other household loans in 2004–2005.

These proxies are certainly biased in a way or another. A major problem may be that the concentration of housing loans between households may have changed since end 2004. On the other hand, using data for 2004 as such would certainly be even more problematic in forward looking analysis.

¹² See Bank of Finland Financial Markets Statistical Review, Table 7.4

Appendix 3

Proxies for dwelling prices

The original data contains no proxy for the price of the dwelling. Instead, there is some information on the size, geographic location and type of building.

Using data on the geographic location, it is possible to identify dwellings in the following cities and towns: Helsinki, the rest of Helsinki metropolitan area, Tampere, Turku, Oulu, Porvoo, Kuopio and Jyväskylä. Most of these localities were chosen because of the relatively large number of inhabitants. Porvoo, however, was chosen because the price level of real estate is clearly higher than in other towns of the same size.

As to the price of flats and row house apartments, the source of price information was Statistics Finland release ‘House Prices’ (8 May 2006). In the case of apartments in the above mentioned localities, the proxy for the price for each dwelling was calculated by multiplying the size of the dwelling by the average price of a square metre. In the case of ‘Helsinki metropolitan area except Helsinki’, the price was calculated as the average of prices in Espoo-Kauniainen and Vantaa, the price for each locality being weighted by the number of inhabitants. As to the rest of Finland, the price EUR 1265 per square meter was used. The average price level in the rest of the country must be very close to this approximation; otherwise the national average outside the Helsinki metropolitan area could not be EUR 1327 reported by Statistics Finland. Again, the number of inhabitants was assumed to be the proper weight for each locality.

As to one-family detached houses, there is less price information available. The calculations are based on Statistics Finland data¹³. Average prices per square meter are available for Helsinki region (€ 1978/m²), other cities with more than 100 000 inhabitants (€ 1790/m²) and the whole country except Helsinki region (€ 1136/m²). With these data and population statistics it can be calculated that the average price in municipalities with less than 100 000 inhabitants must be about € 931 per square meter.

In surprisingly many cases the housing loan exceeds the price of the dwelling. This is not credible because dwelling prices have been increasing for many years and because the loan to value ratio should be less than 100%. There are many possible explanations to these cases, and based on the data, one cannot evaluate which of them is the typical correct explanation. In some cases the loan has been taken to buy a dwelling that has already been sold. It is also possible that some of the loans have been taken to buy a dwelling under construction, or still occupied by the previous owner. Some households may have several dwellings. Some

¹³ Release 16 Jun 2006.

dwellings may be very expensive relative to other dwellings in the same locality because of reasons such as an expensive local district or an exceptionally high standard. It was assumed that the housing wealth of no household can be lower than the housing loan, provided the household is living in an owner-occupied dwelling. Dwelling prices are adjusted accordingly.

Appendix 4

Controlling for non-linear marginal tax rates

Net income after tax should be a mechanistic and deterministic function of the different components of gross income and a few other factors. This relationship, however, is complicated because of the complexity of the tax system. Wages and capital income are taxed differently. Pensions and unemployment benefits are taxable, some other income transfers are not. The marginal tax rate for the wage income is higher for high income groups, and the distribution of wage income between family members affects the total amount of income taxes; if both spouses earn the same salary, the total income tax will be lower than in the case of another family with the same total gross income but only one breadwinner. The local tax rate varies between municipalities. Members of the two major religious communities (Lutheran and Orthodox churches) pay church taxes that depend on the locality. There is no special tax rate on entrepreneur income, but the tax system divides this source of income into earned and capital income. It might be possible to code the whole tax system in the simulator, including all the above mentioned details, but this would be extremely time consuming.

The second possibility is to use a rough statistical approach; the impact of gross income on net income is estimated with a simple statistical method. It would be possible to assume that the net disposable income is a linear function of the gross income and to run a regression with only one explanatory variable. This parsimonious approach would be simple and maybe even elegant, but it would make a very poor approximation of reality.

A more complicated statistical approach was taken. Different income components were tested as explanatory variables and, when appropriate, interactions and transformations of these income components were also included in the analysis. The data consists of households of the Statistics Finland income distribution survey, and the observation unit is a household, not a person. In the final specification, statistical significance was a central criterion for the choice of variables. Wage income is allowed to affect the net income in a particularly complicated way because of three reasons.

- The income tax system treats earned income in a particularly complicated way.
- Wages are the most important type of income among debtor households.
- Unemployment shocks are a central factor in the various scenarios.

Not much structure is assumed on the final specification. However, the regression was run with no constant term. A non-zero constant term in the equation would

imply that a household can have some net income even if there is no gross income, not even transfers from the government. The explaining variables are PaTu = gross wage income of the whole household; PaTu2 = PaTu squared; PaTu3 = PaTu to the power of three; PaYr = wage income multiplied by household entrepreneur income; PaTuKe = wage income multiplied by $\Sigma(\text{wage income of the person}/\text{wage income of the household})^2$; PaTuKe2 = wage income multiplied by $[\Sigma(\text{wage income of the person}/\text{wage income of the household})^2]^2$; TuSi = received transfers; OmTu = capital income; YrTu = entrepreneur income; YrTu2 = entrepreneur income squared.

Table 8.

The determination of net monetary income

	B	Std.Err.	t(11188)	p-level of B
PaTu	8.1E-01	4.0E-03	205.8	0.000000
Patu2	-1.4E-06	2.5E-08	-57.2	0.000000
Patu3	1.6E-12	3.8E-14	42.2	0.000000
PaYr	-1.4E-06	9.4E-08	-15.3	0.000000
PaTuKe2	6.1E-04	1.8E-04	3.4	0.000606
YrTu	7.8E-01	2.9E-03	263.7	0.000000
YrTu2	-3.2E-07	1.4E-08	-22.2	0.000000
TuSi	7.2E-01	2.9E-03	246.8	0.000000
OmTu	6.7E-01	7.4E-04	901.9	0.000000
Patuke	-1.0E-01	5.1E-03	-20.1	0.000000

Regression Summary for Dependent Variable: Net Monetary Income
R=.995 R²=.990 Adjusted R²=.990
F(11,11188)=1051E2 p<0.0000 Std.Error of estimate:
4296.0

This blurred equation should not be considered an analysis of the tax system. Its aim is not to test whether certain factors affect the net income or not; we know that all of them do. It certainly suffers from various biases, including biases caused by omitted variables (such as the wealth tax and real estate tax), and it is unable to take into account the marginal tax rate discontinuities in the central government income tax. However, this equation is probably relatively useful for its intended purpose. This intended purpose is very narrow and specific. The equation is aimed at calculating a satisfactory proxy for the marginal tax rate in various income intervals. If the wage income of a single person with no other income is close to zero, the marginal tax rate is about 19%. If a single person has an annual wage income of EUR 30 000, the marginal tax rate would be 37%. If the wage income is EUR 80 000, the marginal tax rate is 48%. These marginal tax rates are relatively close to the ones implied by state, municipal and church taxes in most localities. The marginal tax rate on capital income implied by the equation

(32%) is somewhat higher than the actual tax rate (28%), possibly because capital income correlated positively with the wealth tax.

The equation is aimed at analysing the impact of a decline in one component of gross income on net income. The final net income to be used in the various scenarios is not the one predicted by this arbitrary statistical model as such. Instead, the equation is used only as a method to calculate a proxy for the change in the net income at the household level in the various scenarios. If the net income predicted by the equation decreases by one euro, it is assumed that the net income in the scenario declines by one euro, when compared to the original situation. If there is no change in gross income, there is no change in net income either. To be more precise

$$\text{NetInc scenario} = \text{NetIncOriginal situation} + (\text{NetIncEquation, scenario} - \text{NetInc Equation, original situation}).$$

In case of unemployment, the unemployment benefit of persons who are assumed to lose their jobs is not treated as an income transfer but as wage income. This choice is based on the fact that unemployment benefits are treated as earned income by the tax law. Because no shocks are assumed to affect capital income and transfers, the coefficients of these variables are irrelevant to the final net income in the various scenarios. These variables are included in the equation only because it may be important to control for them.

This equation is not used to calculate the tax deductions of housing and student loans in the various scenarios. These deductions are calculated separately according to the rather simple and mechanistic rules stipulated in tax legislation.

Appendix 5

Unemployment benefits

There are three different statutory unemployment benefits in Finland

- Earnings-related allowance
- Basic allowance
- Labour market subsidy

A detailed description of these systems is available on the Social Insurance Institution web page (<http://www.kela.fi/in/internet/english.nsf/NET/081101150015EH?OpenDocument>). The three systems can be briefly summarised as follows.

The earnings-related allowance is paid if the person is member of an unemployment fund and satisfies certain conditions concerning past employment. This benefit consists of two components. The daily basic amount is EUR 23.50 for five days a week. The earnings-related amount equals 45% of the difference between the basic amount and previous daily salary. If the monthly salary before unemployment exceeded EUR 2115, the earnings-related part is somewhat lower. The benefit cannot exceed 90% of the previous salary. Beneficiaries with dependent children get additional increases. This benefit is not paid after 500 working days of unemployment.

If the unemployment has lasted for more than 500 working days, or if the person does not satisfy the previous employment condition, the labour market subsidy can be paid. This benefit equals EUR 23.50 per day for five days a week. Beneficiaries with dependent children get additional increases. Spouse income can reduce this benefit.

If the person is not member of an unemployment fund but satisfies the criteria of past employment, the basic allowance can be paid. In many cases the basic allowance equals the labour market subsidy.

All the benefits are conditional on certain criteria. The applicant must be registered as an unemployed person, willing and able to accept employment, be between 17 and 64 years of age etc. As of 31 December 2004, there were 138 000 persons receiving the earnings-related allowance, 143 000 persons receiving the labour market subsidy and 22 000 receiving the basic allowance. (See the statistical yearbook 2004 of the Social Insurance Institution, p. 227)

Because the benefits are determined in a simple and mechanistic way, the rules determining the allowances are entered as such in the simulator. There is one exception to this, namely the type of benefit to be paid. In the data, there is no information on membership in an unemployment fund, implying that one cannot

say who would get the earnings-related allowance in case of unemployment. In the simulations, it is simplistically assumed that persons who lose their jobs in the scenarios will get the earnings-related allowance provided they have been employed for at least ten months on a full time basis during the previous year. This ten months criterion is a rough approximation of the previous employment criterion of the unemployment benefit system. In reality, not all of them would be entitled to the earnings-related allowance because not all of them are members of an unemployment fund. Those who do not satisfy the condition of ten months of previous employment are assumed to be paid the labour market subsidy.

The tax system treats unemployment benefits as earned income. Hence, the simulator treats unemployment as a reduction in wage income and replaces the original wage income with the calculated unemployment benefit to recalculate the after tax net income.

The simulator is not used to calculate the unemployment benefits of persons who were unemployed in the original situation. As to them, the original income was used as such.

**BANK OF FINLAND RESEARCH
DISCUSSION PAPERS**

ISSN 0785-3572, print; ISSN 1456-6184, online

- 1/2007 Timo Korkeamäki – Yrjö Koskinen – Tuomas Takalo **Phoenix rising: Legal reforms and changes in valuations in Finland during the economic crisis.** 2007. 39 p. ISBN 978-952-462-346-9, print; ISBN 978-952-462-347-6, online.
- 2/2007 Aaron Mehrotra **A note on the national contributions to euro area M3.** 2007. 25 p. ISBN 978-952-462-348-3, print; ISBN 978-952-462-349-0, online.
- 3/2007 Ilmo Pyyhtiä **Why is Europe lagging behind?** 2007. 41 p. ISBN 978-952-462-350-6, print; ISBN 978-952-462-351-3, online.
- 4/2007 Benedikt Goderis – Ian W Marsh – Judit Vall Castello – Wolf Wagner **Bank behaviour with access to credit risk transfer markets.** 2007. 28 p. ISBN 978-952-462-352-0, print; ISBN 978-952-462-353-7, online.
- 5/2007 Risto Herrala – Karlo Kauko **Household loan loss risk in Finland – estimations and simulations with micro data.** 2007. 44 p. ISBN 978-952-462-354-4, print; ISBN 978-952-462-355-1, online.

Suomen Pankki
Bank of Finland
P.O.Box 160
FI-00101 HELSINKI
Finland



.2343