BLAME THE BANKERS? AN EMPIRICAL STUDY OF CYCLICAL CREDIT QUALITY

draft 21/4/2009

Risto Herrala[1]

Abstract

We test the hypothesis that credit quality deteriorates during credit booms. The test case is a pronounced cycle in connection with a banking crisis, ranked among the most extreme in international studies. A unique data set allows us to rate household borrowers during the cycle, and aggregate the ratings to the macroeconomic level. The post-crisis period is also studied to find changes in credit quality.

The method reveals significant variation in average ratings of household borrowers during the crisis cycle and its aftermath. We find that 'Point-In-Time' ratings, calculated with realised data, do not indicate deterioration in average credit quality during the credit boom. In contrast, 'Through-The-Cycle' ratings, constructed by using data that is cleaned from cyclical variation, behave in accordance with the hypothesis. By all measures used, a significant improvement in the average quality of borrowers is found during the post crisis period.

Keywords: credit quality, credit rating, credit cycle, financial crisis household borrowing

JEL: D14, E32, E51, G21

[1] Author: Risto Herrala, Economist, Monetary Policy and Research Department, Bank of Finland, Risto.Herrala@bof.fi.

1 Introduction

Many blame the ongoing crisis on deterioration of credit quality during the pre-crisis credit boom. An angry public outcry has forced urgency among specialists and decision makers to build a new financial architecture to stop this from ever happening again. Discouragingly, the financial fix has been tried many times before! The 'blame game' has been played during numerous crisis episodes amid heated calls for improved banking practises. Bankers have been fired, banking supervisory regimes renewed, and banking laws rewritten. But here we are, again.[2]

Berger and Udell (2004) explain the recurrence by an 'institutional memory hypothesis', whereby credit policies are sharpened after crises when the fallout is in sight. As time passes, important lessons fade from memory and the cycle ends with a 'toxic' credit boom. Other, more formal, models have been presented about why credit quality may deteriorate during credit booms by Minsky (1986), Rajan (1994), Ruckes (2004), and Dell'Ariccia and Marquez (2006) from alternative theoretical foundations.

From the point of view of the present paper, it is noteworthy that the blame game is fuelled by the ex-post fallout, rather than established econometric evidence about ex-ante credit risk. Very little econometric evidence exists about accumulation of credit risk during credit booms. Rajan (1994) reports that banks tended to 'cover up' their policies during a credit boom in New England. Jiménez and Saurina (2006) report that loans to firms were allocated to riskier sectors and were less often collateralised during boom periods that during other periods in

[2] See 'the de Larosière report' at http://ec.europa.eu/commission_barroso/president/pdf/statement_20090225_en.pdf, the 'Turner review' at http://www.fsa.gov.uk/pubs/other/turner_review.pdf and the communiqué of the G20 meeting in London http://www.g20.org/Documents/g20_communique_020409.pdf. For the earlier discussion, see Jimenez and Saurina (2006) and Kindleberger (2005) and references therein.

Spain 1984-2000. The jury is still out about the US sub-prime crisis, as demonstrated by the ongoing work by Demyanyk and Hemert (2008) and Bhardwaj and Sengupta (2008).

The thinness of evidence on an issue of this calibre is due to severe data problems. A variety of rating methods exist for analysis of credit quality but the data that is best suited for such analysis resides in banks, neither standardised nor public. Jiménez and Saurina (2006), Demyanyk and Hemert (2008) and Bhardwaj and Sengupta (2008) use public credit registers, which cover either only a small segment of the credit market, or give only limited info about borrowers. The data issue is not only regrettable from the point of view of the econometrician. If credit quality could be overseen at the macroeconomic level, this might undermine toxic credit booms as depositors, banking supervisors and central bankers could react before a crisis escalates.

Our main contribution to the debate is to promote the idea that statistical surveys of borrowers can (and should) be used in credit market oversight. We construct a data set, based on existing household surveys by Statistics Finland, which allows the application of rating techniques to statistical samples of household borrowers, and valid aggregation of the results to the macroeconomic level. The data covers different cyclical conditions: a credit boom, the subsequent period of credit contraction in connection with a banking crisis, and two post crisis years. We are, thus, able to study the average quality of household borrowers during a major credit cycle in connection with a banking crisis, ranked among the most extreme in international studies.[3]

In line with good banking practise, we calculate two sets of ratings: Point-In-Time (PIT) ratings reflect actual and Through-The-Cycle (TTC) ratings cyclically

[3] The crisis is among the 'five most catastrophic cases' in Reinhart and Rogoff (2008) , it ranks fourth in terms of loss of GDP in Boyd, Kwak and Smith's (2005) list, and among the top 10 in terms of restructuring costs in IMF's (1998) list.

adjusted data. We find that PIT ratings do not show pronounced cyclicality: when unadjusted data is used, the average ratings of household borrowers are not much different during the boom and the bust. Counter-cyclical changes in credit quality are revealed in TTC ratings: when data is adjusted to remove cyclicality in interest rates, asset prices and incomes, then the average quality of borrowers during the boom appears poor. By all measures used, an improvement in the average quality of borrowers is found in the post crisis sample that is furthest from the crisis.

The estimation results highlight a crucial difficulty faced by borrowers, bankers, and policy makers alike when trying to navigate an economic cycle. Now, with the benefit of hindsight, we can filter out cyclicality in the relevant variables, and thus identify the credit risks that prevailed during the credit boom. However, this task was considerably more challenging during the period of observation when the Finnish economy had just undergone a profound regime shift due to financial liberalization, the relevance of historical evidence was questionable, and the future was still in the making. More generally, our results suggest that the accumulation of credit risk during credit booms may in part be related to challenges in proper management of cyclicality in interest rates and asset values.

The rest of the paper is organized as follows. The following section gives an overview of the rating samples. We then move to explain the rating method: how the 'rating components' Loss-Given-Default (LGD) and Probability-of-Default (PD) are calculated for all households in the rating samples, and how a 'scoring function' needed to calculate the PD:s is estimated. Subsequently, the average ratings in the rating samples are presented and discussed. A summary and our views on the agenda for further research conclude.

2 Rating samples

We measure average credit quality during a pronounced cycle and its aftermath by rating household borrowers that participated in household surveys by Statistics Finland. The complex surveys are cross sections of the Finnish household sector (see appendix 1). From the surveys, we select household borrowers to four 'rating samples' during the years 1988 (credit boom), 1995 (credit contraction), 1999 and 2004 (two post crisis years). A household is selected into a rating sample if we observe an increase in the household loan stock during the rating year. The number of household borrowers in each rating sample varies between 1764 (in 1988) and 940 (in 2004).

The cyclical conditions vary widely across the rating samples (chart 2). Liberalization of credit markets in 1986 was followed by a display of symptoms of economic boom such as high lending growth and booming stock and housing markets. The rating year 1988 marks the peak of lending growth. After the boom, the real economy dived into a deep recession, a collapse of the stock- and housing markets and a systemic banking crisis.[4] In rating year 1995 the real economy was recovering, but the financial contraction still continued. In 1999 both real and financial growth had returned to positive territory. In 2004, loan growth had accelerated considerably. Fears of a looming banking crisis had, again, surfaced in the public debate.

{please insert chart 1 around here}

[4] For a broader review, see Honkapohja (2009), Nyberg and Vihriälä (1994), Koskenkylä and Vesala (1994), Herrala (1999).

3 Rating methodology

In accordance with good banking practise,[5] we calculate for all households in the rating samples two rating components, 'probability-of-default' (PD) and 'loss-given-default' (LGD). PD gives the probability of default of a borrower during the post rating year and LGD gives the estimated loss to the lender as a percentage of exposure, if default occurs.

LGD is the easier of the two rating components to calculate. In this paper, LGD equals the negative net wealth of the borrower at the time of rating. No attempt is made here to estimate realisation costs since they are unlikely affect the comparison between rating years. The market value of the net wealth of households can be reasonably approximated from information given in the surveys. The calculation is explained in appendix 2. We report average 'censored LGD', where positive values of net wealth are set to zero.

Assignment of PDs is more complicated because default probabilities are not directly observable. In some studies (particularly related to the sub-prime crisis, see eg. Demyanyk and Hemert (2008) and Bhardwaj and Sengupta (2008)), PD:s are statistically estimated based on observed ex post default rates in the rating samples. In our case this approach is not feasible because we do not have information about ex post default rates in the rating samples. In this sense our situation resembles the situation that prevails when banks rate new borrowers: ratings must be based on anticipated rather than observed defaults of the borrowers.

The solution to this dilemma is to estimate the relationship between PD and observable household characteristics (a 'scoring function') statistically from a separate estimation sample. The estimation sample is unique in that it

[5] These are outlined in Basel Committee of Banking Supervision (BCBS 2005).

includes information about household distress, which we use as a proxy for default, as well as the factors that most likely affect the probability of default. The estimated scoring function is then used to calculate estimates of PD in the rating samples.

The formulation of a scoring function is based on the insight from the theory of household default that households choose between default and non-default based on the expected income streams associated with these two regimes (Foote et. al. 2008). In Finland (in contrast with the Anglo - Saxon countries) default typically leads to a loan recovery process, in which a household is allocated a tight budget to allow for minimal consumption.[6]

In accordance with this institutional feature, we start from the premise that loan default occurs when a shortage of liquid funds elevates the marginal utility of consumption. Under this premise, the probability of default is a decreasing function of the available liquid funds of the household. Functional form is determined by the marginal utility function. Formally, the following type of 'scoring function' is assumed:

$$PD_i = Pl\{-u'[\beta_1 Y_i + \beta_2 W_i]\} + \varepsilon_i \quad (1)$$

where *i* refers to household, *Pl* is probability from the logistic distribution, *u'* is marginal utility, W is wealth, *Y* is consumable income, β are parameters from the unit line, and ε is the residual.

The standard practise is to use binary indicators of household payment difficulties (household distress) to proxy default events, and estimate the scoring function by logit. This approach is also followed here. According to good banking

[6] The Finnish Consumer Agency (2008), http://www.kuluttajavirasto.fi/Page/61249094-01c2-417d-a45d-423cbaefb86b.aspx.

practise, the most suitable proxy of default varies across loan products, but the most usual default proxy is the 90 day non-service threshold. In the absence of information on non-service periods, we use available survey information of household distress to construct alternative proxies of default.

In general, the main difference in our approach compared with standard practise in banks when rating new borrowers relates to data: we use household surveys, while banks generally rely on internal credit registers. In terms of data quality, bank registers contain better data on debt characteristics and repayment patters. Our data is likely to be superior in terms of the scope of coverage of assets, liabilities and income. In household surveys by Statistics Finland, assets, liabilities and income are covered very comprehensively, while in loan negotiations income and wealth may be covered in a more limited manner (ie whether the value of the asset used as collateral is sufficient for the loan, and whether labour income is stable enough and sufficient to cover the repayment cost). For our purposes, a major advantage of using survey data is statistical sampling, which enables us to aggregate the ratings to the macroeconomic level and thus test hypotheses about credit quality in the underlying population of household borrowers.

In line with good banking practise, we calculate Point-In-Time and Through-The-Cycle PD:s and LGD:s. They differ in that PIT ratings are calculated with realised data and TTC ratings with adjusted data. The aim of the adjustment is to clean cyclical components from asset values, interest rates and income. To control for the ambiguity related to the calculation of TTC ratings, we use two alternative methods of cyclical adjustment. The calculation of variables for PIT and TTC ratings are explained in detail in appendix 2.

4 Estimation of the scoring function

As explained, before we can calculate PD:s (one of the two rating components) in the rating samples we need to estimate statistically a 'scoring function' (1). The scoring function gives the relationship between PD and household characteristics that are observable in the rating samples. The scoring function is estimated from a separate 'estimation sample', which includes a proxy variable for household default, and the main explanatory variables of default, which are consumable income and wealth.

The estimation sample includes 3685 households who participated in the wealth survey of 1998, and the income distribution surveys of 1998 and 1999. In the estimation, wealth is measured at the end of 1998, which is taken to be the point of rating in the estimation sample. Income and distress are measured in 1999. In the estimation stage we do not (and cannot) differentiate between household borrowers and non-borrowers. This is not possible as we do not have information about the loan market behaviour of the households in 1998. Inclusion of both borrowers and non-borrowers in the estimation stage is in line with the interpretation that default is simply illiquidity.

Three estimated variants of the scoring function are reported in table 1. In the simplest model, *D*, distress probability is regressed by consumable income and wealth only. The default indicator, ‘distress’, used in this model gets value unity if the household signals over-indebtedness or any problems with paying bills or loan service. The preferred model for distress, *D**, allows for age and group effects, and non-linearity in wealth and consumable income. We also report a *SD* model of 'severe distress', where the default indicator gets value unity if the household reports over-indebtedness, very often problems paying bills, or more than once problems with loan repayment.

The steps that lead from the *D* to the *D** model can be described as follows. A series of standard statistical tests demonstrated that age, group effects, and higher –powered terms of wealth and consumable income would be significant additions to the *D* model. As a second step, a model which includes all possible age- and group effects, and higher powered terms was estimated. Since not all group variables used in the estimations have been harmonized across the rating samples, we used standard statistical tests to simplify the extended model to the manageable dimensions of the *D** model.

{please insert table 1 around here}

The estimated signs of consumable income and wealth conform with expectations, in light of previous discussion. From the estimated parameters it can be inferred that in the *D* model, a 1000 € increase in consumable income reduces distress probability by about 0.8 percentage points, while an equal increase in wealth reduces default probability by 0.4 percentage points at median values. Estimated marginal effects are larger for serious distress than for distress, while the marginal effects of wealth are about equal in the two cases.

The good fit of the *D** model relies to a large extent on higher powered terms of wealth and consumable income. The inclusion of the higher powered terms affects the way in which marginal effects change with the level of consumable income and wealth. In both the *D* and the *D** model, the marginal effect from an extra euro is lowest for those who are very poor, and those who are very rich, and highest 'in the middle'. The presence of higher powered terms amplifies the marginal effects at levels close to zero, and dampens them at negative and high positive levels in the *D** model vis a vis the *D* model.

We find that age decreases distress probability in square. Various authors have reported that the square of age is a significant regressor also in econometric studies of consumption with micro data. This suggests that the observed non-linear effect may be related to changes in the shape of the marginal utility function with age. Under this interpretation, the estimated negative sign would imply that the marginal utility of consumption decreases with age in square.[7]

The most significant group effects relate to professional group 6 ('industrial, mining, construction, etc. labor'), socioeconomic group 6 ('students'), and area group 1 ('other town like communities'). It is beyond the scope of this study to make a proper assessment about why industrial, mining, etc. workers in particular show significantly elevated default risk, and students a dampened default risk. It also remains unexplained why wealth shows a larger marginal effect in other town like communities than in towns and in the countryside.

The forecast ability of the *D* and *D** models was verified five years ahead, in 2004. During this interval, the point estimate of the level of distress had dropped from 24.6% (in 1999) to 20.3% (in 2004). For the forecast accuracy test, the scoring function was applied to the 2004 wealth survey to calculate a forecast for average distress at that time. The forecast is compared with the actual level of distress as measured from the year 2004 survey of income distribution. This test was not performed on the *SD* model as the grouping of the indicator for payment difficulties had changed to some extent during the interval (one additional group).

It is observed from chart 1 that both the *D* and the *D** forecasts are well within the 95% confidence interval of realised distress at that time. The confidence interval reflects aggregation error.

[7] Campbell and Cocco (2007), Attanasio and Jappelli (2001) and references.

{please insert chart 2 around here}

5 Rating results

Table 2 gives average ratings for households that increased borrowing during the rating years 1988, 1995, 1999 and 2004. To explain the table, LGD is the average (censored) net wealth of the borrowers, while PD is the estimate produced by the *D** model. The tabulated average ratings are population estimates, obtained by using analysis weights given by the data provider. Aggregation error is about 1.5 percentage points in either direction at 95 % confidence.

{please insert table 2 around here}

Developments in the average quality of borrowers during the boom-bust cycle between 1988 and 1995 can be inferred from columns 5 and 6 of table 2. It is observed from the second row, that the average PIT PD for household borrowers during the credit boom year 1988, 38.4 %, was significantly below the tabulated 39.9 % for the credit crunch year 1995. In contrast, TTC PD:s, reported in rows 3-6, develop in all cases pro-cyclically, as do the LGD:s.

All in all, the PIT ratings do not show strong signs of counter-cyclical credit quality: the PD:s develop contrary to the counter-cyclical credit quality hypothesis, and the improvement in LGD during the bust was relatively mild. Greater cyclical changes are observed in connection with TTC ratings. The TTC PD:s and LGD were significantly elevated during the credit boom, and declined during the bust.

It is observed from columns 7 and 8 of table 2 that average credit quality has improved markedly during the post-crisis period, based on both PIT

and TTC ratings. In all scenarios, we find the lowest PD:s and LGD:s in the latest sample.

The robustness of these results with respect to model selection and definition of distress were studied by re-calculating the ratings under various alternative definitions of income and wealth, and with the alternative models *D* and *SD*. We find that the results discussed above are robust to such changes. We show in table 3 the PIT credit scores calculated with the alternative models. While the levels of the scores vary across models, it is observed that in none of the cases does the unstressed PD during the credit boom rise above the level that prevailed during the credit crunch.

{please insert table 3 around here}

6 Discussion and concluding remarks

We test the hypothesis by various that the average quality of borrowers develops counter-cyclically. Our test case is a pronounced banking crisis cycle in Finland. The study also sheds light on how average credit quality developed after the crisis.

We find significant variation in average ratings of household borrowers during the crisis cycle and its aftermath. Pronounced counter-cyclical credit quality is revealed by TTC ratings and not in PIT ratings. Furthermore, we find an improvement in the average quality of borrowers in the post crisis sample that is furthest from the crisis.

The caveats apply that the test results are specific to one sector (households) and selected dates of one historical period. More case studies are called for to increase our understanding cyclical changes in credit quality.

Apart from testing the credit quality hypothesis, we also hope to promote the idea that publicly available surveys of borrowers can give policy relevant information about credit quality for oversight purposes. Even existing surveys, which are not specifically tailored for this purpose, can be used to estimate LGD:s and credit scoring models that both fit and forecast well. In our view, also credit institutions could benefit from the use of survey data in estimation and calibration of their models. Surveys of good quality cover a broad range of relevant variables and a representative sample of potential borrowers.

References

Attanasio, O – Jappelli, T (2001) **Intertemporal Choice and the Cross-Sectional Variance of Marginal Utility.** The Review of Economics and Statistics, February 2001, 83(1):13–27.

Basel Committee of Banking Supervision (2005) **Studies in the Validation of Internal Rating Systems.** Bank of International Settlements, Working Paper No. 14.

Berger, A – Udell, G (2004) **The Institutional Memory Hypothesis and the Procyclicality of Bank Lending Behavior.** Journal of Financial Intermediation 13, 458-495.

Bhardwaj, G – Sengupta, R (2008) **Where's the Smoking Gun? A Study of Underwriting Standards for US Subprime Mortgages.** manuscript (our version October 2008).

Boyd, J – Kwak, S – Smith, B (2005) **The Real Output Losses Associated with Modern Banking Crises.** Journal of Money, Credit, and Banking, Vol. 37, No. 6, December 2005, 977–999.

Campbell, J – Cocco, J (2007) **How do House Prices Affect Consumption? Evidence from Micro Data.** Journal of Monetary Economics, Vol. 54, Issue 1, April 2007, 591–621.

Dell'Arccia, G – Marquez, R (2006) **Lending Booms and Lending Standards.** The Journal of Finance, Vol. LXI, No. 5, 2511-2546.

Demyanyk, Y – Hebert, O van (2008) **Understanding the Subprime Mortgage Crisis.** manuscript (our draft 5.12.2008) .

Foote, C – Gerardi, K – Goette, L – Willen, P. (2008) **Subprime Facts: What (We Think) We Know about the Subprime Crisis and What We Don't.** Federal Reserve Bank of Boston, Public Policy Discussion Papers No. 08-2.

Herrala, R (1999) **Banking Crises vs. Depositor Crises: The Era of the Finnish Markka 1865-1998** Scandinavian Economic History Review, Vol. 47, No. 2, 5–22.

Herrala, R – Kauko, K (2007) **Household Loan Loss Risk in Finland - Estimations and Simulations With Micro Data** Bank of Finland Discussion Papers 5/2007.

Honkapohja, S (2009) **The 1990's Financial Crises in Nordic Countries.** Bank of Finland Research Discussion Papers 5/2009.

Hyytinen, A – Määttänen, N – Johansson, E (2006) **Omistusasunnon ja asuntolainan rooli kotitaloudessa (the role of own house and mortgage in a household** in Finnish), Kansantaloudellinen aikakauskirja, Vol. 102.

International Monetary Fund (1998) **Financial Crises: Characteristics and Indicators of Vulnerability.** In IMF: World Economic Outlook.

Jiménez, G – Saurina J (2006) **Collateral, Type of Lender and Relationship Banking as Determinants of Credit Risk.** Journal of Banking and Finance 28, 2191-2212.

Kindleberger, C (2005) **Manias, Panics, and Crashes A History of Financial Crises.** Fifth edition, Wiley.

Koskenkylä, H.– Vesala, J. (2004) **Finnish Deposit Banks 1980-1993: Years of Rapid Growth and Crisis.** Bank of Finland Discussion Paper 16/94.

Minsky, H. (1986) **Stabilizing an Unstable Economy.** McGraw-Hill, our edition from 2008, foreword by H. Kaufman.

Nyberg, P.– Vihriälä, V. (2004) **The Finnish Banking Crisis and its Handling.** Bank of Finland Discussion Paper 7/94.

Rajan, R. (1994) **Why Bank Credit Policies Fluctuate: A Theory and Some Evidence.** Quarterly Journal of Economics, Vol. 109, No. 2, 399–441.

Reinhart, C. – Rogoff, K. (2008) **Is the 2007 U.S. Sub-Prime Financial Crisis so Different? An International Historical Comparison.** National Bureau of Economic Research, Working Paper 13761.

Ruckes, M. (2004) **Bank Competition and Credit Standards.** Review of Financial Studies, Vol. 17, No. 4, 1073–1102.

Appendix 1 The surveys

The surveys are complex surveys by Statistics Finland. Sampling clusters are households, and stratification is by income class. Imputations of missing variables and reweighting to account for missing observations are by the data provider.[8]

rating year	data
1988	The surveys for saving and indebtedness of 1987 and 1988 were merged to obtain a sample of 5248 households. The number of households that increased borrowing in 1988 is 1764. 'poiminta8' was used as the indicator for stratification, and 'pkor8' as the probability weight.
1995	The wealth survey of 1994 and the income distribution surveys for years 1994 and 1995 were merged to obtain a sample of 4951 households. The number of households that increased borrowing in 1995 was 1082. 'poimluok' was used as the indicator for stratification except that stratification classes 11 and 13 were combined for lack of observations in the latter class. 'pkor' from the 1995 survey was used as the probability weight.
1999	This is the estimation sample. This is also the earliest sample in which distress indicators were surveyed. The wealth survey of 1998 and the income distribution surveys for years 1998 and 1999 were merged to obtain a sample of 3685 households. The number of households that increased borrowing in 1999 was 1072. 'poimluok' was used as the indicator for stratification, except that the stratification classes 11 and 13 were merged for lack of sufficient observations on the latter class. 'pkor' from the 1999 survey was used as the probability weight.
2004	The wealth survey of 2004 of 3455 households was used to calculate the PD:s and LGD:s. The number of households that increased borrowing in 2004 is 940.'poimluok' is used as the indicator for stratification, except that the stratification classes 7 and 8 have been merged for lack of sufficient observations on the latter class. 'koro04' is used as the probability weight. In contrast with the earlier surveys, the wealth survey of 2004 and the surveys for income distribution for 2003 and 2004 cannot be merged at household level. The survey of income distribution of 2004, consisting of 11229 households, was used to calculate the level of distress in 2004 for the forecast accuracy test. 'poimluok' is used as the indicator for stratification, except that the stratification classes 11 and 13 have been for lack of sufficient observations on the latter class. 'pkor' is used as the probability weight.

[8] Hyytinen, Määttänen and Johansson (2006) and Herrala and Kauko (2007) have previously used some of the surveys to study household finance.

Appendix 2 Varibles

The basic economic unit in the analysis is the household.

2.1 Estimation

In estimation, we use wealth for year 1998 (the evaluation year), and consumable income and distress for the next year 1999.

- A household is classified **distressed** if it signals over-indebtedness, problems with paying bills or problems with loan service in the 1999 survey of income distribution. A household is classified as **severely distressed** if it signals over-indebtedness, very often problems with paying bills or more than once problems with loan service.
- **Consumable income** is monetary income in 1999 diluted by loan service costs. These include interest payments and instalments from housing loans (from 1999 survey), and estimated interest payments from other loans, obtained by multiplying the stock of other loans (the 1998 survey of income distribution) by the average interest rate of other loans (Bank of Finland loan statistics).
- **Wealth** is calculated by diluting the value of household debt from household gross wealth in 1998.

2. 2 PIT rating

Households are selected to the rating samples by a positive change in the loan stock during the rating year. Unstressed (Point-In-Time) ratings are calculated with consumable income during the rating year and net wealth at the end of the rating year.

- **Consumable income** is monetary income during the rating year diluted by estimated interest payments of loans and instalments from housing loans. Interest payments from loans are estimated by multiplying the loan stock at the end of the rating year by the average interest rate of loans (Bank of Finland interest rate statistics). Instalments from housing loans equals the loan stock at the end of the rating year times the estimated average maturity of housing loans, obtained from various sources.
- **Wealth** (gross wealth diluted by debt) in 1988, 1995 and 1999 is estimated by adding to the market value of wealth in 1987, 1994 and 1998 (respectively) given in surveys a proportionality factor times the change in taxable wealth between 1987-1988, 1994–1995 and 1998-1999. The proportionality factor is estimated from the data by a linear regression of wealth on taxable wealth. For 2004, we use wealth given in the survey for that year.
- **Group variables** are in the main included in the surveys, or can be constructed exactly. In 2004, the professional group 6 is approximated by using information on socioeconomic groups.

In all simulations, consumable income and wealth are divided by consumption units, given in the respective surveys. They are deflated by the CPI index (base year 1999).

Table 2.1 Population means of exogenous variables

group	exogenous variable	unit	1987	1994	1998	2003
Households that have increased borrowing	Consumable income	real, per cu, 1000€	10.2	10.5	12.4	14.9
	Wealth	real, per cu, 1000€	22.8	17.1	26.0	35.8
	socioeconomic group 6: 'students'	%	5.9	5.2	6.6	6.5
	area group 1: 'other town-like communities'	%	44.9	37.3	44.6	47.8
	prof. group 6:'workers in agriculture, forestry etc.'	%	5.4	3.6	3.5	4.0
	age	years	40.0	43.3	42.7	45.0

results.xslx

2.3 TTC rating

Stressed (Through-The-Cycle) ratings are calculated with adjusted values of wealth and consumable income. The data adjustment aims at projecting wealth and consumable income to the cyclical conditions that prevailed during a selected benchmark year. The analysis thus answers the question of what the ratings would have looked like the cyclical conditions that prevailed during that benchmark year.

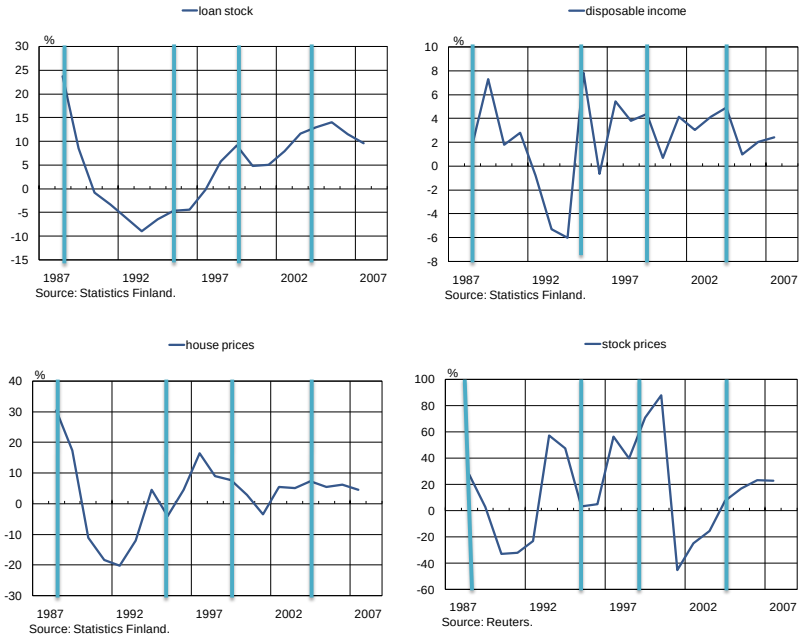
Denote by $x_{i,t}$ the observed (unstressed) value in the data for household *i* at time t, and by $x_{i,t}^{adj}$ the adjusted (stressed) value. Denote by X_t the aggregate value of $x_{i,t}$, and by $\bar{X}_t$ the trend of X_t. Two alternative trends are used in the estimation, a Hodrick-Prescott (HP) -trend and a linear trend.

- **Consumable income:** the adjusted (stressed) value with baseline year T is calculated by the formula:

$$x_{i,t}^{adj} = x_{i,t}\left(\frac{X_T}{\bar{X}_T} * \frac{\bar{X}_t}{X_t}\right)$$

- **Wealth** adjustment is made for housing wealth and stock market wealth. Other components of wealth are not changed. The formula is as above but X_t refers to the aggregate price index of the asset and $\bar{X}_t$ its trend.
- **Interest rates of loans** are adjusted by diluting from the observed loan rates the difference in the 3 month Helibor (money market rate) between the observation date and the benchmark year.

Chart 1 Annual change in the household loan stock, disposable income, house prices and stock prices, 1987–2007 (estimation years are marked with a blue vertical line).



Note: Nominal changes have been deflated by changes in the CPI index. The volatility in the stock index after 1995 reflects to a significant degree changes in the value of the Nokia corporation. Data sources: loan stock, disposable income and house prices from Statistics Finland, stock prices from Reuters.

Table 1. **Estimated parameters, discrimination and goodness-of-fit in *D* and *D****

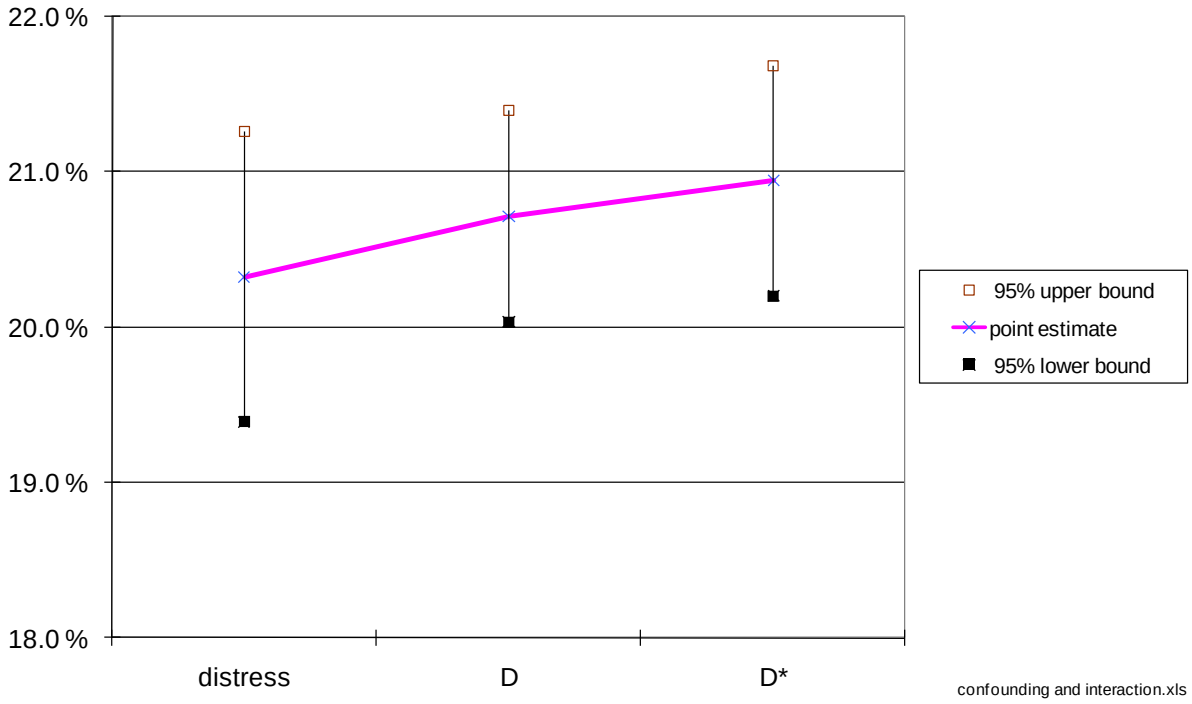
-models for distress and a *SD* model for severe distress.

(**/ * refer to 5% / 1% significance levels)**

Variable	*D*	*D**	*SD*
Y	-.063***	-.12***	-.1***
W	-.027***	-.048***	-.028***
Y^2		.0006***	
W^2		.00043***	
W^3		-10-6***	
age^2		-.00027***	
professional group 6		.61***	
W*(area group 1)		-.0085**	
socioeconomic group 6		-1.3***	
_cons	0.29	1.5***	-.81***
Area under ROC	77 %	81 %	79 %
Goodness-of-fit	0 %	84 %	8 %
confounding and interaction.xls			

Note: Estimation by the command 'svy: logit' in Stata. The dependent variables are distress in *D* and *D** models, and severe distress in the *SD* model. Y is consumable income per consumption unit in 1999, W is wealth per consumption unit in 1998, professional group 6 refers to the professional grouping of Statistics Finland (paaammy=6: industrial, mining, construction etc. workers); area group 1 refers to statistical grouping of communities (taajama=1: other town –like communities); socioeconomic group 6 refers to grouping of socioeconomic status (paasoss=6: students). W*(area group 1) refers to the group indicator times wealth. ^2 and ^3 refer to second and third powers respectively. Significance levels calculated by t –tests. Discrimination gives the area under receiver operators characteristics (ROC) curve, and goodness-of-fit the F –adjusted test by Archer and Lemeshow. Both tests are included in Stata. The number of observations is 3685.

Chart 2. **Forecast ability of *D* and *D** models five years ahead.**



Note: 'distress' refers to the average level of distress in the survey of household income distribution survey of 2004. The *D* and *D** model forecasts are the average distress probabilities of households in the wealth survey of 2004, calculated by the respective models. The point estimates and confidence intervals have been obtained by the 'svy:mean' command in Stata.

Table 2. Average ratings of households that increased borrowing.

measure	rating type	stressing method	base year	1988	1995	1999	2004
PD	PIT			38.4 %	39.9 %	36.7 %	28.6 %
	TTC	HP -trend	1999	37.5 %	35.8 %	36.7 %	28.8 %
	TTC	HP -trend	1994	43.4 %	41.2 %	42.6 %	35.1 %
	TTC	linear trend	1999	41.4 %	35.8 %	37.5 %	31.3 %
	TTC	linear trend	1994	46.4 %	40.6 %	42.5 %	36.4 %
LGD	PIT			41.5 %	39.6 %	39.9 %	27.0 %
	TTC	HP -trend	1999	42.6 %	36.4 %	39.9 %	27.0 %

results.xlsx

Note: PD is probability of default and LGD loss given default (in absolute value). PIT refers to Point-In-Time and TTC to Through-The-Cycle ratings. Stressing method identifies whether the method of data adjustment in TTC ratings was based on HP -trends or linear trends. Base year identifies the baseline year used in data adjustment in TTC ratings.

Table 3 Unstressed PD:s from alternative econometric models.

model	1988	1995	1999	2004
D	33 %	34 %	32 %	26 %
*D**	38 %	40 %	37 %	29 %
SD	12 %	12 %	11 %	8 %

results.xslx

Note: The dependent variables are distress in *D* and *D** models, and severe distress in the *SD* model.