Essays on Modeling and Analysis of Mortgage Loan Pools and the Delphi Method in Forecasting of Financial Variables

Peter Palmroos
Essays on Modeling and Analysis of Mortgage Loan Pools and the Delphi Method in Forecasting of Financial Variables

Peter Palmroos
Main dissertation advisor
Professor Timo Kuosmanen, Aalto University, Finland

Co-dissertation advisor
Doctor Jouko Vilmunen, Bank of Finland, Finland

Opponent
Professor Lasse Koskinen, University of Tampere, Finland
Abstract

This thesis studies censored and truncated distributions in the modeling of correlation between Probability-of-Default (PD) and Loss Given Default (LGD) in mortgage loan pools, and also the use of the Delphi method in economical and financial forecasting.

The first article derives the closed form solution for the moments and correlation of the bivariate, once or twice truncated and/or censored log-normal distribution. As a generalization the formula can also be used for distributions where one or both tails have been partly censored and partly truncated. The central scientific contribution of this paper is the closed form equations for correlation and moments.

The model presented in the second article allows the estimation of the heterogeneous mortgage loan pool structure using only the publicly available data. As a result the model gives three matrices of homogenised loan cohorts. Two of these matrices include information on the number of loans and the total remaining principals for all cohorts. The third matrix includes the Current Loan-to-Values (CLTV) for the same cohorts. These CLTVs can be converted to expected LGDs. The presented model is dynamic and can be used for forecasting the future pool structures. The scientific contribution of the paper is the method that can be used to generate reliable estimates of pool structures, when only publicly available data can be used.

The third article studies the effects of a mortgage loan pool structure on the observed correlation between PD and LGD. The paper applies the same methods and models presented in the first two essays. Both homogeneous and heterogeneous pool structures are studied. Furthermore, the paper analyses the effects of both including and excluding zero-loss-defaults has on the observed sample correlation. The scientific contribution of the paper is to prove that the observed sample correlations are sometimes so biased that they shouldn’t be used as such.

In the fourth article forecasting power of the two expert opinion models, Delphi and Face-to-Face meetings, has been tested. We also present two post-survey methods to correct for possible forecast errors. The first model adjusts the perseverance bias caused by overly strong self-confidence of expert panellists, and experts willingness to prefer the trustworthiness of their own estimates over the estimates of other panellists. The second method tests post-survey adjustment of forecasts using conditionalised forecasts. In theory the latter method can be used to increase the accuracy of forecasts when the forecasted variable is dependent on the explanatory variable whose development is hard to forecast. The key scientific contributions of the paper are the support found for the existence of perseverance bias, and the indication that the mechanical post survey corrections of forecasts are possible.

Keywords  Truncated and censored distributions, Closed-form solution, Mortgage pool structure, Correlation between PD and LGD, Delphi method

ISSN-L 1799-4934 ISSN (printed) 1799-4934 ISSN (pdf) 1799-4942
Location of publisher Helsinki Location of printing Helsinki Year 2016
Pages 160

Tekijä: Peter Palmroos

Väitöskirjan nimi: Esseitä asuntoluottosalkkujen mallintamisesta ja analysoinnista sekä Delfoi-menetelmästä taloudellisia muuttujia ennustettaessa

Julkaisija: Kauppakorkeakoulu

Yksikkö: Tieto- ja palvelutalouden laitos

Sarja: Aalto University publication series DOCTORAL DISSERTATIONS 229/2016

Tutkimusala: Talousmatematiikka ja -tilastotiede

Kieli: Englanti

Monografia Artikkeliväittöskirja Esseeväittöskirja

Tiivistelmä: Tämä väitöskirja tutkii katkaistujen ja sensuroitujen jakaumien käyttöä asuntoluottojen tapiotodennäköisyyden ja tappio-osuuden välisen korrelaation mallintamisesta, sekä Delfoi-menetelmän käyttöä talouteen ja rahoituksen liittyvissä ennusteissa.

Ensimmäisessä artikkelissa johdetaan suljetut ratkaisut kaksiulotteisen, yhdelta tai molemmilta puolilta katkaistun ja/tai sensuroidun log-normaalim jakauman momenteille ja korrelaatiolle. Yleistyksenä on jakauma, jonka toinen tai molemmat hännät on osittain sensuroitu ja osittain katkaistu. Artikkelin kontribuutio on momenttien ja korrelaation suljettu ratkaisu.

Toisessa artikkelissa mallinnetaan heterogeenisten asuntoluottosalkkujen rakennetta julkisesti saatavilla olevan tiedon perusteella. Estimoinnin tuloksena saadaan kolme homogeenisistä lainakohorteista koostuvaa matriisia. Näistä kaksi sisältää lainojen kapppalemääriä ja jäljellä olevat pääomat kohorteittain. Kolmas matriisi osoittaa kunkin kohortin tarkasteluhetken luototuusasteen, joka on tarvittaessa muutettavissa myös odotetuksia tappio-osuudeksi. Esi-
tetty malli on dynaaminen ja sitä voidaan käyttää myös lainasalkun tulevan rakenteen ennustamiseen. Artikkelin tieteellinen kontribuutio perustuu heterogeenisen salkun rakenteen yksinkertaiseen ja tarkkaan mallintamiseen tilanteissa, joissa käytettävissä on vain rajoitetustainojen koskevaa tietoa.

Kolmannessa artikkelissa hyödynnetään edellä esitettyjä menetelmiä vertailtaessa luotto-
salkun rakентeen vaikutusta havaittavun tappiotodennäköisyyden ja tappio-osuuden väliseen korrelaatioon. Vertailussa käytetään sekä homogeenisia että heterogeenisia lainasalkkuja. Salkun rakenteen vaikutusta lisäksi artikkelissa vertailaa tappiotä aiheuttamattomien maksukyvyttömystapauksien havaitsemisen vaikutusta otoskorrelaatioon. Artikkelin keskei-
nen kontribuutio on sen osoittaminen, että otoskorrelaatit ovat joissain tapauksissa niin harkisia, ettei havaittua korrelaatioita voida sellaisenaan käyttää esimerkiksi simuloinneissa.

Neljännestä artikkelissa vertailaa Delfoi-menetelmällä ja Face-to-Face-kokouksen avulla kerättyjen asiantuntijaestimmaattien ennusteekykyää, sekä esitellään kaksi eri menetelmää mah-
dollisten ennusteividen korjaamiseksi. Ensimmäisellä menetelmällä pyritään korjaamaan ns. vahvistusvoinonaa, joka aiheutuu asiantuntijoiden liiallisesta itseluottamuskesta ja halusta pitää omia ennusteita muiden ennusteita luotettavampina. Toisen menetelmän avulla testa-
taan ehdollisten ennusteiden hyödyntämistä ennusteividen korjaamisessa. Näistä jälkimäisen avulla pyritään parantamaan ennusteiden tarkkuutta tilanteissa, joissa ennustettava
tuutaja on riippuvainen toisen, vaikeasti ennustettavan muuttujan arvosta. Artikkelin kes-
keiset tieteelliset kontribuutiot ovat tulosten antama tuki sekä vahvistusvoinonnan olemassa-
ololle että ennusteiden mukaan siselle tarkentamiselle jälkipätevä.

Avainsanat: Katkaistut ja sensuroitud jakaumat, suljettu ratkaisu, Asuntoluottosalkun rakennus, PD ja LGD välinen korrelaatio, Delfoi-menetelmä


ISSN-L: 1799-4934 ISSN (painettu): 1799-4934 ISSN (pdf): 1799-4942


Abstract

This thesis studies censored and truncated distributions in the modeling of correlation between Probability-of-Default (PD) and Loss Given Default (LGD) in mortgage loan pools, and also the use of the Delphi method in economical and financial forecasting.

Modeling and analysis of mortgage loan pools are based on simulated loan pools. When a heterogeneous pool structure is used, the parameters used have been estimated using end of 2010 cross-sectional data reported by Finnish banks.

The first article, which has been written together with Dr. Jouko Vilmunen, derives the closed form solution for the moments and correlation of the bivariate, once or twice truncated and/or censored log-normal distribution. As a generalization the formula can also be used for distributions where one or both tails have been partly censored and partly truncated. The central scientific contribution of this paper is the closed form equations for correlation and moments, which haven’t been presented, as far as we know, as such in the literature. In applications the formula can be used to solve the parameters of restricted distribution, which makes the simulations of restricted samples either unnecessary or at least easier.

The model presented in the second article allows the estimation of the heterogeneous mortgage loan pool structure using only the publicly available data. As a result the model gives three matrices of homogenised loan cohorts. Two of these matrices include information on the number of loans and the total remaining principals for all cohorts. The third matrix includes the Current Loan-to-Values (CLTV) for the same cohorts. These CLTVs can be converted to expected LGDs. The presented model is dynamic and can be used for forecasting the future pool structures, e.g. as part of impact studies of regulatory changes that may affect the supply of mortgage loans. The scientific contribution of the paper is the method that can be used to generate reliable estimates of mortgage loan pool structures, when only publicly available data can be used. The estimated structure can also be used with already existing credit risk models to make the forecasts and esti-
mates more reliable and accurate when pool structures are heterogeneous. Similar methods for modeling the heterogeneous mortgage loan pool structure cannot be found in the literature.

The third article studies the effects of a mortgage loan pool structure on the observed correlation between PD and LGD. The paper applies the same methods and models presented in the first two essays. Both homogeneous and heterogeneous pool structures are studied. Furthermore, the paper analyses the effects of both including and excluding zero-loss-defaults, i.e. default events where LGD is zero, on the observed sample correlation. The scientific contribution of the paper is to prove that the observed sample correlations are sometimes so biased and sensitive to the estimated parameters that they shouldn't be used as such. Instead the paper recommends modeling of the unrestricted distribution behind the loss events, and using the resulting moments and correlations in all simulations and risk calculations.

In the fourth article, that has been written together with Dr. Karlo Kauko, forecasting power of the two expert opinion models, Delphi and Face-to-Face meetings, has been tested. We also present two post-survey methods to correct for possible forecast errors. The first model adjusts the perseverance bias caused by overly strong self-confidence of expert panelists, and experts willingness to prefer the trustworthiness of their own estimates over the estimates of other panelists. The second method tests post-survey adjustment of forecasts using condition-ised forecasts. In theory the latter method can be used to increase the accuracy of forecasts when the forecasted variable is dependent on the explanatory variable whose development is hard to forecast. The key scientific contributions of the paper are the support found for the existence of perseverance bias, and the indication that the mechanical post survey corrections of forecasts are possible. The later result may be used to save both money and time in situations where the input of expert panelists and new forecasting rounds were needed previously.
Tiivistelmä

Tämä väitöskirja tutkii katkaistujen ja sensuroitujen jakaumien käyttöä asuntoluottojen tappiotodennäköisyyden (Probability-of-Default, PD) ja tappio-osuuden (Loss Given Default, LGD) välisen korrelaation mallintamisessa, sekä Delfoi-menetelmän käyttöä talouteen ja rahoitukseen liittyvissä ennusteissa.

Asuntoluottosalkkujen mallintamisessa ja analysoinnissa käytetään simuloitua lainasalkkua, joka on mallinnettu vastaamaan ominaisuusilla mahdollisimman hyvin suomalaisen pankkien yhteenlasketun asuntoluottosalkun rakennetta. Parametrit on estiomoitu vastaamaan vuoden 2010 lopun tilannetta, jolloin myös tarvittavat poikkileikkausaineistot on kerätty.

Ensimmäisessä artikkelissa, joka on kirjoitettu yhteistyössä Jouko Vilmusen kanssa, johdetaan suljetut ratkaisut kaksiulotteisen ja/tai molemmilta katkaistun ja/tai sensuroidun log-normaalisen jakauman momenteille ja korrelaatiolle. Yleistyneenä on jakauma, jonka toinen tai molemmat hännät on osittain sensuroitu ja osittain katkaistu. Artikkelin tieteellinen kontribuutio perustuu momenttien ja korrelaation suljettuun ratkaisuun, jota ei tiettävästi aiemmin ole kirjallisessa esitetty. Käytännön sovelluksissa kaavan avulla voidaan helpottaa, tai joissakin tapauksissa korvata, katkaistujen ja sensuroitujen jakaumien simulointea.

Tiivistelmä

kennetta voidaan käyttää esimerkiksi jo olemassa olevien mallien ja ennusteiden tarkentamiseen. Vastaavia malleja, jotka approksimoivat sekä luototusastetta että salkun lainojen määrää ja kokoa, ei ole aiemmin esitetty kirjallisuudessa.


Neljännässä artikkelissa, joka on kirjoitettu yhdessä Karlo Kaukon kanssa, vertailaan Delfoi-menetelmällä ja Face-to-Face-kokouksen avulla kerättäjien asiantuntijaestimaattien ennustetyyppästä, sekä esitellään kaksi eri menetelmää mahdollisten ennustevirheiden korjaamiseksi. Ensimmäisellä menetelmällä pyritään korjaamaan ns. vahvistusvinoumaa (perseverance bias), joka aiheutuu asiantuntijoiden liiallisen itseluotamuksesta ja halusta pitää omia ennusteita muiden ennusteita luotettavampina. Toisen menetelmän avulla testataan ehdollisten ennusteiden hyödyntämistä ennustevirheiden korjaamisessa. Näistä jälkimmäisen avulla pyritään parantamaan ennusteiden tarkkuutta tilanteissa, joissa ennustettava muuttuja on riippuvainen toisen, vaikeasti ennustettavan muuttujan arvosta. Artikkelin keskeiset tieteelliset kontribuutiot ovat tulosten antama tuki sekä vahvistusviinouman olemassaolon. Toista ennusteiden korjausmenetelmää voidaan hyödyntää tilanteissa, jotka aiemmin ovat vaatineet kalliita ja aikaa vieviä uusia ennustekierroksia ja panelistien läsnäoloa.
Acknowledgements

I would like to thank Professor Timo Kuosmanen and Professor Emeritus Pekka Korhonen for their help and forbearance; my mentor, instructor, co-writer, and former superior Jouko Vilmunen; my co-writer Karlo Kauko; and also Esa Jokivuolle and Professor Matti Virén for their help and advice.

A big thank you to my opponent and pre-examiner, Professor Lasse Koskinen, and pre-examiner, Professor Mikael (Mika) Linden, for the very good and constructive comments I have received. I also want to thank all the referees who have reviewed my papers submitted for publication. The observations have made the papers better and more readable.

Without my research colleague and friend, Doctor Vesa Ronkainen encouraging and urging me on, the whole thesis may have remained a never-ending project. Those numerous intellectual discussions on various topics on the embankment of the river Main in Frankfurt made the heavy burden of the task easier to carry. Thank you for your support.

I am grateful for the insight and ideas received from my former colleagues Matti Koivu, Tarja Sirén, and Laura Koskela. Special thanks go to Meri Rimmanen who shared the agony of doctoral studies carried out alongside working fulltime in the Financial Supervision Authority of Finland.

Without the working hours my former superiors Vesa Hänninen, Katja Taipalus, Jaana Rantama, and Jukka Vesala allowed me to use for my academic studies and research, finishing this project might have taken much longer. The grammar checking done by Glen Harma, Louise Park and Enda Kelly made the text as readable as it is now, so not only I, but also all the possible readers of these papers are grateful to you.

Warm thanks also to my good friends Petteri Louhikoski, Antti Virtanen, and Timo Kanninen for their support and making my scarce free time feel like short but relaxing vacations.

Finally, I would like to thank my parents, Jan-Henrik and Ritva Palmroos, for
their support during my struggle to achieve this ambitious goal. Words fail me in expressing my thanks to my darling wife Johanna Palmroos who has encouraged me, comforted me in the numerous moments of despair, and also spent time studying with me. Thank you also to my children Rosa and Arthur – your love makes all the hard work worthwhile.

The author reports no conflicts of interest and is alone responsible for the content and writing of the papers. In case of shared papers, the best effort has been done to separate author’s own input, whenever it has been possible. During the writing process the author has worked in the Bank of Finland, Financial Supervision Authority of Finland, and Single Supervisory Mechanism, which is a part of the European Central Bank. All the views presented in these papers are those of the authors and do not necessarily represent the views of the previously mentioned organizations.

Frankfurt am Main, 30 April, 2016
Peter Palmroos
# Table of Contents

Abstract .................................................................. 1  
Tiivistelmä ............................................................. 3  
Acknowledgements .................................................. 5  
List of Publications .................................................. 10  
Author’s Contribution ............................................... 11  

**Part I**

Introduction and synthesis ........................................... 15  
1. Introduction ........................................................ 15  
2. Background ........................................................ 16  
   2.1 Mortgage loan markets and sub-prime loans as sources of crises .... 16  
   2.2 Finnish mortgage loan markets .................................... 18  
   2.3 Credit risks of mortgage loans ................................... 19  
   2.4 Truncated and censored distributions ......................... 22  
   2.5 Unrestricted and restricted sample correlations ............... 23  
   2.6 Mortgage loan pool structure ................................... 24  
   2.7 Structured expert opinion based forecasting ................... 24  
3. Summary of essays ............................................... 26  
   3.1 Closed Form Solution of Correlation of Once or Twice Truncated or Censored Bivariate Log-Normal Distribution ............... 26  
   3.2 Modeling the Current Loan-to-Value Structure of Mortgage Pools without Loan Specific Data ........................................ 27  
   3.3 Correlation between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools ............. 28  
   3.4 The Delphi method in forecasting financial markets - An experimental study ...................................................... 30  
4. Conclusion and discussion ....................................... 32  
References ............................................................. 34
Part II

Essay 1: Closed Form Solution of Correlation of Once or Twice Truncated or Censored Bivariate Log-Normal Distribution .................................39
Essay 2: Modeling the Current Loan-to-Value Structure of Mortgage Pools without Loan Specific Data .............................................59
Essay 3: Correlation between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools ..............83

Part III

Essay 4: The Delphi method in forecasting financial markets -
An experimental study ..........................................................117
This doctoral dissertation consists of an introduction, including a summary of articles/essays, and of the following publications:


Palmroos, Peter *Correlation between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools*. Submitted for publication.


---

1 The published paper is a previous version of the paper included in this dissertation. The name of the latest version is ‘Closed Form Solution of Correlation of Once or Twice Truncated or Censored Bivariate Log-Normal Distribution’.
Author’s Contribution

**Essay 1:** *Closed Form Solution of Correlation of Once or Twice Truncated or Censored Bivariate Log-Normal Distribution*

Palmroos Peter: The original research idea was developed by him as a part of study presented in the Essay 3, and is mainly responsible for writing of the paper. Derivation of the closed form solutions for censored and mixture of truncated and censored distributions. He is responsible for all the calculations and designing of the analysis, examples, and presented graphs.

**Essay 2:** *Modeling the Current Loan-to-Value Structure of Mortgage Pools without Loan Specific Data*

Sole author.

**Essay 3:** *Correlation between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools*

Sole author.

**Essay 4:** *The Delphi method in forecasting financial markets - An experimental study*

Palmroos Peter: Albeit the idea of doing this research came from Karlo Kauko, Peter Palmroos was, independently and without knowing of these plans, studying expert opinion methods and developing Delphi based method that could have been used for the forecasting of risk, P&L, and balance sheet developments of the Finnish banks.

He is responsible for development of the conditional forecasting method that can be used in Post Hoc correction of dependent forecasts. He took care of the supervision and collection of answers of the Delphi group during the experience.
Part I
Introduction and synthesis

1 Introduction

Discussion on the relevance of risks of mortgage loans to banks’ profits and capital adequacy came to prominence in Finland after the recession at the beginning of 1990s. This same discussion started again, now world-wide, after the US sub-prime crisis and its impacts on the European finance sector.

Most of the methods used for mortgage loan risk modeling have been based on models developed for the credit risk management of bonds and loans in wholesale markets, although the products, contracts used, as well as the counterparts in the mortgage loan markets differ a lot from the ones in the wholesale markets. Thus modeling of retail loans with scaled-down wholesale market methods possibly miss some properties of mortgage loans and may result to incorrect estimation of the risks. This may induce mispricing of products in relation to the risks, and thus e.g. fuel the unhealthy borrowing booms and housing bubbles in the markets.

Development of more reliable mortgage loan credit risk models requires identification of the different properties of wholesale and mortgage loans and markets. It is also necessary to understand what kind of limitations and possible biases there may be in the available loan and market data. This is important when observation-based estimates are used for the calibration of models and risk simulations.

Articles and essays included in this thesis are divided into two groups. First group (labeled as Part II in the thesis) studies the effects of the mortgage loan pool structures on the observed correlations between Probability-of-Default (PD) and Loss Given Default (LGD). Furthermore the effects of both including and excluding zero-loss-defaults in the sample, i.e. default events where LGD is zero, have been analysed. Second part (labeled as Part III) presents how the Delphi method can be used in financial and economical forecasting, and two post-survey methods of how the possible biases of these forecasts can be corrected.
The observed sample correlation has been studied using simulated homogeneous, and more realistic heterogeneous loan pools. Because the information on pool structures are not available outside the credit institutions, a method to approximate them using publicly available data has been presented.

A closed form solution to calculate moments and correlations of restricted bivariate log-normal samples has also been derived and presented. The closed form solution makes the analysis of restricted samples, where the value of the explanatory variable affects the possibility of observation of the default event, more accurate and easier to perform.

This thesis is organised as follows. Part one includes four sections and introduces the research topics and gives some background information on each research topic. Section 2 introduces the most essential concepts used in the presented research papers. The target of this chapter is to give some background information to help readers to understand the presented research problems, methods used, and findings. Section 3 presents short summaries of the papers including the key research questions and the results. Also the contributions these findings add to the scientific debate are presented, as well as some practical applications. Section 4 concludes and presents some open questions as possible topics of new research. Part two includes three essays on mortgage loan pool modeling and analysis on correlation. Part three presents an essay on how the Delphi method can be used in financial forecasting and also on post-survey correction methods for forecasts.

2 Background

This section presents some background information on the key topics and concepts used in the thesis. More information and reference material can be found from the introduction and method sections of the articles.

2.1 Mortgage loan markets and sub-prime loans as sources of crises

Even before the sub-prime crisis, that hit the markets at the end of 2007, growing house prices and increasing mortgage loan pools of banks induced speculation on housing price bubbles and possible direct and indirect impacts on capital adequacy and profitability of banks.

The forecasting of crises has always been hard, if not impossible. Kindleberger and Aliber (2005, pp. 280) comment on the idea of dividing shocks in to predictable and unpredictable as: ”Several of these shocks were true surprises but several were 'predictable'; a 'predictable shock' seems like an oxymoron since by
definition a shock is not predictable.’’ Whether all crisis can be seen as shocks, as well as the predictability of crises, are topics left out from this thesis.

Again ‘everything is simple as soon you know the answer’, but on March 2007, a few months before the crisis hit the markets at the end of 2007, Geitner (2014, pp. 112) elicit that the Fed was unable to see an increase in risks of mortgage loan markets. In his book Geitner continues that “the prevailing assumption that housing prices would not slump nationwide, though widely shared and backed by seven decades of history, also turn out to be wrong.” (pp. 113).

Even when the shock was unpredictable, the progress of the crisis was not: All the same phases explained by Kindleberger and Aliber can be found also from the sub-prime crisis.

It is widely accepted that the most immediate sources of the crisis were the fast growing and inadequately regulated sub-prime markets, albeit there are multiple views on the primary source triggering the crisis. In many of those theories moral hazard has been mentioned, if not as the main trigger, then at least as one of the key factors (e.g. Hellwig, 2009).

Another catalyst of the crisis, which is closely related to the topic of this thesis, is the pricing error of mortgage loans caused by the underestimation of credit risks. The effects of incorrect pricing have been explained already tens of years before current crisis. Some of these effects on market equilibrium price have been analysed e.g. by Stiglitz and Weiss (1981).

The crisis caused a debt deflation in US markets, which was like the one Finland experienced during the first years of 1990’s. For example as stated in Federal Reserve Bank of San Francisco (FRBSF) Economic Letter (Krainer and LeRoy, 2010) in some states, e.g. in California, Florida, and Nevada, more than 20% of mortgages had remaining principals higher than the estimated values of collaterals. At the same time delinquency rate of single-family mortgages increased from less than 2% at 2007\(^2\) to approx. 11.5% at 2010.

The sub-prime crisis started to spread over the world, including Europe. Naturally the contagion channel was not mortgage loans, but instruments like mortgage backed securities (MBS), collateralised debt obligations (CDO), and in some cases also credit default swaps (CDS). These instruments worked as an efficient contagion channel between banks and thus also from country-to-country.

After the crisis started to spread, the Bank for International Settlement (BIS), European Systemic Risk Board (ESRB), and other bodies reacted and started to develop recommendations for regulation to prevent these risks and to reduce the

\(^2\)Rate has fluctuated between 1.5 and 2.5 the whole period between 1994 and 2007. Source: Board of Governors of the US Federal Reserve System (research.stlouisfed.org).
effects of this kind of crisis. For example new Capital Requirements Directives (CRD II/III/IV) have been set to increase the capital adequacy, and thus also to increase the shock resistance of banks.

In addition to the new capital requirements, the ESRB’s recommendations on macro-prudential tools, and its recommendation to regulate granting of mortgage loans, also affected mortgage loan markets. This was one of the motivators inspiring the writing of this thesis.

2.2 Finnish mortgage loan markets

Finland experienced its debt deflation crisis over a quarter of century ago. The reporting of banks was then partly insufficient and thus giving any reliable estimates on mortgage loan losses was not possible. Based on the information available on the total loan losses from the household sector, the new write-offs during 1991-1992 of loans granted for households, was approx. EUR 192 million, i.e. a little bit over 5% from all write-offs (Pensala and Solttila, 1993). Kajanoja (2012) presents that this was still under 1% from all the loans granted to households.

This surprisingly low loss level, compared e.g. to effects observed in US markets during the sub-prime crisis, was not a coincidence. There are multiple reasons why the creditors are more robust against turbulences in Finnish mortgage loan markets, and why Finnish markets react differently to shocks compared to those markets that have recently generated large credit losses.

In some markets, e.g. at least in some parts of the US, as soon as the house value sinks below the amount of the remaining loan, it is a potential trigger for default. Defaulting can be seen as a voluntary and rational behaviour in case the personal bankruptcy legislation allowing the borrower to get rid of his liabilities by assigning the collateral to the lender. This kind of practice increases the amount of defaults when the asset values start to decline, and thus feeds the negative spiral.

Finnish legislation doesn’t support asset value based defaults. In the case where the borrower defaults and the collateral needs to be realized, the residual amount of loan that the collateral doesn’t cover stays as a borrower’s liability. Thus decreasing asset values do not encourage borrowers to voluntarily default. Instead, in Finnish markets retail borrowers default because of insolvency. Thus the number of defaults is more dependent on other economic conditions, especially on the number of new unemployed, than directly from the asset values.

There are no sub-prime mortgage loan markets in Finland. That kind of market is typical only in the US, and hasn’t been adopted in any European country, albeit the normal mortgage loan markets have mimicked some of the practices, e.g.
amortization-free first years, from sub-prime markets.

Old loan agreements do not allow margin calls for retail mortgage loans. This means that customers can be sure that the interest part of prepayments won’t increase due to banks’ unilateral decisions. Among creditors there is a tendency to change the new contracts to allow adjustment of margins in some special cases, e.g. when the financing costs of creditors increase significantly.

Loans that use floating rates as a reference rate may in theory, and in very extreme cases, run borrowers into insolvency. In practice these seem to be only individual cases, albeit comprehensive statistical data on the causes of defaults from Finnish mortgage markets cannot be found.

The majority of loans are amortizing and have a pre-set maturity or installments. The amortizations decrease principals and CLTV levels, which also decrease the risk of loss causing defaults, and thus decrease the credit risk of banks.

Finnish banks are both able and willing to negotiate with customers that have run into trouble, e.g. to give amortization-free periods to help customers to avoid defaults. This kind of behaviour lowers the default rate, but also makes the concept of default hazy and hard to observe. Banks’ practices on registering these may vary and also the banks’ interpretation may be dependent on when the customer starts the negotiations, i.e. the interpretation may be different depending on whether the latest payment is already overdue or if the previous payment history is clean.

Differences between countries, and also between banks, in eagernessness to negotiate and give relief to otherwise defaulting customers may complicate also the comparisons of PD levels between regions and over the borders. This should be noticed in all applications where parameters of risk models have been estimated from another area to where they will be used. This also supports the idea of replacing the observed default data with modeled defaults derived from the frequency of the triggering events.

2.3 Credit risks of mortgage loans

Dividing the expected loss in to components helps to make the credit loss predictions and models more structured. A very typical, and also intuitive, way to make the split is to divide it in to: Probability-of-Default (PD), or default frequency for ex-post analysis; amount of loss in case of default, i.e. Exposure-at-Default

---

3Main use for bullet loan, i.e. loan without amortizations and typically with short maturity, in mortgage lending is the so-called bridge loan used in home swap situations when a new apartment/house needs to be paid for before the money from the sold apartment/house has been received. In addition to these it is possible that some investors utilise longer bullet loans in speculative housing investments.
(EAD); and to Loss Given Default (LGD), i.e. to the proportion of the loan assumed to be lost in case of default or not covered by the collateral.

Basel II formula, without the maturity adjustment parameter which is not used for mortgage loans, can be written as

$$ EL = PD \times LGD \times EAD $$

(1)

In some papers Recovery rate (RR) has been used instead of LGD in modeling. Conversion between these is easy, because $LGD = 1 - RR$. It should be noticed thus that $\text{Corr}(PD, LGD) = -\text{Corr}(PD, RR)$.

There are limitations on what kind of values these variables can have. E.g. $EAD > 0$, because when $EAD = 0$ the loan has already been paid back and the insolvency of the (previous) borrower doesn’t cause any losses to the lender. As a probability $0 \leq PD \leq 1$, where value $PD = 1$ can be interpreted as a certain default. Also for loss given default $0 \leq LGD \leq 1$, where $LGD = 0$ means that the collateral covers the whole loan principal, accrued interests, and all the other costs of defaults. In case $LGD = 1$, the collateral has no value or it doesn’t exist.

Values of the above-mentioned risk components can change over time: EAD decreases because of amortizations or increases because of unpaid accrued interest, PD is dependent e.g. on changes in the unemployment rate, and LGD depends on the remaining principle and on the development of collateral values. Because values of these components are dependent partly on the same explanatory variables, such as development of unemployment, the dependences between components can be observed as correlations.

As is well known, equation $E(XY) = E(X)E(Y)$ is true only if $X$ and $Y$ are orthogonal, i.e. independent. Otherwise $E(XY)$ needs to be corrected using covariance, and the correct formula of the equation is $E(XY) = E(X)E(Y) + \text{cov}(X,Y)$.

Multiple studies have found support and presented explanations on positive dependence between PD and LGD. Positive correlation between PD and LGD in the mortgage loan pools has been found e.g. in studies of Dimou et al. (2005) and VanOrder (2007). Frye (2000) started to call the situation where both PD and LGD increase as a ’Double hit’.

In the case where PD and LGD are dependent, Equation 1 is not true and as explained, EL is biased. To correct this bias, the effect caused by the correlation needs to be added to the model. This adjustment works only if the dependence between the variables is linear, otherwise copula functions, or in more complex

---

4 Originally ’Double misfortune’ in Frye (2000).
5 Or PD and EAD or LGD and EAD, but this thesis concentrates only to dependence between PD and LGD.
cases, simulations need to be used.

The majority of mortgage pool credit risk models are based on the models used with bonds and wholesale loans. In the literature these models have been traced to two different families: options-theoretic-based structural models, and intensity and hazard rate-based reduced form models. In some papers credit scoring models have been mentioned as a third type, but as scoring models concentrate only on PD modeling, these cannot be seen as real credit risk models in this context.

Structural models compare asset prices to loan values. Where the asset value sinks below the total amount of loans, the company defaults. One of the first structural models can be found from the paper of Merton (1974), and e.g. commercial products like KMV (Kealhofer Merton Vasicek -model) and RiskCalc (Moody’s) are based on Merton’s structural model and expansions derived from that.

Reduced form credit risk models lean on hazard rates, or to hazard processes, and intensity models widely used in actuarial applications, i.e. in insurance risk models. These models assume default rates to be continuous random variables. The history of reduced form models in credit risk calculations starts from the papers Jarrow and Turnbull (1995), Jarrow et al. (1997), and Duffie and Singleton (1999). Thus in some contexts these models are also called Jarrow-Turnbull models. Two commercial products using this kind of modeling are CreditRisk+ of Credit Suisse First Boston (CSFB\textsuperscript{6}) and Kamakura (Kamakura Corporation).

The third group, credit scoring models, are based on indicators assumed or tested to reveal the creditworthiness and probability of default of counterparties. These models don’t have similar kinds of general theory behind the models used like structural and reduced form models. One commercial application using multiple financial ratios is the FICO score of Fair Isaac Corporation.

The assumptions, markets and legislation of mortgage loans differ a lot from wholesale loans and markets. For example, events that trigger default of the mortgagor differ from the triggers used in wholesale structural models. Default triggering events from the retail customers’ perspective have been presented in the paper Cairns and Pryce (2005). Because of these differences, defaults of retail loans cannot be analysed solely by downscaling the models used in wholesale loan risk analysis.

The model presented in this thesis\textsuperscript{7} cannot be classified directly within either of the two first mentioned model families. The modeling of the insolvencies can be seen as a derivative of a reduced form model. As soon as the insolvency events

\textsuperscript{6}In some sources also Credit Suisse Financial Products (CSFP).
\textsuperscript{7}In Essay 3: ‘Correlation between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools’
are divided into observed and unobserved defaults depending on the current LGD, the model also starts to adopt features typical for the structural models.

2.4 Truncated and censored distributions

Let’s assume that a variable follows a known continuous distribution for which at least two first moments exist. Sometimes events outside some limits cannot be observed or events outside the limits can be observed equal to the limit.

Samples where events larger or smaller than the set limit cannot be observed are called truncated. Similarly if these events outside the limits can be observed with a value equal to the size of the limit, the samples are called censored. These kinds of restricted distributions are widely observed in medical and actuarial samples, but can also be found in financial and economic data.

A distribution may also be doubly truncated or censored, meaning left and right truncated or censored, or even a combination of these. A Left Truncated Right Censored (LTRC) distribution is widely used, see e.g. papers Gijbels and Wang (1993) and Zhao et al. (2011), in insurance modeling due to the minimum compensation limit at the left tail and reinsurance of largest losses or defined maximum compensation at the right tail.

Both truncation and censoring affect the moments of observed distributions compared to the moments of unrestricted distributions. Before these moments can be used for analysis, or as parameters in simulations, the moments of unrestricted distributions need to be solved, and thereafter these solved moments should be used as parameters for either unrestricted distributions, or if possible (e.g. in case of bivariate normal or log-normal distributions), also for restricted distributions.

In some applications it is possible to face datasets where one or both tails are partly truncated and partly censored. This may be the case when the sample has been collected from the same population, but the observers collecting the data have different abilities to observe events outside the limits. Also in these cases, when the moments of a population are known, the moments of a restricted distribution can be calculated using a closed form solution. This closed form solution has been presented in the paper ‘Closed Form Solution of Correlation of Once or Twice Truncated or Censored Bivariate Log-Normal Distribution’.

Also multivariate distributions can be truncated or censored. Where one variable is truncated, it also has an effect on the marginal distributions of other variables. Due to the requirements of the presented applications, this thesis concentrates only on the bivariate distributions where one variable is restricted and the other one is unrestricted.
2.5 Unrestricted and restricted sample correlations

Correlation measures statistical dependence between two random variables. There are several different correlation coefficients\(^8\) that measure the dependence using different assumptions or usable when variables follow different scales of measure. In the case of more complex dependencies a common way is to use copula functions.

In the papers of this thesis the term ‘correlation’ always refer to Pearson’s product-moment correlation coefficient, which measures the linear dependence between two random variables. Theory and use of copula functions are not covered by this thesis.

The correlation coefficient defines only the linear dependence between two variables. It can be calculated for any two variables for which the two first moments exist and the moments are finite.

The Pearson’s product-moment correlation coefficient estimates linear dependence between variables, and thus works best with elliptical distributions. In this thesis the correlation coefficients have been calculated for bivariate log-normal distributions, especially where in a case of negative dependence, it can be easily seen that the dependence is not linear, and the distribution is not elliptical. It is important to understand what the correlation coefficient measures and how the estimated correlations can be interpreted in these cases.

For an asymptotic restricted and unrestricted bivariate log-normal sample the minimum and maximum correlations can be calculated. This makes the interpretation of observed sample correlations more meaningful. It should be noticed that the calculated asymptotic minimum and maximum correlations do not limit the sample correlation of smaller samples.

When only small samples are available, large confidence intervals of the correlation estimates add uncertainty to simulations. For example Lai et al. (1999) have studied confidence intervals of correlation in bivariate Normal distributed samples as a function of the number of observations. Sometimes the samples are so small, like natural catastrophes in insurance models, that correlations cannot be calculated, or at least the correlation estimates are extremely unreliable. If the correlation levels cannot be derived from the theory, one possibility is to use in those cases expert opinion models to approximate the correlations required for modeling.

\(^8\)E.g. Pearson’s product-moment correlation, Kendall tau rank correlation, and Spearman’s rank correlation, just to mention some of the most widely known ones. There is also a wide number of different correlation coefficient useful only with ordinal, nominal, and dichotomous data and their combinations.
2.6 Mortgage loan pool structure

Mortgage loan pools are heterogeneous by nature, which means that one or all of the features of loans are different, i.e. they may have different granting time, maturity, amount of principal, and original LTV. Also the reference rate used, margin, amortization type and schedule, as well as the amount and type of collaterals used vary. Additionally the borrowers are demographically different, and thus also the PDs of borrowers are not the same.

As already explained, the risk parameters of loans also change over time. E.g. the PD of borrowers is dependent on the economic cycle and especially on the demand for labour, i.e. changes in the level of new unemployed. Amortizations decrease the remaining principal, and changes in asset values, like the prices of apartments and houses, have an effect on collateral values.

Because all the loans are different, and thus also the risk parameter levels and changes over the time are different, the modeling of credit risks of the whole mortgage loan pool is inaccurate without information on the structure of the pool.

A bank has precise information on its own loans and thus also granular information on the risks of the whole pool. Outside banks the amount of information is much more limited. Central banks and national supervisory authorities collect data on granted mortgage loans and pools, but in most cases these data are not granular enough for loan level analysis. Thus credit risk analysis of the whole mortgage markets needs to be performed on an aggregate level and using assumptions and approximations on the pool structure.

Using e.g. average LTV in risk modeling may give a very biased and wrong view on risks. The riskiness of the pool is dependent on the amount of loans with very high LTVs, e.g. higher than 100%. Without the LTV distribution, or at least information on variance, the average LTV tells very little on the number and amount of loans that may generate losses in case of defaults.

For reliable analysis supervisors or central banks need to find a way to approximate the pool structure as accurately as possible using the information available. The article ‘Modeling the Current Loan-to-Value Structure of Mortgage Pools without Loan Specific Data’ presents one method to do so.

2.7 Structured expert opinion based forecasting

The Delphi method was introduced in the 1950s at the RAND Corporation (see Helmer, 1964). It aims to maintain the advantages of an interacting group without potentially counterproductive group dynamics, such as dominant individuals who
may not be the best experts. Ideally, panellists used are experts in the same field, but with somewhat different backgrounds.

The traditional version of the method is based on a multi-round survey. Respondents are asked to answer a number of questions in writing. Answering is anonymous; other respondents do not know who answered what. In most cases the answers are either numeric estimates, ratings on a scale, or yes or no. Often the respondents also have the opportunity to write comments on the issues raised in the questionnaire. Thereafter statistics on answers and related comments are distributed to respondents, but this information is anonymous and no respondent can identify who answered what. Each respondent is allowed to modify his own answers, and possibly to add more comments.

After a few rounds, some convergence in answers is normally observed due to a group opinion building process, leading to less variance in answers and more agreement within the panel. The number of rounds can be either predetermined or it may depend on the criteria of consensus and stability.

The final answer of the group is defined as the mean or median of the individual answers. In many cases even the questions to be answered are proposed and selected by group members themselves before the first answering round. Forecasting accuracy of the group normally improves over Delphi rounds, and the Delphi method works better than staticized groups, i.e. simple one round surveys; this finding is reported by Parenté et al. (2005), Helmer (1964), Dalkey (1968), Graefe and Armstrong (2011), Song et al. (2013) and in various studies reviewed by Rowe and Wright (1999).

The FTF meeting is the simplest and probably most often used method; group members sit in the same room and discuss the issues until they either reach a consensus, or at least that the majority backs a view. According to Kerr and Tindale (2011), such meetings are good for pooling information, mutual error checking and motivation enhancement. On the other hand, they may be particularly vulnerable to ‘tyranny of the majority’, dominance of powerful individuals, inattention to unshared information, or group overconfidence. Other potential problems of FTF meetings include the bandwagon effect (tendency of ideas to spread among people like fads), the underdog effect (tendency of some people to vote for losing candidates or views), and the halo effect (tendency to weight an opinion according to a general impression of the person who expresses it).

According to Ang and O’Connor (1991, pp. 142), the Delphi method combines both mathematical and behavioural approaches, with an ‘aim to improve behavioural aggregation by substituting the dysfunctional aspects of achieving consensus with a mathematical process of achieving the final group judgement’.
In the best case, the method helps to eliminate a number of problems of FTF meetings, such as the influence of dominant individuals and the unwillingness of many people to defend unorthodox views, even well founded ones.

Updating the forecasts made and making forecasts more reliable requires new meetings of experts or more survey rounds, which makes the process expensive and time consuming. Thus researchers have tried to find more cost-efficient methods to adjust the forecasts. The use of post survey methods to increase the accuracy of forecasts is not a new idea, e.g. Armstrong (2006) has listed and evaluated evidence on numerous post hoc correction methods.

3 Summary of essays

The next four subsections summarise the key findings of the papers, the central scientific contribution, and present some applications. A short synthesis of these findings will be presented in Section 4. Mathematical and more detailed explanations can be found in the articles and essays themselves.

3.1 Closed Form Solution of Correlation of Once or Twice Truncated or Censored Bivariate Log-Normal Distribution

This article, written together with Dr. Jouko Vilmunen, derives the closed form solution of the moments and correlation of bivariate, once or twice truncated and/or censored log-normal distribution. The presented formula is a generalization, and can be used for bivariate log-normal distributions where one or both tails of one variable have been partly censored and partly truncated.

The same formula works with all kind of restriction combinations. It is assumed that all the events in one tail are censored or truncated. The proportion of censored events can be marked with \( D \), where \( 0 \leq D \leq 1 \). The proportion of truncated events is then \( 1 - D \). If \( D = 1 \), all the events of that tail are censored. Contrary, if \( D = 0 \), the tail is fully truncated. In all the other cases where \( 0 < D < 1 \), the tail is partly censored and partly truncated. The limit value and the type of restriction, i.e. the value of \( D \), for both tails can be parametrised individually.

The central scientific contribution of this paper is the closed form equation. Although solutions to some specific cases have been presented in the literature, see e.g. Kotz et al. (2000) and Campbell et al. (2008), this kind of generalisation has not been presented as such. In literature most of the solutions are based on the bivariate Normal distribution, and the restrictions of tails have been fixed. The most common of this kind of solution are the previously mentioned LTRC -models.
In applications the formula can be used to substitute or supplement simulations. For example in cases where the parameters of the population or the unrestricted sample are known, the impacts of regulation, like LTV restrictions as a part of macroprudential tools, on the moments and correlation of regulated markets can be calculated.

The presented formula can be used also another way around. When the parameters of the restricted sample can be estimated from the data, the parameters of the unrestricted sample can be solved. Unfortunately a closed form solution for those can not be derived but the test made indicates that the parameters can be approximated unequivocally using numerical methods\(^9\).

3.2 Modeling the Current Loan-to-Value Structure of Mortgage Pools without Loan Specific Data

The defaults of mortgagors that cause the largest losses for banks are the ones with the highest CLTVs. In practice, if realisation costs are omitted, only loans with CLTV higher than 100% are able to generate losses. To approximate the number of this kind of loans banks have in their pool, they need to know the current pool structure. For banks that is easy - they have all the loan-level information required for these calculations. For supervisors and central banks the situation is much more challenging. To approximate the potential losses and to assess the capital levels of banks, they need to find a way to estimate the current pool structures without granular data, e.g. to approximate the loss levels when values of collaterals decrease and the number of insolvency events increase.

The model presented in this article allows the estimation of the heterogeneous mortgage loan pool structures using only publicly available data.

Granular, loan level data required for the accurate modeling of the pool structure and mortgage loan credit risk analysis, is normally available only for creditors. Thus outside banks, and when the whole mortgage loan markets are analysed, a method to replicate the granular data needs to be developed.

The model is based on the formula that emulates the loan stock dynamics using the stock, inflows, i.e. granting of new loans, and calculated outflows, i.e. amortizations and prepayments. The CLTV calculation utilises approximated collateral value development, e.g. using Housing Price Index (HPI), and the effect of amortizations on the remaining principal.

In the model the heterogeneous pool has been divided in to a large number of

\(^9\)A new research project on the topic has been started at the end of 2014. The study tries to find a method to test if the tail correlation truly differs from the unrestricted correlation, or whether the observed difference between correlations is only caused by truncation.
internally (approximately) homogeneous cohorts, with the same granting period and original LTV. When even more exact analysis is needed, the original pool can be divided into multiple sub-pools using some other characteristics of loans, e.g. maturity at origination or type of amortization, and the same calculations can be performed for all these sub-pools.

The data needed for the calculations includes time series on the total amount of the remaining pool, new granted loans, as well as an LTV distribution of the loans originated. To calculate CLTV a time series on historical development of collateral values, e.g. HPI, is also needed. The method can also use more granular data instead of making assumptions, when that kind of data is available. Thus the more granular data available, the more accurate is the structure estimate.

As a result the model gives three matrices which dimensions equal to the number of maturity buckets\(^{10}\) times the number of LTV groups. Two of these matrices include information on the number of loans and total remaining principals per cohort, and the third shows the calculated CLTVs for the same cohorts. For credit risk analysis the estimated CLTVs can be converted to expected LGDs.

The presented model is dynamic and it can be used to approximate not only the current structure but also for forecasting the future development of pool structures. The forecasting feature is useful e.g. in impact analysis and evaluation of regulatory changes.

The presented method is the scientific contribution of this paper. In practical applications the estimated structure can be used with already existing credit risk models to make the forecasts and estimates more reliable and accurate.

3.3 Correlation between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools

The third article studies how the mortgage loan pool structure and creditors ability to observe zero-loss defaults affect the observed sample correlation. In the analysis both homogeneous and heterogeneous pool structures, e.g. a single loan cohorts as well as the whole mortgage loan pools, have been studied.

This paper is more applied compared to the two first papers that concentrated on presenting and deriving new methods and models. The paper utilises the findings and presented methods of these theoretical papers.

Homogeneous pool structures have been used to analyse the effects of LTV of granted loans on the observed correlation, and also to compare the effects

\(^{10}\)The number of periods included to average maturity. For computational reasons an extra column needs to be added.
creditors’ different abilities to observe zero-loss defaults cause.

The results of analysis of homogeneous pools can be used for single cohorts, but for the whole mortgage loan pools the assumptions used are too strict. The homogeneity assumption of pools is far from the reality and thus does not allow the generalization of the results to the whole mortgage loan pools.

To make the analysis more applicable in the analysis of real mortgage loan markets, the modeled heterogeneous pool structure has been used. The modeled structure uses the parameters estimated from the markets, and as presented in the second paper, the modeled structure is a good approximation to true pool structure.

The amount of different kind of heterogeneous pool structures is vast, even when the structure has been divided into relatively small numbers of heterogeneous cohorts. Taking into account the number of different parameters (number of loans and total remaining principal, and CLTV) and large number of cohorts, e.g. in this paper the number of cohorts is 558, neither comprehensive analysis on effects of heterogeneity nor meaningful ceteris paribus sensitivity analysis can be made. Thus the paper settles for presenting an analysis on a single heterogeneous pool using different market assumptions, i.e. changes in variance of collateral values and different correlations between the rate of new unemployed and collateral values, and comparing properties of homogeneous and heterogeneous pools.

This paper contributes to the discussion on double-hits, started by Frye (2000), where the economic cycle affects simultaneously both PDs and LGDs. In most of the wholesale and bond credit risk models this effect has been noticed, and the adverse effect has been corrected for using correlation or copula based adjustments. Compared to those, retail and mortgage loan products and markets are very different. Thus similar kinds of methods, where observed correlation have been used, cannot be used to make the estimations more reliable.

The scientific contribution of the paper is related to the double-hit correction of the mortgage loan credit risk models. The paper reveals that the observed correlation is a poor estimate, because of the bias and instability, for the correlation that needs to be used in the modeling and simulations. Thus some other way to adjust the bias needs to be found.

One solution which avoids the use of an observed correlation that is sensitive to the structure of the pool, has been presented in the paper. Starting the modeling from the distribution of collateral value and new unemployment rates and the correlation between these variables allows the derivation of PD and LGD values. The testing of the effectiveness of this approach has been left as a topic of later studies.
The presented findings can be used in mortgage loan credit risk applications to increase the reliability of models in a changing environment where default rates, collateral values, and the variance of collateral values may change over time. The presented results also help in the comparisons of results collected from different markets where e.g. the structures of the pools or observability of defaults vary.

3.4 The Delphi method in forecasting financial markets - An experimental study

In banking and finance most of the forecasts are based on econometric and time series models. Sometimes, e.g. when the forecasted events are very rare, there are not enough data for reliable forecasts, or some explanatory variables are qualitative and very hard, if not impossible, to quantify. The first one is typical e.g. in natural catastrophe models, and the second one e.g. when the development of operational risks needs to be forecasted. In these cases expert opinion based forecasts may be used. Unfortunately expert opinions may be heavily biased and tendentious. To make the forecasts more reliable some structured expert opinion-based forecasting methods have been developed.

In this article, written together with Dr. Karlo Kauko, forecasting power of the two expert opinion models, Delphi and Face-to-Face meetings, has been tested, and two post-survey methods to correct the forecast errors have been presented. The experiment has been carried out using two similar size of expert panels from Bank of Finland and Finnish Financial Supervision Authority. The panellists were asked to present forecasts on 15 financial market variables.

The first post-survey forecast correction model aims to correct the perseverance bias caused by the self-confidence of panellists. The hypothesis was that over the forecasting rounds the panellists will change their forecasts to the right direction, but because of strong self-confidence, less than they should. This kind of conservativeness toward their own forecasts may reflect their position as experts and experience on topics forecast.

The bias caused by overconfidence has been measured comparing the distances between original, i.e. first round, forecasts and realized values, and distance between first round forecast and the final, i.e. third round, forecasts. After that it was analysed how much more the panellists should have changed the forecast.

Results of the experiment gave support to the hypothesis of too small corrections. In most cases the correction between first and third round was to the right direction, but as assumed, too little. In the sample data the changes should be multiplied by four (with two decimals 4.39) to get the most accurate forecast.

The second method tests correction of forecast error using conditionalised fore-
casts. Where the forecast variables are dependent on the value of another variable, the forecast error in that explanatory variable affects also the forecasts of these dependent variables. Where the forecasting of that explanatory variable is very unreliable, it may also ruin the reliability of the other forecasts dependent on it.

In the method presented, panellists are first asked to forecast the value of the explanatory variable. Conditional to that forecast, panellists give their forecasts of dependent variables, and also forecasts of dependent variables if the value of the explanatory variable differs a certain amount from their original forecast.

This kind of set-up can be used for two purposes. The first one has been presented in the article, i.e. in theory using the method the accuracy of forecasts can be improved as soon as a more reliable forecast for explanatory variable is available, i.e. the forecasts of dependent variables can be adjusted using the sensitivity of the forecast to the change in the explanatory variable, and the difference between the more reliable estimate and the original one.

Secondly the method can be used to reveal the source of forecast error, and also the experts’ ability to model the phenomenon, i.e. dependence between explanatory variable and dependent variable. That can be done only after the realised values of variables are available. If the adjusted forecast using the realisation of the explanatory variable is better than the original forecast, it can be assumed that the main source of error was the forecast error in the explanatory variable. This kind of outcome implicates that the experts are able to model the dependencies between variables and understand the behaviour of the phenomenon. If the adjusted forecast is not better than the original one, it indicates that the modelling of dependence hasn’t been done correctly, probably because the experts do not recognise the dependencies correctly.

Unfortunately the used data didn’t give any clear indications whether the method works or not. In only five cases out of fourteen the adjusted forecast was better than the original. The sample size was so small that the statistical analysis of the result is not reliable enough to make any conclusions.

The key scientific contributions of the paper are the finding of support to the existence of perseverance bias, and indication that the post survey correction of forecasts is possible.

In the applications the result may help to save both money and time, because the corrections to the forecasts can be made without gathering panelists together and spending time in new forecasting rounds.
4 Conclusion and discussion

The third paper can be seen as a summarising paper of the three first studies. The methods used and results presented are based on the application of the models and methods presented in the two first papers. The synthesis of these studies has been shortly presented in the Section 3.3 and can be found in more detail from the paper itself.

The existence of the double hit is not unique to loan markets only. Similar kinds of effects can be found e.g. from the insurance markets. In actuarial models these dependencies have been taken into account similarly as in banking applications, i.e. using correlations and copulas, see e.g. McNeil et al. (2015).

In the insurance environment in some cases the events are so rare or different that either correlations cannot be calculated or the confidence levels of estimated correlations are so wide that the estimates cannot be used for calculations. In those cases expert opinion models have been used to modify the estimates and possibly to add information to the estimation that cannot be included using only statistical methods and empirical observations.

The fourth paper presents an example how the Delphi method can be used in forecasting. The paper also explains some known reasons why expert opinion methods may result in biased estimates and present two methods how these biases can be fixed.

The definition of the research topic rules out multiple closely related research topics and applications of presented results. These are potential subjects of new research and some of these are presented here.

The method presented in the first paper is applicable only when the parameters of the unrestricted sample or population are known. In the reverted case when the parameters of the restricted sample are known, the formula cannot be used as such to solve the parameters of the unrestricted sample. Instead of using a closed form solution some numerical method needs to be used. The test made indicate that standard optimization methods used to solve systems of equations can be used in solving and the numerical methods seem to converge to the correct results even when the starting values of the parameters are far from the final ones. The writing of this paper has already been started, but the testing of different optimization algorithms needs more work.

Another research project which has been already started studies the confidence intervals of restricted sample correlations. The study analyses the effect of different properties of the restricted sample, like sample size and truncation and 11In the tests Matlab function fsolve.
censoring limits in the confidence interval.

As already mentioned, the deriving of PD and LGD using the properties and development of the triggering variables, e.g. changes in the number of new unemployed and collateral values, needs to be studied more carefully using empirical data.

The fourth paper also left some questions open. E.g. can the conditional forecasting method be used to explain sources of forecast errors and to increase the reliability of forecasts? That said the method seemed to fail in this task according the presented study. Also although the paper found support that the perseverance bias exists and its adverse effects can be adjusted in the correct direction, the paper didn’t study how stable and predictable these effects are, and whether the adjustment factors can be estimated or solved before the final results are known. These results define the applicability of the perseverance bias correction in practical applications.
References


Introduction and synthesis


Vilmunen, Jouko; Palmroos, Peter *Closed form solution of correlation in doubly truncated or censored sample of bivariate log-normal distribution.* Bank of Finland Research Discussion Papers 17/2013. 2013. ISBN: 978-952-6699-31-8, ISSN: 1456-6184.\(^{12}\)

\(^{12}\)The published paper is a previous version of the paper included in this dissertation. The name of the latest version is 'Closed Form Solution of Correlation of Once or Twice Truncated or Censored Bivariate Log-Normal Distribution'.
Essay 1
Closed Form Solution of Correlation of Once or Twice Truncated or Censored Bivariate Log-Normal Distribution

Peter Palmroos\textsuperscript{1}, Jouko Vilmunen\textsuperscript{2}

\textsuperscript{1}Financial Supervisory Authority, PO Box 103, 00101 Helsinki, Finland
\textsuperscript{2}Bank of Finland, PO Box 160, 00101 Helsinki, Finland

Abstract

This paper presents a closed form solution for the moments and correlation of bivariate log-normal distribution which one or both tails are partly or fully truncated or censored. The rate between censored and truncated tail events can be set for both tails separately. Partially censored distributions may be observed when data has been gathered from several sources using non-standardized observation or registration methods. The closed form solution, and knowing the dependence between the correlation of unrestricted distribution and the correlation of bounded distribution, helps the building of more accurate and efficient simulations. The solution works also with bivariate log-normal left truncated and right censored (LTRC) data, which is typical in insurance risk calculations.

1 Introduction

Truncated and censored distributions are common\textsuperscript{1} in follow-up and survey studies in economics and biomedicine, as well as in actuarial applications e.g. with surveillance data. Furthermore, the prudential requirements and modeling options embedded in the Solvency II and Basel II frameworks, which are

\textsuperscript{1}List of LTRC studies can be found e.g. from the paper Gijbels and Wang (1993).
applied in regulating insurance and banking sectors, have greatly increased
the interest in models connected to dependencies in extreme market condi-
tions, e.g. during deep recessions, as well as in measures that capture these
dependencies in different samples.

It is known that the correlation of the bounded\(^2\) distribution differs from
the correlation of the unbounded one, and that the placing of boundaries have
effect on the correlation. One example of these effects on correlation can be
found from the paper Lien (1985), where correlations of split distributions
have been compared.

In this paper we present the comprehensive and easy to apply closed form
solution for the moments, covariance, and correlation of two log-normally dis-
tributed random variables, when the distribution is once or twice truncated or
censored. Our method covers all the possible single and double censoring and
truncation combinations, including the partial censoring and truncation cases.
The rate between censored and truncated tail events can be set for both tails
separately. The method allows also calculation of minimum and maximum
boundaries on bounded correlations. These solved minimum and maximum
correlations help to clarify the interpretation of estimated correlations in used
samples.

Both truncation and censoring have effect on observation and registration of
tail events in distribution, and thus also to moments and correlation. In case
of left truncation, events below the set truncation point can not be observed,
i.e. values of truncated variable are cut away below the truncation point. Idea
of left tail truncation can be seen in Figure 1a. With insurances this kind
of lower boundary may be caused by self-insured retentions. Similarly when
right truncation is used, events beyond the set boundary can not be observed.

In case of right censoring all events larger than the set censoring point are
observed, but with value equal the censoring point. This can be seen in Figure
1a. A commonly used example on censoring is demand of tickets to the football
games. If the demand is larger than the size of the football arena, the true
demand can not be observed, but what can be observed is that in those cases
all the seats of the arena are sold out. In insurance data right censoring can
be caused by re-insurance or maximum limits of insurance compensation. A
more comprehensive definition of truncation and censoring can be found e.g.
from Greene (2003, pp. 756–764) and Wooldridge (2010, pp. 778–780 and

\(^2\)To make the text more readable, bounded distribution will be used as a synonym to
to all doubly of single truncated and/or censored distributions. Analogously, ‘bounded
variable’ is used to refer to a truncated or censored random variable.
In this paper we present also a partial censoring case, where a known proportion of events in a tail is censored and the rest are truncated. In practice partially censored and truncated data may be observed when data is gathered from multiple sources that utilize different methods to observe or register events. For example when mortgage loan default data is gathered from the whole banking sector, some of the banks may be able to register all zero-loss defaults (censoring), while some can observe only loss causing defaults (truncation). Similar situation is also possible in time series data when the regulation of registration of zero-loss defaults changes over the time. The idea of partial censoring can be seen in Figure 1b.

Figure 1. a) Left truncated and right censored (LTRC) bivariate log-normal sample, and b) Partly censored and otherwise truncated right tail. Cross on the observation means that event can not be observed, and dotted arrows indicate how the censored events can be observed.

In tail simulations and in other applications when bounded distributions are used, the distribution parameters and correlation of the unbounded distribution needs to be known. The closed form solutions to the parameters for a single truncated distributions have been derived by Johnson and Kotz (1972) and Lien (1985). Kotz et al. (2000) has derived the solution to the moments and correlation for random variables from a truncated bivariate normal distribution. Zhao et al. (2011) lists some papers that apply or study left truncation and right censoring (LTRC) cases.

Even when these special cases have been studied, and closed for solutions have been presented, none of the papers give the general closed form solution for moments and correlation that can be used with bounded bivariate log-normal distributions, and especially with distributions where one or both tails are partially censored and truncated.

The next section presents the closed form solution. Section three gives some
examples on the effects truncation and censoring have to the correlation of the restricted sample. The last section concludes.

2 The closed form solution

Assume that the random vector \( \mathbf{u} = (x, y)^T = (\ln X, \ln Y)^T \) is jointly normally distributed with a known mean vector and variance-covariance matrix:

\[
\begin{pmatrix}
x \\
y
\end{pmatrix}
= \begin{pmatrix}
\ln X \\
\ln Y
\end{pmatrix}
\sim N_2
\begin{pmatrix}
\mu_x & \mu_y \\
\sigma_x \sigma_y \rho & \sigma_y^2
\end{pmatrix}
\]

Using the Cholesky decomposition of variance-covariance matrix, the components of the random vector \( \mathbf{u} \) can be represented as

\[
x = \mu_x + \sigma_x z_x, \quad z_x \sim N(0, 1) \\
y = \mu_y + \rho \sigma_y z_x + \sigma_y \sqrt{1 - \rho^2} z_y, \quad z_y \sim N(0, 1), \quad z_x \perp z_y
\]

Pearson’s correlation coefficient between \( x \) and \( y \), conditional on bounds imposed on the distribution of \( X \) is defined in the usual way as

\[
\rho_{Q,D} = \frac{\text{cov}(X, Y|Q, D)}{\sqrt{\text{Var}(X|Q, D) \text{Var}(Y|Q, D)}}
\]

Here \( Q \) is defined as \( Q = (L, U) \), where \( L \) and \( U \) denote, respectively, the value of the lower and the upper boundary of \( X \), and the variable \( D = (D_L, D_U) \) indicates the proportion of censored events in tail. As an example \( D_L = 0.4 \) means that 40% of events in the left tail can be observed as a border events and the rest 60% of events have been truncated. Because \( D_L \) and \( D_U \) are proportions, the condition \( 0 \leq D \leq 1 \) should hold.

The first moments of the bounded variable \( X \) are equal in the bivariate and one dimensional cases. The closed form solution to the doubly truncated one dimensional distribution can be found, e.g. from the paper of Bebu and Mathew (2009).

In moment and correlation calculations the proportions of events in to the tails, and also in the center part of the distribution, are needed. These are marked as \( T_L, T_C \) and \( T_U \), where \( L \) indicates the lower (left) tail, \( C \) the mid part, and \( U \) the upper (right) tail.

\[
\begin{align*}
T_L &= \frac{D_L \Phi(L_1)}{D_L \Phi(L_1) + [\Phi(U_1) - \Phi(L_1)] + D_U [1 - \Phi(U_1)]} \\
T_U &= \frac{D_U [1 - \Phi(U_1)]}{D_L \Phi(L_1) + [\Phi(U_1) - \Phi(L_1)] + D_U [1 - \Phi(U_1)]} \\
T_C &= \frac{\Phi(U_1) - \Phi(L_1)}{D_L \Phi(L_1) + [\Phi(U_1) - \Phi(L_1)] + D_U [1 - \Phi(U_1)]} \\
&= 1 - T_L - T_U
\end{align*}
\]
where Φ is the cumulative distribution function of the standardized normal distribution, and where

\[
U_1 = \frac{\ln U - \mu_x}{\sigma_x}; \quad L_1 = \frac{\ln L - \mu_x}{\sigma_x}
\]

Now the variance of \(X\), conditional on the bounds, can be calculated via the well known decomposition

\[
\text{Var}(X|Q, D) = E(X^2|Q, D) - \left[ E(X^1|Q, D) \right]^2
\]

The first moments of the bounded \(X\), \(E(X^k|Q, D)\) where \(k\) indicates the number of the moment \(k = (1, 2)\), can be expressed as

\[
E(X^k|Q, D) = T_L L^k + T_U U^k + T_C e^{k\mu_x + \frac{k^2}{2} \sigma_x^2} \left[ \frac{\Phi(U_{2,k}) - \Phi(L_{2,k})}{\Phi(U_1) - \Phi(L_1)} \right]
\]

where

\[
U_{2,k} = \frac{\ln U - (\mu_x + k\sigma_x^2)}{\sigma_x}; \quad L_{2,k} = \frac{\ln L - (\mu_x + k\sigma_x^2)}{\sigma_x}
\]

The solution to the first \(k\) moments of conditional \(Y\), on the other hand, takes the following form

\[
E(Y^k|Q, D) = e^{k\mu_y + \frac{k^2}{2} \sigma_y^2} \left\{ T_L \left[ \frac{\Phi(L_{3,k})}{\Phi(L_1)} \right] + T_C \left[ \frac{\Phi(U_{3,k}) - \Phi(L_{3,k})}{\Phi(U_1) - \Phi(L_1)} \right] + T_U \left[ \frac{1 - \Phi(U_{3,k})}{1 - \Phi(U_1)} \right] \right\}
\]

where

\[
U_{3,k} = \frac{\ln U - (\mu_x + k\rho\sigma_x\sigma_y)}{\sigma_x}; \quad L_{3,k} = \frac{\ln L - (\mu_x + k\rho\sigma_x\sigma_y)}{\sigma_x}
\]

When both of the first moments have been solved, the variance of \(Y|Q, D\) can be calculated similarly as in Equation (6).

The calculation of the covariance between \(X|Q, D\) and \(Y|Q, D\) follows the standard definition of covariance:

\[
\text{Cov}(X,Y|Q, D) = E(XY|Q, D) - E(X|Q, D)E(Y^1|Q, D)
\]

The only missing term in Equation (3) consequently is \(E(XY|Q, D)\). This can be obtained from the moment generating function of the bivariate normal distribution, and after a couple phases can be written as

\[
E(XY|Q, D) = e^{\mu_x + \frac{1}{2} \sigma_x^2} \left\{ T_L L \left[ \frac{\Phi(L_{3.1})}{\Phi(L_1)} \right] + T_U U \left[ \frac{1 - \Phi(U_{3.1})}{1 - \Phi(U_1)} \right] + T_C e^{\mu_x + \frac{1}{2} \sigma_x^2 + \rho\sigma_x\sigma_y} \left[ \frac{\Phi(U_4) - \Phi(L_4)}{\Phi(U_1) - \Phi(L_1)} \right] \right\}
\]
where

\[ U_4 = \ln U - \left( \frac{\mu_x + \sigma_x^2 + \rho \sigma_x \sigma_y}{\sigma_x} \right) \]

\[ L_4 = \ln L - \left( \frac{\mu_x + \sigma_x^2 + \rho \sigma_x \sigma_y}{\sigma_x} \right) \]

(13)

At this stage we have all the terms needed in the correlation Equation (3). The derivation of these formulas in more comprehensive format can be found from Appendix A.

It can be easily seen that when \( L \to 0 \) and \( U \to \infty \), regardless of the values of \( D_L \) and \( D_U \), the correlation of bounded variables approach the correlation of unbounded variables:

\[ \rho_{Q,D} \stackrel{U \to \infty}{\longrightarrow} \rho_{XY} \]

This is shown in Appendix B.

If \( \rho = 0 \), then the correlation between \( X|Q,D \) and \( Y|Q,D \) also equals zero. The prove can be found from Appendix C.

3 The numerical examples

The truncation and censoring points may have a remarkably effect on the observed correlation. Depending on the relative sizes of the tails that have been truncated or censored, the difference between the correlations of the bounded and unbounded distributions can be large. In addition to the population parameters, the types and places of the boundaries also affect the possible minimum and maximum correlations\(^3\) of the bivariate log-normal distributions. These calculated minimum and maximum correlations hold only for infinite or large samples. With small samples the observed correlation can naturally be outside these borders. Those outside-the-borders cases cause problems to estimation of correct distribution parameters of unrestricted sample.

Table 1 presents these bounded and unbounded correlations and asymptotic minimum and maximum correlations in all four combinations of double truncation and censoring, when the parameters of bivariate normal distribution are the ones presented at and below the table. As can be seen from the presented correlation values at Table 1, the difference between sample and population correlations can be considerable. Differences are so large that omitting the effects of censoring or truncation will make the results of simulations imprecise and misleading.

The bounded correlation can be smaller or larger than the correlation of the unconstrained distribution. This is true for all combinations of truncation

\(^3\)Boundaries are calculated placing -1 and 1 to \( \rho \) in Equations (10), (12), and (13), and keeping all the other parameters unchanged.
### Table 1. Calculated correlations with minimum and maximum boundaries

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivariate normal distribution ((x, y))</td>
<td>-0.791</td>
<td>1.000</td>
</tr>
<tr>
<td>Bivariate log-normal distribution ((X, Y))</td>
<td>-0.700</td>
<td>0.909</td>
</tr>
<tr>
<td>Doubly truncated distribution</td>
<td>-0.448</td>
<td>0.991</td>
</tr>
<tr>
<td>Doubly censored distribution</td>
<td>-0.500</td>
<td>0.907</td>
</tr>
<tr>
<td>Left truncated and right censored distribution</td>
<td>-0.529</td>
<td>0.907</td>
</tr>
<tr>
<td>Left censored and right truncated distribution</td>
<td>-0.467</td>
<td>0.954</td>
</tr>
</tbody>
</table>

When \(x \sim N(0.094, 0.045)\), \(y \sim N(-0.117, 0.652)\), \(\Phi(L_1) = 50\%\), and \(\Phi(U_1) = 95\%\).

or censoring. As shown in the Appendix C, in case of a zero unbounded correlation, the bounded correlations will also be zero in all cases.

![Figure 2](image1.png)

**Figure 2.** Difference between a) the population correlation and the correlation of the doubly truncated and the LTRC samples and b) the population correlation and the correlation of the doubly censored and the LCRT samples as a function of population correlation

Figures 2a and 2b present the difference between the correlation coefficient of the unbounded and bounded distributions \((\rho_{Q,D} - \rho)\) as a function of \(\rho\), when the other parameters are as in Table 1. If the absolute value of the bounded correlation is smaller than the population correlation, the line will stay above the horizontal axis when \(\rho < 0\) and under when \(\rho > 0\).

The calculated correlations of the double partially censored distribution as a function of \(DL\) and \(DU\) are presented in Figure 3. The values at the corners of the presented plane equal the correlation values presented in Table 1, e.g. in the corner where \(DL = 1\) and \(DU = 0\), i.e. the upper right corner that means LTRC case, the value equals the second last row of the Table 1. It should be noticed that correlation of partially censored distribution can get correlation values outside the corner values.
Figure 3. Correlation of the double partly censored distribution as a function of $D_L$ and $D_U$, when $x \sim N(0.094, 0.045)$, $y \sim N(-0.117, 0.652$), $\rho_{x,y} = -0.791$, $\Phi(L_1) = 50\%$, and $\Phi(U_1) = 95\%$.

4 Conclusion

Log-normality is a common assumption that researchers and practitioners, both in banking and actuarial applications, impose in models of risk measurement. Quite commonly these models use the Pearson’s product-moment correlation to model linear dependence between the variables under scrutiny, even when it is known that the variables are not normally distributed.

The validity of log-linearity can be debated, but even if one can justify its use, the possibility of non-linear dependencies between log-normally distributed variables complicates the interpretation of linear correlation measures. Moreover if the distribution of one of the variables is truncated or censored, the correlation coefficient may not be the most natural choice for a measure of dependence and, indeed, it may be very difficult to interpret.

As presented in the Section 3 researchers should pay attention to the results derived in applications, where samples from the underlying bounded distribution are analyzed. For example practitioners should be cautious when performing simulations and interpreting results when truncated or censored multivariate distributions have been used, or when correlation matrices from truncated or censored data have been estimated.
5 Acknowledgements

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Authors work as a full time researchers in Financial Supervision Authority and Bank Of Finland respectively, and didn’t get funding from any other sources.

References


Appendix A. Derivation of the moments

Appendix A.1. Doubly truncated normal distribution

Conditional correlation
\[ \rho_Q = \frac{\text{cov}(x, y | Q)}{\sqrt{\text{var}(x | Q) \cdot \text{var}(y | Q)}} \]  

(14)

where \( Q = (L, U] \), and
\[ \begin{pmatrix} x \\ y \end{pmatrix} \sim N_2 \left( \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma^2_x & \sigma_x \sigma_y \rho \\ \sigma_x \sigma_y \rho & \sigma^2_y \end{pmatrix} \right) \]  

(15)

Through Choleskey decomposition of
\[ \Sigma^{-1} = \begin{pmatrix} \sigma^2_x & \sigma_x \sigma_y \rho \\ \sigma_x \sigma_y \rho & \sigma^2_y \end{pmatrix}^{-1} \]  

(16)

and after relabeling the components of the vector random variable \( u = (x, y)^T \),
\( x \) and \( y \) can be represented as
\[ \begin{align*} 
    x &= \mu_x + \sigma_x \tilde{x}_x, \quad \tilde{x}_x \sim N(0, 1) \\
    y &= \mu_y + \rho \sigma_y \tilde{x}_x + \sigma_y \sqrt{1 - \rho^2} \tilde{z}_y, \quad \tilde{z}_y \sim N(0, 1), \quad \tilde{x}_x \perp \tilde{z}_y 
\end{align*} \]  

(17)

Since
\[ \begin{align*} 
    L < x \leq U & \iff L' < z_x \leq U' \\
    L' &= \frac{L - \mu_x}{\sigma_x} \quad \text{and} \quad U' = \frac{U - \mu_x}{\sigma_x} 
\end{align*} \]  

(18)

using \( \phi(z) = \exp(-0.5z^2) \), hence e.g. \( z\phi(z) = (-1) \frac{d\phi(z)}{dz} \), and then
\[ \begin{align*} 
    E \left[ z_x | L' < z_x \leq U' \right] &= \int_{L'}^{U'} z_x \phi(z_x) \frac{dz_x}{\Phi(U') - \Phi(L')} \\
    &= \int_{L'}^{U'} (-1) d\left[ \phi(z_x) \right] \frac{\Phi(U') - \Phi(L')}{\Phi(U') - \Phi(L')} \\
    &= \frac{\phi(U') - \phi(L')}{\Phi(U') - \Phi(L')} 
\end{align*} \]  

(19)

and also
\[ \begin{align*} 
    E \left[ z_x^2 | L' < z_x \leq U' \right] &= \int_{L'}^{U'} z_x^2 \phi(z_x) \frac{dz_x}{\Phi(U') - \Phi(L')} \\
    &= \int_{L'}^{U'} (-z_x) d\left[ \phi(z_x) \right] \frac{\Phi(U') - \Phi(L')}{\Phi(U') - \Phi(L')} \\
    &= \frac{(-z_x) \phi(z_x)}{\Phi(U') - \Phi(L')} + \int_{L'}^{U'} \phi(z_x) \frac{dz_x}{\Phi(U') - \Phi(L')} \\
    &= 1 - \left[ \frac{U' \phi(U') - L' \phi(L')}{\Phi(U') - \Phi(L')} \right] 
\end{align*} \]  

(20)
Hence

\[
Var(z_x | L' < z_x \leq U') = E \left[ z_x^2 | L' < z_x \leq U' \right] - \left\{ E \left[ z_x | L' < z_x \leq U' \right] \right\}^2
\]

\[
= 1 - \left[ \frac{U'\phi(U') - L'\phi(L')}{\Phi(U') - \Phi(L')} \right] - \left[ \frac{\phi(U') - \phi(L')}{\Phi(U') - \Phi(L')} \right]^2
\]

Consequently

\[
Var(x | Q) = \sigma_x^2 Var(z_x | L' < z_x \leq U')
\]

An expression for the truncated covariance \( cov(x, y | Q) \) can be solved using

\[
E \left[ \left( x - \mu_{x|Q} \right) \left( y - \mu_{y|Q} \right) \bigg| Q \right]
\]

When \( P_\Phi(Q') = \Phi(U') - \Phi(L') \) and \( P(Q) = Pr\{Q\} \) under Normal(\( \mu, \sigma^2 \)) are measured, then the following decomposition can be used

\[
E \left[ \left( x - \mu_{x|Q} \right) \left( y - \mu_{y|Q} \right) \bigg| Q \right] = E \left[ \left( x - \mu_{x|Q} \right) \left( y - \mu_{y|Q} \right) \bigg| Q \right] - \left( \mu_{x} - \mu_{z|Q} \right) \left( \mu_{y} - \mu_{y|Q} \right)
\]

Now

\[
\mu_{x|Q} = E(x | Q) = \frac{1}{P(Q)} \int_{\mathbb{R}} 1_{Q} x dF(x), x \sim NormalF
\]

\[
= \frac{1}{P_\Phi(Q')} \int_{L'}^{U'} (\mu_x + \sigma_x z_x) d\Phi(z_x)
\]

\[
= \mu_x + \sigma_x \frac{1}{P_\Phi(Q')} \int_{L'}^{U'} d\phi(z_x)
\]

\[
= \mu_x - \sigma_x \left[ \frac{\phi(U') - \phi(L')}{\Phi(U') - \Phi(L')} \right]
\]

Also

\[
\mu_{y|Q} = \int_{\mathbb{R}} y dF(y | Q), y \sim NormalF
\]

\[
= \mu_y + \rho \sigma_y \frac{1}{\Phi(Q')} \int_{L'}^{U'} z_x d\Phi(z_x) + \sigma_y \sqrt{1 - \rho^2} \int_{\mathbb{R}} z_y d\Phi(z_y)
\]

\[
= \mu_y - \rho \sigma_y \left[ \frac{\phi(U') - \phi(L')}{\Phi(U') - \Phi(L')} \right]
\]
Next
\[
E \left( (x - \mu_x) (y - \mu_y) \mid Q \right)
\]
\[
= \frac{1}{P(Q)} \int_{\mathbb{R}^2} 1_Q (x - \mu_x) (y - \mu_y) dF(x, y), \quad (x, y)^T \sim \text{Normal}_2 F
\]
\[
= \frac{1}{P(Q)} \int_{\mathbb{R}^2} \int_{L}^{U} (x - \mu_x) (y - \mu_y) f(x, y) dx dy
\]
\[
= \frac{1}{P_\Phi(Q')} \int_{L'}^{U'} \int_{L'}^{U'} \sigma_x z_x \left( \rho \sigma_y z_x + \sigma_y \sqrt{1 - \rho^2} z_y \right) \phi_x(z_x) \phi_y(z_y) dz_x dz_y
\]
\[
= \frac{\rho \sigma_x \sigma_y}{P_\Phi(Q')} \int_{L'}^{U'} z_x^2 \phi_x(z_x) dz_x
\]
\[
= \rho \sigma_x \sigma_y \left[ \text{Var}(z_x \mid Q') + \{E[z_x \mid Q']\}^2 \right]
\]  
(26)

Hence
\[
E \left( (x - \mu_x \mid Q) (y - \mu_y \mid Q) \right) = \rho \sigma_x \sigma_y \left[ \text{Var}(z_x \mid Q') + \{E[z_x \mid Q']\}^2 \right] - \rho \sigma_x \sigma_y \left[ \frac{\phi(U') - \phi(L')}{\Phi(U') - \Phi(L')} \right]^2
\]
\[
= \rho \sigma_x \sigma_y \text{Var}(z_x \mid Q') = \rho \frac{\sigma_y}{\sigma_x} \text{Var}(x \mid Q)
\]  
(27)

Finally, since
\[
\text{Var}(y \mid Q) = \text{Var} \left( \rho \sigma_y z_x + \sigma_y \sqrt{1 - \rho^2} z_y \mid Q \right)
\]
\[
= \rho^2 \frac{\sigma_y^2}{\sigma_x^2} \text{Var}(x \mid Q) + \sigma_y^2 \left( 1 - \rho^2 \right) \text{Var}(z_y \mid Q)
\]
\[
= \rho^2 \frac{\sigma_y^2}{\sigma_x^2} \text{Var}(x \mid Q) + \sigma_y^2 \left( 1 - \rho^2 \right)
\]  
(28)

It should be noticed that \(z_y\) is not restricted. So finally
\[
\rho Q = \frac{\rho}{\sqrt{\rho^2 + \sigma_y^2 \left( 1 - \rho^2 \right) \frac{\sigma_y^2}{\text{Var}(x \mid Q)}}}
\]  
(29)

Appendix A.2. Doubly truncated log-normal distribution

Now, assume that the random vector \( \mathbf{u} = (x, y)^T = (\ln X, \ln Y)^T \) is jointly normally distributed as
\[
\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \ln X \\ \ln Y \end{pmatrix} \sim N_2 \left( \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \sigma_x \sigma_y \rho \\ \sigma_x \sigma_y \rho & \sigma_y^2 \end{pmatrix} \right)
\]  
(30)

The correlation of the doubly truncated log-normal distribution can be computed using
\[
\rho Q = \frac{\text{cov}(X, Y \mid Q)}{\sqrt{\text{Var}(X \mid Q)} \sqrt{\text{Var}(Y \mid Q)}}
\]
where \( Q \) is defined as in bivariate normal case, \((Q = (L, U))\), and

\[
x = \mu_x + \sigma_x z_x, \quad z_x \sim N(0, 1)
\]

\[
y = \mu_y + \rho \sigma_y z_x + \sigma_y \sqrt{1 - \rho^2} z_y, \quad z_y \sim N(0, 1), \quad z_x \perp z_y
\]

It should be noted that the Moment generating function, \( M_z(t) \) of a standard normal variate \( z \) is

\[
M_z(t) = E \left( e^{tz} \right) = \frac{1}{\sqrt{2\pi}} \int_R e^{t(z - \frac{1}{2} z^2)} dz
\]

\[
= \frac{1}{\sqrt{2\pi}} \int_R e^{-\frac{1}{2}(t^2 - 2tz + z^2) + \frac{1}{2}t^2} dz
\]

\[
= e^{\frac{1}{2}t^2} \frac{1}{\sqrt{2\pi}} \int_R e^{-\frac{1}{2}(z-t)^2} dz
\]

\[
= e^{\frac{1}{2}t^2} \quad \text{(31)}
\]

since \( \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} \) is the pdf of a Normal\((t, 1)\) distribution, so that

\[
\frac{1}{\sqrt{2\pi}} \int_R e^{-\frac{1}{2}(z-t)^2} dz = 1.
\]

Once we note that

\[
x = \mu + \sigma z,
\]

then \( M_x(t) \) is

\[
M_x(t) = E \left( e^{tx} \right) = E \left( e^{t(\mu + \sigma z)} \right)
\]

\[
= e^{t\mu} E \left( e^{t\sigma z} \right) = e^{t\mu} M_z \left( t\sigma \right) = e^{t\mu + \frac{1}{2}t^2\sigma^2} \quad \text{(32)}
\]

It should be noticed that for \( \log X = x \sim \text{Normal} \left( \mu, \sigma^2 \right) \), and then \( M_x(t) \) is

\[
E \left( X^k \right) = E \left( e^{kx} \right) = M_x \left( k \right) \quad \text{(33)}
\]

After all

\[
\text{Var} \left( X \mid Q \right) = E \left( X^2 \right| Q \right) - \left[ E \left( X \right| Q \right)^2
\]

Now

\[
E \left( X \mid Q \right) = E \left( e^x \mid Q \right) = e^{\mu_x} E \left( e^{\sigma_x z_x} \mid Q' \right)
\]

\[
= \frac{e^{\mu_x + \frac{\sigma_x^2}{2}}}{\Phi(Q')} \int_{L_2}^{U_2} d\Phi(0, 1)
\]

\[
= e^{\mu_x + \frac{\sigma_x^2}{2}} \left[ \Phi(U_2) - \Phi(L_2) \right] \left[ \Phi(U_1) - \Phi(L_1) \right] \quad \text{(34)}
\]

and

\[
E \left( X^2 \mid Q \right) = E \left( e^{2x} \mid Q \right) = e^{2\mu_x} E \left( e^{2\sigma_x z_x} \mid Q' \right)
\]

\[
= \frac{e^{2\mu_x + 2\sigma_x^2}}{\Phi(Q')} \int_{L_3}^{U_3} d\Phi(0, 1)
\]

\[
= e^{2\mu_x + 2\sigma_x^2} \left[ \Phi(U_3) - \Phi(L_3) \right] \left[ \Phi(U_1) - \Phi(L_1) \right] \quad \text{(35)}
\]

As can be seen, Equations (34), (35) can be rewritten as a function of \( k \)

where \( k \) indicates the level of moment function. The generalized equation is
presented in Equation (7), assuming $T_L = T_U = 0$. Similarly formulas of $U_3$, $L_2$, $U_3$, and $L_3$ presented in (37) can be converted to formulas presented in Equation (8).

Also

$$E(XY|Q) = E(e^{x+y}|Q)$$

$$= e^{\mu_x+\mu_y}E(e^{(\sigma_x+\rho\sigma_y)z_x+\sigma_y\sqrt{1-\rho^2}z_y}|Q')$$

$$= e^{\mu_x+\mu_y}E(e^{(\sigma_x+\rho\sigma_y)z_x}|Q')E(e^{\sigma_y\sqrt{1-\rho^2}z_y}|Q')$$

$$= e^{\mu_x+\mu_y+\frac{1}{2}[(\sigma_x+\rho\sigma_y)^2+\sigma_y^2(1-\rho^2)]}E(U_4)\Phi(L_4)=\Phi(U_1)\Phi(L_1)$$

(36)

where

$$U_1 = \frac{\ln U - \mu_x}{\sigma_x}; \quad L_1 = \frac{\ln L - \mu_x}{\sigma_x}$$

$$U_2 = \frac{\ln U - (\mu_x + \sigma_x^2)}{\sigma_x}; \quad L_2 = \frac{\ln L - (\mu_x + \sigma_x^2)}{\sigma_x}$$

$$U_3 = \frac{\ln U - (\mu_x + 2\sigma_x^2)}{\sigma_x}; \quad L_3 = \frac{\ln L - (\mu_x + 2\sigma_x^2)}{\sigma_x}$$

$$U_4 = \frac{\ln U - (\mu_x + \sigma_x^2 + \sigma_x\sigma_y\rho)}{\sigma_x}; \quad L_4 = \frac{\ln L - (\mu_x + \sigma_x^2 + \sigma_x\sigma_y\rho)}{\sigma_x}$$

(37)

The Equation (36) is equal with the Equation (12), when $T_L = T_U = 0$.

Also

$$E(Y|Q) = E(e^y|Q)$$

$$= e^{\mu_y}E(e^{\rho\sigma_y z_x+\sigma_y\sqrt{1-\rho^2}z_y}|Q')$$

$$= e^{\mu_y+\frac{1}{2}\sigma_y^2(1-\rho^2)}E(e^{\rho\sigma_y z_x}|Q')$$

$$= e^{\mu_y+\frac{1}{2}\sigma_y^2}\Phi(U_3)\Phi(L_3)$$

(38)

and

$$E(Y^2|Q) = E(e^{2y}|Q)$$

$$= e^{2\mu_y+2\sigma_y^2}\Phi(U_6)\Phi(L_6)$$

(39)

where

$$U_5 = \frac{\ln U - (\mu_x + \rho\sigma_x\sigma_y)}{\sigma_x}; \quad L_5 = \frac{\ln L - (\mu_x + \rho\sigma_x\sigma_y)}{\sigma_x}$$

$$U_6 = \frac{\ln U - (\mu_x + 2\rho\sigma_x\sigma_y)}{\sigma_x}; \quad L_6 = \frac{\ln L - (\mu_x + 2\rho\sigma_x\sigma_y)}{\sigma_x}$$

(40)

Similar with the moments of $X|Q$, the moments of $Y|Q$ can be merged to single function presented in Equation (9), where $T_L = T_U = 0$. As well, the formulas $U_5$, $L_5$, $U_6$ and $L_6$ presented in Equation (40) can be rewritten to the form presented in Equation (10).

54
Appendix A.3. Partially censored and truncated log-normal distribution

The group average $\bar{x}_T$ can be calculated when averages $\bar{x}_1$ and $\bar{x}_2$ and number of observations $n_1$ and $n_2$ of the two subgroups are known. The group average formula can be written using proportions of observations as weights:

$$\bar{x}_T = \frac{n_1}{n_1 + n_2}\bar{x}_1 + \frac{n_2}{n_1 + n_2}\bar{x}_2$$  \hspace{1cm} (41)

This is also true for the expected values of the squared observations.

Similarly the $E(X^k|Q,D)$ can be calculated using tail group averages and tail group proportions $T_L$ and $T_U$. In case of censoring, or partial censoring, the tail group averages $L^k$ and $U^k$ are equal with the boundary values $e^L$ and $e^K$, or in case when $k = 2$, squares of these.

$$E(X^k|Q,D) = T_L L^k + T_U U^k + T_C e^{k\mu_y + \frac{1}{2}k\sigma_y^2} \left[ \frac{\Phi(U_2,k) - \Phi(L_{2,k})}{\Phi(U_1) - \Phi(L_1)} \right]$$  \hspace{1cm} (42)

For the $E(Y^k|Q,D)$ the calculation follows the calculation of the $E(X^k|Q,D)$, except that the averages of censored or partially censored tails should be calculated. Then

$$E(Y^k|Q,D) = T_L e^{k\mu_y + \frac{1}{2}k\sigma_y^2} \left[ \frac{\Phi(L_{3,k}) - \Phi(-\infty)}{\Phi(L_1) - \Phi(-\infty)} \right]$$

$$+ T_C e^{k\mu_y + \frac{1}{2}k\sigma_y^2} \left[ \frac{\Phi(U_{3,k}) - \Phi(L_{3,k})}{\Phi(U_1) - \Phi(L_1)} \right]$$

$$+ T_U e^{k\mu_y + \frac{1}{2}k\sigma_y^2} \left\{ T_L \left[ \frac{\Phi(L_{3,k})}{\Phi(L_1)} \right] + T_C \left[ \frac{\Phi(U_{3,k}) - \Phi(L_{3,k})}{\Phi(U_1) - \Phi(L_1)} \right] \right\}$$  \hspace{1cm} (43)

Again the proportions are used as weights in the calculation of $E(XY|Q,D)$. Then

$$E(XY|Q,D) = T_L E(XY|Q_L,D_0) + T_C E(XY|Q,D_0)$$

$$+ T_U E(XY|Q_U,D_0)$$  \hspace{1cm} (44)

where $E(XY|Q_L,D_0)$ and $E(XY|Q_U,D_0)$ are expected values of the product $XY$ in left and right tails respectively and $Q_L = (0,L)$, $Q_U = (U,\infty)$ and $D_0 = (0,0)$.

Because in censored tails $E(X|Q_L,D_0) = L$ or $E(X|Q_U,D_0) = U$ also the correlation between $X$ and $Y$ in censored tails is zero. Then

$$E(XY|Q_L,D_0) = L \cdot E(Y^1|Q_L,D_0)$$  \hspace{1cm} (45)
and then the Equation (44) can be rewritten as

\[
E(XY|Q, D) = T_L[L \cdot E(Y^1|Q_L, D_0)] + T_U[Y \cdot E(Y^1|Q_U, D_0)]
\]

Using the Equation (38), the Equation of \(E(XY|Q, D)\) can be rewritten as

\[
E(XY|Q, D) = T_LLe^{\mu_y + \frac{1}{2} \sigma_y^2} \left[ \frac{\Phi(L_{3,1}) - \Phi(-\infty)}{\Phi(L_1) - \Phi(-\infty)} \right] + T_Ue^{\mu_y + \frac{1}{2} \sigma_y^2} \left[ \frac{\Phi(\infty) - \Phi(U_{3,1})}{\Phi(\infty) - \Phi(U_1)} \right] \\
+ \left\{ T_CE^{\mu_y + \frac{1}{2} \sigma_y^2} \left[ (\sigma_y + \rho \sigma_x)^2 + \sigma_x^2 (1-\rho^2) \right] \right. \\
+ \left. \frac{1 - \Phi(U_{3,1})}{1 - \Phi(U_1)} \right\}
\]

\[
e^{\mu_y + \frac{1}{2} \sigma_y^2} \left\{ T_LLe^{\mu_x + \frac{1}{2} \sigma_x^2} \Phi(L_{3,1}) \left[ \Phi(L_1) - \Phi(-\infty) \right] + T_Ue^{\mu_x + \frac{1}{2} \sigma_x^2} \left[ \frac{1 - \Phi(U_{3,1})}{1 - \Phi(U_1)} \right] \right. \\
+ \left. \frac{1 - \Phi(U_{3,1})}{1 - \Phi(U_1)} \right\}
\]

\[
= \frac{T_CEe^{\mu_x + \frac{1}{2} \sigma_x^2} \left[ (\sigma_x + \rho \sigma_y)^2 + \sigma_y^2 (1-\rho^2) \right] \left[ \Phi(U_4) - \Phi(L_4) \right]}{\Phi(U_1) - \Phi(L_1)}
\]

\[
+ T_CEe^{\mu_x + \frac{1}{2} \sigma_x^2} \left[ \frac{1 - \Phi(U_{3,1})}{1 - \Phi(U_1)} \right] \left\{ T_LLe^{\mu_y + \frac{1}{2} \sigma_y^2} \Phi(L_{3,1}) \left[ \Phi(L_1) - \Phi(-\infty) \right] + T_Ue^{\mu_y + \frac{1}{2} \sigma_y^2} \left[ \frac{1 - \Phi(U_{3,1})}{1 - \Phi(U_1)} \right] \right. \\
+ \left. \frac{1 - \Phi(U_{3,1})}{1 - \Phi(U_1)} \right\}
\]

\[
+ T_CEe^{\mu_x + \frac{1}{2} \sigma_x^2 + \rho \sigma_x \sigma_y} \left[ \frac{\Phi(U_4) - \Phi(L_4)}{\Phi(U_1) - \Phi(L_1)} \right]
\]

Now all the moments used in the correlation Equation (3) have been derived.

### Appendix B. Correlation when \(L \to 0\) and \(U \to \infty\)

When \(L \to 0\) and \(U \to \infty\), also \(T_L \to 0\), \(T_U \to 0\) and \(T_C \to 1\). Now

\[
Var(X|Q) = E\left(X^2|Q\right) - [E(X|Q)]^2
\]

\[
= e^{2\mu_x + 2\sigma_x^2} \left[ \Phi(U_3) - \Phi(L_3) \right]^2 - e^{2\mu_x + \sigma_x^2} \left[ \Phi(U_3) - \Phi(L_2) \right]^2
\]

\[
= e^{2\mu_x + \sigma_x^2} \left( e^{\sigma_x^2} \frac{\Phi(U_3) - \Phi(L_3)}{\Phi(U_1) - \Phi(L_1)} \right) - \left[ \frac{\Phi(U_3) - \Phi(L_2)}{\Phi(U_1) - \Phi(L_1)} \right]^2
\]

\[
\approx \frac{2\mu_x + \sigma_x^2}{U \to \infty, L \to 0} \left( e^{\sigma_x^2} - 1 \right)
\]

similarly

\[
Var(Y|Q) = E\left(Y^2|Q\right) - [E(Y|Q)]^2
\]

\[
= e^{2\mu_y + \sigma_y^2} \left( e^{\sigma_y^2} \frac{\Phi(U_5) - \Phi(L_5)}{\Phi(U_1) - \Phi(L_1)} \right) - \left[ \frac{\Phi(U_6) - \Phi(L_6)}{\Phi(U_1) - \Phi(L_1)} \right]^2
\]

\[
\approx \frac{2\mu_y + \sigma_y^2}{U \to \infty, L \to 0} \left( e^{\sigma_y^2} - 1 \right)
\]
and
\[ \text{Cov}(X, Y|Q) = E(XY|Q) - E(X|Q)E(Y|Q) \]
\[ = e^{\mu_x + \mu_y + \frac{1}{2}(\sigma_x^2 + \sigma_y^2)} \left\{ e^{\rho \sigma_x \sigma_y} \frac{\Phi(U_4) - \Phi(L_4)}{\Phi(U_1) - \Phi(L_1)} \right\} \]
\[ - \left\{ \frac{\Phi(U_2) - \Phi(L_2)}{\Phi(U_1) - \Phi(L_1)} \right\} \]
\[ \div e^{\mu_x + \mu_y + \frac{1}{2}(\sigma_x^2 + \sigma_y^2)} \Phi_1 \]
\[ \xrightarrow{U \to \infty} e^{\mu_x + \mu_y + \frac{1}{2}(\sigma_x^2 + \sigma_y^2)} (e^{\rho \sigma_x \sigma_y} - 1) \quad (50) \]

Finally we see that
\[ \rho Q \xrightarrow{U \to \infty} \frac{(e^{\rho \sigma_x \sigma_y} - 1)}{\sqrt{(e^{\sigma_x^2} - 1)} \sqrt{(e^{\sigma_y^2} - 1)}} \quad (51) \]

**Appendix C. If \( \rho = 0 \), then \( \rho_{Q,D} = 0 \)**

The \( \rho_{Q,D} = 0 \) only if
\[ \text{Cov}(X, Y|Q, D) = 0 \]
\[ \iff E(XY|Q, D) - E(X|Q, D)E(Y|Q, D) = 0 \]
\[ \iff E(XY|Q, D) = E(X|Q, D)E(Y|Q, D) \quad (52) \]

When \( \rho = 0 \), functions \( U \) and \( L \) simplify to formulas
\[ U_1 = U_3 = \frac{\ln U - \mu_x}{\sigma_x} ; \quad L_1 = L_3 = \frac{\ln L - \mu_x}{\sigma_x} \]
\[ U_2 = U_4 = \frac{\ln U - (\mu_x + \sigma_x^2)}{\sigma_x^2} ; \quad L_2 = L_4 = \frac{\ln L - (\mu_x + \sigma_x^2)}{\sigma_x^2} \quad (53) \]

Now the \( E(XY|Q, D, \rho = 0) \) can be written as
\[ E(XY|Q, D, \rho = 0) \]
\[ = e^{\mu_x + \frac{1}{2} \sigma_y^2} \left\{ T_L L \left[ \frac{\Phi(L_3)}{\Phi(L_1)} \right] + T_U U \left[ \frac{1 - \Phi(U_3)}{1 - \Phi(U_1)} \right] \right\} \]
\[ + \left\{ T_C e^{\mu_x + \mu_y + \frac{1}{2}(\sigma_x^2 + \sigma_y^2)} \frac{\Phi(U_4) - \Phi(L_4)}{\Phi(U_1) - \Phi(L_1)} \right\} \]
\[ = e^{\mu_y + \frac{1}{2} \sigma_y^2} (T_L L + T_U U) \]
\[ + \left\{ T_C e^{\mu_x + \mu_y + \frac{1}{2}(\sigma_x^2 + \sigma_y^2)} \frac{\Phi(U_2) - \Phi(L_2)}{\Phi(U_1) - \Phi(L_1)} \right\} \]
\[ = e^{\mu_y + \frac{1}{2} \sigma_y^2} \left\{ (T_L L + T_U U) \right\} \]
\[ + T_C e^{\mu_x + \frac{1}{2} \sigma_x^2} \left[ \frac{\Phi(U_2) - \Phi(L_2)}{\Phi(U_1) - \Phi(L_1)} \right] \quad (54) \]

The first moment of variables \( X \) and \( Y \) get a forms
\[ E(X|Q, D, \rho = 0) \]
\[ = (T_L L + T_U U) + T_C e^{\mu_x + \frac{1}{2} \sigma_x^2} \left[ \frac{\Phi(U_2) - \Phi(L_2)}{\Phi(U_1) - \Phi(L_1)} \right] \quad (55) \]
and

\[
E(Y|Q, D, \rho = 0) = e^{\mu_y + \frac{1}{2} \sigma_Y^2} \left\{ T_L \left[ \frac{\Phi(3_1)}{\Phi(L_1)} \right] + T_C \left[ \frac{\Phi(U_3, k) - \Phi(3_1)}{\Phi(U_1) - \Phi(L_1)} \right] + T_U \left[ \frac{1 - \Phi(U_3, k)}{1 - \Phi(U_1)} \right] \right\} = e^{\mu_y + \frac{1}{2} \sigma_Y^2} \left( TL + TC + TU \right)_{=1}
\]

(56)

Now it can be seen that

\[
E(XY|Q, D, \rho = 0) = E(X|Q, D, \rho = 0)E(Y|Q, D, \rho = 0) \quad (57)
\]

is true with all \( Q \) and \( D \). So when the correlation of the unrestricted bivariate log-normal distribution equals zero, the correlation of the bounded distribution, \( \rho_{Q,D} \), is also zero.

Copyright ©2016 Incisive Risk Information (IP) Ltd.

Reprinted with permission.
Modeling the Current Loan-to-Value Structure of Mortgage Pools without Loan Specific Data

Peter Palmroos

Financial Supervisory Authority, PO Box 103, 00101 Helsinki, Finland

Abstract

This study presents a method for approximating the Current Loan-to-Value (CLTV) and remaining principal structures of heterogeneous mortgage loan pools. The method uses widely available public aggregate loan data instead of loan-specific data, the availability of which is highly restricted outside lenders. The model is based on a simple matrix equation for the pool’s in- and outflows, and on a division of the pool into multiple homogeneous cohorts. The estimated structure is compared with the true structure as reported by Finnish banks. This comparison indicates the method is accurate. The resulting CLTV and remaining principal structures help to improve the accuracy of mortgage loan credit risk models and enable a reliable approximation of the pools’ Expected Loss-Given-Defaults (ELGD).

1 Introduction

The Current Loan-to-Value (CLTV) has been found to be the most important factor in the modeling of loan loss severities and Loss Given Defaults (LGD) for the mortgage loan pools, as can be seen eg, from the papers of Park and Bang (2014) and Qi and Yang (2009).

Unfortunately the availability of the CLTV data is very restricted outside the lenders, and the other loan-specific properties needed for CLTV calcula-
tions, such as the time of origination and remaining principal, are not publicly available either. Instead of CLTV Deng et al. (2000) and Lekkas et al. (1993) used the Loan-to-Value (LTV) at origination in their models. Even though both papers reported good results, the LTV ignores the effects of amortizations and developments in collateral values between origination and default, and thus the use of LTV instead of CLTV may reduce the accuracy of these models.

This paper presents a simple but accurate matrix difference equation-based model to approximate the CLTV and remaining total principal structures of mortgage loan pools, when loan-specific data is not available.

Before the model can be used, the otherwise heterogeneous pool needs to be divided into numerous homogeneous loan cohorts, using LTV at origination and age of the loan in classification. The calculations were done using the publicly available time series of total amounts of originated mortgages and the sizes of mortgage loan pools.

The model takes into account the recurrent amortizations of loans, prepayments, defaults and developments in collateral values. Depending on the data available, both the amounts and numbers of loans per cohort can be estimated. The model allows one not only to replicate the current and previous pool structures, but also to forecast the development of structures in the future, and eg, to simulate the effects of LTV regulations to the sizes of cohorts and thus also to the credit risks of the banks.

The modeled heterogeneous mortgage loan pool structure makes it possible for researchers outside creditors to analyze how changes in markets or regulation affect the dynamics of mortgage loan pools over time. This makes it possible to compare the credit risks of different pool structures, eg risks with and without implementation of the LTV regulation.

Knowing the structures makes it possible to approximate the credit risks of loans exceeding the selected CLTV levels and to estimate the Expected Loss Given Default (ELGD) levels of both the whole pool and single cohorts.

The estimated heterogeneous pool structures can also be used to improve the accuracy of existing mortgage loan credit risk models.

As has been presented in the ESRB (2014) publication, LTV caps are used as...
a part of larger macro-prudential tool set to prevent excessive credit growth, and to limit substantial risk accumulation to financial system, during credit booms. To recognize the effects of LTV regulation, impact analysis needs to be performed.

The model presented in this paper has originally been developed to support the impact analysis of regulation restricting LTV levels in Finnish banking sector. The performed analysis concentrated on credit risks and structural effects on mortgage loan pools. As a part of these analysis stochastic simulations were used to estimate market reactions, and responses of potential borrowers to different LTV limits. This information was used to approximate changes in pool structure and effects on risks over time. The dynamic analysis were needed to recognize how long it takes before the impacts of made policy actions can be observed in supervisory reporting, and how long it takes before the impact will be fully realized. Results of these analysis help in anticipation and scheduling of timely regulatory actions.

In the literature, the modeling of Probability-of-Default (PD) has been studied much more than the modeling of LGD, as Park and Bang (2014), Qi and Yang (2009) and Allen et al. (2004) note in their papers. The first two papers also report that most of the LGD research has concentrated on wholesale loans. This seems to be true not only in the LGD research, but the focus of credit risk modeling has been more on wholesale than on retail loans (see Dermine and Neto de Carvalho, 2006). One pragmatic explanation for this may be that properties of single loans can only be observed by the lenders, which noticeably restricts the public availability of data.

The Internal Ratings-Based Approach (IRBA) assumes the modeled mortgage loan pools to be homogeneous and the number of small loans in pools to be large (see BIS, 2001a, pp. 55). Despite this Deng et al. (2000) highlight the importance of the heterogeneity of loans and borrowers in the risk modeling of mortgage loans. Allen et al. (2004) note that heterogeneity is one reason why models for wholesale loans cannot be used as such for mortgage loans. The method presented in this paper in some cases enables the using of wholesale credit risk models also with mortgage loan pools.

This paper is organized as follows. Section 2 introduces the mathematical setup. The model is introduced here in a simple difference equation format. In Section 2.1 the model is developed to include cohorts and to adapt a vector approach. Section 2.2 adds a new dimension to the model: the LTV distribution at origination. 2.3 presents the calculation of CLTVs of cohorts. Comparisons between estimated results and empirical values reported by Finnish banks are
presented in Section 3. The last section concludes.

2 Modeling of the pool structure

The model is similar in principle to the simple warehouse models used in logistics: the loan pool can be seen as a reservoir, with current period originations as inflows, and amortizations, prepayments and defaults as outflows. The dynamics can then be written as

\[ L_{t-1}^* + l_t^* - a_t^* - u_t^* = L_t^* \]  \hspace{0.75cm} (1)

where \( L_{t-1}^* \) and \( L_t^* \) are the mortgage loan pools at the end of periods \( t-1 \) and \( t \), \( l_t^* \) is the amount of new loans originated during period \( t \), \( a_t^* \) is the amount of the planned amortizations and \( u_t^* \) is the amount of prepayments and mortgages gone to default. For most applications there is no need to separate the effects of defaults from prepayments; they have similar effects on the structure of the pool, and thus both can be included in \( u_t^* \). The asterisk in subscript is used to help to separate the scalar variables, used in the Equations (1) and (2), from the elements of vectors in the later equations.

Values of \( L_t^* \) and \( l_t^* \) are part of the regular reporting of banks and can be found from public sources. But the values of \( a_t^* \) and \( u_t^* \) are not as easily available. One can obtain the sum \( a_t^* + u_t^* \) when all the other variables are known, albeit without the separate values of \( a_t^* \) and \( u_t^* \). Reliable estimation of \( a_t^* \) requires information on the remaining principals and ages of loans in the pool. The method for solving the value of \( a_t^* \) is presented in Section 2.1, when the time-of-origination -dimension will be added to the model.

Equation (1) can be rewritten, for later purposes, using multipliers \( \alpha_t^* \) and \( \upsilon_t^* \) instead of \( a_t^* \) and \( u_t^* \), as

\[ L_{t-1}^* + l_t^* - \alpha_t^* L_{t-1}^* - (1 - \alpha_t^*) \upsilon_t^* L_{t-1}^* = L_t^* \]

\[ \Leftrightarrow (1 - \alpha_t^* - (1 - \alpha_t^*) \upsilon_t^*) L_{t-1}^* + l_t^* = L_t^* \]  \hspace{0.75cm} (2)

where \( \alpha_t^* \) is the proportion of planned amortizations and \( \upsilon_t^* \) is the proportion of prepayments and defaults after the amortizations.

The structures and dynamics of the pool can be calculated for both the numbers of loans and loan amounts. When the numbers of loans are used, in addition to prepayments, only the last planned amortization reduces the number of loans in the pool. Thus the number of loans behaves like a set of bullet loans, even when the principals of loans are amortized.

In this paper all comparisons between estimated and reported structures are
presented using only remaining principals, but in practice most applications require estimations of both the loan amount and the number of loan structures.

2.1 Estimation of the pool structure using cohorts

Before the model can be converted to use loan cohorts and vector presentation, some assumptions are needed. First, all new loans have the same original maturity. Secondly, amortization frequency is assumed to be equal to the used period length. And finally, loans can be prepaid, and defaults can be observed only in connection with amortization, i.e., at the end of the period. Restrictions due to the first assumption can be bypassed by dividing the pool into several subpools with similar maturities. The effects of the next two can be reduced by shortening the length of used period.

Solving the structures of the current pool is a recursive and iterative process. The used method needs a starting pool with $m + 1$ cohorts, where $m$ is the maturity of loans in number of periods. An excess cohort is needed for the time dynamics and gets value zero. Especially when the time series of pool size development and new originated loans are short, the authenticity of the starting structure is important. The longer the time series, the less sensitive the results are to the changes in the starting pool structure. When time series lengths are one to two times the used maturity, the estimated pool structures are relatively robust to the used starting pool structure. This sensitivity should always be tested with available time series, e.g., using different starting pool structures.

A method based on averaging and available time series to calculate an approximation for the starting pool will be presented in Appendix A. An alternative method is to assume that the starting pool follows the structure of the current pool. In this case the structure can be calculated recursively, using the estimated structure as a new starting pool for the next estimation round, until the resulting structure is stable.

The lengths of the vectors used in the calculations are $m + 1$, equaling the number of time cohorts. The origination and observation periods of the loans included in the cohorts needs to be specified. Thus in all the following formulas, the observation period will be indicated at the superscript and origination period at the subscript. One exception for this is the subscript $RS$, which indicates a right shift vector operation. The right shift operation drops out the rightmost component of the horizontal vector, and adds new component.

\footnote{Like the right logical shift in bitwise operations of binary numbers.}
equaling zero, in to the leftmost end of the vector. In this case the rightmost component equals zero, and thus the sum of the components stays unchanged.

The Hadamard product is used in the component-by-component multiplication of vectors and matrices. This can be written using a vector presentation as

\[ x \circ y = [x_1y_1 \ x_2y_2 \ \ldots \ x_ny_n]' \]

where \( x \) and \( y \) are vectors of equal size.

The vector presentation of the Equation (2) can then be written as

\[ L_{t-1}^{-1}RS + l^t - \alpha^t \circ L_{t-1}^{-1}RS - ((1 - \alpha^t) \circ \nu^t) \circ L_{t-1}^{-1}RS + l^t = L^t \]

\[ \Leftrightarrow (1 - (\alpha^t + (1 - \alpha^t) \circ \nu^t)) \circ L_{t-1}^{-1}RS + l^t = L^t \]

\[ \Leftrightarrow (1 - \gamma^t) \circ L_{t-1}^{-1}RS + l^t = L^t \]

where cohort vectors \( L_{t-1}^{-1}RS \) and \( L^t \) are the pool structures at the end of periods \( t-1 \) and \( t \). It should be noted that \( 1' L_{t-1}^{-1}RS = 1' L^t-1 \). Vector \( l^t \) includes the new loans originated during period \( t \). The outflow multiplier vector \( \gamma^t \) includes the effects of both amortizations \( \alpha^t \) and prepayments and defaults \( \nu^t \). 1 is the unit vector of length \( m+1 \).

The cohort vectors can be written as

\[ L^t = [l^t_t \ l^t_{t-1} \ \ldots \ l^t_{t-(m-1)} \ 0]' \]

\[ L_{t-1}^{-1}RS = [0 \ l^t_{t-1}^{-1} \ l^t_{t-2}^{-1} \ \ldots \ l^t_{t-m}^{-1}]' \]

\[ l^t = [l^t_t \ 0 \ \ldots \ 0]' \]

In all of these vectors \( l^t_{i-j} \) identifies the remaining total principal values per cohort. Again the subscript identifies the period of origination and the superscript the period of observation. For example \( l^t_{i-j} \) is the remaining amount of principals of loans originated at period \( t-i \) and observed at the end of period \( t-j \), ie, after \( i-j \) amortizations. In case \( i=j \), the remaining principal equals the amount of originated loans at the end of the origination period. Loans originated during the period \( t \) become the first component of the vector \( l^t \), and all other components are zeros.

The outflow multiplier vector \( \gamma^t \) and its factors can be written as

\[ \gamma^t = \alpha^t + (1 - \alpha^t) \circ \nu^t \]

\[ \alpha^t = [0 \ \alpha^t_{t-1} \ \ldots \ \alpha^t_{t-m}]' \]

\[ \nu^t = [0 \ \nu^t_{t-1} \ \ldots \ \nu^t_{t-m}]' \]

where the \( \alpha^t_{t-i} \) equals the rate between the amortization and remaining principal at time \( t \) for loans originated \( i \) periods ago. Vector \( \nu^t \) includes both
the effects of prepayments and defaults. The component \( v^t_{t-i} \) is a ratio where the sum of prepayments and defaults is divided by the remaining principal at the period \( t \) on mortgages of age \( i \). Due to the equality of single period and amortization frequencies, by definition, there cannot be prepayments before the first amortization, and thus the first component of \( \gamma^t \) is zero.

The value of \( \alpha^t_{t-i} \) depends on the type of amortization. The cohort specific amortization multipliers can be calculated using the formula

\[
\alpha^t_{t-i} = \frac{a^t_{t-i}}{p^t_{t-i}} = \begin{cases} 
0 & \text{when } i = 0 \\
\frac{1}{m-i+1} & \text{when } i \in \{1, \ldots, m\}, \text{ fixed amortization loans} \\
\frac{r_{t-i}}{(1+r_{t-1})^{m-i+1}} & \text{when } i \in \{1, \ldots, m\}, \text{ annuity loans}
\end{cases}
\]

where \( a^t_{t-i} \) equals the amortization at the end of period \( t \) for the loans granted at \( t-i \), \( p^t_{t-i} \) is the remaining principal preceding the amortization, and \( r_{t-1} \) is the interest rate used for accrued interests calculations including a customer margin. In all cases \( \alpha^t_{t-m} = 1 \), because the loan with maturity \( m \) at origination, will be totally repaid after \( m \) periods. Derivation of amortization formulas presented in Equation (6) can be found from Appendix B.

Using the Equations (1) and (2) and the values of vector \( \alpha^t \), one can solve the value of \( \upsilon^t \) and thus also the value \( u^t_* \).

If the distribution of \( v^t_{t-i} \) in vector \( v^t \) is known, it can be written as

\[
v^t = \begin{bmatrix} 0 & v^t_{t-1} & \ldots & v^t_{t-m} \end{bmatrix}^t \\
= \begin{bmatrix} 0 & \phi_{t-1} & \ldots & \phi_{t-m} \end{bmatrix}^t \xi^t_\star
\]

(7)

where the relative size of \( \phi_{t-i} \) indicates the proportion of prepayments and defaults dependent on the age of the loan, and variable \( \xi^t_\star \) is the only unknown. The absolute values of \( \phi_{t-i} \) do not affect the calculation, because the solving of \( \xi^t_\star \) scales the multipliers \( v^t_{t-i} \) in vector \( v^t \).

If the \( v^t_{t-i} \) distribution is unknown, ie, there is no empirical data or theory based approximation for it, the simplest approximation is to assume the prepayment and default ratios to be independent on the age of the loans. Thus vector \( v^t \) in Equation (7) can be written as

\[
v^t = \begin{bmatrix} 0 & 1 & 1 & \ldots & 1 \end{bmatrix}^t \upsilon^t_\star
\]

(8)

In Section 3 it can be seen that even this simple uniform distribution assumption gives results in line with the true data.

\footnote{A single payment includes two parts: amortization of the loan and accrued interest. In this paper only the amortizations have been included to calculations. In case of a fixed amortization loan the amortization stays the same over the time and is independent on interest rate. Contrary to that, the amortization of the annuity loan is dependent on interest rate and will change over the time.}
Once the components of vector $a^t$ are calculated, and $v^t$ structure is known, one can solve the value of $v^t$ and the structure of vector $L^t$. The keys for this solving process are the equivalences $1' L^t \equiv L^t_s$ and $1' L^{t-1} \equiv L^{t-1}_s$.

### 2.2 Using the original Loan to Value distribution

In case all the loans have been granted with the same LTV, the vector $L^t$ solved in Section 2.1, includes the whole structure required for the CLTV distribution estimation of the pool. Normally banks grant loans with multiple different LTVs, depending on the needs of customers. This section represents a method to include LTV distribution of granted loans to the model.

In a vector presentation all the loans in a single cohort have the same origination period. When the discretized distribution of LTVs of originated loans per period is known, the loans in cohort vector $L^t$ can be divided into smaller cohorts. These smaller cohorts have not only the same origination period but also the same original LTV.

The proportions of loans with the same LTV create a discrete distribution vector with $k$ components. This vector can be written as

$$\pi_{t-i} = [\pi_{t-i}^1, \pi_{t-i}^2, \ldots, \pi_{t-i}^k]$$  \hspace{1cm} (9)

where the proportion of loans, originated at $t-i$ and LTVs belonging to group $j$, are denoted by $\pi_{t-i}^j$. Because the $\pi$’s are proportions, the vector should fulfill the conditions $1' \pi_{t-i} = 1$ and $\pi_{t-i}^j \geq 0$ for all $j$. These $\pi_{t-i}^j$ vectors can be written as a matrix

$$\Pi^t = \begin{bmatrix} \pi_t & \pi_{t-1} & \ldots & \pi_{t-m} \\ \pi_1^t & \ldots & \pi_m^t \\ \vdots & \ddots & \vdots \\ \pi_1^k & \ldots & \pi_k^m \end{bmatrix}$$  \hspace{1cm} (10)

The cohort composition matrix $S^t$, where every cell includes the amount of loans per cohort, can be calculated by multiplying matrices $\Pi^t$ and $1'(L^t)'$ component by component.

$$S^t = \Pi^t \circ 1'(L^t)' = \begin{bmatrix} \pi_1^1 l_t^1 & \ldots & \pi_1^1 l_{t-m}^1 \\ \vdots & \ddots & \vdots \\ \pi_k^1 l_t^k & \ldots & \pi_k^k l_{t-m}^k \end{bmatrix}$$  \hspace{1cm} (11)

Again the condition $1'S^t1 \equiv L^t_s$ holds. The elements of $L^t$ and $L^t_s$ are the same already presented in Section 2.1.

---

Notice that the superscript of $\pi$ describes the LTV group, not the observation period.
In case the history of yearly LTV distributions is not available, one possibility is to use the latest one known, and assume that the discrete distribution of the LTV proportions has remained stable over the history, ie, $\pi_{t-i} = \pi$. In those cases the matrix $S^t$ can be calculated simply as

$$S^t = \pi(L^t)'$$

(12)

2.3 Current Loan to Values of the mortgage cohorts

The remaining total principals of every cohort have been solved using methods presented in Sections 2.1 and 2.2. This section presents a method to calculate CLTVs of these cohorts.

The method for approximating CLTV when original LTV and the development of the collateral value are known is trivial but surprisingly rarely used in mortgage literature. One example of this kind of calculation can be found from Nystöm and Skoglund (2006).

The CLTV of the loan at the end of period $t$, originated at period $t-i$ with original Loan-to-Value $LTV_{t-i}$, is denoted as $CLTV_{t-i}^t$, and can be calculated using the formula

$$CLTV_{t-i}^t = LTV_{t-i} \cdot \lambda_{t-i}^t$$

(13)

where $\lambda_{t-i}^t$ is the relative change in LTV between $t-i$ and $t$.

There are two factors affecting $\lambda_{t-i}^t$. First is the development of the collateral value. A method to approximate this development can be found eg, in the paper by Leow and Mues (2012, pp. 186). Secondly, in the case of amortizing loans, the loan principal decreases over time.

The multiplier $\lambda_{t-i}^t$ can then be calculated using formula

$$\lambda_{t-i}^t = \frac{p_{t-i} / p_{t-i}^c}{c_t / c_{t-i}} = \frac{p_{t-i} / p_{t-i}^c}{c_t / c_{t-i}}$$

(14)

where $p_{t-i}^c$ is the original principal of the loan granted at $t-i$, and $p_{t-i}^c$ is the remaining principal at time $t$ of the loan granted at $t-i$. Similarly $c_{t-i}$ and $c_t$ are the collateral values at the same periods, respectively.

When housing property is the only collateral, as in Leow and Mues (2012), the collateral value can be assumed to track developments in the housing price index (HPI). If any indices are available that describe the development of housing prices more accurately than HPI, eg local HPI in case of regionally restricted pool, those could be used to increase the accuracy of the model.

---

6For some applications also the number of loans in every cohort may be solved (calculation method explained in the last two paragraphs of Section 2).
The denominator $c^t/c^{t-i}$, can then be calculated using time series of the HPI, and the numerator, $p^t_{t-i}/p^{t-i}_{t-i}$, as

$$\frac{p^t_{t-i}}{p^{t-i}_{t-i}} = \begin{cases} \frac{m-i}{m} & \text{for the fixed amortization loans} \\ \prod_{h=0}^{i-1} (1 - \alpha_{t-i}^h) & \text{for the annuity loans} \end{cases}$$

(15)

where $\alpha_{t-i}^h$ are the historical values of $\alpha_{t-i}^t$ solved in Equation (6), and $m$ is the maturity of loans (see Section 2.1). Derivation of $p^t_{t-i}/p^{t-i}_{t-i}$ formulas can be found from Appendix C.

The solved multipliers can be written as a vector

$$\Lambda^t = \begin{bmatrix} \lambda^t_t & \lambda^t_{t-1} & \cdots & \lambda^t_{t-m} \end{bmatrix}'$$

(16)

The LTVs of the cohorts (see vector $\pi_{t-i}$ in Equation (9)) can be placed in a vector:

$$B = [b^1 \ldots b^k]'$$

(17)

where $b^j$ is the LTV of the classification group $j$. Depending on the application, the $b^j$ might be for example the upper limit or the mean of the LTV in classification group $j$. The LTV groups stay the same over the time and thus $B$ is time independent, ie vector of constants.

The matrix $V^t$ of CLTV’s of all cohorts at time $t$ can be calculated by vector multiplication:

$$V^t = B(\Lambda^t)’$$

(18)

This paper models the structure of the remaining pool, and thus it should be noticed that matrix $V^t$ includes CLTV values of loans remaining in the pool after amortizations, ie loans that have not defaulted.

The matrix presentation helps finding and matching of all the cohort specific information. The CLTV and remaining amount of loans belonging to the cohort originated in $t-i$ and LTV at origination belonging to group $j$, can be found from the same components of matrices $V^t$ and $S^t$, respectively.

3 Comparison between estimated and observed structures

The Financial Supervisory Authority of Finland asked the banks in Finland to report the age structure of mortgages in pools at the end of the year 2010. The reported data covers over 99% of Finnish mortgage loans originated by the banks.

7The CLTVs of defaulted loans can be calculated similarly, but $p^t_{t-i}/p^{t-i}_{t-i}$ should be replaced with the remaining principal of previous period, ie $p^{t-1}_{t-i}/p^{t-i}_{t-i}$.
The reported maturity structure is compared with the estimated structure calculated using mortgage loan stock and origination data from years 1993 to 2010 (see Figure 1b). The used maturity assumption at origination is 18 years, which is the rounded average of maturities reported at the publication of Federation of Finnish Financial Services (2013, pp. 29). In this paper it is assumed that the proportions of repayments or defaults are independent of the age of the loan, so that all of the prepayment multipliers in vector $\nu^t$ are equal. This assumption does not restrict changes in $v^t_{t-i}$ from period to period.

The result of the estimation process explained in Section 2.1 and the reported data are both presented in Figure 1a. The figure indicates how closely the estimated structure, scaled to sum to 100%, ie, $L^t(1'L)^{-1}$, follows the observed structure. The difference between these two curves is less than 0.5 percentage point in all periods except for the current year 2010. The larger value for 2010 estimate is due to two factors. First, the one year frequency of periods is longer than the actual mean amortization frequency of mortgage loans. The reported amount of originated loans is larger than the amount of the new mortgages in the loan stock data, where the same loans have already been partly amortized. Secondly, the data on originated loans include the bridge loans, ie, short term loans used in home swapping, paid back before the reported end-of-the-year stock data.

Besides the age of the mortgages, the FSA Finland asked banks to report the LTV distribution of mortgages originated during a shorter time span in 2010. The average maturity of new loans fluctuated between 17.4 to 18.8 years during the reported period.

Notice that the compared series are discrete. A line chart has been used instead of a histogram to show the differences more clearly.
In this reporting the LTV was divided into 31 buckets of equal 5% width. Due to the lack of data of LTV distribution history, the 2010 distribution is assumed to approximate the distributions of all the previous periods.

The calculations presented in Sections 2.2 and 2.3, give us two $19 \times 31$ matrices. The matrix $S^t$ indicates the remaining principal amount of loans per cohort, and the matrix $V^t$ the CLTVs of these same cohorts. The components of these matrices are presented in Figures 2a and b respectively. The components of $S^t$ have been scaled to sum to 100%, i.e., the values equal $S^t(1'S^t1)^{-1}$.

In the calculation of $V^t$ the HPI of Finland is used to approximate the value changes of collaterals and all loans are assumed to be annuity loans. At the end of 2012 approximately 84% of all mortgage loans were annuity loans and 10% were fixed amortization loans.

The estimated pool structure can be used in loan loss simulations and with existing models like those of Leow and Mues (2012) and the mortgage loan-specified modification of Jokivuolle and Peura (2003). The estimated structure and CLTV data can also be used for calculating the number of mortgages facing collateral value deficit conditional on a decrease in HPI, as well as in ELGD and Expected Loss (EL) simulations.

Figure 3a indicates the proportion of loans with $LGD > 0$ as a function of collateral value. The values presented in Figures 3a and b have been calculated using the matrices $S^t$ and $V^t$. From Figure 3a can be seen that if the collateral value decreases 10%, i.e., collateral value equals 90% of the current value, the portion of loans with $LGD > 0$ is 18.7%.

\(^{10}\)The development of HPI in Finland can be found at the website of Statistics Finland www.stat.fi.

\(^{11}\)For simplicity, in this paper the LGD is assumed to equal to the proportion of collateral deficit, and thus e.g., the realization costs of collateral have been omitted.
Differences between the practices, legislation and default-trigger events in the American and European mortgage loans markets are presented in a paper of Campbell (2013). Briefly, in both of these markets the CLTV is the most important explanatory variable for LGD. One difference between the US and European markets is in the CLTVs importance as a trigger event of default. In US markets when the level of CLTV rises above 100%, it is a potential trigger of a default. In Europe, where possibilities of personal bankruptcies are more strictly limited by legislation, the level of CLTV does not act as a default trigger. In European markets the main driver of defaults is the borrowers’ insolvency.

In case the insolvency is assumed to be the only trigger for default, loans with $CLTV < 100\%$, ie, $LGD = 0$, can also default. As can be seen from the Figure 3a, when the collateral value is 90%, 81.3% of all defaulted loans have $LGD = 0$. The solid line in Figure 3b shows the ELGD as a function of collateral value in this kind of environment.

In the pure US-like mortgage loan environment, where the value of collateral sinking below the amount of remaining loan can be assumed to be the only trigger event of default, defaults occur only if $LGD > 0$. The dotted line in Figure 3b shows ELGD in that kind of environment. The difference between the lines depends on whether or not the amounts of $LGD = 0$ cases are included in the denominator. The amount of loss, ie, the numerator, is the same in both cases.

A 10% decrease of collateral value leads to an ELGD of 2.71% in European style markets and to ELGD of 16.04% in US style markets, as Figures 3a and b reveal. How these different ELGD’s affect the EL cannot be analyzed using only the method presented, because the different trigger events also have effect
4 Conclusion

The method presented in this paper makes it possible to approximate the CLTV, age and remaining principal structures of a mortgage loan pool when loan-specific data is not available. A numerical method is based on solving a matrix equation of loan pool development with all inflows and outflows. One of the solved matrices includes the total remaining principals of loans divided into cohorts with the same origination period and original LTV, and the other matrix gives the CLTVs of these same cohorts. If the numbers of new loans and loans in the pool are available, a matrix including the remaining numbers of loans per cohort can also be generated.

In this study the empirical data, ie, time series used for calculations and true reported LTV and remaining loan amount structures, have been collected from the Finnish banking sector. A comparison between reported structure and structure estimate indicates that the method is accurate.

The solved structure makes it possible to analyze changes in the dynamics of a mortgage loan pool over time, and also to compare the risks of different pools. One of the most important applications of the model is that it can be used to supplement the existing mortgage loan credit risk models, and thus to improve the accuracy of risk and loss estimates.

5 Acknowledgements

The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper.

References


Basel Committee on Banking Supervision (2001) The New Basel capital ac-


**Appendix A. Method to estimate a starting pool**

Before construction of the starting pool structure, the geometric average growth rates of the originated mortgages $l_t^*$ and the total amounts, or the number, of loans in the pool, $L_t^*$, should be calculated. These growth rates are denoted $g_l$ and $g_L$ respectively.

The sum of remaining loans in the starting pool is

$$\hat{L}_{t-1} = \frac{L_t^*}{1 + g_L}$$

(19)

where $L_t^*$ is the amount of loans in the pool at the end of first observation period and thus $\hat{L}_{t-1}$ is the amount one period earlier.

The amount of loans in cohorts, originated at $t-i$, where $i \in [1, 2, \ldots, m]$ are denoted as $\hat{l}_{t-i}$. The amount of prepaid loans is unknown. Thus:

$$\hat{l}_{t-i} = f(i, a) = \frac{l_t^*}{(1 + g_l)^i} \cdot \frac{m - (i - 1)}{m} \cdot (1 - a)^{(i-1)}$$

(20)

The starting pool $\hat{L}_{t-1}$ can be written as

$$\hat{L}_{t-1} = \left[ \hat{l}_{t-1} \ \hat{l}_{t-2} \ \ldots \ \hat{l}_{t-m} \ 0 \right]$$

$$= \left[ f(1, a) \ f(2, a) \ \ldots \ f(m, a) \ 0 \right]$$

(21)

The prepayment parameter $a$ can be solved using Equation (22) with $g(a) = 0$.

$$g(a) = 1' \hat{L}_{t-1} - \frac{L_t^*}{1 + g_L}$$

(22)

The solution for $a$ can be used as a parameter when the starting structure is calculated using formulas (20) and (21).

**Appendix B. Derivation of amortization formulas in Equation (6)**

The formulas in Equation (1) are used to calculate amortizations as a percentages from the remaining principal. The derivation of these equations is explained in this appendix.
As explained before, this paper presents a method to approximate the remaining mortgage pool structures, not LGDs of the defaulting loans. Thus accrued interests can be excluded from the model, and only dynamics of the principals are included.

The amortization formulas have been presented in Equation (6) as

\[
\alpha_{t-i} = \frac{a_{t-i}}{p_{t-i}} = \begin{cases} 
0 & \text{when } i = 0 \\
\frac{1}{m-i+1} & i \in \{1, \ldots, m\}, \text{ fixed amortization loans} \\
\frac{r_{t-i}(1+r_{t-i})^{m-i-1}}{(1+r_{t-i})^{m-i-1} - 1} & i \in \{1, \ldots, m\}, \text{ annuity loans}
\end{cases}
\]  

(23)

As Equation (1) defines the pool dynamics, amortizations of loans granted during period \( t-1 \), will be started at period \( t \). Thus if the maturity of the loan is \( m \), the last amortization will be paid when the granting period is \( t-m \) and current period is \( t \).

Every loan payment includes amortization and accrued interest parts. With fixed amortization loans, the amortization stays the same over time, ie \( a_{t-i} = a \) when \( 1 \leq i \leq m \) (otherwise zero). Thus the fixed amortization \( a \) can be written as \( a = \frac{p}{m} \), and after \( i \) amortizations the remaining principal will be \( p_{t-i} = (m-i)a \).

For the loan that has been granted \( i \) periods ago, the remaining principal at time \( t \) before amortization, ie \( p_{t-i}^{t-1} \) can be written as

\[
p_{t-i}^{t-1} = (m-i+1)a
\]  

(24)

So, for the loan that has been granted one period ago, ie at \( t-1 \), the remaining principal is \( p_{t-1}^{t-1} = (m-1+1)a = ma \), and for the loan that has been granted \( m \) periods ago, ie at \( t-m \), the remaining principal is \( p_{t-m}^{t-1} = (m-m+1)a = a \).

The ratio between amortization and remaining principal, ie the multiplier used to for calculations, can be written as

\[
\alpha_{t-i} = \frac{a}{p_{t-i}} = \frac{a}{(m-i+1)a} = \frac{1}{m-i+1}
\]  

(25)

This means that the size of the first amortization equals \( \frac{1}{m} \) from the original principal, and the last amortization, after \( m \) periods, equals \( \frac{1}{1} \), which is always 100% from the remaining principal.

Similarly the ratio between amortization and remaining principal can be derived to annuity loans. The basic formula for the remaining principal of
annuity loan, as derived eg in the book Brealey et al. (2011, see p. 29), can be written as:

\[ PV = C \left( \frac{1}{r} - \frac{1}{r(1+r)^t} \right) \]  

(26)

where \( PV \) is the remaining principal, \( C \) is the fixed payment (including both accrued interests and amortization), and \( r \) is the interest rate including loan marginal.

This can be written with the same variable names and formulation used in Equation (6)

\[ c_{t-i} = p_{t-i}^{t-1}r_{t-1} + a_{t-i} \]

where \( c_{t-i} \) is the fixed payment. Equation (26) can be written as

\[ p_{t-i}^{t-1} = c_{t-i} \left( \frac{1}{r_{t-1}} - \frac{1}{r_{t-1}(1+r_{t-1})^{m-i+1}} \right) \]

\[ \Leftrightarrow p_{t-i}^{t-1} = c_{t-i} \left( \frac{(1+r_{t-1})^{m-i+1} - 1}{r_{t-1}(1+r_{t-1})^{m-i+1}} \right) \]

(27)

\[ \Leftrightarrow c_{t-i} = p_{t-i}^{t-1} \frac{r_{t-1}(1+r_{t-1})^{m-i+1}}{(1+r_{t-1})^{m-i+1} - 1} \]

and after that

\[ c_{t-i} = p_{t-i}^{t-1} \left( r_{t-1} \left( \frac{(1+r_{t-1})^{m-i+1} - 1 + 1}{(1+r_{t-1})^{m-i+1} - 1} \right) \right) \]

\[ = p_{t-i}^{t-1} \left( r_{t-1} \left( 1 + \frac{1}{(1+r_{t-1})^{m-i+1} - 1} \right) \right) \]

\[ = p_{t-i}^{t-1} \left( r_{t-1} + \frac{r_{t-1}}{(1+r_{t-1})^{m-i+1} - 1} \right) \]

(28)

\[ = p_{t-i}^{t-1}r_{t-1} + p_{t-i}^{t-1} \frac{r_{t-1}}{(1+r_{t-1})^{m-i+1} - 1} \]

\[ = p_{t-i}^{t-1}r_{t-1} + a_{t-i} \]

Thus

\[ a_{t-i} = \frac{a_{t-i}}{p_{t-i}^{t-1}} \]

\[ = \frac{p_{t-i}^{t-1}}{p_{t-i}^{t-1}} \left( \frac{r_{t-1}}{(1+r_{t-1})^{m-i+1} - 1} \right) \]

(29)

\[ = \frac{r_{t-1}}{(1+r_{t-1})^{m-i+1} - 1} \]

Even when the annuity formula is not defined when the interest rate is zero, the derived result should approach fixed amortization formula in case the interest rate approach zero. This is also true, ie

\[ \lim_{r_{t-1} \to 0} \frac{r_{t-1}}{(1+r_{t-1})^{m-i+1} - 1} = \frac{1}{m - i + 1} \]

(30)
Appendix C. Derivation of period by period principal changes presented in Equation (15)

At the end of granting period the remaining principal equals the granted principal. In case of fixed amortization loans the original principal can be written as $p_{t-i} = ma$, and until the loan is fully paid back, the size of the amortization will stay the same $a$. Thus the remaining principal after $i$ amortizations is $p_{t-i} = ma - ia$, and the ratio between these can be written as

$$\frac{p_{t-i}}{p_{t-i}} = \frac{ma - ia}{ma} = \frac{m - i}{m} \quad (31)$$

After $m$ periods ($i = m$) the loan has been fully paid back, and the LTV calculation has no longer meaning.

For the annuity loans the ratio can be calculated using the same principle as in Equation (31). In this case the size of the amortization is dependent on interest rate (including marginal). To calculate the remaining principal, the sum of past amortizations should be subtracted from the original principal, ie

$$p_{t-i} = p_{t-i} - \sum_{h=0}^{i-1} a_{t-i}^{t-h} \quad (32)$$

where $a_{t-i}^{t-h}$ is at time point $t-h$ calculated amortization for loans granted $t-i$ periods ago using interest rate $r_{t-i}$. These amortizations can be rewritten using multipliers $a_{t-i}^{t-h}$, which have been already calculated for vector $a^t$ in Equation (5), using Equation (6).

Using $a_{t-i}^{t-h}$ multipliers the equation can be written as

$$\frac{p_{t-i}}{p_{t-i}} = \frac{p_{t-i} - \sum_{h=0}^{i-1} a_{t-i}^{t-h}}{p_{t-i}} = \frac{p_{t-i} - \sum_{h=0}^{i-1} p_{t-i}^{t-h} a_{t-i}^{t-h}}{p_{t-i}}$$

The remaining principal after the first amortization is

$$p_{t-i+1} = p_{t-i} - p_{t-i} a_{t-i+1}^{t-i+1}$$

$$p_{t-i+1} = p_{t-i}(1 - a_{t-i+1})$$

and similarly, using the previous result, the remaining principal after the second amortization can be written as

$$p_{t-i+2} = p_{t-i}(1 - a_{t-i+1}^{t-i+2}) - p_{t-i+1} a_{t-i+2}^{t-i+2}$$

$$p_{t-i+2} = p_{t-i}(1 - a_{t-i+1}) - p_{t-i}(1 - a_{t-i+1}) a_{t-i+2}^{t-i+2}$$

$$p_{t-i+2} = p_{t-i}(1 - a_{t-i+1})(1 - a_{t-i+2})$$
and so on. Using these results in Equation (33) can be rewritten as a product of sequences

\[
\frac{p_{t-i}^t}{p_{t-i}^{t-1}} = \frac{p_{t-i}^{t-1} \prod_{h=0}^{i-1} (1 - \alpha_{t-i}^{-h})}{p_{t-i}^{t-1}} = \prod_{h=0}^{i-1} (1 - \alpha_{t-i}^{-h})
\]

which is the same formula presented in Equation (15).
Palmroos, Peter *Correlation between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools*. Submitted for publication.
Correlation between Probability of Default and Loss Given Default on Homogeneous and Heterogeneous Mortgage Loan Pools

Peter Palmroos

Financial Supervisory Authority, PO Box 103, 00101 Helsinki, Finland

Abstract

This paper studies how the creditor’s ability to observe and register zero-loss mortgage loan defaults affects the observed correlation between Probability of Default (PD) and Loss Given Default (LGD), when homogeneous and heterogeneous mortgage pool structures are assumed. The results reveal that both the Loan-to-Value and maturity structure of the pool, and the way the zero-loss defaults are registered, affect the observed correlation. The presented results indicate that the observed correlation may be biased enough to cause significant errors on credit risk models, and that observed correlation cannot be used in calculations or simulations as such. The presented findings increase our knowledge of the behavior of mortgage loan risk components with different pool structures and thus help to create more reliable and accurate mortgage loan credit risk models.

1 Introduction

Since the Basel II internal ratings-based approach (IRBA) models were presented, the independence assumption between Probability of Default (PD) and Loss Given Default (LGD) has provoked discussion, both within the groups of practitioners and of researchers. Studies, like Altman et al. (2001), Di-
mou et al. (2005) and VanOrder (2007), have found evidence of correlation between these credit risk factors when ex-post point-in-time (PIT) estimates have been used. One explanation for this correlation is that both PD and LGD are dependent on the same macroeconomical variables.

The effects of dependencies between explanatory variables are relatively easy to analyze for single loans or homogeneous loan pools. But virtually all the mortgage loan pools of banks are large and heterogeneous, which means that the use of a homogeneity assumption predisposes the presented results to criticism. The heterogeneity of pools means that loans have been granted at different times; with unequal principals, Loan-to-Values (LTV), and maturities; and also the debtors have different probabilities to default. Due to the different characteristics of loans, depending on which mortgagors default, the observed LGDs, and thus also the realized losses, vary remarkably.

Not only the sample of debtors in default affects the observed LGD, but also the LGDs of the defaulting loans may affect creditors’ ability to observe these defaults. This may be the case when a debtor faces an insolvency in a moment when the loan is fully covered by the collateral, and after a successful realization of collateral, the debtor prematurely pays back the loan. From the debtor’s point of view the insolvency event was no doubt a default, but from creditors perspective the case is not so clear.

The realization of the collateral needs acceptance of the creditor, but the debtor doesn’t need to inform the creditor of the reason for the realization and repayment of the loan. The process and outcome from the creditors view is the same whether it was caused by the normal house exchange event or zero-loss default caused by the insolvency of the customer. Thus, the creditor is not able to separate these two events from each other. The insolvency of the debtor can be observed for sure only if either the collateral doesn’t cover the remaining loan and accrued interests, or the collateral cannot be realized before the next amortization date.

The proportion of zero-loss insolvency events may affect the number of observed defaults. If the delay in realization is omitted, the current LTV, or LGD, measures whether the insolvency of the debtor covers the loss to the debtor or not. Thus the current LTV structure of the pool has an effect on LGDs of the loans. Thus also the structure of the pool has an effect on the identification and registering of defaults, i.e. PD, especially in case of zero-loss

1In this paper the PD denotes both ex-ante probability-of-default, as well as ex-post calculated default rate. This decision has been made because, due to the topic, the risk of confusion is minimal.
defaults. This effect of pool structures on correlation between PD and LGD has been almost completely omitted in the academic studies and models.

This paper studies the effects of two properties of mortgage loan pools on the observed correlation between PD and LGD: first, how the different recognizing and observing assumptions affect the observed correlation, and secondly whether the different mortgage pool structures have an effect on observable correlations between PD and LGD.

The paper is organized as follows. Section 2 presents some theoretical background and findings presented in the literature related to the topic of this paper. Section 3 introduces the assumptions made and the mathematical setup using a single loan. In Section 4 the model is developed to cover homogeneous loan pools and cohorts. Section 5 explains how the heterogeneous pool structures are estimated using publicly available data. Section 6 presents the correlation calculation methods in heterogeneous pools when a large number of homogeneous cohorts are used. Results of example calculations and simulations are presented in Section 7. The last section concludes.

2 Background and literature

The IRBA models (see e.g. Basel Committee, 2003), used for banks’ credit risk measurement for capital adequacy calculations, assume independence of PD and LGD, albeit the possibility of existence of correlation is recognized. The effect of possible correlation has been covered using a long-term (over-the-cycle, OTC) and downturn PD and LGD estimates in capital requirement calculations. This OTC assumption might be justified for capital requirement calculations, however when IRBA-like models are used for credit risk prediction models, the PIT estimates of PD and LGD are needed. As with OTC estimates, with PIT estimates of the pool the independence assumption can be found unacceptable, as will be seen.

Support for the dependence between PD and LGD hypothesis has been presented e.g. in papers Chabaane et al. (2004) and Schuermann (2004). Frye (2000b), Altman et al. (2002), and Das (2007) have found positive correlation between the number of defaults and LGD in corporate loan and bond data. This is in line also with the paper of Altman et al. (2001), that found both PD and LGD to behave stochastically and also to be partially dependent. The same paper also presents a list of other studies supporting this hypothesis.

The mentioned papers have studied corporate loans and bonds, whose practices and assumptions differ from the ones used in mortgage loan markets.
Nonetheless similar results can be found also when residential mortgage exposure data have been studied. Dimou et al. (2005) and VanOrder (2007) present evidence on positive correlation between PD and LGD also within mortgage loan pools.

Support have been found for the hypothesis that macro variables, or economic cycles, explain this dependence. E.g. Erlenmaier and Gersbach (2001), Gross and Souleles (2002), and Allen and Saunders (2003) have found higher number of defaults during recession.

Like the number of defaults, the values of houses and apartments, which are the most widely used collateral on mortgage loans, and thus also the most important factor affecting the LGD, have been found to be dependent on macroeconomic variables. At least Frye (2000b), Allen and Saunders (2003), Frye (2003), and Schuermann (2004) have presented evidence of higher LGDs in recessions than during expansions, and e.g. Frye (2000a) explains the higher LGD levels during recessions with consequences of decreasing asset values.

The effect of the omitted correlation between PD and LGD might be drastic for the loss forecasts, and thus increase the risk of unexpected losses. For example Frye (2000a) and Altman et al. (2002) highlight the importance of correlation and remind that a severe downturn might bring a double hit on banks. In the papers of Frye the double hit has been explained as simultaneous increase of both PD and LGD.

The trigger events causing defaults of debtors vary. Reasons for this are the differences between the economical and juridical environments in housing and mortgage loan markets, as well as in the market practices. To take these into account, theoretical models have been categorized also on the basis of these trigger events.

Whitley et al. (2004) categorize the theories into two groups: 'equity theories' and 'ability-to-pay' -theories based on the way rational debtors are assumed to react to trigger events.

In the equity-theories framework, when other costs caused by the default have been omitted, a rational debtor defaults as soon as the value of collateral sinks below the remaining mortgage loan principal. In that kind of situation voluntary defaulting is rational, only if the legislation allows personal bankruptcy, and the mortgage loan can be fully cleared by assigning the collateral to the creditor. One paper assuming equity-theory is Kau et al. (1992). The idea of the equity theories lie on structural-form models, originally presented for corporate loans and bonds by Merton (1974). Later on different variants of this kind of models are presented by e.g. Altman et al. (2001) and
Jokivuolle and Peura (2003). More comprehensive reviews can be found e.g. from Altman et al. (2002).

If a debtor is, after the realization of the collateral, still liable for the remaining loan, a voluntary default is irrational. This is the case in environments where legislation doesn’t allow the personal bankruptcy of private persons. Thus in most of the European countries the equity theories are not suitable to explain the defaults of the borrowers. Instead of decreasing property prices, it is assumed that defaults are an undesired result of the insolvencies of debtors, instead of rational choices. Whitley et al. (2004) refer to these approaches as 'ability-to-pay' theories.

List and explanations of triggering events that may cause insolvency of the customer in 'ability-to-pay' environment have been presented in the paper of Cairns and Pryce (2005).

In this paper the ability-to-pay framework is assumed. The presented models are indifferent on the causes which trigger insolvencies of the customers, albeit in calculations presented in this paper the distribution parameters of insolvencies are estimated using Finnish unemployment data\textsuperscript{2}.

Basel II instructs to include realization and other default-related costs to calculations when LGD and EAD are measured. The correlation calculations are not sensitive to this kind of relatively small LGD adjustments. Thus in the correlation calculations between PD and LGD, only the remaining loan principal is taken into account, and all other costs and accrued interests are omitted. Although the analysis presented in this paper can also be performed including these additional costs when needed.

3 Model and assumptions

In this paper defaults have been analyzed from the creditor’s point of view. To make the terminology consistent, all the insolvency cases of debtors can be classified either as loss causing defaults, in short only defaults, or as zero-loss defaults.

The paper assumes a single borrower per loan, and wages as the only source of income. With these assumptions unemployment is a clear trigger event causing insolvency. If multiple borrowers per loan are allowed the model becomes more complex. In that case the default is dependent on whether the remaining incomes of the household, after the unemployment of a single bor-

\textsuperscript{2}The available sources of the data are Statistics Finland and the Ministry of Employment and the Economy.
rower, are large enough to cover amortizations of the mortgage loan after the mandatory living costs. For finding an answer to the key research question this kind of analyses are not necessary and thus have been considered out of the scope.

A borrower can become insolvent at any time during the amortization period, but insolvency can cause a default only at the scheduled amortization payment date.

It is also assumed that housing markets are efficient and fully liquid, i.e. at a single point in time there is only one price for housing property, and it can be realized using that price without any delays. Instead of the 90 day rule of default, as has been defined in Basel II documentation, the default is assumed to occur at the moment when the borrower is unable to take care of the planned amortization.

The number and nature of collateral have been restricted to one residential real estate collateral per loan. Restricting the number of pieces of collaterals to one simplifies the LGD calculations when correlations between changes in the values of different kinds of collaterals can be omitted. This assumption can be done without losing the generality of the model.

With only one residential real estate as a collateral, the LGD at the end of amortization period is dependent on the LTV at the beginning of the period, and from the development of collateral value during the period. The relative change of collateral value during the observation period can be written as

\[ \gamma_t = \frac{C_t}{C_{t-1}} \]  

where \( \gamma_t \) is the ratio of collateral values changes between times \( t - 1 \) and \( t \), and \( C_{t-1} \) and \( C_t \) are collateral values at these times respectively.

The amortization period is assumed to be equal in length to a single observation period, and thus the principal of the loan stays the same from the beginning of the period until the amortization at end of the period.

At the beginning of the period, either immediately after the loan has been granted or after the previous amortization, the LTV is

\[ LTV_t = \frac{P_t}{C_t} \]  

where \( LTV_t \) is the Loan-to-Value of the loan at time \( t \), \( P_t \) is the remaining loan principal after the planned amortization or granting of the loan, and \( C_t \) the value of collateral at the same time.

At the end of the period, just before the planned amortization, the current
Loan-to-Value (CLTV) is

\[ CLTV_t = \frac{P_t + a_t}{C_t} = \frac{P_{t-1}}{C_t} = \frac{P_{t-1}}{\gamma_t C_{t-1}} = \gamma^{-1} LTV_{t-1} \quad (3) \]

where \( CLTV_t \) is the Current Loan-to-Value at period \( t \), just before the planned amortization \( a_t \). \( P_{t-1} \) and \( P_t \) are the remaining loan principal after the amortizations at time \( t-1 \) and \( t \), and \( C_{t-1} \) and \( C_t \) are the values of the collateral at time \( t-1 \) and \( t \). \( LTV_{t-1} \) is the LTV after the previous amortization, i.e. at the beginning of the current period.

CLTV is the ratio that measures whether the collateral covers the remaining loan or not in the case of insolvency. The only difference between \( CLTV_t \) and \( LTV_t \) is that the first is calculated just before, and the latter right after the amortization \( a_t \) at time \( t \), as can be seen comparing Equation (2) and the first part of Equation (3).

When the \( CLTV_t \) is known, \( LGD_t \) of the loan can be calculated. This functional dependence between \( LGD \) and \( CLTV \) can be written as

\[
LGD_t = \max \left( 0, 1 - CLTV_t^{-1} \right)
\]
\[
= \max \left( 0, 1 - \frac{C_t}{P_{t-1}} \right)
\]
\[
= \max \left( 0, 1 - \frac{\gamma_t C_{t-1}}{P_{t-1}} \right)
\]
\[
= \max \left( 0, 1 - \gamma_t LTV_{t-1}^{-1} \right)
\quad (4)
\]

LGD can only get values greater or equal to zero and less or equal to one\(^3\). This is because it describes the effect of default for the creditor. Thus, no matter how much higher the value of collateral is than the remaining principal, the loss caused to the creditor is zero, i.e. \( LGD = 0 \).

The inequality between insolvencies faced by the debtors and defaults observed by the creditors is the key assumption of this paper. For the debtor the insolvency automatically means his or her default, and no doubt he or she is able to observe it. But for the creditor this is not so simple. When the collateral does not cover the remaining principal of the loan, the creditor observes the default. But in the case where the value of the collateral is higher than or equal to the remaining loan principal, observing defaults or separating zero-loss defaults from the common home exchange situations may be impossible.

In this paper the effects on correlation caused by different observation and registration rules and LTV or LGD structures are analyzed. These effects have

\(^3\)Usually percentages are used and the values can vary between 0% and 100%, including the starting and end points.
been analyzed using three different scenarios with homogeneous and heterogeneous pool structures.

In the first scenario, when rationally behaving borrowers and efficient markets are assumed, the insolvency of the borrower can be observed by the bank only if $LGD_t > 0$, which is equal with $CLTV_t^{-1} = C_t/P_t^{-1} < 1$. In that case the remaining principal is larger than the value of the collateral, and thus, in the case of insolvency, the borrower is not able to repay the loan after the realization of the collateral.

The second scenario assumes that the creditor is able to observe and register all insolvency cases, including zero-loss defaults. This kind of outcome is possible e.g. when the rationality hypothesis of the borrower is rejected, i.e. when the borrower does not even try to realize the collateral after the insolvency.

The third scenario assumes, similar to scenario one, that zero-loss defaults cannot be observed, but contrary to scenario one, the zero-default and zero-loss periods are registered. In this scenario the creditor falsely assumes that the number of observed defaults, i.e. zero, is also the actual number of all defaults.

As the assumptions used in the scenarios indicate, the observed defaults do not comprise an independent random sample from the population of insolvency cases. This is in line with the findings of Ambrose et al. (1997), Chabaane et al. (2004), and Dimou et al. (2005).

The effects of the first two scenarios on the correlation between PD and LGD will be presented using both homogeneous and heterogeneous pool structures. The effects of the third scenario will be analyzed only with homogeneous pools.

4 Homogeneous pool and single cohort

A homogeneous pool or cohort is defined to include only loans with the same LTV at the beginning of every observation period. For all the loans residential real estates are the only collateral, and value of this collateral is assumed to be fully correlated. Thus also the CLTV at the end of the period is the same for all loans of the pool. The PD is assumed to be stochastic, i.e. vary from period to period, but in a single period the same for all borrowers. In estimating the correlation between PD and LGD, other properties of the loans, e.g. maturity, age, and size, need not be identical.

The model assumes that both relative changes in values of collaterals, and insolvency rate, follow log-normal distributions. This assumption of log-normality gets support both from the literature, e.g. papers Chabaane et al. (2004) and
Dimou et al. (2005), and also from the empirical data. The unemployment indicator used and the variable approximating changes in collateral prices have been presented, and the log-normality has been tested, in Appendices B and C respectively. It is also assumed, in accordance with the results of research papers mentioned at Section 2, that the correlation between these variables is negative or zero.

These assumptions can be written as

\[ I_t \sim \text{LogN}(\mu_I, \sigma_I) \]
\[ \gamma_t \sim \text{LogN}(\mu_\gamma, \sigma_\gamma) \]
\[ \text{Corr}(I, \gamma) \leq 0 \]  

(5)

where \( I_t \) is the insolvency rate, and \( \gamma_t \) is the ratio of collateral values between periods \( t - 1 \) and \( t \).

When the results of different scenarios are analyzed, the paper uses the unrestricted "true correlation" between PD and LGD as a reference. To define it, the unrestricted LGD, marked as \( \text{LGD}_t^* \), needs to be defined. That can be written as

\[ \text{LGD}_t^* = 1 - \frac{CLTV_t - 1}{C_{t-1}} = 1 - \gamma_t \frac{C_{t-1}}{P_{t-1}} = 1 - \gamma_t \frac{LTV_{t-1}^{-1}}{P_{t-1}} \]

(6)

\( \text{LGD}_t^* \) is otherwise the same as \( \text{LGD}_t \) as defined at Equation (4), except that it is allowed to become negative. As a concept this is the LGD as a debtor can see it: in the case of insolvency, if the CLTV is less than one, the debtor get the surplus of the value of collateral after the repayment of the loan. Although the negative LGDs are not meaningful for the creditors, they are relevant for the debtors.

It should be noticed that when \( \gamma \) is assumed to be log-normally distributed, use of Equation (6) makes the \( \text{LGD}_t^* \) follow a mirrored log-normal distribution⁴.

If it is assumed that all the insolvency cases can be observed and registered as defaults, we can write

\[ PD_t = I_t \]

(7)

Where \( PD_t \) is the rate of observed defaults in a group of all mortgage loan owners, and \( I_t \) is the rate of insolvency cases in the group of debtors.

⁴Also known as 'mirror-log-normal distribution' and 'inverted log-normal distribution'.
It can be seen that \( \text{Corr}(PD, LGD^*) = -\text{Corr}(I, \gamma) \). The absolute value of correlations stays the same, but the sign of the correlation changes. This result is derived in Appendix A. In short, the minus sign is caused by the mirrored, or inverted, distribution of \( \gamma_t \). Later in this paper \( \text{Corr}(PD, LGD^*) \) will be written as \( \rho_{LN} \).

The three scenarios presented in the previous section lead to three different truncated or censored distributions of PD and LGD.

Scenario one, where the bank is unable to observe defaults when \( LGD = 0 \), generates a left-truncated bivariate distribution, where the truncation point is \( LGD = 0 \). Following these assumptions the variables and the correlation can be written as

\[
PD_{T,t} = I_t |_{LGD_t > 0} \\
LGD_{T,t} = LGD_t |_{LGD_t > 0} \\
\rho_T = \text{Corr}(PD_T, LGD_T)
\] (8)

where \( T \) indicates truncation of the distribution and subscript \( t \) identifies a single, period \( t \) element of the observation vector. So only observation periods with \( LGD_t > 0 \) are included in to the vectors \( PD_T \) and \( LGD_T \). Again \( I_t \) is the insolvency ratio of borrowers.

The second scenario assumes that all insolvency cases can be observed, including zero-loss defaults. Thus the bivariate distribution of PD and LGD is left censored from the point \( LGD = 0 \). This can be written as

\[
PD_{C,t} = I_t \\
LGD_{C,t} = LGD_t \\
\rho_C = \text{Corr}(PD_C, LGD_C)
\] (9)

where \( C \) indicates the censoring of the distribution, and subscript \( C, t \) identifies a single, period \( t \) element of the observation vector. It should be noticed that the censored \( LGD_t \) has been used, not the uncensored \( LGD_t^* \).

In the third scenario the periods without observed defaults are included in the sample as zero values. The distribution of LGD is again left-censored, but unlike as in the previous scenario, the observed number of defaults is also affected by the censoring. It means that when \( LGD_t = 0 \) the observed rate of defaults is also zero. This can be written as

\[
PD_{Z,t} = \begin{cases} I_t & \text{if } LGD_t > 0 \\ 0 & \text{if } LGD_t = 0 \end{cases} \\
LGD_{Z,t} = LGD_t \\
\rho_Z = \text{Corr}(PD_Z, LGD_Z)
\] (10)
where $Z$ indicates that zero-loss defaults are included as zero values, and subscript $Z,t$ identifies a single, period $t$ element of the observation vector.

In all the presented three cases, when the population parameters of $I$ and $\gamma$ are known, as well as the $\rho_{\gamma,I}$, there is a closed form solution of observed correlation for all the presented restricted distributions. The derivation of the closed-form solutions of $\rho_T$ and $\rho_C$ is presented in the paper Vilmunen and Palmroos (2013).

5 Modeling heterogeneous pool structure using publicly available information

When loan-by-loan data is not available, the pool structure may only be approximated using the publicly available information. In this paper the structure of the mortgage loan pool is modeled following the method presented in the paper Palmroos (2016).

The method used classifies loans to a large number of homogeneous cohorts. Hereafter the remaining numbers of loans and principals of the loans are calculated for every cohort, based on a simple stock and flow model. The model uses time series of mortgage loan pool size, and amounts\(^5\) of granted loans, which are widely available e.g. via statistics published by central banks. As soon as the numbers and amounts of loans per cohorts have been calculated, the cohort-by-cohort CLTVs can be solved using the LTV distribution of the granted loans, estimated amortizations, and the development of collateral values.

![Figure 1](image-url)  
Figure 1. a) Distribution of proportions of remaining loan principals per cohorts, b) Current Loan-to-Value (CLTV) of the same cohorts.

\(^5\)Both number of loans and amounts of principals are used and needed for pool structure modeling and calculations.
The pool structure used in simulations of this paper has been presented in Figures 1a and b. The values presented in these figures have been calculated using publicly available data\(^6\) of Finnish banks and economy.

In modeling, a two-sided exponential distribution, fitted to reported data, has been used instead of the reported LTV distribution. While the two-sided exponential distribution won’t approximate the data very nicely, the assumption has been made to make the results of the correlation calculations easier to understand. If the smoothly behaving two-sided exponential distribution were not used, the large fluctuations between adjacent LTV groups would cause arbitrary looking correlation estimates when the variance of \(\gamma\) is small. Otherwise the results calculated using fitted and original distributions are very similar in nature. Both the fitted and reported LTV distributions can be found in Appendix D. Due to the missing historical LTV distributions of originated loans\(^7\), the distribution has been assumed to stay the same over time.

Figure 1a presents the estimated proportions of the remaining principals per cohorts at time \(t\). The peak at 100% LTV means that when only residential housing property has been used as collateral, granting new loans to the full value of the house has been the mode\(^8\).

As can be seen, the amount of remaining loans is smaller the more time has elapsed since the granting. Two explanations can be found: first is amortizations and repayments, and secondly the granted amount of single loans has increased over time due to rising housing prices.

Figure 1b presents the estimated CLTV structure at the end of 2010 for all the cohorts. Both the decreasing original LTV and increasing age of the loans decrease the CLTV. The amortized principals (numerator) and positive development in housing prices (denominator) explain this phenomenon.

6 Heterogeneous pool with multiple cohorts

Section 4 defined the calculation methods for correlations of restricted distributions when homogeneous mortgage loan pools were assumed. This section presents a method for correlation calculations when homogeneity assumption

---

\(^6\) Data source Financial Supervision Authority of Finland (Fin-FSA) and Bank of Finland

\(^7\) LTV distribution of the granted loans has been reported by Finnish banks at the end of 2010 as a part of the mortgage loan sample study of Fin-FSA.

\(^8\) After 2010 both regulation of Fin-FSA and legislation have set limits to LTV levels. Also the uncertainty in the economy has made both creditors and debtors more cautious, and thus decreased the LTV levels of new loans.
do not hold, i.e. loans in the pools have been granted at different times, with
different LTVs, maturities, and principals, i.e. when the pools structures are
heterogeneous.

In a homogeneous pool it is assumed that all the loans have the same LGD,
and after every observation period, depending on the CLTV, either all or none
of the insolvencies of the debtors can be observed. With a heterogeneous
structure both the number of observable insolvency cases and the expected
LGD of the whole pool are dependent on $\gamma$ and the current pool structure.

The functional dependence between $\gamma$ and the whole pool risk parameters can
be easily explained with a three loan example. Let the LTVs of the otherwise
similar loans be 90%, 70%, and 50%. When the value of $\gamma$ is higher than 0.9,
the LGDs of all loans equal zero. Assuming that zero-loss defaults cannot be
observed, the Proportion of Observable Insolvencies, POI, equals zero. When
the value of $\gamma$ decrease below 0.9, first the LGD of the highest LTV loan will
become larger than zero, and when $\gamma$ decreases below 0.7, the same happens
also with the second loan. It should be noticed that as soon as $\gamma$ decrease
below 0.9, and continues decreasing, the LGD of the higher LTV loan starts
to increase, and it continues increasing all the time as $\gamma$ decreases. At the
same time also the number of insolvency cases able to be observed increases.
Table 1 presents the development of these factors as functions of $\gamma$.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>LGD of the Loans with</th>
<th>The whole pool</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LTV 90%</td>
<td>LTV 70%</td>
</tr>
<tr>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>80%</td>
<td>11.1%</td>
<td>0%</td>
</tr>
<tr>
<td>60%</td>
<td>33.3%</td>
<td>14.3%</td>
</tr>
<tr>
<td>40%</td>
<td>55.6%</td>
<td>42.9%</td>
</tr>
<tr>
<td>20%</td>
<td>77.8%</td>
<td>71.4%</td>
</tr>
<tr>
<td>0%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 1. LGDs and POI of the three loan sample pool as a function of collateral value
development $\gamma$, when PD = 100% is assumed.

Calculations as in Table 1 can be performed on all pools on a loan-by-loan
basis. This is possible only for creditors who have all the required information.

In the heterogeneous pool models presented in this paper a single cohort
includes the same loans$^9$ from the granting until all the loans in a cohort
have been fully amortized or prematurely paid back. This is contrary to the

$^9$The number of loans may decrease due to repayments, but new loans cannot appear
in cohorts. Also only full premature repayments are assumed.
homogeneous pool model, where the loans included in a cohort have been reselected at the beginning of every period to keep the characteristics of the cohort stable over time.

The closed form correlation calculations presented in Section 4 can be performed for every single homogeneous cohort, but the heterogeneity and large number of cohorts result in no closed form solution for the whole pool correlation. Thus, simulation methods need to be used in solving for whole pool correlations.

When the granting period and LTV at origination have been used as classification rules, the different cohorts create a matrix of size $m \times n$, where the original maturity of all loans is $m$ periods, and the number of LTV clusters at origination equals $n$.

The ratio of loss causing defaults is called an Expected Proportion of Observable Insolvency cases, EPOI, and can be written as

$$EPOI_T(\gamma) = \frac{\sum_{i=1}^{k} (L_i | LGD_i > 0)}{\sum_{i=1}^{k} L_i}$$

(11)

where $EPOI_T(\gamma)$ is the expected proportion of observable insolvency cases when the change of value of collaterals equals $\gamma$, $k$ is the number of cohorts, i.e. $k = m \times n$, $L_i | LGD_i > 0$ is the number of loans in cohort $i$ with $LGD_i > 0$ (otherwise zero). The denominator equals the total number of loans in the whole pool. As explained in Equation (4), LGD of the cohort is a function of $\gamma$, and thus $EPOI_T$ is also a function of $\gamma$. Subscript $T$ indicates that truncation of zero-loss defaults has been assumed.

In a population with a limited size, and when $0% < I < 100\%$, the insolvency cases create a random sample of loans from the population of all mortgage loans. This means that insolvency cases can distribute between cohorts without following exactly the proportions of loans per cohorts. In this paper the variance caused by random sampling has been omitted, and $EPOI_T(\gamma)$ is used as an estimate for the proportion of observable insolvency cases. In correlation calculations the effect of omitting the variance is negligible, but for credit risk calculations and model calibrations the possible effect needs to be recognized.

The expected value of LGD will be called Expected Loss Given Default (ELGD). In line with the $EPOI$ calculation, the variance caused by random sampling is omitted, and only the expected value is used. $ELGD$ can be calculated as

$$ELGD_T(\gamma) = \frac{\sum_{i=1}^{k} (P_i | LGD_i > 0 \times LGD_i | LGD_i > 0)}{\sum_{i=1}^{k} (P_i | LGD_i > 0)}$$

(12)
where \( ELGD_T(\gamma) \) is the Expected Loss-Given-Default of the whole pool, where zero-loss defaults are omitted, \( \gamma \) is the percent change of the value of collaterals, and \( k \) is the number of cohorts. \( P_i|_{LGD>0} \) is the total sum of remaining principals in the cohort \( i \), when the cohort’s \( LGD > 0 \), otherwise \( P_i = 0 \). Subscript \( T \) indicates truncation of zero-loss defaults.

When it is assumed that zero-loss defaults can not be observed, i.e. following Scenario one, the correlation between PD and LGD can be calculated as

\[
\begin{align*}
PD_{HeT,t} &= EPOI_T(\gamma_t) \cdot I_t \\
LGD_{HeT,t} &= ELGD_T(\gamma_t) \\
\rho_{HeT} &= Corr(PD_{HeT}, LGD_{HeT})
\end{align*}
\] (13)

where subscript \( HeT \) indicates truncation of zero-loss events. This is the heterogeneous pool version of the homogeneous pool calculation presented in Equation (8).

In the second scenario it is assumed that all the insolvency cases can be observed, including zero-loss defaults. Then Equations (11) and (12) get different formulations, and in this case EPOI can be written as

\[
EPOI_C(\gamma) = \frac{\sum_i^n L_i}{\sum_i^n P_i} = 1
\] (14)

In this case \( EPOI_C(\gamma) \), where subscript \( C \) indicates censoring, always equals one, and indicates what is assumed; 100% of insolvency cases can be observed as defaults.

Using the same assumption \( ELGD_C(\gamma) \) can be written as

\[
ELGD_C(\gamma) = \frac{\sum_i^n (P_i \times LGD_i(\gamma))}{\sum_i^n P_i}
\] (15)

where \( P_i \) is the total sum of principals in cohort \( i \). So, under the given assumption, the ELGD of the whole pool is a weighted average of the cohorts’ LGDs, using the total of principals per cohort as a weight.

The correlation between PD and LGD, when zero-loss defaults are included, can be written as

\[
\begin{align*}
PD_{HeC,t} &= EPOI_C(\gamma_t) \cdot I_t = I_t \\
LGD_{HeC,t} &= ELGD_C(\gamma_t) \\
\rho_{HeC} &= Corr(PD_{HeC}, LGD_{HeC})
\end{align*}
\] (16)

where subscript \( HeC \) indicates censoring.

The value of \( EPOI_T \) as a function of \( \gamma \) is presented in Figure 2a. The same downward sloping structure as in Table 1 can be seen; when \( \gamma = 0 \), the value of
\(EPOI_T(\gamma) = 1\), and when the value of \(\gamma\) increases past the point where for all cohorts \(LGD = 0\), also \(EPOI_T(\gamma) = 0\). The intuitive explanation of the curve is that when the value of collaterals approaches zero, i.e. \(\gamma \to 0\), \(LGD \to 100\%\) for all loans, and thus no matter which loans have been included in the sample of insolvency cases, every single one of them result in losses near the value of remaining principal. When the value of \(\gamma\) increases, the proportion of loans with \(LGD = 0\) also starts to increase. When the value of \(\gamma\) increases past the highest LTV in the pool, all the insolvency cases result in zero-loss defaults.

The values of \(ELGD_T\) and \(ELGD_C\) as functions of \(\gamma\) can be seen in Figure 2b. As with Equation (11), again the same features presented in Table 1 can be observed: when \(\gamma = 0\), the value of \(ELGD_T(\gamma)\) and \(ELGD_C(\gamma)\) equal one. When the value of \(\gamma\) grows past the point where for all cohorts \(LGD = 0\), both ELGDs equal zero.

\[\begin{align*}
\text{Figure 2.} & \quad \text{a) Expected Proportion of Observed Insolvencies, EPOI, as a function of collateral value change,}
& \text{b) ELGD of the whole pool as a function of changes in collateral values, calculated with and without } LGD = 0 \text{ cases.}
\end{align*}\]

7 Findings and implications

To analyze the effects of pools structure and creditors’ ability to observe defaults ceteris paribus, the other factors that might have an effect on correlation need to be defined and fixed.

With homogeneous pools the selected exogenous factors are: 1) the correlation between the unrestricted insolvency cases and collateral values, i.e the ”true correlation”, and 2) the LTV of the cohort, i.e. the point of truncation or censoring.

In all the analysis and graphs, the unrestricted correlation between insolvency rate and changes in collateral values, \(\rho_{LN}\), is used as reference. Without
censoring or truncation the value of the correlation between PD and LGD is the same as the correlation between \( \gamma \) and \( I \), except that the sign will change. The change in sign can be avoided using Recovery Rate (RR) instead of LGD. The LGD is more widely known and used in research, and is thus used in this paper instead of RR.

The difference between observed correlation and \( \rho_{LN} \) is presented using graphs (i.e. in Figures 3, 4, and 5). In this paper it is assumed that both unrestricted and restricted correlations are negative, and the difference between them is presented using the formula \( \rho_{\text{Restricted}} - \rho_{LN} \). A positive difference means that the observed correlation is closer to zero, i.e. absolute value of it is smaller than \( \rho_{LN} \). In the hypothetical case when the restricted correlation equals the unrestricted one in all cases, the difference function would be a straight plane at the level of zero.

In all the calculations the distribution parameters of insolvency cases used have been picked out to approximately meet the ratio of new unemployment cases divided by the previous year’s number of employed in Finland. This reflects the assumption that only employed people are able to get mortgage loans and that unexpected unemployment is the only trigger for insolvency. Because the number of new unemployed includes also events so short that the mortgagor won’t become insolvent and other similar cases where the unemployment won’t cause immediate insolvency, e.g. loan has multiple mortgagors, the used unemployment indicator overestimates the number of insolvency events. Thus, because this paper doesn’t try to estimate the correct correlations or other parameters, but reveal that the observed correlations shouldn’t be used in calculations as such, the estimated rate has been divided by two\(^{10} \) to get a ratio better in line with reality.

The parameters of the distribution of \( \gamma \) have been estimated from the time series of the Finnish HPI\(^{11} \). Data used for both \( I \) and \( \gamma \) estimations can be found in Appendices B and C.

Figure 3a presents the observed correlation as a function of correlation of the unrestricted distribution and the point of truncation, i.e. the LTV of the pool at the beginning of the observation period. Figure 3b reveals the difference between the correlation of the restricted sample and \( \rho_{LN} \).

As can be seen, the larger the LTV at the beginning of the period, the closer the values of observed correlation and the correlation of the unrestricted

---

\(^{10}\)Defining the correct size of the divider needs more research. Two is only a working assumption.

\(^{11}\)Price index of old dwellings in housing companies: Whole country. Source: Statistics Finland.
Figure 3. a) Truncated correlation ($\rho_T$) as a function of unrestricted correlation ($\rho_{LN}$) and $LTV_t$, b) The difference between the truncated and unrestricted correlations. Distribution parameters used in calculations are: $E_I = 0.045$, $S_I = 0.016$, $E_C = 1.02$, and $S_C = 0.105$.

distribution are to each other. That is because the smaller the portion of insolvency cases that cause zero loss defaults, the closer the sample of observed defaults is to the population of all defaults. Explanation for that can be seen by comparing the correlation Equations (6) and (8) without restriction. The same is true also with Equations (9) and (10), as can be seen in Figure 4.

The restricted correlation is equivalent or close to the value of the unrestricted correlation also in cases when the unrestricted correlation is zero or close to its minimum\(^\text{12}\).

The difference between the correlation of the censored sample, where all the zero-loss defaults can be observed, and the correlation of the unrestricted correlation, is presented in Figure 4a.

Similarly as with the truncated distribution, when the LTV of the pool is very high, the observed sample starts to be close to the population of all defaults, and the correlation of the restricted distribution starts to approach $\rho_{LN}$. Where the LTV of the pool is far below the expected value, almost all of the insolvencies are zero-loss defaults and in the censored part of the distribution. Thus the observed correlation is near zero. Also when $\rho_{LN} = 0$, the censoring has no effect on observed correlation, and $\rho_C = \rho_{LN}$. Also similarly as with the truncated distribution, with the censored distribution the observed correlation is higher, i.e. the correlation is closer to zero.

The correlation difference caused by the third scenario, where the zero loss\(^\text{12}\)It should be noted that with the bivariate log-normal distribution the minimum correlation is always $> -1$. The closed form solution of the minimum has been presented in the paper Vilmunen and Palmroos (2013)
Insolvency cases cannot be observed, but zero-default years are registered as zero-default and zero-LGD events, is presented in Figure 4b. In this scenario the lower the LTV, the larger is the portion of observations at the origo. In this case the restricted correlation is higher than the absolute value of the unrestricted correlation. With this scenario the observed correlation can indicate quite large negative correlations even if the PD and LGD are not correlated, i.e. $\rho_{LN} = 0$.

Comparison of Figures 3 and 4 show that the way zero-loss defaults are observed and registered has a large effect on the observed correlation when homogeneous pools have been analyzed. In almost all the cases the observed correlation differs from the correlation of the unrestricted distribution. This means that the observed correlation cannot be used as such in simulations.

The way zero-loss defaults are registered is important information especially if correlations of different banks or countries has been compared.

Testing the effect of pool structure is performed comparing the observed correlations of a single structure against the correlation between $\rho_{LN}$. In this case the pool structure includes a large number of cohorts with different LTVs, amounts of loans, and total principals. Now the observed correlation can not be calculated as a function of LTV, i.e. the truncation/censoring point, because every cohort has its own. Instead of that, the LTV has been replaced by the previously mentioned variance of $\gamma$. Again the other exogenous variable is $\rho_{LN}$ as in the analysis of the homogeneous pools.

Figures 1 and 2 provide a picture on the mortgage pool structure used.
Figure 5. Observed correlations, when a) only LGD > 0 events have been included in calculations, and b) all cases, i.e. also LGD = 0, have been included. Distribution parameters used in calculations are: $E_I = 0.05$, $S_I = 0.02$, and $E_\gamma = 1.02$. $S_\gamma = \sigma_\gamma$.

Figure 5a presents the case where zero-loss insolvency cases are omitted, i.e. it is similar to Scenario one used with homogeneous pools. The effects of the heterogeneous structure are remarkable. In particular the differences between $\rho_{HeT}$ and $\rho_{LN}$ when the later is close zero, are worth noticing.

As can be seen, the observed correlation is very high even when there is no correlation between $\gamma$ and $I$. The effect of $EPOI$ is so strong that the model gives similar kinds of results to the case when $I$ is constant. Explanation for this can be found from the functions of $EPOI_T(\gamma)$ and $LGD_{HeT}$, and it can be seen from Figure 2.

In Figure 5b all insolvency events have been registered and used in calculations. In this case the observed correlation and the correlation of the unrestricted distribution are close to each other in all variable values except with the lowest standard deviations of $\gamma$. With the structure used it seems that the observed correlation is relatively close to the $\rho_{LN}$.

The result presented in Figure 5b is dependent on the LTV structure over the cohorts. If the LTVs are very close to each other, the pool structure begins to approach a homogeneous pool structure, and thus also the observed correlation starts to approach the one presented in Figure 4.

The observed correlations $\rho_{HeT}$ and $\rho_{HeC}$ are sensitive not only to $\rho_{LN}$ and $\sigma_\gamma$, but also to the structure of the mortgage loan pool. It should be noted that the structure of the pool is sensitive to the latest changes in collateral values, development of granted loans, and to changes in premature repayments.

The correlations of samples where $LGD = 0$ events have also been included, $\rho_{HeC}$, are less sensitive to changes in all these factors, and thus are better
estimates for unrestricted correlations $\rho_{LN}$, than $\rho_{HT}$.

To get the more reliable correlation estimates for calculations and simulations the unrestricted correlation, $\rho_{LN}$, needs to be estimated using other methods. If the heterogeneous mortgage loan pool is large enough, data of one or more homogeneous sub-pools can be used to estimate the unrestricted correlation. In the case that is not possible, or the estimated correlations are not reliable enough, the unrestricted correlation can be estimated using information on distributions and the correlation of $I$ and $\gamma$. In that case information on the usage of different collaterals, and correlations between them needs to be known, as well as understanding possible trigger events causing insolvencies.

8 Conclusion

This paper studied how the different recognizing and observing assumptions affect the observed correlation between PD and LGD, and whether the different mortgage pool structures have an effect on correlation.

The presented results indicate that both the Loan-to-Value and maturity structure of the pool, and the way the zero-loss defaults are registered, have a strong effect on the correlation. The presented results also indicate that the difference between observed and true correlation may be large enough to cause significant errors on credit risk models, and that observed correlation cannot be used in calculations or simulations as such. The findings increase our knowledge on the behavior of mortgage loan risk components with different pool structures and default registration assumptions, and thus help to create more reliable and accurate mortgage loan credit risk models.

In the case where the observed correlation is between the minimum and maximum boundaries of the population, the effects of restricted observation to correlation with homogeneous mortgage loan pools can be corrected using closed form formulas of correlation in censored and truncated bi-variate log-normal distribution. With heterogeneous pools this is not possible.

When the unrestricted correlation between PD and LGD needs to be known, there are at least two possible solutions. First, if the heterogeneous pools are large enough, a homogeneous sub-pool or multiple sub-pools can be selected, and those sub-pools can then be used for correlation calculations. Secondly, the correlation can be approximated using correlation between collateral values and insolvency cases. The second method needs information on the collaterals used and understanding of the causes of insolvencies.
The results presented in this paper can be used to develop more accurate mortgage loan credit risk models and to find reliable correlation parameters for the current models.

9 Acknowledgements

The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper.

References


Appendix A. Correlations between $\gamma$ and $I$, and between $LGD^*$ and $PD$

The correlation between $\gamma$ and $I$ can be written as

$$Corr(\gamma, I) = \frac{Cov(\gamma, I)}{Std(\gamma)Std(I)}$$

(17)

The Equation (6) illustrates the relationship between unrestricted $LGD$ and $\gamma$. The correlation between $LGD^*$, i.e. in the case where LGD can also become negative, and PD, where $PD = I$, the correlation between $PD$ and $LGD^*$ can written as

$$Corr(LGD^*, PD) = Corr(1 - \lambda \gamma, I)$$

$$= \frac{Cov(1 - \lambda \gamma, I)}{Std(1 - \lambda \gamma)Std(I)}$$

(18)

where $\lambda = LTV^{-1}$ and thus constant for a single cohort.

There are some well known variance, standard deviation, and covariance rules:

$$Var(a + x) = Var(x)$$

$$Var(ax) = a^2 Var(x)$$

$$Std(ax) = \sqrt{Var(ax)} = |a| Std(x)$$

$$Cov(a + x, b + y) = Cov(x, y)$$

$$Cov(ax, by) = ab Cov(x, y)$$

Thus, if $\lambda \neq 0$, then

$$Corr(LGD^*, PD) = \frac{Cov(1 - \lambda \gamma, I)}{Std(1 - \lambda \gamma)Std(I)}$$

$$= \frac{-\lambda Cov(\gamma, I)}{|\lambda| Std(\gamma)Std(I)}$$

(20)

$$= \begin{cases} 
\frac{Cov(\gamma, I)}{Std(\gamma)Std(I)}, & \text{if } \lambda < 0 \\
\text{not defined}, & \text{if } \lambda = 0 \\
-\frac{Cov(\gamma, I)}{Std(\gamma)Std(I)}, & \text{if } \lambda > 0 
\end{cases}$$

(21)

If $\lambda = 0$, then $1 - \lambda \gamma = 1$, and the correlation between a constant and any variable equals zero.

If $\lambda > 0$, as is the case in this paper, it can be seen that

$$Corr(LGD^*, PD) = -Corr(\gamma, I)$$

(22)

which is in line what has been presented in Section 4.
Appendix B. Finnish new unemployed data

The unemployment data has been collected from the tables presented in the Finnish Labour Review (Työpoliittinen aikakauskirja) 2/2016 vol. 59 (ISSN 0787-510X) published by Ministry of Employment and the Economy.

The ratio of new unemployed, i.e. the ratio of new insolvency cases \( I \), has been calculated using the formula

\[
I = \frac{UE_t}{E_{t-1}}
\]  

(23)

where \( UE_t \) is the number of new unemployment cases and \( E_{t-1} \) is the number of employed one year earlier. The number of new unemployed has been approximated using the number of unemployed whose unemployment period has been shorter than 52 weeks.

The data used is presented at Table 2.

<table>
<thead>
<tr>
<th>Year</th>
<th>UE(_t) ((UE))</th>
<th>E(_{t-1}) ((E))</th>
<th>(I) ((\%I))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>2 328</td>
<td>92.6</td>
<td>4.52</td>
</tr>
<tr>
<td>1981</td>
<td>2 353</td>
<td>106.4</td>
<td>4.71</td>
</tr>
<tr>
<td>1982</td>
<td>2 377</td>
<td>111.9</td>
<td>4.83</td>
</tr>
<tr>
<td>1983</td>
<td>2 390</td>
<td>108.4</td>
<td>4.97</td>
</tr>
<tr>
<td>1984</td>
<td>2 413</td>
<td>116.5</td>
<td>4.42</td>
</tr>
<tr>
<td>1985</td>
<td>2 437</td>
<td>121.2</td>
<td>4.97</td>
</tr>
<tr>
<td>1986</td>
<td>2 431</td>
<td>113.8</td>
<td>4.68</td>
</tr>
<tr>
<td>1987</td>
<td>2 423</td>
<td>107.2</td>
<td>4.42</td>
</tr>
<tr>
<td>1988</td>
<td>2 507</td>
<td>90.3</td>
<td>3.71</td>
</tr>
<tr>
<td>1989</td>
<td>2 504</td>
<td>90.3</td>
<td>3.60</td>
</tr>
<tr>
<td>1990</td>
<td>2 375</td>
<td>175.9</td>
<td>7.02</td>
</tr>
<tr>
<td>1991</td>
<td>2 266</td>
<td>293.3</td>
<td>12.35</td>
</tr>
<tr>
<td>1992</td>
<td>2 071</td>
<td>356.0</td>
<td>16.14</td>
</tr>
<tr>
<td>1993</td>
<td>2 054</td>
<td>340.6</td>
<td>16.45</td>
</tr>
<tr>
<td>1994</td>
<td>2 099</td>
<td>315.0</td>
<td>15.34</td>
</tr>
<tr>
<td>1995</td>
<td>2 127</td>
<td>303.0</td>
<td>14.44</td>
</tr>
<tr>
<td>1996</td>
<td>2 170</td>
<td>276.3</td>
<td>12.99</td>
</tr>
<tr>
<td>1997</td>
<td>2 222</td>
<td>251.6</td>
<td>11.59</td>
</tr>
<tr>
<td>1998</td>
<td>2 296</td>
<td>240.7</td>
<td>10.83</td>
</tr>
<tr>
<td>1999</td>
<td>2 335</td>
<td>224.9</td>
<td>9.80</td>
</tr>
<tr>
<td>2000</td>
<td>2 367</td>
<td>210.5</td>
<td>9.01</td>
</tr>
<tr>
<td>2001</td>
<td>2 372</td>
<td>204.9</td>
<td>8.66</td>
</tr>
<tr>
<td>2002</td>
<td>2 365</td>
<td>203.6</td>
<td>8.58</td>
</tr>
<tr>
<td>2003</td>
<td>2 365</td>
<td>204.3</td>
<td>8.64</td>
</tr>
<tr>
<td>2004</td>
<td>2 365</td>
<td>204.3</td>
<td>8.64</td>
</tr>
<tr>
<td>2005</td>
<td>2 401</td>
<td>193.7</td>
<td>8.19</td>
</tr>
<tr>
<td>2006</td>
<td>2 444</td>
<td>184.1</td>
<td>7.67</td>
</tr>
<tr>
<td>2007</td>
<td>2 492</td>
<td>164.6</td>
<td>6.73</td>
</tr>
<tr>
<td>2008</td>
<td>2 531</td>
<td>160.7</td>
<td>6.45</td>
</tr>
<tr>
<td>2009</td>
<td>2 457</td>
<td>223.9</td>
<td>8.85</td>
</tr>
<tr>
<td>2010</td>
<td>2 447</td>
<td>211.6</td>
<td>8.61</td>
</tr>
<tr>
<td>2011</td>
<td>2 474</td>
<td>187.3</td>
<td>7.65</td>
</tr>
<tr>
<td>2012</td>
<td>2 483</td>
<td>193.0</td>
<td>7.80</td>
</tr>
<tr>
<td>2013</td>
<td>2 457</td>
<td>221.2</td>
<td>8.91</td>
</tr>
<tr>
<td>2014</td>
<td>2 447</td>
<td>235.9</td>
<td>9.60</td>
</tr>
<tr>
<td>2015</td>
<td>2 437</td>
<td>243.6</td>
<td>9.96</td>
</tr>
</tbody>
</table>

For the sample available \( \text{Avg}(I) = 0.0846 \) and \( \text{Std}(I) = 0.0360 \), and
Avg[Log(I)] = −1.1113 and Std[Log(I)] = 0.1889.

Testing of normality of Log(I) using three different tests gives the results presented in Table 3. In all the test rejecting \( H_0 \) indicates non-normality. The confidence level used is 95%.

**Table 3. Testing normality of Log(I)**

<table>
<thead>
<tr>
<th>Test</th>
<th>( H_0 )</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>Stays</td>
<td>0.5422</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>Stays</td>
<td>0.3587</td>
</tr>
<tr>
<td>Lilliefors</td>
<td>Stays</td>
<td>0.1277</td>
</tr>
</tbody>
</table>

None of the tests rejects \( H_0 \). Based on this result it can be assumed that Log(I) is Normally distributed and I Log-normally distributed.
Appendix C. Finnish Housing Price Index data

The Housing Price Index data used has been collected from the tables presented at the web page of Statistics Finland (www.stat.fi). The time series used to approximate the development of collateral values is ‘Price index of old dwellings in housing companies: Whole country’.

The relative change in collateral values \( \gamma \) has been calculated using the formula presented in Equation (1), assuming that \( C_t = HPI_t \).

The data used is presented in Table 4. The initial time period used is the year 1983, i.e. index value of HPI was 100 at year 1983.

Table 4. Housing Price Index using old dwellings and calculated \( \gamma \)

<table>
<thead>
<tr>
<th>Year</th>
<th>HPI ((C))</th>
<th>(HPI_{t-1}) ((\gamma))</th>
<th>Year</th>
<th>HPI ((C))</th>
<th>(HPI_{t-1}) ((\gamma))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>184.2</td>
<td></td>
<td>2002</td>
<td>237.0</td>
<td>1.0743</td>
</tr>
<tr>
<td>1989</td>
<td>225.2</td>
<td>1.2226</td>
<td>2003</td>
<td>252.1</td>
<td>1.0637</td>
</tr>
<tr>
<td>1990</td>
<td>212.5</td>
<td>0.9436</td>
<td>2004</td>
<td>270.6</td>
<td>1.0734</td>
</tr>
<tr>
<td>1991</td>
<td>182.6</td>
<td>0.8593</td>
<td>2005</td>
<td>287.0</td>
<td>1.0606</td>
</tr>
<tr>
<td>1992</td>
<td>150.3</td>
<td>0.8231</td>
<td>2006</td>
<td>308.4</td>
<td>1.0746</td>
</tr>
<tr>
<td>1993</td>
<td>138.6</td>
<td>0.9222</td>
<td>2007</td>
<td>325.4</td>
<td>1.0551</td>
</tr>
<tr>
<td>1994</td>
<td>146.6</td>
<td>1.0577</td>
<td>2008</td>
<td>327.2</td>
<td>1.0055</td>
</tr>
<tr>
<td>1995</td>
<td>141.4</td>
<td>0.9645</td>
<td>2009</td>
<td>326.3</td>
<td>0.9972</td>
</tr>
<tr>
<td>1996</td>
<td>149.0</td>
<td>1.0537</td>
<td>2010</td>
<td>354.8</td>
<td>1.0873</td>
</tr>
<tr>
<td>1997</td>
<td>175.2</td>
<td>1.1758</td>
<td>2011</td>
<td>364.3</td>
<td>1.0268</td>
</tr>
<tr>
<td>1998</td>
<td>193.1</td>
<td>1.1022</td>
<td>2012</td>
<td>370.3</td>
<td>1.0165</td>
</tr>
<tr>
<td>1999</td>
<td>210.1</td>
<td>1.0880</td>
<td>2013</td>
<td>376.2</td>
<td>1.0159</td>
</tr>
<tr>
<td>2000</td>
<td>222.5</td>
<td>1.0590</td>
<td>2014</td>
<td>374.1</td>
<td>0.9944</td>
</tr>
<tr>
<td>2001</td>
<td>220.6</td>
<td>0.9915</td>
<td>2015</td>
<td>371.0</td>
<td>0.9917</td>
</tr>
</tbody>
</table>

For the sample available \( \text{Avg}(\gamma) = 1.0296 \) and \( \text{Std}(\gamma) = 0.0838 \), and \( \text{Avg}[\log(\gamma)] = 0.0259 \) and \( \text{Std}[\log(\gamma)] = 0.0833 \).

Testing of normality of \( \log(\gamma) \) using three different tests gives the results presented in Table 5. In all the test rejecting \( H_0 \) indicates non-normality. The confidence level used is 95%.

None of the tests rejects \( H_0 \). Based on this result it can be assumed that \( \log(\gamma) \) is Normally distributed and \( \gamma \) Log-normally distributed.
Table 5. Testing normality of $\log(\gamma)$

<table>
<thead>
<tr>
<th>Test</th>
<th>$H_0$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>Stays</td>
<td>0.4934</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>Stays</td>
<td>0.0780</td>
</tr>
<tr>
<td>Lilliefors</td>
<td>Stays</td>
<td>0.0960</td>
</tr>
</tbody>
</table>

Appendix D. Fitted and reported Loan-to-Value structures used for calculations

![Fitted and reported LTV distributions](image)

Figure 6. Original reported and fitted LTV distributions aggregated over all reporting banks.
Part III
DOI: 10.1016/j.ijforecast.2013.09.007

Copyright ©2013 International Institute of Forecasters.
Published by Elsevier B.V.
Reprinted with permission.
The Delphi method in forecasting financial markets – an experimental study

Karlo Kauko¹, Peter Palmroos²

¹Bank of Finland, PO Box 160, 00101 Helsinki, Finland
²Financial Supervisory Authority, PO Box 103, 00101 Helsinki, Finland

Abstract

Experts were used as Delphi panellists and asked to present forecasts on financial market variables in a controlled experiment. We found that respondents with the least accurate or least conventional views were particularly likely to modify their answers. Most of these modifications were in the right direction but too small, probably because of belief-perseverance bias. This paper also presents two post survey adjustment methods for Delphi method based forecasts. First, we present a potential method to correct for the belief perseverence bias. The results seem promising. Secondly, we test a conditional forecasting process, which unexpectedly proves unsuccessful.

1 Introduction

The Delphi method was introduced in the 1950s at the RAND Corporation. It aims to maintain the advantages of an interacting group without potentially counterproductive group dynamics, such as dominant individuals who may not be the best experts.

In short, the traditional version of the method is based on a multi-round survey. Respondents are asked to answer a number of questions in writing.
Answering is anonymous; other respondents do not know who answered what. In most cases the answers are either numeric estimates, ratings on a scale, or yes or no. Often the respondents also have the opportunity to write comments on the issues raised in the questionnaire. Thereafter statistics on answers and related comments are distributed to respondents, but this information is anonymous and no respondent can identify who answered what. Each respondent is allowed to modify his own answers, and possibly to add more comments.

After a few rounds, some convergence in answers is normally observed due to a group opinion building process, leading to less variance in answers and more agreement within the panel. The number of rounds can be either predetermined or it may depend on the criteria of consensus and stability. In the papers reviewed by Rowe and Wright (1999), the number of respondents in Delphi panels varied from 3 to 98. Ideally, respondents are experts in the same field, but with somewhat different backgrounds.

Linstone and Turoff (2011) emphasised the role of communication in judgemental forecasting, and argued that the Internet will have a large impact on how comparable methods will be used in the future; the number of potential participants will be much larger than in traditional Delphi panels. In recent years, some studies have used the real-time version of the method (see Gordon and Pease, 2006), which is normally an on-line application that allows respondents to modify their answers at any time, up until the end of the answering process (for validation of the method, see Gnatzy et al., 2011).

The final answer of the group is defined as the mean or median of the individual answers. In many cases even the questions to be answered are proposed and selected by group members themselves before the first answering round. Forecasting accuracy of the group normally improves over Delphi rounds, and the Delphi method works better than staticized groups, i.e. simple one round surveys; this finding is reported by Parenté et al. (2005), Helmer (1964), Dalkey (1968), Graefe and Armstrong (2011), Song et al. (2013) and in various studies reviewed by Rowe and Wright (1999). In some experiments respondent groups have been asked to provide estimates on almanac events, i.e. issues related to the past and present (see e.g. Graefe and Armstrong, 2011); whereas in other cases they have been asked to forecast the future (Parenté et al., 1984; Parenté et al., 2005). More detailed descriptions can be found in Linstone and Turoff (1975), Parenté and Anderson-Parenté (1987) or Rowe et al. (1991).

In light of most of the earlier experimental research, the Delphi method seems either substantially (Basu and Schroeder, 1977; Riggs, 1983) or some-
what better (Graefe and Armstrong, 2011) than Face-to-Face (FTF) meetings, although some authors have found the differences to be negligible (Brockhoff et al., 1975). Findings by previous authors have been summarised by Woudenberg (1991, Table 3) and Rowe and Wright (1999, Table 4); according to both surveys, most previous contributions had found that, with some exceptions, the Delphi method had outperformed traditional meetings.

So is it reasonable to use sophisticated structured processes if there is no unambiguous evidence that they would yield significantly better forecasts than a simple FTF meeting? As will be seen in this paper, structured techniques have at least one clear advantage: they can be improved and the forecasts honed.

In our study, the FTF meeting was used as the benchmark case. This is the simplest and probably most often used method; group members sit in the same room and discuss the issues until they either reach a consensus, or at least that the majority backs a view. According to Kerr and Tindale (2011), such meetings are good for pooling information, mutual error checking and motivation enhancement. On the other hand, they may be particularly vulnerable to ‘tyranny of the majority’, dominance of powerful individuals, inattention to unshared information, or group overconfidence. Other potential problems of FTF meetings include the bandwagon effect (tendency of ideas to spread among people like fads), the underdog effect (tendency of some people to vote for losing candidates or views), and the halo effect (tendency to weight an opinion according to a general impression of the person who expresses it).

According to Ang and O’Connor (1991, pp. 142), the Delphi method combines both mathematical and behavioural approaches, with an ‘aim to improve behavioural aggregation by substituting the dysfunctional aspects of achieving consensus with a mathematical process of achieving the final group judgement’. In the best case, the method helps to eliminate a number of problems of FTF meetings, such as the influence of dominant individuals and the unwillingness of many people to defend unorthodox views, even well founded ones. Different biases in the Delphi method have also been studied. Ecken et al. (2011) discuss the desirability bias: the general tendency of respondents to over-estimate the probability of events they consider desirable. Unfortunately, no direct test on the ability of the proposed correction to improve forecast accuracy was presented, but evidence was given on the existence of the bias.

This paper has two main objectives. First, we address the development of individual answers during the process. Secondly, we try to develop the Delphi method further, using observations on panellists’ behaviour and findings from
existing research in psychology.

The use of post survey methods to increase the accuracy of forecasts is not a new idea. Armstrong (2006) listed and evaluated evidence on numerous post hoc methods. Our methods however differ from those presented by Armstrong. Instead of assigning different weights to the panellists depending on, for example, previous forecasting performance, we corrected the assumed undersized adjustments the panellists had made between the first and last rounds.

This paper is organised as follows. In the next section we provide a number of hypotheses that are subsequently tested empirically. Section 3 details the experiment. Section 4 focuses on forecasting accuracy and Section 5 on examining the dynamics of individual forecasts in the Delphi context. Section 6 examines post-forecast correction methods, and the last section concludes.

2 Hypotheses development

Instead of comparing the Delphi and FTF meeting methods, the main aim of this paper is to study whether the post survey methods we used can reduce forecast errors. This was done by testing a number of hypotheses with data from a controlled laboratory experiment. We present three hypotheses on the modification of individual answers.

Hypothesis 1: Individual answers improve more than average if they originally differed from the group average.

To express the same idea in statistical terminology, there is a negative correlation between changes in the absolute value of a respondent’s forecast error \( |ψ_{j,p,\text{final}}| - |ψ_{j,p,1}| \) and the original deviation from the group mean \( |X_{j,p,1} - \overline{X}_{j,1}| \), where \( ψ_{j,p,n} \) is a measure of respondent \( p \)'s answering inaccuracy in round \( n \), question \( j \), \( X_{j,1} \) is the 1st round group answer (mean of individual answers) to question \( j \), and \( X_{j,p,1} \) is respondent \( p \)'s 1st round answer to question \( j \).

Hypothesis 2: Individual answers improve more than average if they were originally more inaccurate than average.

Using statistical terminology, there is a negative correlation between changes in the absolute value of the error \( |ψ_{j,p,\text{final}}| - |ψ_{j,p,1}| \) and size of the original error. \( |X_{j,p,1} - Y_j| \), \( Y_j = \) correct answer (observed later).

Hypothesis 3: Being an outlier and being inaccurate have an interacting effect on improvements in answers – being simultaneously an outlier and inaccurate predicts an improvement in accuracy that exceeds the sum of the two separate effects.
Hypotheses 1, 2 and 3 follow from previous research and common sense. The experimental findings at RAND in the early years of the method indicated that, at least for almanac events, most improvements happen because deviant members move towards the centre (Dalkey, 1968). The findings of Rowe and Wright (1996) also indicate that those who give the least accurate forecasts will modify their opinions more than others. If an answer is either an outlier or remarkably inaccurate, it may be more random and based on much less consideration and much less expertise than typical answers. The respondent is aware of this and adjusts the forecast, partly in response to others’ views. In a different kind of experiment, respondents even adjusted their answers closer to an intentionally fed false group consensus (Scheibe et al., 1975). To take an example, a payment system specialist may understand that she does not know much about banks’ loan losses. The respondent begins with an estimate that is out of line with the rest of the group and will turn out to be a very poor forecast. She understands that her initial forecast was erroneous, which increases the willingness to adjust the forecast in line with the views of the rest of the group. The Delphi method and feedback from other respondents may not work in the expected way unless the probability of adopting a new opinion correlates with both the initial error and the deviation from others’ views. Moreover, it is reasonable to believe that these two effects produce an interaction effect, at least if the method works as it should. Studying others’ answers is an essential part of the Delphi method, and this feedback is useful if and only if it causes reactions that improve accuracy. Normally, a reaction is a move towards group average, which is useful if and only if the others are generally more accurate. In an ideal situation relatively accurate answers are not modified in response to anything. If the only respondent who provides a relatively accurate forecast modifies her views towards an inaccurate group average, the Delphi method will be counterproductive.

We also test a hypothesis with direct applicability in real-world forecasting.

Hypothesis 4: Most modifications in individual forecasts will be in the right direction but too small.

The unwillingness to admit the unreliability of one’s original forecast would be a manifestation of a well known phenomenon documented by psychologists, known as the belief perseverance or confirmatory bias. The human mind does not normally update beliefs like a true Bayesian would; instead, there is a strong tendency to underreact to new information and even to ignore information that is not consistent with prior beliefs. This phenomenon is closely related to overconfidence or egocentric discounting, the general tendency of
people to overestimate the reliability of their 'private signals' (see e.g. Fellner and Krügel, 2012). Nickerson (1998) even concludes that 'The evidence supports the view that once one has taken a position on an issue one’s primary purpose becomes that of defending or justifying that position'. Empirically, when one must draw conclusions based on information acquired over time, information acquired at earlier stages is more heavily weighted in final conclusions (Sherman et al., 1983). Testees have been observed to stick with intentionally fed misinformation even when told that the so-called 'research results' were pure fiction (Anderson et al. 1980). Lord et al. (1979) found that people may even interpret the same new information differently and systematically in favour of their own previous beliefs. More psychological literature on this phenomenon was reviewed in detail by Nickerson (1998). The potential elimination of this bias in meta-analysis has been briefly discussed by Rabin and Schrag (1999). This psychological phenomenon may even carry fundamental macroeconomic consequences by affecting inflation and output dynamics (Adam, 2007).

We also tested two post-correction methods.

The first method is directly derived from Hypothesis 4. If respondents typically modify their answers in the right direction but by too little, one can obtain better forecasts by scaling up modifications in individual forecasts ex post.

The second method can be used when new information on causally relevant background variables becomes available. Because our panellists were qualified professionals in finance and economics, we assumed that they would understand how the variables to be forecasted depend on the interest rate. Therefore, we expected that a substantial portion of the forecast errors would be due to unexpected shocks to the market interest rate, so that eliminating this error would lead to much more accurate forecasting.

### 3 The experimental design

#### 3.1 Preparation of the experiment

We organised a regimented experiment with two similar expert panels. The middle management of the Bank of Finland\(^1\) (BoF) and the Financial Supervisory Authority of Finland (FSA) showed interest in experimental research on

\(^1\)The Bank of Finland is the national central bank.
forecasting, possibly because these organisations routinely produce different kinds of financial and macroeconomic forecasts. We were thus able to arrange the experiment as an in-house project in which experts were asked to forecast developments in financial markets. For getting the full benefits, the Delphi method requires a sufficiently large number of panellists. According to Rowe and Wright (2001) and references therein, ten persons should be sufficient; there is no evidence that the accuracy improves by increasing the size of the panel beyond 7–10 persons. In order to have two similar panels of equal size, we needed 20 participants, which we were able to obtain. This was probably close to the maximum of scarce human resources we would have been able to recruit. The respondents were asked to participate by their direct supervisors, but no one was forced to take part.

Most previous experiments (such as Parenté et al., 1984; Bolger et al., 2011; and Graefe and Armstrong, 2011) have used non-experts as respondents. There may be practical reasons for the use of e.g. students instead of experienced professionals in experimental research. However, in real life, non-specialists are seldom relied on when significant policy and business decisions are to be made. Results obtained with panels consisting of first-year students answering questions related to headline news may not extend to expert groups (see e.g. Rowe et al., 1991).

In our panels all the respondents were experts in finance, working either for the FSA or for the BoF’s financial markets department. Each respondent professionally monitors one or more sectors of domestic and international financial markets. With one exception, each participant had several years of experience in the field. All of them have at least an MSc degree in a relevant field, such as economics, finance, law or business administration, and many have doctorates or licentiate degrees.

According to previous studies (Hussler et al., 2011, p 1650; Rowe and Wright, 1996) experts seem to behave differently from non-experts in Delphi surveys; at least they are less likely to modify their answers between rounds. One of the few forecasting experiments with experienced professionals in finance was carried out in Germany in the 1970s; differences in the accuracy of Delphi panels and FTF meetings were minor (Brockhoff et al., 1975). A recent contribution was presented by Song et al. (2013) who examined how forecasts based on statistical models could be improved by Delphi panels of experts, but they did not test different ways of deriving the collective subjective assessment of a group.

Our twenty experts were randomly separated into two groups. One group
answered questions on Finnish financial markets in a Delphi setting while the other group arranged an FTF meeting to answer the same questions.

The expert groups, initially named Group One (later FTF group) and Group Two (later Delphi group), were formed randomly, except that both FSA employees and BoF employees were divided evenly between the two groups. Moreover, there were specific 'gender quotas': both groups originally included two women and eight men. However, one male participant was replaced by a female in Group One, which eventually ended up being the FTF group. This replacement took place before the start of the experiment. Most of the respondents knew each other before the experiment, and had worked together on numerous occasions. Before the first forecasting round all participants were informed of the outcome of the grouping.

Even though answering in the Delphi group was basically anonymous, we were able to follow the development of individual answers. Each respondent was asked to choose a pseudonym that would not reveal her identity and not to disclose the pseudonym to anyone. These pseudonyms ranged from humorous to poetic, and from meaningless character strings to enigmatic descriptions of the respondent. Respondents were asked to write the pseudonym on each questionnaire to be returned during the process. This method enabled us to follow the development of individuals’ assessments while largely retaining the anonymity requirement of the Delphi method. Needless to say, a respondent may be somehow emotionally involved with his pseudonym.

### 3.2 The questionnaire

The number of variables to be forecasted was limited to 15 by the organisers. This was a compromise between overburdening our respondents and having the problems of too little data. Each question was related to the future of domestic financial markets, except for the last question on the development of the Euribor rate, which is the interbank money market rate for the whole euro area. Unlike in the experiments by e.g. Bolger et al. (2011), each forecast was numeric, so that perfectly accurate answers could not be realistically expected. The original questions (translated from Finnish), realised values of variables (not known by anyone when the questions were answered in May 2011) and some statistics on answers are presented in the Appendices A and B. Each answer was a simple point estimate – there were no questions on confidence intervals.

We used a paper-and-pencil\(^2\) form questionnaire. Unlike in many Delphi

\(^2\)Using the categorization of Parenté and Anderson-Parenté (1987), our method falls
studies, the topic and precise questions were chosen by the experiment organisers. This approach was used partly to ease the burden on respondents, but the idea was also to ensure that we could avoid certain problems in the following stages of the project, such as questions with no objectively correct answer or questions on excessively highly correlated phenomena. If there are many close dependencies between variables, the composing of internally consistent forecasts might turn out to be impossible (Turoff, 1972). It would also be meaningless to ask for forecasts of variables like the stock of deposits in consecutive months, because not much additional evidence can be collected by asking virtually the same question several times. With the exception of the Euribor rate, the intention was to choose questions that would fall within the field of expertise of the respondents, and at the same time would be as independent of each other as possible. It is difficult to evaluate in an objective way how unrelated such questions are, but at least one observation indicates that the variables are not merely different manifestations of one or two underlying factors; correlations between the 20 first round answers to each question were not alarmingly high; there were only seven pairs of questions for which the 20 original answers were statistically significantly correlated at the 5% level; the expected value of the number of such correlations would be $0.05 \times (15 \times 15 - 15)/2 = 5.25$. Based on the binomial distribution, we cannot reject the null hypothesis that the answers are non-correlated.

Participants were given a document containing some statistics on past developments of each variable to be predicted. The background data did not include information on any variable other than those to be forecasted. These data were given in order to ease the burden on respondents. In principle, each respondent would also have otherwise had access to the same data, with some effort.

### 3.3 First round and the lottery

The first round started when questionnaires were distributed by email on May 18th 2011. The participants had a week to answer the questions and return the printed forms by May 25th in internal mail. Each participant was given the same questions and was asked to answer every one.

The respondents and organisers met on Monday, May 30th 2011 at 9:30 am. It was checked that each participant knew which group he was in. The first step of the meeting on May 30th was a lottery by which the groups’ forecasting methods were randomly assigned. Group Two was chosen to continue the in the Conventional Delphi rather than the Paper-and-pencil Delphic polls category.
Delphi process, and so Group One was asked to discuss the questions in an FTF meeting. Since one member of the Delphi group (former Group Two) was prevented from participating, the two remaining rounds of the experiment were carried out with nine respondents.

3.4 Second and third Delphi rounds and the Face-to-Face meeting

The Delphi group had two additional forecasting rounds. No convergence criteria were used to determine the number of iterations. Consensus can be measured in different ways (von der Gracht, 2012), and at least some previous research indicates that three rounds suffice for stability (Rowe and Wright, 1999) and accuracy (Parenté et al., 2005). The Delphi group was provided with information on previous-round answers by the group itself, including the answers of each respondent, the group average and the median for each question. Moreover, all written comments were distributed to participants. Rounds 2 and 3 of the Delphi process were conducted on the same day (May 30th), with a short break between rounds.

Group One was now the FTF meeting group. To be precise, this was an 'estimate-talk' exercise, not a pure FTF meeting because members had already answered the same questions in writing without circulating answers among themselves. No information on first round answers was disclosed to any participants of the meeting group; feedback on first round answers is an essential part of the Delphi method, but is not part of the standard FTF meeting as we understand it. Anonymous voting was specifically forbidden by the organisers because the FTF meeting does not, by definition, derive its output from any anonymous methods. Most participants were relatively active in the discussions, and there seemed to be no dominant individual who would have pressured others to accept her views.

Immediately after the third round of the Delphi process each respondent was asked to fill out a post-survey form. Respondents had not been told before that they would be asked to do this. The form included questions on the process; on how reliable (on a scale of 1–5) the participant considers his own forecasts; and on how much attention the respondent had paid to others’ numeric forecasts and written comments. These meta questions were asked in respect of each test question separately, and they are specific to the Delphi method. The answers could shed some light on the role of feedback from other respondents and respondents’ propensity to modify their answers.

Figure 1 outlines the conduct of the experiment.
1st round
Each respondent answers the same 15 questions

Lottery
The 2nd group chosen to continue the Delphi process

2nd round
Respondents see each others’ anonymised 1st round answers, answer the same 15 questions again

3rd round
Respondents see each others’ anonymised 2nd round answers, answer the same 15 questions again

Post survey questionnaire

Group 1
FTF meeting

Meeting
The same 15 questions answered again; no feedback on 1st round answers; answers decided in a Face-to-Face meeting

Group 2
Delphi

Figure 1. The forecasting process
3.5 Conditional forecasting and conditionalised questions

Normal unconditional forecasts can be seen as static point estimates, and once made, forecasts are hard to update even if values of exogenous variables change. Point forecasts can be replaced by forecasts conditional on some observable macro environmental variables, such as the phase of the economic cycle or the federal funds rate. For example, Waggoner and Zha (1999) have developed a conditional forecasting method for econometric modelling, and they cite the updating of forecasts as one application.

The idea of conditional forecasts is based on the assumption that every respondent uses an implicit personal model; the forecasted variable is a function of exogenous variables. Forecast errors may be due to unexpected developments in exogenous variables or model inadequacies. Respondents may share the same dysfunctional representations (Kerr and Tindale, 2011) and make similar errors in forming expectations of the environment, which can cause systematic errors in forecasts. This, however, can be corrected post hoc if new information on the exogenous variable becomes available.

It would be a significant burden to update the forecast by arranging a new full-scale forecasting round once the economic situation of the markets changes. However, in many cases updated forecasts of some background variables, such as interest rates, are always readily available in derivatives markets. If we know how experts update their forecasts in response to such new information, their original forecast could be updated at very little cost.

The Euribor rate is nearly exogenous to the Finnish economy and is assumed to affect most sectors of the economy. For this reason we chose it in advance as the causally relevant exogenous variable for conditioning.

To find the forecasted variables that panellists assume to be the most sensitive to changes in the designated exogenous variable, we asked everyone in round 1 to select the five most interest rate-sensitive variables. The seven most voted for variables were used for conditional forecasting. The interest rate may not affect all the variables, and we did not want to overburden our respondents by testing the method for every question. Hence, this correction method was tested with these seven variables only, the number being somewhat arbitrary.

As can be seen in Table 1, after round 1, both groups were unanimous in the selection of the most sensitive variables, albeit some respondents chose less than five variables. Questions 1, 2 and 9 each got three votes, but only one of them was arbitrarily chosen for conditional forecasting by the organisers.

In rounds 2 and 3, respondents in the Delphi group were asked to provide
Table 1. Number of respondents who thought the variable would be particularly sensitive to market rates and variables for conditional forecasting

<table>
<thead>
<tr>
<th>j</th>
<th>Question</th>
<th>Group 2 Selected</th>
<th>Delphi</th>
<th>FTF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Big banks' loan losses</td>
<td>x</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Stock market turnover</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>BoF claims on credit institutions</td>
<td></td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>M2 deposits</td>
<td>x</td>
<td>4</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>Interest in corporate loans</td>
<td>x</td>
<td>5</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>Net interest income of small banks</td>
<td>x</td>
<td>8</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>Serial bonds</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>Households' mutual fund shares</td>
<td>x</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>Stock of corporate bonds</td>
<td></td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>Banks' foreign net assets</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>Mean deposit rate of term deposits</td>
<td>x</td>
<td>8</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td>Interest in equities</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>Foreign ownership in Nokia</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>Stock of housing loans</td>
<td>x</td>
<td>7</td>
<td>7</td>
<td>14</td>
</tr>
</tbody>
</table>

Total 7 41 39 80

Some respondents chose less than five variables.

two different forecasts, on different background assumptions, for the selected seven variables. The baseline forecast assumed that the interest rate would develop as expected by the respondent. The respondents were also asked to provide a forecast based on the assumption that the interest rate would be 100 basis points (one percentage) higher than in the baseline forecast. For this alternative forecast, respondents were asked to take into account all the direct and indirect effects the change in the exogenous variable can cause or reflect. In round three, respondents were given statistical summaries on both types of answers from round 2, and they submitted new answers. The meeting group was then asked to provide conditional forecasts of the same interest rate-sensitive variables. Assuming linear dependence between variables, this approach enabled us to calculate the derivatives of the seven selected variables with respect to the market rate.

### 3.6 Penalty points and independence of questions

In papers such as Rowe and Wright (1996), changes in forecast errors have been measured using mean absolute percentage error (MAPE) and z-scores. We introduced a penalty point method to evaluate each answer, be it an individual answer in a questionnaire or the output of a forecasting group.
The penalty points for each forecast were equated to the absolute value of the difference between forecast and outcome values divided by the standard deviation of first round answers:

\[ \psi_{j,p,n} = \frac{|x_{j,p,n} - y_j|}{StDev(x_{j,1})} \]  

(1)

where \( \psi_{i,p,n} \) is the standardized penalty points of respondent or group \( p \), question \( j \), round \( n \), and \( x_{j,p,n} \) is the forecast; \( y_j \) the actual outcome, \( StDev(x_{j,1}) \) the standard deviation of all 20 round 1 answers to question \( j \), and \( |x_{j,p,n} - y_i| \) the absolute value of the forecast error.

The measure calculated according to Equation (1) is robust to the scale used; if the participants had to forecast temperatures, the choice between Fahrenheit, Celsius and Kelvin scales would not affect the amount of penalty points, unless the choice of scale affects respondents’ expectations. Respondents could even be asked to forecast growth rates of whichever variable instead of its future values without affecting the penalty points\(^3\). If a question is particularly difficult, the above measure is more lenient towards inaccuracies because, as reported already by Dalkey (1968), respondents’ answers probably differ significantly, making the denominator larger. If one divides the forecast error by the realised value, which has sometimes been done in forecasting experiments, the measurement scale would matter, and problems would arise if the realised value turned out to be negative or zero. It would be possible to use the squared value of the difference instead of its absolute value, but this would place disproportionate weight on questions with the most surprising outcomes and, as will be seen, differences between methods would depend almost exclusively on the third question.

It would also be possible to calculate the percentage reduction in error\(^4\) instead of the reduction in penalty points. This would not reverse the results; for the Delphi group, the correlation across questions between the change in standardized penalty points of the group and the percentage reduction in the error was 0.73 (N=15). This alternative method would almost ignore question

\(^3\)Consider the following fictional example. Three respondents were told that the population of France was 63.601 million in January 2007. They were asked about the population in January 2012. Their answers were 65.71, 65.90 and 65.45 million and the correct answer 65.35 million. The standard deviation of forecasts is 0.226, and consequently respondents’ penalty points are 1.59, 2.44 and 0.44. If they had been asked about the growth in percent, and if reframing the question had not affected expectations, their answers would have been 3.31%, 3.62% and 2.91% and the standard deviation 0.00355. Respective penalty points are 1.59, 2.44 and 0.44, i.e. precisely equal to those mentioned above.

\(^4\)\[ 100 \times \left( \frac{|x_{j,p,n} - y_j|}{|x_{j,p,n} - y_j|} - 1 \right)\% \]
three; the initial error was so huge for both groups that no change in forecast error would make much of a difference relative to the original error.

4 Group performance

In evaluating group performance, the average of individual answers for any particular question is taken as the first round group answer, irrespective of how the group continued its forecasting exercise. The two groups got roughly equal penalty scores in the first round, which should not be a surprise (see Table 2). In most cases the groups presented relatively accurate forecasts, but they both failed completely to foresee the large increase in claims on credit institutions on the balance sheet of the Bank of Finland (there is no scaling error for question 3 in Table 2). Using standardised penalty points, the Delphi group outperformed the FTF meeting at the subsequent stages of the process, but the difference in final penalty points was negligible. Moreover, the results depend on the definition of performance. Based on the average percentage reduction in error (see footnote 3), the FTF meeting improved its accuracy slightly more than the Delphi group (15.6%, vs. 11.3%). The surprisingly accurate final answer of the FTF meeting for question 4 strongly affects this result. These findings were similar to those of Graefe and Armstrong (2011) and Brockhoff et al. (1975): performance differences between a Delphi survey and an FTF meeting are minor, at least in a laboratory experiment of this kind. Needless to say, real-world forecasting may differ from artificial experiments.

The Delphi group improved from round 1 to round 3 in 13 of the questions. If we assume that improvements and deteriorations are equally likely, the binominal distribution implies that the probability of at least 13 improvements out of 15 would be about 0.5 per mill.

In the post-survey questionnaire, respondents in the Delphi group were asked how much attention they paid to each others’ numeric answers. This was done using a scale of 1–5, where 1 meant ‘not at all’ and 5 ‘to a very large extent’. The average varied only a little across the questions, from 2.9 to 3.25. When questions conditioned on interest rates were ignored, we founded that in about nine per cent of all answers (12 cases) a respondent said that she had paid no attention to others’ numeric answers; one might question whether the Delphi method is of any use if respondents behave this way. Respondents were not particularly eager to provide written comments. Only three respondents in the Delphi group gave written comments in round one, and in total there were 26 comments. In round 2, only two group members used the opportunity to
Table 2. Penalty points by question

<table>
<thead>
<tr>
<th>j</th>
<th>Question</th>
<th>Group 1</th>
<th>Group 1</th>
<th>FTF</th>
<th>Change in ( \psi )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>round 1</td>
<td>FTF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Big banks’ loan losses</td>
<td>0.98</td>
<td>0.80</td>
<td>-0.17</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Stock market turnover</td>
<td>1.08</td>
<td>0.34</td>
<td>-0.74</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>BoF claims on credit institutions</td>
<td>21.09</td>
<td>22.14</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>M2 deposits</td>
<td>0.69</td>
<td>0.09</td>
<td>-0.60</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Interest in corporate loans</td>
<td>3.27</td>
<td>3.59</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Net interest income of small banks</td>
<td>0.14</td>
<td>0.11</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Serial bonds</td>
<td>0.99</td>
<td>0.47</td>
<td>-0.52</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Households’ mutual fund shares</td>
<td>2.35</td>
<td>2.22</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Stock of corporate bonds</td>
<td>0.86</td>
<td>0.80</td>
<td>-0.06</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Banks’ foreign net assets</td>
<td>6.63</td>
<td>6.84</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Mean deposit rate of term deposits</td>
<td>0.38</td>
<td>0.41</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Interest in equities</td>
<td>4.56</td>
<td>4.36</td>
<td>-0.20</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Foreign ownership in Nokia</td>
<td>0.80</td>
<td>0.67</td>
<td>-0.13</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Stock of housing loans</td>
<td>0.85</td>
<td>0.62</td>
<td>-0.23</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>12-month Euribor</td>
<td>2.05</td>
<td>3.06</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>3.12</td>
<td>3.10</td>
<td>-0.01</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>j</th>
<th>Question</th>
<th>Group 2</th>
<th>Group 2</th>
<th>Delphi</th>
<th>Change in ( \psi )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>round 1</td>
<td>Delphi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Big banks’ loan losses</td>
<td>1.14</td>
<td>1.01</td>
<td>-0.13</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Stock market turnover</td>
<td>1.40</td>
<td>0.99</td>
<td>-0.40</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>BoF claims on credit institutions</td>
<td>21.649</td>
<td>21.648</td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>M2 deposits</td>
<td>0.87</td>
<td>0.65</td>
<td>-0.22</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Interest in corporate loans</td>
<td>3.30</td>
<td>3.14</td>
<td>-0.16</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Net interest income of small banks</td>
<td>0.85</td>
<td>0.62</td>
<td>-0.23</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Serial bonds</td>
<td>2.14</td>
<td>2.18</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Households’ mutual fund shares</td>
<td>0.41</td>
<td>0.38</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Stock of corporate bonds</td>
<td>7.34</td>
<td>7.22</td>
<td>-0.12</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Banks’ foreign net assets</td>
<td>0.36</td>
<td>0.41</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Mean deposit rate of term deposits</td>
<td>4.12</td>
<td>4.12</td>
<td>-0.01</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Interest in equities</td>
<td>0.45</td>
<td>0.41</td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Foreign ownership in Nokia</td>
<td>0.38</td>
<td>0.22</td>
<td>-0.16</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Stock of housing loans</td>
<td>1.99</td>
<td>1.69</td>
<td>-0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>3.13</td>
<td>3.00</td>
<td>-0.12</td>
<td></td>
</tr>
</tbody>
</table>
### Table 3. Correlations between group forecast improvements and paying attention to others’ answers

<table>
<thead>
<tr>
<th></th>
<th>NumForec</th>
<th>VerbCom</th>
<th>Δ Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention paid to others’ numeric answers</td>
<td>1.000</td>
<td>0.647 ***</td>
<td>-0.641 **</td>
</tr>
<tr>
<td>Attention paid to others’ verbal comments</td>
<td>0.674 ***</td>
<td>1.000</td>
<td>-0.527 **</td>
</tr>
<tr>
<td>Change in penalty points</td>
<td>-0.641 **</td>
<td>-0.527 **</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**FIRST-ORDER PARTIAL CORRELATIONS**

<table>
<thead>
<tr>
<th></th>
<th>NumForec</th>
<th>VerbCom</th>
<th>Δ Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention paid to others’ numeric answers</td>
<td>1.000</td>
<td>0.723 ***</td>
<td>-0.650 **</td>
</tr>
<tr>
<td>Attention paid to others’ verbal comments</td>
<td>0.723 ***</td>
<td>1.000</td>
<td>-0.643 **</td>
</tr>
<tr>
<td>Change in penalty points</td>
<td>-0.650 **</td>
<td>-0.643 **</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Controlled for the rank order of the original forecast error
(1 = worst error, 15 = least serious error)

** = 5 % significance, *** = 1 % significance

Nevertheless, as can be seen in Table 3, paying attention to each others’ answers made forecasts more accurate. The correlations between studying others’ answers and improving in the forecasts were quite high. These correlations cannot be explained away by saying that both paying attention to others’ answers and accuracy improvement were due to the joint dependence of these factors on the initial error (controlling for the rank order of the initial error⁵), as the partial correlations were about as high as the zero-order correlations (see Table 3).

This finding is consistent with previous experimental research; the possibility to examine others’ answers would seem to be the most important aspect of the Delphi method (Parenté et al., 2005). On the other hand, it is also possible that respondents simply paid more attention to others’ views if they invested a lot of time and effort in finding a good answer, in which case the reported attention is a mere proxy for effort.

The members of the Delphi group were also asked to evaluate the reliability of their own forecasts for each question separately in the same post-survey questionnaire. There was a strong negative Spearman rank order correlation

---

⁵ Because most of the variation in \( \psi \)'s is due to the massive forecast errors concerning question three, the rank of the error is used instead of its value; the rank of the largest error equals 1, rank of the smallest error is 15.
(-0.75, N=15) between the average self-assessed accuracy and the final level of penalty points \( \psi_{j, \text{group}2,3} \) across the 15 questions, indicating that the group was able to assess its own forecasting abilities.

5 Individual answers

Respondents’ answers converged during the process, which is a common result in Delphi surveys (see e.g. Dalkey, 1968). The average standard deviations of answers of the nine respondents involved in rounds 1 to 3 declined by 17%. Only views on the 12-month Euribor rate and the net interest income of small banks diverged during the process, as measured by the standard deviation.

Little evidence was found on the correlation between personal forecasting accuracy and a willingness to write comments, albeit those who provided written comments had slightly more accurate initial forecasts. In the first round, five respondents wrote comments on several questions, and their average of penalty points was 3.14. The average of penalty points of the remaining 15 respondents was 3.36.6 Because only a few respondents commented extensively, we could not reliably distinguish between individual effects and respondent-question combinations.

One respondent did not answer question 15 in round 1, but we obtained complete sets of 15 x 9 answers for round three. In the following, the change from round 1 to round 3 is analysed. Of the 9 x 15 – 1 = 134 respondent-question combinations that could be followed from round 1 through round 3, there were 62 cases where the answer did not change, 19 where the accuracy deteriorated according to the criterion of Equation (1) and 53 where the accuracy improved. Based on the binomial distribution, the probability of 53 or more improvements in 72 modifications is about 0.01 per mill, unless improvements are more likely than deteriorations.

Findings reported by Dalkey (1968) indicate that if a respondent modifies an answer, the modification is very likely to be an improvement if the respondent is otherwise unlikely to change his answers. A ‘stubborn’ respondent is likely to modify an answer if and only if the modification is almost certainly an improvement, a finding that was observed in our data. On average, the number of unmodified answers of a respondent was 8.0. Those whose answers for at least eight questions remained unchanged (Group Holdouts, four respon-

6These average penalty points are higher than those reported in Table 2 because forecasts derived by averaging individual answers are systematically more accurate than individual answers.
dents) were compared to the rest of respondents (Group Opinion Swingers, five respondents). Among Group Holdouts, there were 21 changes in answers; only two of which resulted in an increase in penalty points. The remaining 19 modifications were improvements, so that the probability of a modification being an improvement is about 90%. Group Opinion Swingers changed answers in 51 cases; there were 34 improvements and 17 deteriorations, i.e. only about 67% of modifications were improvements. This difference is statistically significant; in a simple cross-tabulation the chi-squared equals 4.34, which is significant at the 5% level, and the Fisher-exact p-value equals 0.032.

Hypotheses 1–3 were tested using ordinary least squares (OLS). The explained variable was the difference between third and first round penalty points; so that negative values of the explained variable indicated improvements. This measure was regressed on the first round penalty points and a measure of initial disagreement with the rest of Group 2 members (Group Delphi). For respondent $p$ and question $j$, the disagreement with the rest of the group was defined as the difference between the individual forecast and the group average divided by the standard deviation of the 20 (or 19) first round answers.

$$OPPOS_{j,p} = \frac{|x_{j,p,1} - \frac{1}{9} \sum_{i=1}^{9} x_{j,i,1}|}{StDev(x_{j,1})}$$ (2)

where $x_{j,p,1}$ is the answer of respondent $p$ to question $j$ in round 1. Respondent and question-specific dummy variables were used as controls.

Both Hypotheses 1 and 2 were supported by the data; an answer was likely to improve if a respondent’s initial forecast was either particularly inaccurate or differed from the answers of the rest of the group. On the other hand, there was no support for Hypothesis 3; the interaction term of the initial forecast error and the disagreement measure of Equation (2) had no detectable explanatory power (see Table 4).

In the OLS analysis, the explained variable had the value 0 in a disproportionate number of cases. As already seen, not changing the answer was typical for certain individuals, but it may also have depended on the degree of disagreement with the group, and possibly on the initial accuracy of the forecast. As can be seen in Table 5, when controlled for respondent and question-specific dummy variables, a forecast was more likely to be modified if it differed from the group average, but the initial inaccuracy of the forecast is a poor (yet maybe not completely insignificant) predictor of the adoption of a new view. Both findings are consistent with those reported by Bolger et al. (2011); expertise is not closely related to opinion changes but being in a minority within
Table 4. Change in penalty points as a function of initial forecast error and disagreement with others

OLS, N=134
White heteroskedasticity-consistent standard errors and covariance
Dependent variable $\psi_{j,p,3} - \psi_{j,p,1}$

<table>
<thead>
<tr>
<th></th>
<th>Coeff</th>
<th>t-value</th>
<th>Coeff</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.680</td>
<td>5.1 ***</td>
<td>0.670</td>
<td>5.3 ***</td>
</tr>
<tr>
<td>$\psi_{j,p,1}$</td>
<td>-0.200</td>
<td>-4.7 ***</td>
<td>-0.197</td>
<td>-4.6 ***</td>
</tr>
<tr>
<td>OPPOS</td>
<td>-0.230</td>
<td>-3.4 ***</td>
<td>-0.229</td>
<td>-3.4 ***</td>
</tr>
<tr>
<td>$\psi_{j,p,1} \times$ OPPOS</td>
<td>-0.002</td>
<td>-1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.529</td>
<td></td>
<td>0.530</td>
<td></td>
</tr>
<tr>
<td>Mean dep var</td>
<td>-0.177</td>
<td></td>
<td>-0.177</td>
<td></td>
</tr>
<tr>
<td>F-stat</td>
<td>5.107</td>
<td>***</td>
<td>4.873</td>
<td>***</td>
</tr>
</tbody>
</table>

*** = 1 % significance

The sum of coefficients for the nine respondent-specific dummy variables forced to equal 0.
The sum of 15 question-specific dummy variables forced to equal 0. The sum of 9 respondent-specific dummy variables forced to equal 0.

the group is. The interaction term was non-significant. Again, respondent and question-specific dummy variables were used as controls.

When the analysis of Table 4 was re-run with only those cases where the answer was changed, disagreeing with the rest of the group was not found to explain improvements. Instead, having an initially inaccurate forecast had a very strong impact on subsequent improvements. Again, the interaction term was not significant (see Table 6).

Hypothesis 1 was confirmed. Disagreeing with the rest of the group increased the probability of adopting a new opinion, which was usually an improvement. However, as shown in Table 6, if an answer was changed, in response to a group view, this did not always increase the odds that the modification was an improvement.

Hypothesis 2 was confirmed. Being originally more inaccurate than the others may have somewhat increased the odds that the answer would be changed. More importantly, being badly inaccurate implied that the change would almost certainly be a major improvement.

Hypothesis 3 found no support in the data. There was no systematic tendency of inaccurate outlier forecasts to change, let alone improve, more than could be expected as the sum of the two separate effects.

Hypothesis 4 was also confirmed. Most changes in answers were in the right direction, but by too little. The overwhelming majority of forecast changes,
Table 5. Probability of changing the forecast as a function of initial forecast error and disagreement with others

LOGIT, N=134
Dep. variable = 1 if third round answer = first round answer

<table>
<thead>
<tr>
<th></th>
<th>Coeff</th>
<th>z-value</th>
<th>Coeff</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>3.247</td>
<td>2.7 ***</td>
<td>3.112</td>
<td>2.8 ***</td>
</tr>
<tr>
<td>$\psi_{j,p,1}$</td>
<td>-0.637</td>
<td>-1.7 *</td>
<td>-0.571</td>
<td>-1.6</td>
</tr>
<tr>
<td>OPPOS</td>
<td>-1.842</td>
<td>-3.0 ***</td>
<td>-1.854</td>
<td>-2.8 ***</td>
</tr>
<tr>
<td>$\psi_{j,p,1} \times$ OPPOS</td>
<td>-0.046</td>
<td>-1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

McFadden $R^2$ | 0.330 | 0.335 |
LR-stat | 61.01 *** | 61.847 |

* = 10 % significance, *** = 1 % significance

The sum of coefficients for the nine respondent-specific dummy variables forced to equal 0.
The sum of 15 question-specific dummy variables forced to equal 0. The sum of 9 respondent-specific dummy variables forced to equal 0.

Table 6. Change in penalty points, cases where forecast changed

OLS, N=72; Cases where the answer changed
White heterosked.-consistent standard error and covariance
Dependent variable ($\psi_{j,p,3} - \psi_{j,p,1}$)

<table>
<thead>
<tr>
<th></th>
<th>Coeff</th>
<th>z-value</th>
<th>Coeff</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.975</td>
<td>4.8 ***</td>
<td>0.939</td>
<td>4.8 ***</td>
</tr>
<tr>
<td>$\psi_{j,p,1}$</td>
<td>-0.330</td>
<td>-5.8 ***</td>
<td>-0.327</td>
<td>-5.4 ***</td>
</tr>
<tr>
<td>OPPOS</td>
<td>-0.127</td>
<td>-1.4</td>
<td>-0.129</td>
<td>-1.4</td>
</tr>
<tr>
<td>$\psi_{j,p,1} \times$ OPPOS</td>
<td>-0.001</td>
<td>-0.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$R^2$ | 0.684 | 0.684 |
Mean dep var | -0.33 | -0.330 |
F-stat | 4.23 *** | 3.98 *** |

*** = 1 % significance

The sum of coefficients for the nine respondent-specific dummy variables forced to equal 0.
Sum of 15 question-specific dummy variables forced to equal 0. The sum of 9 respondent-specific dummy variables forced to equal 0.
were improvements. Most of these changes (46) were in the right direction but by too little, as hypothesised. In only seven cases was the change in the right direction but by too much, and even in these cases the forecast accuracy improved from the original, as no one changed the answer in the right direction but by so much (e.g. from five to eight when the correct answer was six) as to reduce the accuracy.

6 Post Hoc correction methods

6.1 Eliminating belief perseverance and overconfidence via the Delphi process

The strong evidence in favour of Hypothesis 4 has an obvious implication. If the tendency not to make sufficiently large changes in forecasts can be corrected at the individual level by scaling up belief updates, the forecasting accuracy of the whole group might improve. The change in each answer of each respondent could be multiplied by a suitable constant. The invariance-corrected answer of respondent \( p \) to question \( j \) is then defined as

\[
x'_{j,p} = x_{j,p,1} + \alpha [x_{j,p,3} - x_{j,p,1}]
\]

The group forecast for each question \( j \) is the average of invariance-corrected individual forecasts. Each individual forecast is corrected without information on other respondents’ answers. If \( \alpha = 1 \), there is no difference between the original round 3 forecast and the invariance-corrected forecast; \( x'_{j,p} = x_{j,p,1} \).

\[
x'_{j,Group2} = \frac{\sum_{p=1}^{9} x'_{j,p}}{9}
\]

Unfortunately it is hardly possible to derive a suitable value of \( \alpha \) from theory. Instead, one might experiment with some arbitrary values. Penalty points for an invariance-corrected forecast are calculated according to Equation (5):

\[
\psi'_{j,Group2} = \frac{|x'_{j,Group2} - y_j|}{StDev(x_{j,1})}
\]

As to integers, with our data the best value is four, which is a surprisingly large (see Table 7). If decimal numbers are also considered, the optimal value of \( \alpha \) becomes 4.39. The number of observations is fairly limited (N=15), but in light of the t-test, the change in penalty points differs from zero statistically significantly at the 1% level. If one calculates the differences between
penalty points of the FTF meeting group and the invariance-corrected Delphi panel, the average difference is statistically significant at the 5% level, if alpha is 3, 4 or 5 (see Table 7). At least when alpha is close to 4, one cannot reject the null hypothesis that the improvements over the standard Delphi and the FTF meeting are normally distributed. Shapiro-Wilk tests for normality of differences in penalty points yield $W = 0.957$ for the improvement over the standard Delphi and $W = 0.907$ for the improvement over the FTF meeting ($N=15$). In 14 questions out of 15 the invariance-corrected Delphi outperforms the standard Delphi, and this fact is very robust to different error metrics. The magnitude of the reduction in forecasting error may be more interesting than the statistical significance (see Armstrong, 2007). By applying the value $\alpha = 4$ we reach an 8% reduction in average and a 19% reduction in median penalty points. This improvement is neither dramatic, nor is it negligible.

If this technique works in other contexts, it would enable better forecasting with expert panels. It is difficult to say how this technique would improve accuracy in other Delphi surveys, or whether the suitable value of $\alpha$ would always be roughly the same. Calibrating the parameter is an empirical issue. So far, $\alpha = 4$ is the only choice backed by any evidence.

We have offered three arguments in support of the hypothesis that the de-biasing method could be successfully applied in other Delphi surveys where respondents are asked to provide numeric answers on a continuous scale.

1. Previous psychological literature, reviewed in Section 2, predicts that this de-biasing technique should work; under-reacting is the norm in updating beliefs. The bias itself is well documented and its existence has been confirmed several times.

2. As can be seen in Table 8, de-biasing with a very low value of alpha ($\alpha = 1.1$) resulted in an improved forecast in 14 of 15 cases, question 11 being the sole exception. This result is not based on any in-sample optimisation on the value of alpha. Marginal de-biasing is almost always better than no de-biasing.

3. A small out-of-sample test was carried out for each question. The de-biasing parameter alpha was calibrated 15 times; for each question its value was optimised to minimise penalty points for the other 14 questions. The resulting parameter values ranged from 4.1 to 4.52. As can be seen in Table 8, accuracy improved in 14 cases, and in only one case (question 11) did it
Table 7. Penalty points of invariance-corrected Delphi forecasts relative to original penalty points

<table>
<thead>
<tr>
<th>j</th>
<th>Question</th>
<th>FTF</th>
<th>Delphi</th>
<th>α=2</th>
<th>α=3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Big banks’ loan losses</td>
<td>0.804</td>
<td>1.011</td>
<td>0.967</td>
<td>0.923</td>
</tr>
<tr>
<td>2</td>
<td>Stock market turnover</td>
<td>0.342</td>
<td>0.992</td>
<td>0.565</td>
<td>0.139</td>
</tr>
<tr>
<td>3</td>
<td>BoF claims on credit institutions</td>
<td>22.138</td>
<td>21.648</td>
<td>21.643</td>
<td>21.637</td>
</tr>
<tr>
<td>4</td>
<td>M2 deposits</td>
<td>0.087</td>
<td>0.652</td>
<td>0.466</td>
<td>0.281</td>
</tr>
<tr>
<td>5</td>
<td>Interest in corporate loans</td>
<td>3.589</td>
<td>3.138</td>
<td>2.965</td>
<td>2.791</td>
</tr>
<tr>
<td>6</td>
<td>Net interest income of small banks</td>
<td>0.111</td>
<td>0.131</td>
<td>0.103</td>
<td>0.075</td>
</tr>
<tr>
<td>7</td>
<td>Serial bonds</td>
<td>0.469</td>
<td>0.868</td>
<td>0.588</td>
<td>0.308</td>
</tr>
<tr>
<td>8</td>
<td>Households’ mutual fund shares</td>
<td>2.217</td>
<td>2.179</td>
<td>2.174</td>
<td>2.168</td>
</tr>
<tr>
<td>9</td>
<td>Stock of corporate bonds</td>
<td>0.798</td>
<td>0.384</td>
<td>0.311</td>
<td>0.238</td>
</tr>
<tr>
<td>10</td>
<td>Banks’ foreign net assets</td>
<td>6.840</td>
<td>7.216</td>
<td>7.139</td>
<td>7.062</td>
</tr>
<tr>
<td>11</td>
<td>Mean deposit rate of term deposits</td>
<td>0.415</td>
<td>0.415</td>
<td>0.545</td>
<td>0.676</td>
</tr>
<tr>
<td>12</td>
<td>Interest in equities</td>
<td>4.357</td>
<td>4.115</td>
<td>4.042</td>
<td>3.968</td>
</tr>
<tr>
<td>13</td>
<td>Foreign ownership in Nokia</td>
<td>0.670</td>
<td>0.414</td>
<td>0.407</td>
<td>0.400</td>
</tr>
<tr>
<td>14</td>
<td>Stock of housing loans</td>
<td>0.620</td>
<td>0.218</td>
<td>0.154</td>
<td>0.090</td>
</tr>
<tr>
<td>15</td>
<td>12-month Euribor</td>
<td>3.059</td>
<td>1.746</td>
<td>1.611</td>
<td>1.475</td>
</tr>
</tbody>
</table>

|                | Average penalty points                       | 3.101| 3.008  | 2.912 | 2.815 |
|                | Improvement over meeting (t-value)            | 1.5  | 2.3 ** |
|                | Improv. over non-adapted Delphi (t-value)     | 2.8 **2.8 ** |

<table>
<thead>
<tr>
<th>j</th>
<th>Question</th>
<th>α=4</th>
<th>α=5</th>
<th>α=6</th>
<th>α=7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Big banks’ loan losses</td>
<td>0.879</td>
<td>0.835</td>
<td>0.791</td>
<td>0.747</td>
</tr>
<tr>
<td>2</td>
<td>Stock market turnover</td>
<td>0.287</td>
<td>0.713</td>
<td>1.139</td>
<td>1.566</td>
</tr>
<tr>
<td>3</td>
<td>BoF claims on credit institutions</td>
<td>21.632</td>
<td>21.627</td>
<td>21.621</td>
<td>21.616</td>
</tr>
<tr>
<td>4</td>
<td>M2 deposits</td>
<td>0.095</td>
<td>0.090</td>
<td>0.276</td>
<td>0.462</td>
</tr>
<tr>
<td>5</td>
<td>Interest in corporate loans</td>
<td>2.618</td>
<td>2.444</td>
<td>2.271</td>
<td>2.098</td>
</tr>
<tr>
<td>6</td>
<td>Net interest income of small banks</td>
<td>0.047</td>
<td>0.018</td>
<td>0.010</td>
<td>0.038</td>
</tr>
<tr>
<td>7</td>
<td>Serial bonds</td>
<td>0.028</td>
<td>0.252</td>
<td>0.532</td>
<td>0.812</td>
</tr>
<tr>
<td>8</td>
<td>Households’ mutual fund shares</td>
<td>2.163</td>
<td>2.158</td>
<td>2.152</td>
<td>2.147</td>
</tr>
<tr>
<td>9</td>
<td>Stock of corporate bonds</td>
<td>0.166</td>
<td>0.093</td>
<td>0.020</td>
<td>0.052</td>
</tr>
<tr>
<td>10</td>
<td>Banks’ foreign net assets</td>
<td>6.985</td>
<td>6.908</td>
<td>6.831</td>
<td>6.754</td>
</tr>
<tr>
<td>11</td>
<td>Mean deposit rate of term deposits</td>
<td>0.806</td>
<td>0.937</td>
<td>1.068</td>
<td>1.198</td>
</tr>
<tr>
<td>12</td>
<td>Interest in equities</td>
<td>3.894</td>
<td>3.821</td>
<td>3.747</td>
<td>3.673</td>
</tr>
<tr>
<td>13</td>
<td>Foreign ownership in Nokia</td>
<td>0.392</td>
<td>0.385</td>
<td>0.377</td>
<td>0.370</td>
</tr>
<tr>
<td>14</td>
<td>Stock of housing loans</td>
<td>0.025</td>
<td>0.039</td>
<td>0.104</td>
<td>0.168</td>
</tr>
<tr>
<td>15</td>
<td>12-month Euribor</td>
<td>1.340</td>
<td>1.205</td>
<td>1.069</td>
<td>0.934</td>
</tr>
</tbody>
</table>

|                | Average penalty points                       | 2.757| 2.768  | 2.801 | 2.842 |
|                | Improvement over meeting (t-value)            | 2.6 **2.1 **1.6 1.2 |
|                | Improv. over non-adapted Delphi (t-value)     | 3.1 ***3.1 ***2.3 *1.4 |

H₀₀: Average difference = 0

* = 10 % significance, **= 5 % significance, *** = 1 % significance
Table 8. Penalty points of invariance-corrected Delphi method forecasts, different values of alpha

<table>
<thead>
<tr>
<th>j</th>
<th>Question</th>
<th>α = 1</th>
<th>α = 1.1</th>
<th>Varying α</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Big banks’ loan losses</td>
<td>1.001</td>
<td>1.006</td>
<td>0.875</td>
</tr>
<tr>
<td>2</td>
<td>Stock market turnover</td>
<td>0.992</td>
<td>0.949</td>
<td>0.509</td>
</tr>
<tr>
<td>3</td>
<td>BoF claims on credit institutions</td>
<td>21.648</td>
<td>21.647</td>
<td>21.630</td>
</tr>
<tr>
<td>4</td>
<td>M2 deposits</td>
<td>0.652</td>
<td>0.634</td>
<td>0.077</td>
</tr>
<tr>
<td>5</td>
<td>Interest in corporate loans</td>
<td>3.138</td>
<td>3.121</td>
<td>2.600</td>
</tr>
<tr>
<td>6</td>
<td>Net interest income of small banks</td>
<td>0.131</td>
<td>0.128</td>
<td>0.036</td>
</tr>
<tr>
<td>7</td>
<td>Serial bonds</td>
<td>0.868</td>
<td>0.840</td>
<td>0.115</td>
</tr>
<tr>
<td>8</td>
<td>Households’ mutual fund shares</td>
<td>2.179</td>
<td>2.178</td>
<td>2.161</td>
</tr>
<tr>
<td>9</td>
<td>Stock of corporate bonds</td>
<td>0.384</td>
<td>0.376</td>
<td>0.158</td>
</tr>
<tr>
<td>10</td>
<td>Banks’ foreign net assets</td>
<td>7.216</td>
<td>7.208</td>
<td>6.977</td>
</tr>
<tr>
<td>11</td>
<td>Mean deposit rate of term deposits</td>
<td>0.415</td>
<td>0.428</td>
<td>0.873</td>
</tr>
<tr>
<td>12</td>
<td>Interest in equities</td>
<td>4.115</td>
<td>4.108</td>
<td>3.887</td>
</tr>
<tr>
<td>13</td>
<td>Foreign ownership in Nokia</td>
<td>0.4143</td>
<td>0.4135</td>
<td>0.389</td>
</tr>
<tr>
<td>14</td>
<td>Stock of housing loans</td>
<td>0.218</td>
<td>0.212</td>
<td>0.019</td>
</tr>
<tr>
<td>15</td>
<td>12-month Euribor (N=8)</td>
<td>1.746</td>
<td>1.733</td>
<td>1.287</td>
</tr>
</tbody>
</table>

In the last column each forecast is calculated using a question-specific value of alpha which is optimised for data on the other 14 questions.

deteriorate. The probability of having so many improvements is virtually zero, assuming that improvements and deteriorations are equally likely.

Even though the outcome was consistent with previous psychological literature, more research is needed to corroborate these findings. One possibility is that this method works if, and only if, the panel consists of experts; as they may be particularly unwilling to admit to themselves that their round 1 forecasts were widely inaccurate, making them under-react to any influences. At least the findings by Hussler et al. (2011) indicate that a non-specialist would be more inclined to adopt a new answer. It might be better not to tell respondents about this adapting technique before they have returned the last questionnaires because knowing this might affect their answering behaviour.

### 6.2 Conditional forecasts

The forecast errors of the seven variables believed to be interest rate-sensitive (see Section 3.5) comprise two components.

1. The error due to unexpected developments in a causally relevant exogenous variable, in this case the Euribor rate, i.e. variable 15, (ω).
2. The error due to imprecise modelling and changes in other background variables ($\theta$).

We use the following notation: $x_j|_z$ denotes the forecast of variable $j$ given that the market interest rate, as defined in question 15, equals $z$ in December 2011. By definition $x_j|x_{15} \equiv x_j$.

The conditionalised information enables us to calculate a proxy for the first derivative of the forecasted variable with respect to interest rates, assuming the second derivative equals zero:

$$\beta_j = (x_j|x_{15}+100bps) - x_j|x_{15})$$

(6)

The betas for the seven questions were calculated for the FTF meeting group and separately for each Delphi respondent.

$$x_j|y_{15} = x_j|x_{15} + \beta_j(y_{15} - x_{15})$$

(7)

where $x_j|y_{15}$ is the forecast the individual respondent or FTF meeting group would presumably have given if December market rates had been known beforehand, and $y_{15}$ is the interest rate outcome for December 2011. The interest rate-corrected forecast of the Delphi group is defined as the average of individual forecasts. The effect of the forecast error caused by unexpected changes in the market rate is

$$\omega_j = [x_j|y_{15}] - [x_j|x_{15}] = \beta_j\omega_{15}$$

(8)

The error that cannot be explained by unexpected developments in interest rates is denoted $\theta_j$, implying

$$y_j = x_j|x_{15} + \omega_j + \theta_j$$

(9)

The penalty points due to the two components of forecast errors are presented in Table 9.

As can be seen (last column), in nine of fourteen cases forecasting accuracy actually deteriorated when the baseline forecast was replaced by the interest rate-corrected forecast. The result is astonishing and constitutes strong evidence against our assumptions.

This result cannot be explained based on the available data. It is possible that the respondents over-emphasised the direct impact of the interest rate, and underestimated the role of factors that drive both interest rates and other variables.
Table 9. Components of final forecast errors in penalty points

| Group 2 - Delphi | j | Question | $|\omega + \theta|$ | $|\theta|$ | $|\omega|$ | $|\theta| - |\omega + \theta|$ |
|------------------|---|----------|----------------|----------------|----------------|-----------------|
|                  |   |          | $\text{StDev}$ | $\text{StDev}$ | $\text{StDev}$ | $\text{StDev}$ |
| 1                | 1 | Big banks’ loan losses | 1.011 | 0.837 | 0.174 | -0.174 |
| 4                | 4 | M2 deposits | 0.652 | 1.093 | 0.441 | 0.441 |
| 5                | 5 | Interest in corporate loans | 3.138 | 3.306 | 0.168 | 0.168 |
| 6                | 6 | Net interest income of small banks | 0.131 | 0.088 | 0.219 | -0.043 |
| 8                | 8 | Households’ mutual fund shares | 2.179 | 2.376 | 0.197 | 0.197 |
| 11               | 11 | Mean deposit rate of term deposits | 0.415 | 1.617 | 1.202 | 1.202 |
| 14               | 14 | Stock of housing loans | 0.218 | 0.162 | 0.380 | -0.057 |

| Group 1 - FTF | j | Question | $|\omega + \theta|$ | $|\theta|$ | $|\omega|$ | $|\theta| - |\omega + \theta|$ |
|---------------|---|----------|----------------|----------------|----------------|-----------------|
|               |   |          | $\text{StDev}$ | $\text{StDev}$ | $\text{StDev}$ | $\text{StDev}$ |
| 1             | 1 | Big banks’ loan losses | 0.804 | 0.768 | 0.036 | -0.036 |
| 4             | 4 | M2 deposits | 0.087 | 0.728 | 0.642 | 0.642 |
| 5             | 5 | Interest in corporate loans | 3.589 | 3.946 | 0.357 | 0.357 |
| 6             | 6 | Net interest income of small banks | 0.111 | 0.135 | 0.246 | 0.024 |
| 8             | 8 | Households’ mutual fund shares | 2.217 | 2.724 | 0.507 | 0.507 |
| 11            | 11 | Mean deposit rate of term deposits | 0.415 | 1.997 | 1.582 | 1.582 |
| 14            | 14 | Stock of housing loans | 0.620 | 0.495 | 0.125 | -0.125 |

$\text{StDev}$ is the standard deviation of the first round answers, as in Equation (1).

7 Conclusions and discussion

This paper presents the findings from a controlled forecasting experiment. Expert groups were asked to give quantitative forecasts for different variables related to Finnish financial markets. The Delphi method was tested and the results were compared with the performance of a Face-to-Face meeting of a reference group. The forecasting horizon was relatively short, only a few months. The Delphi method and FTF meeting seemed to be roughly equally reliable. In preliminary calculations (not presented in detail) both methods clearly outperformed simple trend extrapolations based on the assumption that growth rates observed in the past will continue in the future.

Because Delphi respondents used pseudonyms when answering, we were able to see how the answers of individual respondents changed from round to round, even though their responses were essentially anonymous. The accuracy of individual answers tended to improve during the process. Improvements were substantial if the initial forecast was either remarkably imprecise or not in line with the rest of the group. In most cases respondents changed their answers in the right direction but by too little.

Experimental results, such as those of Brockhoff et al. (1975), Graefe and
Armstrong (2011) and the findings presented above, indicate that the Delphi method as such may not be significantly more reliable than FTF meetings. However, structured processes have one major advantage. In the case of a large number of individual answers from different rounds, it is easier to apply objectively testable techniques to correct for known cognitive biases. The output from a traditional FTF meeting consists of one answer per question, and it provides much less data for use as inputs to post-hoc calculations.

Two potential ways to improve forecasting accuracy were tested. First, we tried to eliminate the belief perseverance bias reported by psychologists. This was possible with the Delphi method only. Each individual answer was moved further in the direction it had already been moved by the respondent. These corrections led to significant improvements in forecasts. More research and experimentation is needed before we can say how general this result is and how stable the optimal parametrisation is.

Second, respondents were asked to evaluate how their answers would depend on the 12-month money market rate. Using this estimated interest rate-sensitivity, forecasts were adjusted post hoc according to the difference between the forecasted interest rate and the outcome. Perhaps surprisingly, this experiment failed, as if knowing the interest rate outcome many months in advance would not have helped respondents to make accurate forecasts of other variables.

Attempts at improving judgemental forecasting processes should take maximal advantage of findings reported in the psychological literature. By definition, judgemental forecasting is based on the workings of the human mind and is known to be affected by cognitive biases. Indeed, it is difficult to see why being a Delphi panellist would render one immune to psychological phenomena. Delphi panels in real-world forecasting tasks often consist of experts, and experts seem to be, perhaps paradoxically, more affected than laymen at least by belief perseverance bias; they are less likely to change their original answers (Hussler et al., 2011). At least the subadditivity effect, the tendency to misperceive the probability of an event to be less than the sum of probabilities of two mutually exclusive versions of the event, is an obvious candidate in a forecasting exercise (see Tversky and Koehler, 1994). Taking such cognitive biases into account in interpreting and processing survey results would seem to be a promising area for future research.
References


Bolger, F., Stranieri, A., Wright, G., Yearwood, J. (2011) Does the Delphi process lead to increased accuracy in group-based judgmental forecasts or does it simply induce consensus amongst forecasters? Technological forecasting & social change, 78(9), 1671–1680.


Appendix A. Questions, forecasts and the actual realisations of the variables

Questionnaire

1. How much in combined total of loan losses will be reported by Nordea Bank Finland group, Sampo Bank group and the OP Pohjola bank group report for 2011?

2. What will be the share turnover on the Helsinki Stock Exchange in 2011?

3. How large will the item "Lending to euro area credit institutions related to monetary policy operations denominated in euro" be on the Bank of Finland balance sheet as of 31st December 2011?

4. What will be the stock of M2 deposits in Finnish monetary and financial institutions in December 2011?

5. What percent of respondents in the Banking Barometer 4/2011 (of the Federation of Finnish Financial Services) will answer "stronger" to the question on expected near future corporate customers’ loan demand?

6. What will be the combined total of net interest income for Ålandsbanken, Tapiola bank and local cooperative banks in 2011?

7. What will be the outstanding stock of Finnish government serial bonds in December 2011?

8. What will be the outstanding stock of Finnish households’ holdings of mutual fund units in December 2011?

9. What will be the outstanding stock of corporate bonds issued by Finnish companies in December 2011?

10. What will be the outstanding stock of net foreign assets of Finnish monetary and financial institutions (excluding the Bank of Finland) in December 2011? (Negative if liabilities exceed assets)
11. What will be the average rate of interest on households’ and companies’ term deposits with Finnish monetary and financial institutions in December 2011?

12. What percent of respondents in the Banking Barometer 4/2011 (of the Federation of Finnish Financial Services) will answer ‘increasing’ to the question on the general interest in equities and equity funds?

13. What will be the share of foreign ownership in Nokia plc in September 2011?

14. What will be the stock of Finnish households’ housing loans in September 2011?

15. How high will the 12-month Euribor 360 rate be at end-December 2011?

### Appendix B. Answers to the questions

In the table the best final forecast is denoted by a star (*).
<table>
<thead>
<tr>
<th>j</th>
<th>Question</th>
<th>Actual realisation</th>
<th>Group 2 (Delphi) mean</th>
<th>Group 1 (FTF) mean</th>
<th>StDev of all 20 answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Big banks’ loan losses</td>
<td>mEUR 224.7</td>
<td>347.25</td>
<td>329.7</td>
<td>107.5</td>
</tr>
<tr>
<td>2</td>
<td>Stock market turnover</td>
<td>mEUR 137 898</td>
<td>163 545</td>
<td>157 744</td>
<td>18.37</td>
</tr>
<tr>
<td>3</td>
<td>BoF claims on credit inst.</td>
<td>mEUR 2 311</td>
<td>51.0</td>
<td>109.3</td>
<td>104.4</td>
</tr>
<tr>
<td>4</td>
<td>M2 deposits</td>
<td>bEUR 119.504</td>
<td>115.460</td>
<td>116.300</td>
<td>4.633</td>
</tr>
<tr>
<td>5</td>
<td>Interest in corporate loans</td>
<td>% 9</td>
<td>51.3</td>
<td>50.9</td>
<td>12.82</td>
</tr>
<tr>
<td>6</td>
<td>NI of small banks</td>
<td>mEUR 137.9</td>
<td>140.4</td>
<td>140.2</td>
<td>16.423</td>
</tr>
<tr>
<td>7</td>
<td>Serial bonds</td>
<td>mEUR 66 502</td>
<td>62 643</td>
<td>63 340</td>
<td>3.202</td>
</tr>
<tr>
<td>8</td>
<td>Households’ mutual funds</td>
<td>mEUR 10 945</td>
<td>13 590</td>
<td>13 850</td>
<td>1 234</td>
</tr>
<tr>
<td>9</td>
<td>Stock of corporate bonds</td>
<td>mEUR 4 815</td>
<td>4 936</td>
<td>5 072</td>
<td>298.29</td>
</tr>
<tr>
<td>10</td>
<td>Banks’ foreign net assets</td>
<td>bEUR -100.6</td>
<td>-43.1</td>
<td>48.6</td>
<td>7.832</td>
</tr>
<tr>
<td>11</td>
<td>Mean dep. rate of term dep.</td>
<td>% 2.17</td>
<td>2.12</td>
<td>2.12</td>
<td>0.1447</td>
</tr>
<tr>
<td>12</td>
<td>Interest in equities</td>
<td>% 13</td>
<td>56.5</td>
<td>61.1</td>
<td>10.56</td>
</tr>
<tr>
<td>13</td>
<td>Foreign ownership in Nokia</td>
<td>% 80.71</td>
<td>79.42</td>
<td>78.40</td>
<td>2.867</td>
</tr>
<tr>
<td>14</td>
<td>Stock of housing loans</td>
<td>bEUR 80.055</td>
<td>79.597</td>
<td>79.023</td>
<td>1.218</td>
</tr>
<tr>
<td>15</td>
<td>12-month Euribor</td>
<td>% p.a. 1.947</td>
<td>2.44</td>
<td>2.46</td>
<td>0.2494</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>j</th>
<th>Question</th>
<th>Actual realisation</th>
<th>Group 2 Delphi mean</th>
<th>Adapted, $\alpha=4$ mean</th>
<th>Group 1 FTF mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Big banks’ loan losses</td>
<td>mEUR 224.7</td>
<td>333.33</td>
<td>319.2</td>
<td>311.1 *</td>
</tr>
<tr>
<td>2</td>
<td>Stock market turnover</td>
<td>mEUR 137 898</td>
<td>156 111</td>
<td>132 627 *</td>
<td>144 174</td>
</tr>
<tr>
<td>3</td>
<td>BoF claims on credit inst.</td>
<td>mEUR 2 311</td>
<td>51.1</td>
<td>52.8 *</td>
<td>0.0</td>
</tr>
<tr>
<td>4</td>
<td>M2 deposits</td>
<td>bEUR 119.504</td>
<td>116.5</td>
<td>119.06</td>
<td>119.10 *</td>
</tr>
<tr>
<td>5</td>
<td>Interest in corporate loans</td>
<td>% 9</td>
<td>49.2</td>
<td>42.6 *</td>
<td>55</td>
</tr>
<tr>
<td>6</td>
<td>NI of small banks</td>
<td>mEUR 137.9</td>
<td>140.0</td>
<td>138.6 *</td>
<td>139.7</td>
</tr>
<tr>
<td>7</td>
<td>Serial bonds</td>
<td>mEUR 66 502</td>
<td>63 722</td>
<td>66 413 *</td>
<td>65 000</td>
</tr>
<tr>
<td>8</td>
<td>Households’ mutual funds</td>
<td>mEUR 10 945</td>
<td>13 633</td>
<td>13 613 *</td>
<td>13 680</td>
</tr>
<tr>
<td>9</td>
<td>Stock of corporate bonds</td>
<td>mEUR 4 815</td>
<td>4 929</td>
<td>4 864 *</td>
<td>5 053</td>
</tr>
<tr>
<td>10</td>
<td>Banks’ foreign net assets</td>
<td>bEUR -100.6</td>
<td>-44.1</td>
<td>-45.9</td>
<td>-47 *</td>
</tr>
<tr>
<td>11</td>
<td>Mean dep. rate of term dep.</td>
<td>% 2.17</td>
<td>2.11 *</td>
<td>2.05</td>
<td>2.11 *</td>
</tr>
<tr>
<td>12</td>
<td>Interest in equities</td>
<td>% 13</td>
<td>56.4</td>
<td>54.1 *</td>
<td>59</td>
</tr>
<tr>
<td>13</td>
<td>Foreign ownership in Nokia</td>
<td>% 80.71</td>
<td>79.52</td>
<td>79.59 *</td>
<td>78.79</td>
</tr>
<tr>
<td>14</td>
<td>Stock of housing loans</td>
<td>bEUR 80.055</td>
<td>79.789</td>
<td>80.024 *</td>
<td>79.300</td>
</tr>
<tr>
<td>15</td>
<td>12-month Euribor</td>
<td>% p.a. 1.947</td>
<td>2.368</td>
<td>2.281 *</td>
<td>2.710</td>
</tr>
</tbody>
</table>
Discussion on the relevance of risks of mortgage loans to banks’ profits and capital adequacy came to prominence in Finland after the recession at the beginning of 1990s. This same discussion started again, now world-wide, after the US sub-prime crisis and its impacts on the European finance sector.