Statistical Department

COLLECTION OF THESES

Beatrix Oravecz

SELECTIONAL BIAS AND ITS REDUCTION BY CREDIT SCORING MODELS

PhD dissertation

Consultant:
László Hunyadi , Ph.D.
Professor

© Beatrix Oravecz

Content

I. Research history and justification of the topic

In the last 15-20 years there have been revolutionary changes on the market of financial services. Banks started to use automatic decision-making methods and decision support models to be able to speed up credit approval decisions.

An information asymmetry exists between creditor and applicant. One of the greatest risks for banks is the *crediting risk,* which expresses the risk that the credit applicant does not pay, or pays only partly the borrowed capital and/or its interests back, and thus the bank suffers loss. The basic interest of the banks is to gain more information of better quality possible about the customers, and to get more information from them about the liquidity and willingness to pay with the help of different data mining systems. Credit scoring, used for rating serves this aim.

Credit scoring played a very important role in the explosive growth of the stock of consumer credits. Without an accurate and automatic risk analyzing system banks could not have increased their retail placing in such a big measure.

Despite the wide application of credit scoring methods, the methodology still has aspects, which are not paid enough attention to neither in the special literature, nor in practice. The question of representativity of the model building sample is a field like this. Scoring models are usually built on a non-representative sample, as in this case we have a total data base typically only at those customers, who have come through a credit-review process and have been accepted. The credit scoring model, used to accept/ reject the applications by and by looses its actuality, accuracy so needs to be re-built. If the model is not refreshed, it does not follow the forthcoming changes in the population and the effect of explanatory variables, and the original model loses its predictive power. On the other hand, however, if only the data of customers accepted is used to refresh the model, the validity of the new model will be questionable, as the distribution of accepted and rejected will probably differ as a result of the systematic judgment process, so the accepted do not represent all applicants, indicating the whole population.

This phenomenon is called reject bias, or more generally selectional bias.

Model building, using the features of rejected (*reject inference*) may serve an answer for the dilemma. This is actually the estimation of how the rejected applicant would have behaved, if he had been granted the credit.

An often quoted example is the old offender. Old offender applicants are almost always rejected. If all were rejected, then without reject inference this criterion would not appear in the final model. The fact that the majority is rejected often means, that the minority, who is accepted, disposes really special features, and usually does not represent old offenders at all. So, if a model is constructed only on the performance of the accepted, the final model will be too optimistic.

In the dissertation we deal with methods, suitable for the *reduction of selectional bias in case of credit scoring models.* The analysis of the phenomenon is almost completely missing from the Hungarian special literature, and could be met only on the level of mentioning.

Beyond its curiosity, the choice of topic is reasonable because of its *practical importance.* As if the performance of the model can be improved a little bit, it may result in a huge increase in profit or a decrease of risk for the *banks*, as it is about placing of great volumes. The more accurate appraisal of risk is at the same time advantageous *for the customers too,* as it makes the reduction of additional price of risk for good debtors possible, or those, who have been rejected so far, can get a credit of an adequate additional price of risk.

Summarizing former researches in connection to the topic, we can state, that the adoption of rejected during the model building can be a sensible and useful solution only, if *certain conditions come true* for the accepted and rejected population. These solutions may work in practice as assumptions are usually reasonable, or at least show to a good direction. For instance it is a rational assumption that the ratio of bad is higher within the rejected, than within the accepted (even with the same score), even if it cannot be correctly defined numerically how much it is greater.

The benefit of the application of the real and imputated data of the rejected depends on the rejection rate, the distribution within the population and sample and the fulfillment of applied statistical conditions. There are some portfolios, where the ratio of rejected is really low (for example the market of mortgages). In these cases dealing with the rejected may be unnecessary, as their ratio within the population is negligible, so bias caused by them does not need any correction. On the other hand in case of portfolios of greater risk, for example in case of crediting small- and beginner companies, the rejection rate may be really high, so selectional bias can not be neglected.

The best solution, applicable may be occasionally different (by customer groups, products). There is no finished theoretical background in reference, whether the dropout of which conditions cause significant bias in parameter estimations. Such a general principle would be difficult to lay down, as bias is greatly dependent on the database.

According to some statisticians, the problem of conclusion from a non-random sample can be solved with the right imputation of the missing collapse data of the rejected. (Joanes 1993/4, Donald 1995, Copas and Li 1997, Greene 1998). Generatives of scorecard already apply reject inference techniques abroad, in which they are supported by statistical software packages (e.g. SAS). However, they usually work as black boxes, because the underlying principles and assumptions are not clear for the users.

If certain assumptions are acceptable, and the rejected are applied with some kind of imputation, we face a question: how can our model be validated and how can the improvement be measured? Only a few relevant studies were made in this topic, as the majority of data bases, used for testing is not complete, or was simulated (Donald 1995, Feelders 1999, Manning et al. 1987). Hand and Henley (1993/4) revealed that solutions, used in business life are problematic, as they are usually based on really doubtable assumptions.

II. Methods applied

I examined the applicability of some methods in my dissertation, and then chose the one, which seemed to be the best applicable in case of the concrete research data base.

Selectional bias, appearing in case of credit qualifying models is a problem, deriving from missing data, as in case of customers, previously rejected, the value of the variable, describing credit risk (re-payment) is missing (is not observable), so in chapter I. the types and possible methods of handling of *missing data* is taken one after another.

In the next chapter (**II.**), the tasks of *credit scoring,* the most often used methods and indices, suitable to value them are shortly looked over. In practice the application of **logistic regression** is most widespread in case of credit scoring models; therefore I also used logit models during the empirical research to estimate the non-payment possibility of the customers.

In chapter **III.** methods**,** presented in the special literature and *serving the reduction of selectional bias appearing in case of scoring model* are reviewed. All methods use somehow the information available about the rejected.
The effective re-payment information of the rejected is unknown, that's why – as information cannot arise from nothing – if we want to use them for the model building, we need to use *assumptions*, or *additional information* needs to be gathered about their re-payment behavior.
In this chapter the theoretical background of such techniques (reject inference) will be introduced, highlighting the assumptions applied or the method of gathering and using additional information and sum up the practical experiences so far.

The fulfillment of the conditions applied cannot be generally tested, so -after studying the special literature- I came to the conclusion that *the only robust and* *effective way to eliminate bias, is to credit a part of rejected and their behavior and possible collapse is observed in this way.*

It is undoubted, that the model could be corrected with the use of *additional information,* as this time we lean on more information during model building. This way, however, can not always be realized, because of the money and time need of the solution. The application of ***gate opened for a mouth*** with a kind of cost optimal sample distribution is a possible way to decrease the costs of the procedure.
This means, that all customers, who are otherwise to be rejected have chance to get into the sample, but not with the same possibility. Those, whose expected loss is higher, can get a loan with a lower possibility and with a higher possibility those, where this expected loss is smaller. So we get a stratified sample with a certain sort of cost-optimal sample distribution. Finally a sample, representing the whole population is achieved by **reweighting** without taking charge of huge costs by allowing everybody in.

In terms of *the empirical research* I examined on a real bank data base (retail credit card data) the improvement, costs and benefits attainable with the method of gate opened for a mouth, on the scoring model, built with the help of logistic regression.

III. The results of the dissertation

- Making the special literature known I have introduced methods categorized adequately for missing data mechanisms, suitable for the reduction of selectional bias appearing by credit scoring models

In terms of the ***empirical research*** I examined the improvement, costs and expected benefit of the method of gate opened for a mouth on a real bank data base (on retail credit card data). As a result of the model calculations utilizable recommendations are formed for practical experts.

- During the empirical research we experienced, that *models of lower performance can be built in case of higher rejection rate (strong and not completely random selection),* than in case of a lower-ratio rejection. One of the reasons for this is that in this case only a few bad customers are included in the portfolio, making the recognition of the characteristics of bad for the models more difficult. The other reason is that certain values of otherwise significant variables do not get into the sample as a result of selection; therefore the explanatory variable will not be significant.

- In such cases one method of collecting additional information may help, if new observations are gained from internal source with the application of gate opened for a mouth. We have seen that *the performance of the model was improved by the method of gate opened for a mouth, and as a result the profit, attainable on the product increased.*

- We found that if the aim is to maximize profit, it is *better to use a cutoff value determined theoretically,* opposite to the method of empirical definition, which is in practice widespread.

- According to our results the degree of model improvement and increase of profit was the highest in the first step. So *other customers, close to those customers, who are otherwise to be accepted and only a little bit worse than them worth being allowed in with the help of gate opened for a mouth.*

  This first step, extensive improvement of the model and increase in the profit is probably only the characteristic of the data base, but other general considerations suggest this strategy too. Our estimations are much better close to the acceptance range. The ratio of bad can probably be well estimated here, therefore the costs of the extra sample can be more easily estimated, and are lower, than taking the sample from a further range of the sample.

Finally we can say, that techniques, theoretical- and practical considerations reviewed in the dissertation can be applied not only in the fields of credit scoring, but in case of many data mining problems, including similar sample selectional mechanism.

IV. Main references

E.I. ALTMAN [1968], “Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy”, *Journal of Finance, September 1968 , Vol.23. page 589-609.*

L. ASCH [1995], “How the RMA/Fair, Isaac Credit-Scoring Model Was Built,” *Journal of Commercial Lending*, 10-16.p., June 1995.

D. ASH ÉS S. MEESTER [2002], “Best Practices in Reject Interference,” *Presentation at Credit Risk Modeling and Decision Conference*, Wharton Financial Institutions Center, Philadelphia, May 2002.

T. ASTEBRO ÉS I. BERNHARDT [2001], “Bank Loans as Predictors of Small Start-up Business Survival”, manuscript

T. ASTEBRO ÉS G. CHEN [2000], “Missing Data Analysis for Single Choice and Multiple Choice Survey Questions When Data are Sparse,” *Presented at 2000 Academy of Management Conference*, Symposium entitled “Much ado about missing data,” Canada, Toronto.

S. AZEN ÉS M. VAN GUILDER [1981], “Conclusions regarding algorithms for handling incomplete data” *Proc. Stat. Computing Sec., Am. Statist. Assoc.,*53-56.p

J.B. BANASIK, J.N. CROOK ÉS L.C. THOMAS [2003], “Sample Selection Bias in Credit Scoring Models,” *Journal of the Operational Research Society*, 54. issue 822-832.p.

J. BARNARD, AND D.B. RUBIN [1999], “Small-sample degrees of freedom with multiple imputation”, *Biometrica,Vol.* 86., 949-955.p.

H. BIERMAN AND W. H. HAUSMAN [1970], „The Credit Granting Decision,” *Management Sciences*, Vol. 16.., 519-532.p.

BÓDY SÁNDOR, SULYOK PAP MÁRTA [1997], “Cégminősítés”, *Nemzetközi Bankárképző Központ,* Budapest

W.J. BOYES, D.L. HOFFMAN AND S.A. LOW [1989], “An Econometric Analysis of the Bank Credit Scoring Problem,” *Journal of Econometrics*, Vol. 40. 3-14.p.

G. CHEN AND T. ASTEBRO [2003], "How to Deal with Missing Categorical Data: Test of a Simple Bayesian Method," *Organizational Research Methods*, Vol. 6. Iss. 3. 309-321.p.

G. CHEN AND T. ASTEBRO [2003], "Bound and Collapse Bayesian Reject Inference When Data are Missing not at Random," in T. Astebro, P. Beling, D. Hand, B. Oliver and L.B. Thomas (Eds.]: Mathematical Approaches to Credit Risk Management, Conference Proceedings, Banff International Research Station for Mathematical Innovation and Discovery, 11-16. October 2003.

G. CHEN AND T. ASTEBRO [2006], "A Maximum Likelihood Approach for Reject Inference in Credit Scoring", November 25, 2006, manuscript, available at SSRN: http://ssrn.com/abstract=872541

R.K. CHHIKARA [1989], "The State of the Art in Credit Evaluation", *American Journal of Agricultural Economics*, Vol. 71.Iss 5., 1138-1144.p.

J.B. COPAS AND H.G. LI [1997], "Inference for Non-random Samples (with discussion]," *Journal of the Royal Statistical Society*, B, Vol. 59. 55-95.p.

J. CROOK and J. BANASIK [2002], "Does Reject Inference Really Improve the Performance of Application Scoring Models?" *Working paper 02/3*, Credit Research Centre, University of Edinburg, England.

A.P. DEMPSTER, N.M. LAIRD and D.B. RUBIN [1977], "Maximum Likelihood Estimation From Incomplete Data Via the EM Algorithm," *Journal of the Royal Statistical Society B*, Vol. 39. 1-38.p.

S.G. DONALD [1995], "Two-step Estimation of Heteroskedastic Sample Selection Models," *Journal of Econometrics*, Vol. 65. 347-380.p.

A.J. FEELDERS, S. CHANG and G.J. MCLACHLAN [1998], "Mining in the Presence of Selectivity Bias and its Application to Reject Inference," *Proceedings of the fourth international conference on knowledge discovery and data mining (KDD-98]*, AAAI Press, 199-203.p.

A.J. FEELDERS [1999], "Credit Scoring and Reject Inference with Mixture Models," *International Journal of Intelligent System in Accounting, Finance and Management*, Vol. 8 , 271-279.p.

A.J. FEELDERS [2000], "Credit Scoring and Reject Inference with Mixture Models," *International Journal of Intelligent System in Accounting, Finance and Management*, Vol. 9. 1-8.p.

A.J. FEELDERS [2001], "An Overview of Model Based Reject Inference for Credit Scoring," *Working paper*, Institute for Information and Computing Sciences, Utrecht University, Holland.

D.J. FORGARTY [2005], "Multiple Imputation as a Missing Data Approach to Reject Inference on Consumer Credit Scoring", *Manuscript*, DavidF1967@email.uopx.edu

H. FRIEDMAN, E. ALTMAN AND D.L. KAO [1985], "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress", *Journal of Finance*, March 1985., Vol. 40. Iss.1, 269-291 page.

A. GELMAN, J.B. CARLIN, H.S. STERN and D.B. RUBIN [1995], "Bayesian Data Analysis", *Chapman & Hall*.

W.R. GILKS, S. RICHARDSON, és D. J. SPIEGELHALTER (Eds.] [1996]. „ Markov Chain Monte Carlo in Practice", *Chapman & Hall*, London.

R. GLYNN, N. M. LAIRD and D.B. RUBIN [1986], „Selection modeling versus mixture modeling with nonignorable nonresponse", *In H. Wainer (ed.] Drawing Inferences from Self-Selected Samples,* 119-146. New York: Springer-Verlag.

J.W. GRAHAM and S.I. DONALDSON [1993], "Evaluating Interventions with Differential Attrition: the Importance of Nonresponse Mechanisms and Use of Followup Data," *Journal of Applied Psychology*, Vol. 78. 119-128.p.

W.H. GREENE [1998], "Sample Selection in Credit-scoring Models," *Japan and the World Economy*, Vol. 10. 299-316.p.

HAJDU OTTÓ [2003]: „Többváltozós statisztikai számítások", KSH, Budapest

HAJDU OTTÓ és VIRÁG MIKLÓS [1996], „Pénzügyi mutatószámokon alapuló csődmodell-számítások", *Bankszemle,* 1996/1-2

HÁMORI GÁBOR [2001], „A CHAID alapú döntési fák jellemzői", *Statisztikai szemle,* Vol. 79. Iss. 8, 703-710.p.

D.J. HAND and W.E. HENLEY [1993/4], „Can Reject Inference Ever Work?," *IMA Journal of Mathematics Applied in Business & Industry*, Vol. 5. Iss. 4. 45-55.p.

D.J. HAND and W.E HENLEY [1994], „Inference About Rejected Cases in Discriminant Analysis," *Springer*, 292-299.p., New York.

D.J. HAND [1997], „Construction and Assessment of Classification Rules," *Chichster: Wiley*.

D.J. HAND [1998], "Reject Inference in Credit Operations," *in Credit Risk Modeling: Design and Application* (ed. E. Mays], 181-190.p. AMACOM.

D.J. HAND [2001], "Measuring Diagnostic Accuracy of Statistical Prediction Rules," *Statistica Neerlandica*, Vol. 53. 3-16.p.

J.J. HECKMAN [1979], "Sample Selection Bias as a Specification Error," *Econometrica*, Vol. 47. 153-161.p.

D. HEDEKER. AND R. D. GIBBONS [1997], „Application of random-effects pattern-mixture models for missing data in longitudinal studies"*, Psychological Methods,* 2(1], 64-78.

C.W. HOLSAPPLE et.al. [1988], "Adapting Expert System Technology to Financial Management" *Financial Management*, autumn 1988.Vol. 19. 12-22.p.

D.C HSIA [1978], "Credit Scoring and the Equal Credit Opportunity Act," *The Hastings Law Journal*, Vol. 30. 371-448.p. November 1978.

HUNYADI LÁSZLÓ [2001], "A mintavétel alapjai", *Egyetemi Jegyzet SZÁMALK,* Budapest

HUNYADI LÁSZLÓ és VITA LÁSZLÓ [2002], "Statisztika közgazdászoknak", *Központi Statisztikai Hivatal,* Budapest

T. JACOBSON és K.F. ROSZBACH [1999], "Evaluating Bank Lending Policy and Consumer Credit Risk," *in Computational Finance 1999* [edited by Y.S. Abu-Mostafa et al.] the MIT Press, 2000.

D.N JOANES [1993], "Reject Inference Applied to Logistic Regression for Credit Scoring," *IMA Journal of Mathematics Applied in Business and Industry*, Vol. 5.Iss. 4.szám 35-43.p.

N.M. KIEFER and C.E. LARSON [2003], “Specification and Informational Issues in Credit Scoring”, manuscript

J.O. KIM and J. CURRY [1977], „ The treatment of missing data in multivariate analysis”, *Sociol. Meth. Res.*, Vol. 6. 215-240.p.

P.W. LAVORI, R. DAWSON and D. SHERA [1995], „ A multiple imputation strategy for clinical trials with truncation of patient data”*, Statistics in Medicine*, Vol. 14 , 1913-1925.p.
R.J.A. LITTLE [1979], “Maximum Likelihood Inference for Multiple Regression with Missing Values: A Simulation Study”, *Journal of the Royal Statistical Society*, vol.41., 76-87.p.

R.J.A. LITTLE and D.B. RUBIN [1987], “Statistical Analysis with Missing Data”, *John Wiley & Sons*, New York.

R.J.A. LITTLE and D.B. RUBIN [2002], “Statistical Analysis with Missing Data,” 2. Edition, *John Wiley & Sons*, New York.

R.J.A. LITTLE [1993], “Pattern-mixture Models for Multivariate Incomplete Data,” *Journal of the American Statistical Association*, Vol.88. 125-134.p.

R.J.A. LITTLE and N. SCHENKER [1994], “Missing Data“ in *Handbook for Statistical Modeling in the Social and Behavioral Sciences* [G. Arminger, C. C. Clogg és M. E. Sobel szerk.] New York: Plenum 39-75.p.

G.S. MADDALA [1983], “Limited Dependent and Qualitative Variables in Econometrics,” *Cambridge University Press*, Cambridge, UK.

MÁDER MIKLÓS PÉTER [2005],”Imputálási eljárások hatékonysága”, *Statisztikai Szemle*, Vol. 83.Iss. 7. 628-644.p.

N.K. MALHOTRA [1987], “Analyzing Marketing Research Data with Incomplete Information on the Dependent Variable”, *Journal of Marketing Research*, Vol.24, 74-84.p

O.L. MANGASARIAN [1965], “Linear and nonlinear separation of patterns by linear programming”, *Operation Research*, Vol. 13 , 444-452.p.

E. MAYS [2004], “Credit Scoring for Risk Managers” *South Western Thomson Learning*

G.J. MCLACHLAN and K.E. BASFORD [1988], “Mixture Models, Inference and Applications to Clustering,” *Marker Dekker*, New York.

G.J. MCLACHLAN [1992], “Discriminant Analysis and Statistical Pattern Recognition,” *Wiley*, New York.

R.W. MCLEOD et.al. [1993], “Predicting Credit Risk: A Neural Network Approach”, *Journal of Retail Banking,* Vol. 15. Iss. 3. 37-40.p.

C.L. MENG and P.SCHMIDT [1985], “On the Cost of Partial Observation in the Bivariate Probit Model,” *International Economic Review*, Vol. 26. Iss. 1. szám 71-85.p., February 1985.

M.C. PAIK, R. SACCO és I.F. LIN [2000], “Bivariate Binary Data Analysis with Nonignorably Missing Outcomes,” *Biometrics*, Vol. 56. 1145-1156.p.

D.J. POIRIER [1980], “Partial Observability in Bivariate Probit Model,” *Journal of Econometrics*, Vol. 12. 209-217.p.

P.L. ROTH and F.S. SWITZER III [1995], “A Monte Carlo Analysis of Missing Data Techniques in a HRM Setting,” *Journal of Management*, Vol. 21. Iss 5.1003-1023.p.

D.B. RUBIN [1987], “Multiple Imputation for Nonresponse in Surveys,” *John Wiley & Sons*.

D. B. RUBIN [1996], ”Multiple imputation after 18+ years [with discussion]”, *Journal of the American Statistical Association*, 91, 473-489.p.

D. B. RUBIN [2003], ”Nested Multiple Imputation of NMES via Partially Incompatible MCMC”, *Statistica Neerlandica,* 57, 3-18.p.

RUDAS T. [1998],”Hogyan olvassunk közvélemény-kutatásokat?”, Új Mandátum Könyvkiadó, Budapest

J.L. SCHAFER and J.W. GRAHAM [2002], “Missing Data: our View of the State of the Art,” *Psychological Methods*, Vol. 7. Iss. 2. 147-177.p.

J.L. SCHAFER and M.K. OLSEN [1998],"Multiple imputation for multivariate missing-data problems: a data analyst's perspective", *Multivariate Behavioral Research*, Vol. 33., 545-571. p.

P. SEBASTIANI and M. RAMONI [2000], "Bayesian Inference with Missing Data Using Bound and Collapse," *Journal of Computational and Graphical Statistics*, Vol. 9. Iss. 4. 779-800.p.

K.Y. TAM [1991], "Neural Network Models and the Prediction of Bank Bankruptcy" *Omega. The International Journal of Management Science*, Vol. 19., Iss. 5.429-445.p.

K.Y. TAM and M.Y. KIANG [1992], "Managerial Applications of Neural Networks: The Case of Bank Failure Predictions". *Management Science*, July 1992. Vol. 38.Iss. 7. 926-947p.

L.C. THOMAS [2000], "A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers," *International Journal of Forecasting*, Vol. 16. 149-172.p.

L.C. THOMAS, D.B. EDELMAN and J.N. CROOK [2002], "Credit Scoring and Its Applications", *Society for Industrial and Applied Mathematics, Philadelphia*

F. VELLA [1992], "Simple Tests for Sample Selection Bias in Censored and Discrete Choice Models", *Journal of Applied Econometrics*, Vol. 7. 413-421.p.

VIRÁG MIKLÓS és KRISTÓF TAMÁS [2006], "Iparági rátákon alapuló csődelőrejelzés sokváltozós statisztikai módszerekkel", *Vezetéstudomány*, Vol. 37. Iss, 1. 25-35.p.

D. WEST [2000], "Neural Network Credit Scoring Models", *Computers and Operations Research,* Vol. 27. 1131-1152.p

J.C. WIGINTON [1980], "A note on the comparison of logit and discriminant models of consumer credit behaviour", *Journal of Financial Quantitative Anal.,15,* 757-770.p.

V. List of own publication, in connection with the topic

Articles:

Oravecz Beatrix (2007): Credit scoring modellek és teljesítményük mérése. (Credit scoring models and the measurement of their performance) Hitelintézeti Szemle, Vol. 6. Iss. 6., 607 -627 p.

Oravecz Beatrix : Hiányzó adatok és kezelésük a statisztikai elemzésekben. (Missing data and its handling in statistical analysis) Statisztikai Szemle, expected issue January 2008.

Oravecz Beatrix: Szelekciós torzítás és csökkentése a credit scoring modelleknél. (Selectional bias and its reduction by credit scoring models) Hitelintézeti Szemle, expected issue January 2008.

Article in a book:

Oravecz Beatrix (2004): Imputációs eljárások. Egy reneszánsz statisztikus. Tanulmánykötet Hunyadi László tiszteletére, KSH (Imputation methods. A renaissance statistican. Edition of essay in honour of László Hunyadi)

Others:

Oravecz Beatrix (2000): Hitelmonitoring, Szakdolgozat, BKAE Pénzügy Tanszék (Credit monitoring, thesis, BKAE Financial Department)

Oravecz Beatrix (2002): Dijkstra, W.: Új módszer az interjúk közbeni kölcsönkapcsolatok tanulmányozására. cikkismertetés, Statisztikai szemle, 79. évf. 7. szám, 636-637.o. (New method to study relations during interviews, article review)