ISSN – 1303-099X

EGE AKADEMİK BAKIŞ

Ekonomi, İşletme, Uluslararası İlişkiler
ve Siyaset Bilimi Dergisi

EGE ACADEMIC REVIEW

Journal of Economics, Business Administration,
International Relations and Political Science

Cilt 19 • Sayı 4 • Ekim 2019
Volume 19 • Number 4 • October 2019

Yayınlanma Sıklığı / *Frequency:* **Yılda dört kez** / *Four Times in a Year*

Yayınlayan / *Publisher*
Ege Üniversitesi, İktisadi ve İdari Bilimler Fakültesi
Bornova 35100 İZMİR / TÜRKİYE

E-mail: eab@mail.ege.edu.tr

Basım Yeri: Ege Üniversitesi Basımevi Müdürlüğü **Basım Tarihi:** 1 Ekim 2019
T.C. Kültür ve Turizm Bakanlığı Sertifika No:18679

Ege Akademik Bakış

Ege Akademik Bakış Dergisi, iktisat, işletme, uluslararası ilişkiler ve siyaset bilimi alanlarında çalışan akademisyenler, araştırmacılar ve profesyonellerin görüşlerini paylaştıkları bir forum oluşturmak amacıyla, bu alanlarda yapılmış olan uluslararası çalışmaları kapsamaktadır. Ege Üniversitesi İktisadi ve İdari Bilimler Fakültesi tarafından Ocak, Nisan, Temmuz ve Ekim aylarında olmak üzere yılda dört defa yayınlanan hakemli bir dergi olup, Türkçe veya İngilizce olarak kaleme alınmış tüm çalışmalar dergide yayınlanmak üzere gönderilebilir. Ege Akademik Bakış Dergisi aşağıdaki veri tabanlarınca taranmaktadır:

- EconLit (http://www.aeaweb.org/)
- ULAKBİM, Sosyal ve Beşeri Bilimler Veri Tabanı (http://www.ulakbim.gov.tr/)
- Emerging Sources Citation Index (ESCI)
- Director of Open Access Journals(http://www.doaj.org/)
- EBSCO Publishing (http://www.ebscohost.com/)
- PERO(http://knjiznica.irb.hr/pero)
- Scientific Commons(http://en.scientificcommons.org)
- WorldWideScience(http://worldwidescience.org)
- ProQuest(http://www.proquest.com)
- ASOS Index(http://www.asosindex.com)
- RePEc (http://www.repec.org)

Makaledeki görüşler yazarlarına aittir. Dergide yayınlanan makaleler kaynak göstermeden kullanılamaz.

Ege Academic Review includes international papers about economics, business administration, international relations and political science with the aim of providing a forum for academicians, researchers and professionals interested in these fields. This journal is subject to a peer-review process. Ege Academic Review is published by Ege University Faculty of Economics and Administrative Sciences for four times in a year. Papers written in Turkish and English can all be sent in order to be published in the journal. The articles in Ege Academic Review are indexed/abstracted in:

- EconLit (http://www.aeaweb.org/)
- ULAKBİM, Social Sciences and Humanities Database (http://www.ulakbim.gov.tr/)
- Director of Open Access Journals(http://www.doaj.org/)
- EBSCO Publishing (http://www.ebscohost.com/)
- PERO(http://knjiznica.irb.hr/pero)
- Scientific Commons(http://en.scientificcommons.org)
- WorldWideScience(http://worldwidescience.org)
- ProQuest(http://www.proquest.com)
- ASOS Index(http://www.asosindex.com)
- RePEc (http://www.repec.org)

Authors are responsible for the content of their articles. Papers published in the journal can not be quoted without reference.

Cilt 19 • Sayı 4 • Ekim 2019
Volume 19 • Number 4 • October 2019

İçindekiler / Contents

EGE AKADEMİK BAKIŞ / *EGE ACADEMIC REVIEW*

Cilt 19 • Sayı 4 • Ekim 2019
SS. 389 - 399
Doi: 10.21121/eab.638571
Başvuru Tarihi: 25/5/2019 • Kabul Tarihi: 26/8/2019

Using Panel Data for Macroeconomic Policy Evaluation: A Survey[1]

Ron P. Smith

Abstract

In order to measure the macroeconomic effect of some policy or event, a "treatment", we need to construct a "counterfactual", a prediction of what would have happened in the absence of treatment, which is unobserved. Panel data for countries and regions, where the number of units and time periods are large, potentially provide untreated control groups which can be used to construct the counterfactual. A number of different procedures have been suggested for such policy evaluations, including the synthetic control method, SCM, and the panel data approach, PDA. We survey these and other methods.

Keywords: panel data, counterfactuals, policy evaluation, synthetic controls, macroeconomics.

JEL classification: C18, C54, E65,

Economists often want to measure the effect of some "treatment", a policy or event, on an outcome in some aggregate unit, such as a country or region. For example, the treatment might be joining the euro; the outcome is GDP for a particular country that adopted the euro. In another example, the treatment might be a US state introducing a "Stand Your Ground"[2], SYG, law; the outcome is the murder rate in the state. The objective is to measure the effect on GDP or the murder rate of these policy interventions. There are now many panels available for a large number of countries or regions over long time periods. A number of macro policy evaluation procedures have been developed to take advantage of the fact that such panels potentially provide untreated control groups which can be used to construct counterfactuals and allow the estimation of the effect of policy interventions, like joining the euro or adopting SYG laws, on treated groups.

This is different from the typical macro policy evaluation exercise, which considers, for instance, a monetary policy shock, calculated as a one standard error displacement of the structural disturbance of a policy equation, such as a Taylor rule. The impulse response function (IRF) is the time profile of the deterministic component of the effect of such a displacement. The IRF yields ex-ante information about the way the model responds to such a displacement, not an ex-post evaluation of the effectiveness of a real policy intervention. In addition IRFs ignore the cumulative uncertainty associated with the stochastic component, the post-intervention disturbances. Pesaran & Smith (2018) discuss these issues in more detail. This paper examines procedures that compare the actual realisations of the outcome after the intervention with a counterfactual that predicts what would have happened in the absence of the intervention.

In microeconometrics there is a long tradition of measuring such treatment effects on individuals. Abadie, & Cattaneo (2018) provide a recent review of this program evaluation literature. The macroeconomic interest in measuring the effect on aggregates is more recent. A number of the approaches to policy evaluation have their roots in the microeconometric literature. However, the micro and macro issues are rather different. This paper reviews the various approaches and compares their relative advantages and disadvantages.

Section 1 sets out the framework used in macro policy evaluation and contrasts it with the microeconometric treatment effect literature. Section 2 sets out and compares the two main approaches currently

Birbeck University of London, Department of Economics, Mathematics and Statistics, e-mail: r.smith@bbk.ac.uk

used in macro-policy evaluation: the synthetic control method, SCM, and, the panel data approach, PDA. The relative advantages of these procedures has generated some controversy. Section 3 examines some extensions to those approaches. Section 4 considers a different approach, which like SCM has its origins in the microeconometric literature and uses propensity scores. Section 5 has some concluding comments. It appears that none of the approaches is universally best, so it seems important to use more than one and compare, or average, their estimates.

1. The Framework

All the procedures considered will work with multiple treated units, but for clarity consider the effect on a single unit. Suppose the outcome in unit *1*, in period *t*, is y_{1t}. The effect of the policy is

$$\delta_{1t} = y^1_{1t} - y^0_{1t}$$

the difference between the two potential outcomes: the outcome with treatment, y^1_{1t}, (being in the euro) and the outcome without treatment, y^0_{1t}, (not being in the euro). However, it is impossible to observe both y^1_{1t}, and y^0_{1t}, in period *t*. Country 1 is either in the euro or out of the euro, it cannot be both in and out.

Since only one of the potential outcomes can be observed, one needs a "counterfactual", a prediction of what would have happened in the unobserved case. Suppose unit 1 was actually treated in period T_0, adopted the euro, and post-treatment values, $y^1_{1,T0+h}$, are observed[3]. The estimate of the effect of treatment, in the post-treatment period T_0+h, $h=1,2,\ldots,H$ is

$$\hat{\delta}_{1,T_0+h} = y^1_{1,T_0+h} - \hat{y}^0_{1,T_0+h} \quad (1)$$

where $\hat{y}^0_{1,T_0}+h$ is the estimate of the counterfactual: a prediction of what would have happened to country *1* in period T_0+h had it not joined the euro. Thus we have(i) an estimation problem, finding a model to construct the counterfactual outcome in the absence of the policy intervention and(ii) an inference problem, determining whether the difference between the realized and counterfactual is larger than would have been expected by chance. To the extent that suitable control units are available to provide the counterfactual, there is no need to construct a structural model of how the outcome is determined or worry about identification. In which case policy evaluation can be data-driven and relatively atheoretical.

Panels for countries and regions with data on y_{it}, $i=1,2,\ldots,N$ and $t=1,2,\ldots,T$, where *N* and *T* are large potentially provide untreated control groups which can be used to construct the counterfactual, $\hat{y}^0_{1,T_0+h}$, allowing the estimation of $\hat{\delta}^0_{1,T_0+h}$, the effect of a policy intervention on a treated group, and an evaluation of the policy.

A number of different procedures have been suggested, which are reviewed below, and there has been considerable controversy about their relative effectiveness. The two most prominent methods are the synthetic control method, SCM, and the panel data approach, PDA.

Abadie & Gardeazabal (2003) introduced the SCM to measure the costs of Basque terrorism, and it was subsequently applied by Abadie, Diamond, & Hainmueller (2010) to the California smoking program and Abadie, Diamond, & Hainmueller (2015) to German reunification. Since the package Synth became available on Matlab, R and Stata, SCM has been widely used. The PDA was introduced by Hsiao, Ching & Shui (2012) to measure the effects of Hong Kong's political and economic reunification with China. These studies evaluate the effects of events on single units, but both approaches can be used to evaluate effects on multiple units. Bove, Elias & Smith (2016) use both SCM and PDA to measure the effect on GDP of Civil Wars. Gobillon & Magnac (2016) compare the use of SCM and PDA in regional economics.

While we can learn from the microeconometric literature, the micro and macro issues are rather different. The classic micro method is difference in differences. If $\bar{y}_{C0}$ is the average over the units in the control group, C, in period *0* before treatment and $\bar{y}_{C1}$, in period 1, and similarly the treated group, A, averages are $\bar{y}_{A0}$ and $\bar{y}_{A1}$, then the difference in difference estimator is

$$\delta = \left(\bar{y}_{A1} - \bar{y}_{A0}\right) - \left(\bar{y}_{C1} - \bar{y}_{C0}\right) \quad (2)$$

The first term measures the change in the averages for the treated group, the second term controls for any general trends, assuming that the trends in the control group are parallel to those for the treated group. Defining a dummy for group A, D_A, and a dummy for period 1, D_1, using the original observations it can be written as a two way fixed effect model plus a treatment effect:

$$y_{it} = \alpha + \alpha_A D_A + \alpha_1 D_1 + \delta D_A D_1 + \varepsilon_{it} \quad (3)$$

where the four parameters in (3) are functions of the four means in (2). In more general cases, where one has, for instance, more time periods, covariates or endogenous treatment, (3) is a useful representation.

There are a number of points to note about this macro approach. Firstly, in micro cases *N* is large, *T* is small, often only *T=2*, as in (2) above, and there are major problems associated with endogenous selection into treatment. The endogeneity and sample selection biases that arise in the micro case from heterogeneity correlated with treatment, across the units, are not problems in the macro case. There the focus is on a single unit, and the "policy on/policy off" comparisons are done over time rather than across units. In micro terminology, the parameter of interest in the macro case is the effect of "treatment on the treated", not the average treatment effect over individuals. Because it is primarily a time series problem, the rules for assignment to treatment are not an issue. In terms of the examples given above, it makes no sense to consider either the effect of Hong Kong being integrated with West Germany or of East Germany being integrated with China. In addition, macro panels tend to exhibit cross-section dependence that results from strong factors driving all units. This means that other units can be used to construct controls that can be used to specify the counterfactual in the analysis of a treatment effect or the evaluation of a policy intervention. Unlike difference in differences, the parallel trends assumption is not required.

Secondly, a multi-horizon effect is estimated for each period for this unit.

There is no assumption that the effect is constant over time. One can average $\hat{\delta}_{1t}$ over time to get an average treatment effect, ATE:

$$\hat{\bar{\delta}}_{1H} = \sum_{h=1}^{H} \hat{\delta}_{1,T_0+h} / H \tag{4}$$

but this ATE is quite different from the micro case which compares the average over treated units with the average over untreated units as in (2). The average in (4) is over time for a single unit not over units. Tests on the individual $\hat{\delta}_{1,T_0+h}$ are likely to be sensitive to the distributional assumption of the model, whereas tests on the average, $\hat{\bar{\delta}}_{1H}$, can rely on the central limit theorem if *H* is large to obtain a distribution and may be more robust.

Thirdly, the Lucas critique, which refers *to ex ante* policy evaluation is not a problem. Ex ante policy evaluations compare two predictions, one with the policy and one without and face the problem that the intervention may change the parameters. In estimating (1) the concern is with *ex post* evaluation of a policy intervention, where time series data are available before as well as after the policy change. The comparison is based on the difference between the realisations of the outcome variable of interest and counterfactuals obtained assuming no policy change. The counterfactuals, based on estimates using pre-intervention data, will embody pre-intervention parameters while the realized post-intervention outcomes will embody the effect of the change in the policy parameters and any consequent change in expectations.

Finally, whatever we use to predict $\hat{y}_{1t}^{0}$ must not be influenced by the policy intervention or treatment itself. For major changes, with spillover effects, this can be a very strong requirement: a change in one country, like the re-unification of West Germany, can affect many other countries. Similarly, there must be no large change in the control units that would not have affected the treated units.

The fact that the counterfactual is a prediction, means that we can learn from the forecasting literature about good ways to construct it. For instance, it is well known that averaging over forecasts improves performance and Hsiao & Zhou (2018) suggest averaging over counterfactuals produced by different procedures. Pesaran & Smith (2016) point out that simple parsimonious models tend to forecast better than ones with more parameters, so there are arguments for using simple models to generate the counterfactuals.

Counterfactuals differ from a conventional forecasts in that they are not about the future, they are conditional on observed data and there is no actual against which to compare them. This latter feature presents a major difficulty for the evaluation of the different methods. Because we never observe the truth, we cannot say which method gets us closest to the truth. To get around this problem, evaluation is based on simulations, where by construction we do know the truth. But the results are then dependent on the choice of data generating process, DGP, in the simulation. Gardeazabal & Vega-Bayo (2016) and Wan, Xie & Hsiao (2018) differ on the appropriate way to define the DGP in simulations used to compare the SCM and PDA.

The various approaches can be applied to multiple treated units very easily; for instance where a number of.countries join the euro or a number of US states adopt a particular law. But for ease of exposition consider a single treated unit. Suppose unit *1* is subject to treatment, some intervention at T_0. There is pre-intervention data $t=1,2,...,T_0$; post intervention data $t=T_{0+1},T_{0+2},...,T_0+T_1$; with $T=T_0+T_1$. We assume that there are enough pre-intervention observations to estimate models.

Defining $d_{it}=1$ if $i=1$ and $t>T_0$, zero otherwise we observe either the treated or the untreated, but not both:

$$y_{it} = d_{it}y_{it}^1 + (1-d_{it})y_{it}^0$$

In constructing the estimated counterfactuals, $\hat{y}_{it}^0$ we may use (i) observations on the outcome variables in other units: $y_{2t},y_{3t},..,y_{Nt}$ (ii) observations on a vector of covariates $\mathbf{x}_{it}$ in the treated or untreated units and (iii) lagged values in dynamic models. The models may be data driven (atheoretical or non-parametric), just based on correlations or similarity, or the models may be more theoretical parametric models. The synthetic control method, SCM, treats the other outcomes as providing controls and the panel data approach, PDA, treats them as providing predictors, both are data-driven.

Consider case (i) and suppose that there are *N-1* controls not subject to the intervention and not affected by the intervention in unit 1. The estimated counterfactual is a weighted average of the outcomes in these control or predictor units. The issue is how to choose controls and weights. In a static model the effect of the intervention is measured as

$$\delta_{1,T_0+h} = y_{1,T_0+h} - \sum_{i=2}^{N} w_i y_{i,T_0+h}; \quad h=1,2,...,T_1 \qquad (5)$$

It is important to examine the pre-treatment predictive power of the counterfactual, which is an important diagnostic. The method should predict well before treatment and the estimation errors $\hat{\delta}_{1,t}$ for $t<T_0$ should be small. The standard errors of the post-treatment estimate of $\hat{\delta}_{1,t+h}$ will reflect the pre-treatment fit. If the model does not fit well pre-treatment, the standard errors will be large and tests will have little power to detect the effect of a policy intervention.

It is also important to be precise about the nature of the policy intervention, what it is conditional on, and the plausibility of the counterfatual.

2. SCM and PDA

2.1 Synthetic Control

SCM uses the analogy with microeconometric treatment effect studies, where one chooses controls that are similar in characteristics to those that are treated. One would match patients treated with a drug to untreated controls with similar covariates such as age, sex, and health and compare the outcomes in the two groups. Similarity is usually measured by propensity score, the probability of being treated conditional on the covariates.

To determine the SCM weights w_i let $\mathbf{x}_{1kt}$ be a set of $k=1,2,...,K$ covariates or predictor variables for y_{1t}, with the corresponding variables in the other units given by $\mathbf{x}_{jkt}$, $j=2,3,...,N$. These variables are averaged over the pre-intervention period to get $\bar{\mathbf{x}}_{1k}^{T_0}$ and $\bar{\mathbf{X}}_k^{T_0}$ the *N-1* x *1* vector of predictor *k* in the control group. Then the *N-1* x *1* vector of weights $W=(w_2,w_3,\ldots,w_N)`$ are chosen to minimize

$$\sum_{k=1}^{K} v_k \,(\bar{\mathbf{x}}_{1k}^{T_0} - W'\,\bar{X}_k^{T_0})^2$$

subject to $\sum_{i=2}^{N} w_i = 1,\ w_i \geq 0$, where v_k is a weight that reflects the relative importance of variable *k*. Call the SCM weights $\tilde{w}_i$, many of them will be zero, for countries not included in the synthetic control.

SCM chooses the comparison units to be as similar as possible to the target along the dimensions included in $\mathbf{x}_{ikt}$.The v_k are often chosen by cross-validation, which may be problematic for potentially non-stationary time-series samples. The pre-intervention outcome variable may be included in $\mathbf{x}_{ikt}$; it is argued that matching on the pre-intervention outcomes helps control for the unobserved factors affecting the outcome of interest.

Using the SCM weights the estimate of the counterfactual is

$$\tilde{y}_{1,T_0+h}^0 = \sum_{i=2}^{N} \tilde{w}_i y_{i,T_0+h} \qquad (6)$$

In the case of German Reunification, Abadie et al. (2015), use controls and weights w_i of Austria, 0.42, US, 0.22, Japan 0.16, Switzerland 0.11 and Netherlands, 0.09. The synthetic West Germany is similar to the real West Germany in pre 1990 per capita GDP, trade openness, schooling, investment rate and industry share. As they note there may be spillover effects. Since Austria, Switzerland and Netherlands share bor-

ders with Germany there is a distinct possibility that their post 1990 values may be influenced by German reunification. Those that are geographically the most similar are most likely to show spillover effects.

Given the way the SCM estimate is constructed inference, testing whether the effect of the intervention is significant, is not straightforward. Xu (2018) generalises the SCM in various respects to give generalised SCM, GSCM.

2.2 PDA

Hsiao, Ching & Shui (2012), HCS, measure the benefits of political and economic integration of Hong Kong with mainland China using PDA . They use (5), but choose the weights by regression of y_{1t}, growth in Hong Kong, on y_{jt}, $j=2,3,...,N$, growth in the control countries during the pre-intervention period. Then using the pre-intervention estimates they predict the post-intervention counterfactual as.

$$\hat{y}^{0}_{1,T_0+h} = \hat{\beta}_1^{T_0} + \sum_{i=2}^{N} \hat{\beta}_i^{T_0} y_{i,T_0+h} \tag{7}$$

where the $\hat{\beta}_i^{T_0}$ are the regression coefficients estimated on the pre-intervention period up to time T_0. Because the counterfactual is a forecast from a standard regression, inference is easier and HCS use robust standard errors to allow for serial correlation. The coefficients for most countries will be zero, only a subset of other countries are used to predict Hong-Kong. The subset is chosen by a model selection procedure, but other procedures have been used. HCS emphasize that Hong Kong is too small for the effects of integration with China to influence any of the control countries. The control group, weights chosen by AIC for the period 1993:Q1–1997:Q2 are Japan -0.69, Korea -0.3767, USA 0.8099, Philippines -0.1624, Taiwan 0.6189. They find that the political integration had little effect on the growth rate, but that the subsequent economic integration did; an example of the issues in choosing T_0.

Li & Bell (2017) show that the PDA works for a wider range of data generating processes than those HCS originally assumed; derive the asymptotic distribution of the average treatment effect estimator and propose using lasso to select control units. Lasso will work when the number of controls is greater than the sample size, when model selection methods will not. They argue that lasso is computationally more efficient than model selection methods and leads to better out-of-sample prediction.

The underlying interpretation of (7) rests on a factor model:

$$\begin{aligned} y_{it} &= \mu_i + \sum_{j=1}^{r} \lambda_{ij} f_{tj} + e_{it} \\ y_{it} &= \mu_i + \boldsymbol{\lambda}_i' \mathbf{f}_t + e_{it} \end{aligned} \tag{8}$$

The outcome in a unit is determined by a vector of r common factors, $\mathbf{f}_t$ which have different effects on different countries, reflected in their heterogeneous factor loadings, $\boldsymbol{\lambda}_i$, and an idiosyncratic factor e_{it}. In a macroeconometric context, it is natural to think of very different countries driven by the same common trends: the 2008 crisis hit most countries, though to different degrees. Hsiao et al. (2012) include the US in the controls, not because the US is like Hong Kong, the justification in the SCM procedure, but because US growth is a good predictor of Hong Kong growth. Factors are said to be strong if they influence almost every unit, weak if they only influence a sub-set of units.

2.3 Comparisons

Both equations can be interpreted as regressions, but (6) is a constrained regression, it has no intercept and the weights are non-negative and sum to one; whereas (7) is an unconstrained regression. SCM, just matching on covariates, can be estimated with fewer pre-treatment observations than PDA, which requires T_0 to be large enough to estimate a regression. SCM proponents criticize the fact that regression methods can give negative weights to controls. But this is to be expected if one interprets the procedure as involving prediction using common factors. Suppose Hong Kong before integration is largely driven by global factor A, the US by factors A and B, and Japan largely by factor B; then the US minus Japan provides an estimate of factor A, which drives Hong Kong.

In the case of microeconometric treatment effect studies, when the units are only subject to weak factors, SCM is sensible: choose controls that are similar in characteristics to those that are treated: match patients treated with a drug to untreated controls of similar age, sex, and health. Similarity is usually measured by propensity score, $p(x)$, the probability of treatment given the covariates; though since it is difficult to predict macroeconomic policies or civil wars, macro propensity scores may be less accurate than micro ones. We return to propensity scores below. In

addition the $\mathbf{x}_{jkt}$ are often poor predictors for y_{it}, which is why pre-intervention y_{it} are often used.

It is not clear that SCM is as sensible in macroeconometric time-series contexts, where there are strong common factors driving the y_{it}, so prediction from outcomes in other units y_{jt} may be more sensible than trying to identify units with similar x_{it} .

The SCM procedure requires that the matching variables of the treated units and the control units overlap, in microeconometric terms you have to find someone of the same age, sex and health as the treated person. In statistical terms this is referred to as the need for common support. The support of a variable is the range of values that it takes. Common support means that the range of the matching variables of the treated units is within the range of the matching variables of the control units. This may not hold in micro studies, if, for instance, there is a 100 year old man in the treated group and nobody over 90 in the control group. It is even more of a problem in macro studies. The common support assumption is unlikely to hold for Hong Kong, no other country is like Hong Kong, not even Singapore, the closest comparison. This is not a problem for the prediction method. SCM relies on interpolation, but the PDA can extrapolate. Of course one can question the relative accuracy of extrapolation relative to interpolation.

There is a growing literature on comparing SCM and PDA, which are likely to work well in different circumstances, depending for instance on the size of T_0 and whether one interpolates (the support of the controls covers the treated case) or extrapolates. Gobillon and Magnac (2016) investigate the use of interactive effect, Bai (2009), or linear factor models, difference in differences, (2), and SCM in regional policy evaluation. They show that (2) are generically biased, derive support conditions for SCM and use Monte Carlo to compare the methods, Wan et al. (2018) compare the panel time series approach and the synthetic control method. Hsiao & Zhou (2019) compare parametric, semi-parametric and non-parametric methods. SCM & PDA are non-parametric methods. Since they find that no method dominates in all circumstances they suggest model averaging.

Bove et al (2016) use both SCM and PDA to measure the effect on GDP of Civil Wars. They find that the results from the two methods are similar, perhaps because they both tend to weight the same countries. What makes a large difference is whether the outcome measure modelled is log GDP or the change in log GDP, the growth rate. These give very different results for the costs of civil war to a country. The sensitivity of results to specification is always an issue.

3. Extensions

There are many natural variants or extensions to these procedures, often allowing for exogenous variables and dynamics as well as unobserved factors. For notational simplicity, $\hat{y}^0_{1,T_0+h}$ is used for any estimated counterfactual and $\hat{\delta}^0_{1,T_0+h}$ for any estimated policy effect, not distinguishing between the methods used to generate them.

Suppose y_{it} is determined by an intercept, a vector of exogenous variables, $\mathbf{x}_{it}$, and a single factor f_t:

$$y_{it} = \alpha_i + \beta_i' \mathbf{x}_{it} + \lambda_i f_i + \varepsilon_{it} \tag{9}$$

Average across units to give

$$\bar{y}_t = \bar{\alpha}_t + \bar{\beta}_t' \bar{\mathbf{x}}_t + \bar{\lambda} f_t + \bar{\varepsilon}_t + N^{-1} \sum (\beta_i - \bar{\beta})' x_{it}$$

$$f_t = \bar{\lambda}^{-1} \{ \bar{y}_t - \bar{\alpha} + \bar{\beta}_t' \bar{\mathbf{x}}_t + [\bar{\varepsilon}_t + N^{-1} \sum (\beta_i - \bar{\beta})' x_{it}] \}$$

The term in [..] will average zero under fairly weak assumptions, so the $\bar{y}_t$ and $\bar{\mathrm{x}}_t$ provide a proxy for the unobserved factor, as long as $\bar{\lambda} \neq 0$. This is the basis of the correlated common effect, CCE, estimator of Pesaran (2006), which suggests filtering out the effect of the factors by adding the means of the dependent and independent variables to the regression for each unit

$$y_{it} = \alpha_i + \mathbf{x}_{it}' \beta_i + \gamma_{0i} \bar{y}_t + \gamma_i' \bar{\mathbf{x}}_t + u_{it} \tag{10}$$

Notice that the covariance between $\bar{y}_t$ and ε_{it} goes to zero with *N,* so for large *N* there is no endogeneity problem. The CCE generalises to many factors and lagged dependent variables, but requires *N* and *T* large.

Bai (2009) calls (9) with a homogeneous β and multiple factors an interactive fixed effect model

$$y_{it} = \alpha_i + \beta' \mathbf{x}_{it} + \boldsymbol{\lambda}_i' \mathbf{f}_t + \varepsilon_{it}$$

contrasting it with the usual additive fixed effect model which has $\alpha_i + \alpha_t$. Bai suggests an iterative principal component based method to determine the

r x *1* vector of factors $\mathbf{f}_t$, and to estimate the parameters. One advantage of CCE over alternative principal component based estimates of the factors is that one does not need to determine, *r*, the number of factors, which can be difficult.

Using data up to T_0 the CCE estimates could directly predict the counterfactual as

$$\hat{y}^0_{1,T_0+h}=\hat{\alpha}_1^{T_0}+\mathbf{x}'_{1,T_0+h}\,\hat{\beta}^{T_0}_{1,T_0+h}+\hat{\gamma}^{T_0}_{01}\bar{y}_{T_0+h}+\hat{\gamma}_1^{T_0}\bar{\mathbf{x}}_{T_0+h} \quad (11)$$

For large changes like German or Hong Kong re-unification, or joining the euro, it is not clear that there would be any country specific exogenous variables unaffected by the change. Having the weights specified a priori, equal in this case, avoids the problems associated with model selection or lasso based choices. If there are strong factors, the choice of weights might not matter very much.

Hsiao & Zhou (2019) characterise the Xu (2017) GSCM as a parametric method using covariates. Xu suggests estimating the homogeneous β, interactive fixed effect model

$$y_{it}=\beta'\mathbf{x}_{it}+\boldsymbol{\lambda}'\mathbf{f}_t+\varepsilon_{it}$$

on the data for the control group using all the observations, *t=1,2,...,T*. These estimates of β and $\mathbf{f}_t$ can then be used to estimate the factor loadings $\boldsymbol{\lambda}_1$ for the treated group using the pre-treatment data, up to T_0. The intercept is treated as a factor. The Xu counterfactual is then

$$\hat{y}^0_{1,T_0+h}=\hat{\beta}'\mathbf{x}_{1,T_0+h}+\hat{\boldsymbol{\lambda}}'_1\hat{\mathbf{f}}_{T_0+h}$$

Geng & Zhou (2018) combine the CCE and PDA approaches, to suggest what they call a panel data with exogenous regressors, PDX estimator Geng & Zhou provide asymptotic distributions that allow inference: tests and confidence intervals for the ATE. The method involves getting the CCE residuals and applying the the PDA approach to them. The model adds an exogenous variable, not affected by the treatment, with a homogeneous coefficient to (8) to give

$$y_{it}=d_{it}\delta_{it}+\alpha_i+\mathbf{x}'_{it}\beta+\lambda'_i\mathbf{f}_t+\varepsilon_{it} \quad (12)$$

The first stage is to estimate the CCE regressions for each country on the pre-treatment period, obtain the CCE residuals $y_{it}-\mathbf{x}'_{it}\hat{\beta}^{CCE}$. One can then apply the PDA approach by regressing the CCE residuals from the treated group on those for the untreated groups, using lasso or model selection procedures to identify the relevant controls. These estimates can be used to construct the counterfactual for the post-treatment period, $\hat{y}^0_{1,T_0+h}$, from which $\hat{\delta}_{1t}$ can be calculated.

Compared to SCM or GSCM, Geng & Zhou argue that this approach has the advantages of: that there is no need to impose constraints on both observables and unobservables and the number of parameters to be estimated in the model is greatly reduced. They use their method to measure the effect of US Stand Your Ground (SYG) laws on state murder rates and find adopting SYG laws tends to increase the murder rate. They use a number of different methods to measure the effect and note that the PDX fits the data in the pre-treatment period well compared with the other methods and different methods give very different estimates for the effects of the SYG laws.

Chan & Kwok (2016) also extend the Pesaran (2006) procedure and extract principal components from the control group to form factor proxies.

None of these procedures are fully dynamic. Pesaran & Smith (2016, 2018) explicitly consider testing for the effect of a policy intervention in a dynamic context, either in the case of a parsimonious reduced form or final form equation, P&S (2016), or in the context of a complete DSGE P&S(2018).

They suggest estimating for *t=1,2,...,*T_0 the pre-treatment period, the equation

$$y_{1t}=\alpha_{10}+\alpha_{11}\,y_{1,t-1}+\gamma_1\,y^*_{1t}+u_{it} \quad (13)$$

Where y^*_{1t} is some country specific function of the outcome variables in the untreated units. This function could be freely estimated, as in PDA, the mean as in CCE, or some other weighted average, such as the trade weighted averages of the other countries as used in the Global VAR, GVAR, Chudik & Pesaran (2016). The counterfactual is then the dynamic forecast for y_{1,T_0+h} conditional on the observed $y^*_{T_0+h}$. They consider tests for policy ineffectiveness.

In some cases one wishes to forecast a number of outcomes and ensure that the conditional forecasts are consistent with each other, in the sense that they exhibit their historical association. To that end Akhmadieva, & Smith (2019) estimate, what is called a VARX* in the GVAR literature. This is a vector version of (13) explaining an *m* x *1* vector of variables for a country that joined the euro, say country 1, by a corresponding vector of constructed foreign variables

$$\boldsymbol{y}_{1t} = \hat{\alpha}_{10}^{T_0} + \widehat{\mathbf{A}}_{11}^{T_0}\mathbf{y}_{1,t-1} + \widehat{\mathbf{C}}_{1}^{T_0}\mathbf{y}_{1t}^{*} + \hat{\mathbf{u}}_{it}^{T_0} \quad (14)$$

$\widehat{\mathbf{A}}_{11}^{T_0}$ and $\widehat{\mathbf{C}}_{1}^{T_0}$ are *m* x *1* matrices, where *m=6*. This is estimated on pre-treatment data, up to T_0, experimenting with various dates. One can then construct dynamic forecasts conditional on the $\mathbf{y}_{1t}^{*}$, which will provide the counterfactual. Part of the change associated with joining the euro is that post-treatment different euro policy reaction functions determined interest rates and exchange rates than the pre-treatment national ones. One can distinguish the effect of changing policy rules from other changes by also constructing a counterfactual in which the exchange rate and short interest rate are treated as exogenous. Estimating large unrestricted VARs on relatively short periods can produce unreliable results and some of the counterfactuals did not seem reasonable.

Pesaran & Smith (2016, 2018) point out that there are various types of policy change including discretionary interventions where there is a deterministic change to the policy variable and rule based interventions where one or more parameters of a stochastic policy rule are changed. P&S (2018) consider both a standard case where all variables in the macroeconometric model, including policy variables, are endogenous and a general case where the DSGE model is augmented by exogenous variables. The latter case accommodates interventions that change exogenous policy parameters, such as a fixed money supply target, or when steady states of some of the variables are changed as occurs when the inflation target is altered. They make the point that in stable DSGEs estimated on deviations from steady states, any policy changes which do not change the steady states will only have transitory effects and thus be difficult to detect.

All the procedures discussed so far compare the actuals to a counterfactual. Pesaran, Smith and Smith (2007) in considering what would have happened if the UK, or Sweden, had joined the euro compare two counterfactuals. Since the UK is so large relative to the euro area, it would have changed the behaviour of the euro area. They simulate the GVAR, with and without the constraints that UK interest rates and exchange rates were equal to the euro area ones. Over the period they considered 1999-2005, UK interest rates were similar to euro area ones and the sterling euro rate was very stable, so the effects were not large.

4. Propensity Score Methods

Another recent approach to macro policy evaluation also borrows techniques from the micro literature to obtain an estimate of an average treatment effect. Angrist, Jorda and Kuersteiner (2018), AJK, develop flexible semiparametric time series methods for the estimation of the causal effect of US monetary policy on macroeconomic aggregates. While Jorda and Taylor (2015) use similar procedures to estimate the effect of fiscal policy in a panel of countries. AJK use local linear projection type estimators to measure the average effect of policy changes on future values of the outcome variables (inflation, industrial production, and unemployment), using inverse probability weightings, in a way similar to that used to adjust non-random samples, the probability weights obtained from policy propensity scores. They have a parametric model for the propensity score the probability of policy, changes in the federal funds rate target rate announcements, conditional on daily financial market data. Identification comes from the assumption that information revealed by an announcement, conditional on market rates the day before the announcement, is independent of future potential outcomes.

They say it captures the average causal response to discrete policy interventions in a macrodynamic setting, without the need for assumptions about the process generating macroeconomic outcomes. The proposed estimation strategy, based on propensity score weighting, easily accommodates asymmetric and nonlinear responses. Their main conclusion is about the asymmetric effects: interest rates rises can slow the economy down but cuts do not boost it much.

They rely on outcomes averaged across different (possibly heterogenous) policy episodes whilst the studies surveyed so far consider a single policy intervention and average the counterfactual outcomes over the post intervention sample corresponding to that intervention. AJK do not use a model for outcomes and their analysis is potentially subject to the Lucas Critique. Their approach requires that the underlying parameters are invariant to policy changes, since it is only policy changes within the same regime that are identified in their framework (see AJK, p.373). In addition, matching estimators of this sort require a lot of data whereas macroeconometric samples tend to be data-poor relative to microeconometric samples. This isrejected in the large confidence bands AJK report around the measures of their estimated effects of target rate changes on macro variables in figs 2 and 3.

Terzi (2019) uses a propensity score matching models to construct counterfactuals for the euro area crisis countries (Greece, Portugal,Ireland, Cyprus, Spain) based on over 200 past macroeconomic adjustment episodes between 1960-2010 worldwide. At its trough, between 2010 and 2015 per capita GDP had contracted on average 11 percentage points more in the Eurozone periphery than in the standard counterfactual scenario.

5. Conclusion

Although they have not yet become the standard approach to macroeconomic questions these procedures, which draw on microeconometrics, have been increasingly used in macro policy evaluation. However, many studies show that the estimates of the counterfactual and the treatment effect are sensitive to the method used and to the particular specification choices within methods. Of course, this is also true of more conventional macro methods. Hsiao & Zhou (2019) comment that no method appears capable of dominating all other methods under all different DGPs and different sample configurations of cross-sectional dimension N and pretreatment time dimension T_0. Since the true DGP is usually unknown and the statistical findings could be very different for different situations, they suggest model-averaging as a robust method for generating counterfactuals. They also note that the absolute magnitude of the model average estimates could be very different from the estimates based on a particular method to generate counterfactuals.

This sensitivity to choice of method and detail of implementation raises the possibility that researchers might search over specifications in order to obtain the counterfactual that produces an effect that confirms their prejudices. Again this problem is not confined to these methods but indicates the importance of replication studies to investigate the robustness of results.

REFERENCES

Abadie, A. & M. D Cattaneo (2018). Econometric methods for program evaluation. *Annual Review of Economics*, 10, 465-503.

Abadie, A., A. Diamond, & J. Hainmueller (2010) Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program, *Journal of the American Statistical Association*,105(490) 493--505.

Abadie, A., A. Diamond, & J. Hainmueller (2015) Comparative politics and the synthetic control method, *American Journal of Political Science*, 59(2), 495-510.

Abadie, A. & J. Gardeazabal, Javier (2003) The Economic Costs of Conflict: A Case Study of the Basque Country, *American Economic Review*, 93(1) 113-132.

Akhmadieva, V & R. P. Smith (2019) The macroeconomic impact of the euro, BCAM Working Paper 1903. Birkbeck, University of London, London, UK, forthcoming in *Scientific Annals of Economics and Business*.

Angrist, J. D., O. Jorda & G. Kuersteiner (2018): Semiparametric estimates of monetary policy effects: string theory revisited, *Journal of Business & Economic Statistics*, 36:3, 371-387,

Bai, J. (2009) Panel Data models with interactive fixed effects, *Econometrica*, 77(4) 1229-1279.

Bove, V, L. Elia and R.P. Smith (2017) On the heterogenous consequences of civil war *Oxford Economic Papers*, 69(3) 550-568.

Chan, M.K. & S. Kwok (2016) Policy Evaluation with Interactive Fixed Effects. University of Sydney Economics Working paper 2016-11.

Chudik, A & M.H. Pesaran (2016) Theory and Practice of GVAR Modeling, *Journal of Economic Surveys*, 30(1) 165-197.

Gardeazabal, J., Vega-Bayo, A., (2016). An empirical comparison between the synthetic control method and Hsiao et al.'s panel data approach to program evaluation. *Journal of Applied Econometrics,* 32 (5), 983--1002.

Geng, H & Q. Zhou (2018) Estimation and inference of treatment effects using a new panel data approach with application to the impact of US SYG Law on state level murder rate,

Gobillon, L. & T. Magnac (2016) Regional Policy Evaluation: Interactive Fixed Effects and Synthetic Controls, *Review of Economics and Statistics*, 98(3) 535-551.

Hsiao, C., H.S. Ching, and K. W. Shui (2012), A panel data approach for program evaluation: measuring the benefits of political and economic integration of Hong kong with mainland China, *Journal of Applied Econometrics*, 27(5) p705--740.

Hsiao, C & Q. Zhou (2019) Panel parameteric, semiparametric and nonparametric construction of counterfactuals, *Journal of Applied Econometrics*

Li, K., & Bell, D. (2017). Estimation of average treatment effects with panel data: Asymptotic theory and implementation. *Journal of Econometrics*, 197, 65--75.

Jorda O. & A.M.Taylor (2016) The time for austerity: estimating the average treatment effect of fiscal policy. *Economic Journal*, 126(1) 219-255.

Pesaran, M.H. (2006) Estimation and Inference in Large Heterogeneous Panels with a multifactor error structure, *Econometrica*, 74(4) 967-1012.

Pesaran, M.H., L.V Smith and R.P. Smith (2007) What if the UK or Sweden had joined the Euro in 1999? An empirical evaluation using a Global VAR. *International Journal of Finance and Economics*, 12, 55-87.

Pesaran, M.H., and R.P. Smith (2016) Counterfactual analysis in macroeconometrics: an empirical investigation into the effects of quantitative easing, *Research in Economics*, 70(2), 262-280.

Pesaran M.H. & Smith R.P. (2018) Tests of policy interventions in DSGE models, *Oxford Bulletin of Economics and Statistics*, 80(3), 457-484.

Terzi, A. (2019) The Euro Crisis and Economic Growth:A Novel Counterfactual Approach, CESifo working paper, 7746.

Wan, S-K, Y Xie, C. Hsiao (2018) Panel data approach vs synthetic control method, *Economics Letters* 164, 121-123..

Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects model. *Political Analysis*, 25(1), 57--76.

(Endnotes)

[1]I am grateful to Veronika Akhmadieva for comments on an earlier version.

[2]These laws allow citizens more scope to use deadly force against others in situations like a burglary. Geng & Zhou (2018) analyse this case.

[3]The assumption that we know the date of treatment is not innocuous. One might argue that the relevant date for euro entry was in the early 1990s, after the ERM crisis, when countries began to adjust to meet the entry criteria; in 1999, the formal date; in 2002 when euro notes and coins were introduced; or in 2008 when the euro constraints began to bind.

EGE AKADEMİK BAKIŞ / *EGE ACADEMIC REVIEW*

Cilt 19 • Sayı 4 • Ekim 2019
SS. 401 - 410
Doi: 10.21121/eab.463383
Başvuru Tarihi: 24/9/2018 • Kabul Tarihi: 19/9/2019

Analyzing Financial Insufficiency of Households in Turkey with Multivariate Probit Model

Türkiye'de Hanehalklarının Finansal Yetersizliğinin Multivariate Probit Model ile Analizi

Çiler SİGEZE[1] *https://orcid.org/0000-0001-5329-5066*
Mehmet Sedat UĞUR[2] *https://orcid.org/0000-0002-2562-5299*

ABSTRACT

Permanent or particular financial insufficiencies of households cause crucial interruptions on compulsory payments of households. In this study, the determining factors of financial insufficiency of 11918 households that surveyed in TurkStat's Survey of Income and Living Conditions (SILC) 2012 are analyzed. The financial insufficiency of households can be denoted by several indicators. The responses of household head for three questions on financial conditions of household in survey are treated as indicators for financial insufficiency. Multivariate Probit Model is used to determine the effects of the properties of household head and households as independent variables on three dependent variables about financial conditions and to estimate these models simultaneously. The findings show that household size, being a tenant, changing the job, being unemployment but looking for a job and being chronic illness for household head have positive, income, being a female and being a graduate from university or a higher degree for household head have negative effects on each financial insufficiency indicators.

Keywords: Multivariate probit model, financial insufficiency, households, Turkey

ÖZET

Hanelerin zaman içerisinde sürekli ya da kısmı zamanlı olarak ekonomik anlamda sıkıntıya girmeleri yani finansal durumlarının yetersiz olması hanelerin zorunlu ödemelerinde aksaklıklar meydana getirebilmektedir. Bu çalışmada TÜİK 2012 yılı Gelir ve Yaşam Koşulları Araştırması (GYKA)'nda yer alan 11918 hanenin finansal yetersizliğinin belirleyicileri analiz edilmiştir. Hanelerin finansal yetersizliği çeşitli göstergelerle ifade edilebilmektedir. Çalışmada hanehalkı reisinin, ankette yer alan hanelerin ekonomik durumlarına ilişkin üç soruya verdiği yanıtlar finansal yetersizliğinin birer göstergesi olarak ele alınmaktadır. Hanelerin ekonomik durumları üzerine bilgi sağlayan üç bağımlı değişken için hanehalkı reisi ve hanenin özelliklerinin etkilerinin belirlenebilmesi ve bu modellerin aynı anda tahmin edilmesi amacıyla çalışmada Multivariate Probit Model kullanılmıştır. Elde edilen bulgular incelendiğinde, her bir finansal yetersizlik göstergesi üzerinde hanehalkı büyüklüğünün, hanehalkının kiracı olmasının, hanehakı reisinin iş değişikliği yapmasının, iş arıyor olmasının ve kronik hasta olmasının pozitif, hanenin gelirinin, hanehalkı reisinin kadın olmasının ve hanehalkı reisinin üniversite ve üstü eğitim düzeyinde olmasının ise negatif etkisi bulunmaktadır.

Anahtar Kelimeler: Multivariate probit model, finansal yetersizlik, hanehalkı, Türkiye

1. INTRODUCTION

One of the main determinants of the welfare levels of individuals is to consider whether they finance their regular payments. In that case, delivering the regular payments of households such as house rent, installments, credit card and other debts as well as electricity, water and gas bills are seen as the substantial indicators for their subsistence levels. Also, the incapability to pay these basic payments for many times will cause both the household economy to weaken and

[1]**Çukurova University Faculty of Economics and Administrative Sciences, Department of Econometrics, csigeze@cu.edu.tr**
[2]**Çankırı Karatekin University Faculty of Economics and Administrative Sciences, Department of Economics, sedatugur@karatekin.edu.tr**

the individuals to become socially and psychologically distressed.

There are several studies on analyzing economic conditions of individuals or households which are mostly concerned with the poverty status of the relevant units. In general, the logit and the probit models are used to analyze the determinants of the poverty status of households. Financial insufficiency can be seen as an indicator for households that suffering from poverty. Thus, multiple indicators can be used on determining the financial insufficiency of households. In the case where these indicators are set as dependent variables, it is appropriate to use multivariate probit (MVP) model to examine the affects of independent variables on these indicators by considering the correlation between the dependent variables. Using the MVP model makes it possible to analyze whether the units of analysis finance their basic and compulsory payments and to examine the deprivation conditions of households by considering multiple indicators instead of taking a solely income-based indicator.

In this study, we used three financial condition indicators as dependent variables which are generally used as dimensions of subjective wellbeing in the literature, because they include direct responses of households in relevant survey. The paper will start with a brief background study about the relevant theory and literature and this will comprise the following chapter. On third chapter, we will explain the data and methodology used in our analysis by describing multivariate probit model and our data set in brief. The final two chapters will include the findings and the conclusions of the study.

2. BACKGROUND

Investigating the living standards of the individuals is a long-concerned issue in economics. These studies are mostly encompassed by poverty studies. But poverty, as a quite complex concept in its nature, has many definitions. Income or expenditure poverty (World Bank, 1990; Ravallion, 1996), lack of basic needs (Streeten, Burki, ul Haq, Hicks & Stewart, 1981), relative deprivation (Runciman, 1966; Townsend, 1979), lack of capabilities and functionings (Sen, 1985; Nussbaum, 2000), social exclusion (Rodgers, Gore and Figueiredo, 1995; Bhalla and Lapeyre, 1997), human underdevelopment (UNDP, 1990; ul Haq, 1995), multidimensional poverty (Bourguignon & Chakravarty, 2003; Alkire & Foster, 2009), vulnerability (Morduch, 1994), ill-being (Brock, 1999) and unsustainability of subsistence levels (Chambers and Conway, 1992) are some of these definitions, among others. One other dimension of poverty is related with life satisfaction and happiness of individuals which is called as subjective wellbeing. Orshansky (1969) defines poverty as a value judgment and for Orshansky, poverty, like beauty, lies in the eye of the beholder. In here, poverty is seen as a subjective concept and related with wellbeing. Subjective wellbeing is defined as judging life positively, often feeling joy and experiencing sadness infrequently. That definition is linked with Diener et al (1997)'s definition, as cited by Veenhoven (2008): "a person is said to have high subjective wellbeing if she/he experiences life satisfaction and frequent joy, and only infrequently experiences unpleasant emotions such as sadness and anger".

Subjective wellbeing studies are generally related with the concept of happiness in economics. As Veenhoven (1984) noted, overall happiness is synonymous with life satisfaction and subjective wellbeing. Veenhoven (2000) also defines subjective wellbeing as "four qualities of life" by including both inner and outer qualities of life chances and life results. These qualities of life involve the livability of environment, external utility of life, life-ability of person and inner appreciation of life (Veenhoven, 2000). There are other considerable studies in literature which analyze subjective wellbeing (van Praag, 1971; van Praag and Kapteyn, 1973; Easterlin, 1974; Oswald, 1997; Pradhan and Ravallion, 2000; van Praag et al., 2003; among others). Although we accept that general happiness of life includes all dimensions of subjective wellbeing which may include health, leisure and environment satisfaction, among others; we only use financial indicators to determine financial wellbeing of households in this study. Thus, rather using the concept of financial wellbeing, we use the concept of financial insufficiency in our analysis and the analysis mostly concern with the financial domains of life of households.

There are only a few studies for Turkey which are related with our study. One of them is Selim (2008)'s study which investigates life satisfaction and happiness in Turkey by using an ordered logit model, and they found that employment and health status, in particular, exert a strong influence on life satisfaction. Dal and Sevuktekin (2018) also investigate the factors for life satisfaction with ordered logit model by using the Survey of Income and Living Conditions (SILC) 2013 for Turkey. Their findings show that the satisfaction of individuals in financial conditions has the key

effect on general life satisfaction. This inference shows the priority of financial conditions, among others. To our knowledge, this study is the first attempt in the literature for Turkey using the relevant methodology and data. Çevik and Korkmaz (2014), Eren and Aşıcı (2017) and Timur and Akay (2017) are other related studies that investigate subjective wellbeing in Turkey for specific periods.

3. DATA AND METHODOLOGY

3.1. Data And Variables

We used cross-sectional data set of Turkstat's Survey of Income and Living Conditions (SILC) for 2012. This survey has been annually published since 2006. The size of the sample for our period of analysis is 11918 households. There are some questions in the survey which are asked to the household head and are shown whether the income of household is insufficient. Three of these questions on financial condition of households are as follows: "whether they have the incapability to finance arrears on mortgage, loan repayments or rent payments in the last 12 months", "whether they have the incapability to finance arrears on utility bills in the last 12 months" and "whether they have the incapability to finance arrears on hire purchase installments, credit cards or other loan payments in the last 12 months". The answers are classified as "yes", "no" and "there is no such kind of payment" in the survey. The dependent variables are created as taking on the values of 0 and 1, where 1 means that the answer of the household head is "yes" and 0 is otherwise. Table 1 shows the basic definitions and descriptive statistics of the variables which are used in the study.

Table 1: Descriptive Statistics

Variables	Definition	Mean	Std. Err.
Dependent Variables			
Financial condition 1	incapability to finance arrears on mortgage, loan repayments or rent payments in the last 12 months	0.088	0.002
Financial condition 2	incapability to finance arrears on utility bills in the last 12 months	0.373	0.004
Financial condition 3	incapability to finance arrears on hire purchase installments, credit cards or other loan payments in the last 12 months	0.265	0.004
	Household Head Related Variables		
Independent Variables	**Definition**	**Mean**	**Std. Err.**
Female	1 if the household head is female, 0 otherwise	0.160	0.003
Age	The age of the household head	49.082	0.14
Age ^{2}	The square of the age of the household head	2643.391	14.808
Chronic illness	1 if the household head is chronically ill, 0 otherwise	0.380	0.004
SSI (Social Security Institution)	1 if the household head is registered to social security institutions in the main job, 0 otherwise	0.414	0.004
Job replacement	1 if the household head has changed his/her job in the last 12 months, 0 otherwise	0.082	0.002
Marital Status			
Single	1 if the household head is single, 0 otherwise	0.03	0.001
Married	1 if the household head is married, 0 otherwise	0.823	0.003
Widow or divorced (reference)	1 if the household head is widowed or divorced, 0 otherwise	0.147	0.003
Education			
Illiterate or Literate (reference)	1 if the household head is illiterate or literate but not a graduate, 0 otherwise.	0.164	0.003
Primary-Secondary school	1 if the household head is graduated from primary school or secondary, vocational secondary or primary education school, 0 otherwise.	0.545	0.004
High School	1 if the household head is graduated from high school, or vocational or technical high school, 0 otherwise	0.161	0.003
University	1 if the household is graduated from faculty/university, college or higher education level, 0 otherwise	0.13	0.003
Working Situation			
Employee	1 if the household head is employed, 0 otherwise	0.648	0.004
Unemployed (Looking for a job)	1 if the household head is looking for a job, 0 otherwise	0.028	0.001

Retired	1 if the household head is retired, 0 otherwise	0.175	0.003
Inactive (reference)	1 if the household head is old, permanently disabled and/or unfit to work or fulfilling domestic tasks and care responsibilities, 0 otherwise	0.149	0.003
Employment Type			
Regular	1 if the household head is a regular employee, 0 otherwise	0.355	0.004
Casual	1 if the household head is a casual employee or an unpaid family worker, 0 otherwise	0.065	0.002
Employer	1 if the household head is an employer, 0 otherwise	0.038	0.001
Self-employed	1 if the household head is self-employed, 0 otherwise	0.19	0.003
Household Related Variables			
Household size	The size of the household	3.983	0.019
Tenant status	1 if the household is tenant, 0 otherwise	0.202	0.003
Income	The logarithm of the income of households.	9.893	0.006

We classified our variables as two main categories which one is related directly to the household head and the other one is concerned with the households as a whole. All variables which are related with household head are categorical and dummy variables with the exception of age variable. Since we want to investigate whether female headed households are more vulnerable to financial insufficiency, we accept females as our primal gender variable.

One other important group of variables is marital status of the household head. In here, reference variable is widowed or divorced individuals. Again, we know that educational level and working situation of household head is a key component on determining the welfare of household. Also, we define some other independent variables which are seen as determining factors for financial insufficiency such as household size.

Table 1 shows that the 8.8% of households in the survey do not have the capability to finance arrears on mortgage, loan repayments or rent payments in the last 12 months. On the other hand, 37.3% of households do not have the capability to finance arrears on utility bills in the last 12 months. Finally, 26.5% of the households stated that they do not have the capability to finance arrears on hire purchase installments, credit cards or other loan payments in the last 12 months. Thus, it appears that the households suffer mostly in paying utility bills which includes the basic needs of the individuals such as electricity, water and gas.

3.2. Multivariate Probit Model

The multivariate probit (MVP) model denotes the influence of the set of independent variables on dependent variables by allowing the unobserved error terms to be freely correlated (Golob and Regan, 2002; Greene, 2003) which is also one form of a correlated binary response regression model (Castillo-Manzano, 2010). The general specification for M-equation multivariate probit model can be expressed as (Cappellari and Jenkins, 2003);

$$y_{im}^{*} = \beta_m{}' X_{im} + e_{im} \quad m = 1,2,...,M \tag{1}$$
$$y_{im} = 1 \text{ if } y_{im}^{*} > 0 \text{ and } 0 \text{ otherwise,}$$

where e_{im} represents error terms that have multivariate normal distribution with zero mean and variance -covariance matrix V with values 1 on the leading diagonal and correlations $\rho_{jk} = \rho_{kj}$ as off-diagonal elements. We focus on the trivariate probit case in which M=3, then the log likelihood function for a N sample of observations can be written by;

$$L = \sum_{i=1}^{N} w_i \log \Phi_3(\mu_i, \Omega) \tag{2}$$

where w_i is an optional weight for observation $i = 1,2,...,N$ and Φ_3 is the trivariate standard normal distribution with arguments μ_i and Ω. $\mu_i = (K_{i1}\beta_1' X_{i1}, K_{i2}\beta_2' X_{i2}, K_{i3}\beta_3' X_{i3})$. is a constituent of , $K_{i1} = 2y_{i1} - 1$, $K_{i2} = 2y_{i2} - 1$ and $K_{i3} = 2y_{i3} - 1$. Here Ω consists of $\Omega_{jj} = 1$ for $j = 1,2,3$ and $\Omega_{12} = \Omega_{21} = K_{i1}K_{i2}\rho_{21}$, $\Omega_{13} = \Omega_{31} = K_{i3}K_{i1}\rho_{31}$, $\Omega_{23} = \Omega_{32} = K_{i3}K_{i2}\rho_{32}$ (Cappellari and Jenkins, 2003).

MVP models are estimated by using the simulated maximum likelihood method that uses Geweke-Hajivassiliour-Keane (GHK) smooth recursive conditioning simulator procedure to evaluate the multivariate normal distribution (Cappellari and Jenkins, 2003). The estimations are made by using STATA 14.0 software.

In the study, we calculated eight joint probabilities corresponding to the eight possible combinations of successes ($y_{im}=1$) and failures ($y_{im}=0$) in the trivariate case. For example, the possibility that every outcome is a success with a conditioning upon unobservable variables that are correlated with each other is given by (Cappellari and Jenkins, 2003);

$$P(y_1=1,y_2=1,y_3=1)=P\left(e_1\le\beta_1'X_1,e_2\le\beta_2'X_2,e_3\le\beta_3'X_3\right) \quad (3)$$

$$=P\left(e_3\le\beta_3'X_3\,|\,e_2<\beta_2'X_2,e_1<\beta_1'X_1\right)\times P\left(e_2<\beta_2'X_2\,|\,e_1<\beta_1'X_1\right)\times P\left(e_1<\beta_1'X_1\right) \quad (4)$$

4. EMPIRICAL RESULTS

We estimated a multivariate model with three dependent variables to analyze the determining factors of financial insufficiency of households in the last 12 months. The first model implies the incapability to finance arrears on mortgage, loan repayments or rent payments in the last 12 months that is named as the financial condition 1. The second model specifies the incapability to finance arrears on utility bills in the last 12 months that is named as the financial condition 2. Finally, the third model implies the incapability to finance arrears on hire purchase installments, credit cards or other loan payments in the last 12 months that is named as the financial condition 3. Estimation results for the MVP models are shown in table 2 below. The models 1, 2 and 3 respectively show the estimation results for the dependent variables of financial conditions 1, 2 and 3.

Table 2: Multivariate Probit Model Coefficient Estimation Results

	Model 1		Model 2		Model 3	
	Coef.	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Household Head Related Variables						
Female	-0.153*	0.091	-0.192***	0.056	-0.105*	0.057
Age	0.038***	0.011	0.020***	0.005	0.019***	0.005
Age ^{2}	-0.0005***	0.0001	-0.0003***	0.00005	-0.0003***	0.00005
Chronic illness	0.155***	0.052	0.073**	0.028	0.148***	0.028
SSI (Social Security Institution)	-0.124*	0.070	-0.159***	0.039	-0.043	0.040
Job replacement	0.236***	0.066	0.206***	0.045	0.267***	0.044
Single	-0.016	0.097	-0.146***	0.055	-0.065	0.057
Maried	-0.179	0.139	-0.198**	0.091	-0.159*	0.093
Education						
Primary-Secondary school	0.030	0.084	-0.128***	0.041	-0.075*	0.042
High School	-0.127	0.100	-0.288***	0.053	-0.178***	0.054
University or higher	-0.396***	0.116	-0.610***	0.064	-0.476***	0.064
Working Situation						
Employee	0.265	0.515	0.475	0.350	0.237	0.357
Unemployed (looking for a job)	0.246*	0.131	0.247***	0.086	0.266***	0.083
Retired	-0.073	0.110	-0.242***	0.055	-0.055	0.057
Employment Type						
Regular	-0.387	0.509	-0.512	0.348	-0.216	0.356
Casual	-.262	0.511	-0.367	0.348	-0.144	0.355
Employer	-0.237	0.522	-0.473	0.354	-0.208	0.362
Self-employed	-0.384	0.509	-0.413	0.348	-0.177	0.355

Household Related Variables						
Household size	0.045***	0.126	0.115***	0.006	0.075***	0.006
Tenant status	2.096***	0.051	0.067**	0.032	0.199***	0.032
Income	-0.190***	0.039	-0.360***	0.022	-0.197***	0.022
Constant	-1.102**	0.452	2.966***	0.237	0.947***	0.242
Correlation Coefficient						
rho21	0.379***	0.018				
rho31	0.390***	0.019				
rho32	0.659***	0.010				
Likelihood ratio test	chi2(3)= 2674.84		Prob>chi2=0.000			

*,**,*** denote significance at the %10, %5 and %1 level respectively.

We began with applying the likelihood ratio test to analyze the existence of the relationship between error terms of models and the degree of correlation. The basic (H_0) hypothesis in this test presents that the correlations between models are equal to zero. If the correlations between models are equal to zero, the models will be independent to each other, and then they can be estimated separately. Table 2 shows that H_0 hypothesis is rejected due to the likelihood ratio test. Then, there is a correlation between models and they need to be estimated altogether. So, it is appropriate to use MVP model in the analysis.

According to table 2, coefficients of female, university-related and income variables are statistically significant and negative for all three models. It means that these variables have negative effects for households on having financial insufficiency. Recent studies (Devicenti, 2002; Andriopoulou and Tsakloglou, 2011; Wallelign et al., 2016; Acar and Başlevent, 2014; Dalgıc et al., 2015) also support the result that being a female household head has a negative effect on financial insufficiency of household.

For the variables of educational status of households, we defined "household head is illiterate or literate but not a graduate" variable as a reference variable. According to model 1, the probability of the incapability to finance arrears on mortgage, loan repayments or rent payments in the last 12 months decreases for only the households who graduated university or higher. The findings of model 2 and model 3 show that the probability of the incapability to finance arrears on utility bills in the last 12 months and the incapability to finance arrears on hire purchase installments, credit cards or other loan payments in the last 12 months decreases as educational level of household head increases. The common literature between educational level and financial status also support that view (Tilak, 1989; Easterlin, 2001; Bourguignon and Chakravarty, 2003).

The coefficients of income variable are negative and statistically significant in all three models, as expected. As income increases, the probability of incapability to finance compulsory payments decreases. It is obvious here that the income is a vital variable on financial conditions of households as it is an important indicator for their subsistence levels.

According to table 2, the coefficients of chronic illness, job replacement, unemployed (looking for a job), household size and tenant status variables are positive and statistically significant. It means that if the household head has a chronic illness, he/she has changed job in the last 12 months, he/she is unemployed but looking for a job or he/she is a tenant, it is probable that he/she will be incapable to finance his/her compulsory payments. For the household head, having a chronic illness may cause household to have an irregular income which is consequently resulted as incapability to finance their compulsory payments. Similarly, changing the job in the last 12 months leads household to have an irregular income. As a result, the households are unable to cover their compulsory payments at least once in the relevant period. In that case, unemployment insurances may be useful for these households to avert their income volatilities. Increasing household size causes negative effects in financial conditions of households, as expected. Promoting public awareness on family planning by the state and helping the families on child and elderly care may secure households to get rid of financial insufficiency.

Being a tenant has a positive effect on incapability to finance basic payments at least once in all three

models. But reasonably, the highest coefficient value is seen in model 1 which includes the financial condition variable related to the incapability to finance arrears on mortgage, loan repayments or rent payments in the last 12 months.

One of the remarkable findings of the analysis is that the employment type of the households has no effect on the incapability to finance the relevant payments. It means that being a regular, casual or self-employed worker is not related with relevant financial conditions, according to our models. Furthermore, according to household head who is inactive (old, permanently disabled and/or unfit to work or fulfilling domestic tasks and care responsibilities) in economic lives which is the reference variable, working situation of household as an employee is also statistically insignificant in all three models. Therefore, occupation of the household and his/her status as employed is not solely sufficient for households to make ends meet. This is in fact one other indicator that shows the income of the household also is not solely sufficient at that point. Planning the incomes and expenses well enough to ensure household welfare becomes a crucial issue at that perspective.

In model 2, according to inactive household head variable as reference, working situation of household as retired has negative effect on the probability of the incapability to finance arrears on utility bills in the last 12 months. In that case, being retired means that there is a regular income flow to household and then paying utility bills is delivered as a routine.

For household head, being registered to Social Security Institution (SSI) has a negative effect on incapability to finance relevant payments for model 1 and model 2, but it has no statistically significant effect on model 3.

According to the widowed or divorced household head variable which is the reference variable, being a single and married household head has a negative effect on incapability to finance utility bills in the last 12 months (model 2). Therefore, it is observed that single and married individuals have better financial conditions than the others. Also, being a single household head has a negative effect on incapability to finance arrears on hire purchase installments, credit cards or other loan payments in the last 12 months (model 3). But, marital status of the household head has no statistically significant effect for model 1.

For all three models, the coefficients of the age of the household head are positive, but the coefficients of the age squared of the household head are negative. This means that the probability of incapability to finance compulsory payments increases as the age of the household head increases; but after the mean age, the probability of incapability to finance these payments decreases as the age of the household head increases. The mean of the household age is found 49 years in this study. Here we can say that the job experience along with the continuity in earnings and savings maintain to finance compulsory payments for the households whose ages are above the mean age.

In addition, we predicted the probability of success for each regression. Table 3 shows the summary statistics about the predicted probability of success for relevant variables:

Table 3: Summary Statistics About The Predicted Probability Of Success

Variable	Observation	Mean	Std. Dev.	Min	Max
pall1s	11918	0.038	0.079	1.80e-07	0.714
pall0s	11918	0.538	0.205	0.012	0.992
pmargm1	11918	0.089	0.166	1.02e-06	0.843
pmargm2	11918	0.370	0.185	0.002	0.982
pmargm3	11918	0.264	0.126	0.003	0.818

The “pmarg” and “pall” shows predicted marginal success probability for each model and joint probabilities, respectively. The joint probabilities refer to the probability that each binary outcome = 1 for successes, or that each binary outcome = 0 for failures, for the each observations of the variables (Cappellari and Jenkins, 2003). The variable of pall1s show the probability of the dependent variable for an observation on taking the value of 1 in each models which present the probability of incapability to finance basic payments of households for all three compulsory financial conditions. The mean value of this probability is around 3%. Lower mean for probability of incapability to finance basic payments of households simultaneously reflects the structure of the dataset. On the other hand, the variable of pall0s show the probability of the dependent variable for an observation on taking the value of 0 in each models which present the probability of capability to finance basic payments of households for all three compulsory financial conditions. The mean value of this probability is around 53% and the maximum value is 99.2%. It indicates that the proba-

bility to finance all type of compulsory payments for households is quite high.

The "pmarg" specifies the predicted marginal success probability from each model, and the main findings show that the average marginal success probabilities of the models are around 8%, 37% and 26% respectively. Thus, the model 2 is identified as the most successful model. As a matter of fact, criteria of the maximum values are taken into account; the highest marginal probability of success in the model 2 is higher than the other two models and the value of probability is quite close to 1.

In general, the findings show that the model 2 is the most successful model in our three models on predicting the incapability to finance relevant payments which are taken as the financial insufficiency indicators of households. It means that the incapability to finance utility bills have a key role on determining the financial insufficiency of households.

5. CONCLUSION

This study is aimed to determine whether the households experience financial insufficiency in Turkey by using the data set of TurkStat's SILC 2012 and identify the type of payments for the households that having difficulties to meet when they have financial insufficiency. In that case, we analyzed the determining factors of incapability to finance arrears on mortgage, loan repayments or rent payments, utility bills, hire purchase installments, credit cards or other loan payments in the last 12 months. Because of the dependent variables which present three different payment difficulties are correlated, we used multivariate probit model. Thus, it is important to identify whether the compulsory payments are made and investigate these cases with multiple financial-based indicators rather than using a sole income-based indicator.

The findings of the study show that the main indicator of the financial insufficiency of the households is the incapability to finance utility bills in the last 12 months. Also, being a female household head and faculty or higher-degree graduated household head have a negative effect on incapability to finance compulsory payments in all three models. This means that these variables have positive impacts for households on not having financial insufficiency. Additionally, being chronically ill, changing job in the last 12 months, being unemployed but looking for a job, the size of the household and being a tenant have a positive effect on incapability to finance compulsory payments in all three models. Finally, it is found that being an employee and the employment types do not have any statistically significant effect on the problem of incapability to finance the compulsory payments of the households. Accordingly, the findings show that being an employed household head is not the solely adequate variable itself for the households to make their ends meet. But, the sufficient income flows to households and also decreasing the quantity of payments have a substantial role on welfare of the households.

As a conclusion, the individuals need to endeavor well on planning household economics and the policy-makers have to work quite efficiently on promoting employment and education to individuals, family planning and operating decent social security institutions.

REFERENCES

Acar, A. ve Baslevent, C. (2013) "Examination of the Transition of Turkish Households into and out of Poverty between 2007-2010" (No. 5779), *EcoMod*.

Alkire, S. ve Foster, J. (2009) "Counting and Multidimensional Poverty Measurement" *OPHI Working Paper*. No. 32. December. Oxford Department of International Development, Queen Elizabeth House (QEH), University of Oxford.

Andriopoulou, E. ve Tsakloglou, P. (2011) "The Determinants of Poverty Transitions in Europe and the Role of Duration Dependence" *IZA Discussion Paper* No. 5692.

Bhalla, A. ve Lapeyre, F. (1997) "Social Exclusion: Towards an Analytical and Operational Framework" *Development and Change,* 28: 413-433.

Bourguignon, F. ve Chakravarty, S.R (2003) "The Measurement of Multidimensional Poverty" *The Journal of Economic Inequality*, 1 (1): 25-49.

Brock, K. (1999) "*It's Not Only Wealth That Matters - It's Peace of Mind Too: A Review of Participatory Work on Poverty and Illbeing"* Washington: World Bank.

Cappellari, L. ve Jenkins, S. P. (2003) "Multivariate Probit Regression using Simulated Maximum Likelihood" *The Stata Journal*, 3(3), 278-294.

Castillo-Manzano, J. I. (2010) "Determinants of Commercial Revenues at Airports: Lessons Learned from Spanish Regional Airports" *Tourism Management*, *31*(6), 788-796.

Chambers, R. ve Conway, G. (1992) "Sustainable Rural Livelihoods: Practical Concepts for the 21st Century" *IDS Discussion Paper.* No. 296. Brighton: IDS.

Çevik, N. K. ve Korkmaz, O. (2014) "Türkiye'de Yaşam Doyumu ve İş Doyumu Arasındaki İlişkinin İki Değişkenli Sıralı Probit Model Analizi" Ömer *Halisdemir* Üniversitesi İktisadi *ve* İdari *Bilimler Fakültesi Dergisi*, 7(1), 126.

Dal, S. ve Sevüktekin, M. (2018) "Türkiye'de Öznel İyi Oluş'un Yaşam Alanları Yaklaşımı İle Ölçülmesi" *Uluslararası* İktisadi *ve* İdari İncelemeler *Dergisi*, 2018: 433-446.

Dalgıç, B., İyidoğan, P. V. ve Güven, A. (2015) "Yoksulluk ve Yoksulluk Geçişlerinin Belirleyenleri: Türkiye Örneği", *Sosyoekonomi*, *23*(24):51-70.

Devicienti, F. (2002) "Poverty Persistence in Britain: A Multivariate Analysis using the BHPS, 1991–1997" *Journal of Economics*, 9:307-340.

Diener, E., Suh, E. ve Oishi, S. (1997) "Recent Findings on Subjective Well-being" *Indian Journal of Clinical Psychology*, 24: 25-41.

Easterlin, R.A. (1974) "Does Economic Growth Improve the Human lot? Some Empirical Evidence" In David, P.A., and Reder, M.W. (eds.). *Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz* (pp. 89- 125). Academic Press, NY.

Easterlin, R. A. (2001) "Income and Happiness: Towards a Unified Theory", *The Economic Journal*, 111(473): 465-484.

Eren, K. A. ve Aşıcı, A. A. (2017) "The Determinants of Happiness in Turkey: Evidence from City Level Data", *Journal of Happiness Studies*, 18(3): 647-669.

Golob, T. F. ve Regan, A. C. (2002) "Trucking Industry Adoption of Information Technology: A Multivariate Discrete Choice Model", *Transportation Research Part C: Emerging Technologies*, 10(3): 205-228.

Greene, W. H. (2003) *Econometric Analysis*, Pearson Education: India.

Morduch, J. (1994) "Poverty and Vulnerability" *The American Economic Review*, 84(2): 221-225, Papers and Proceedings of the Hundred and Sixth Annual Meeting of the American Economic Association (May).

Nussbaum, M. (2000) *Women and Human Development*, Cambridge University Press.

Orshansky, M. (1969) "How Poverty is Measured" *Monthly Labor Review*, 92(2): 37-41.

Oswald, A. J. (1997) "Happiness and Economic Performance" *The Economic Journal*, 107(445): 1815-1831.

Pradhan, M. ve Ravallion, M. (2000) "Measuring Poverty using Qualitative Perceptions of Consumption Adequacy" *Review of Economics and Statistics*, 82(3): 462-471.

Ravallion, M. (1996) "Issues in Measuring and Modelling Poverty" *The Economic Journal*. 106(438): 1328-1343.

Rodgers, G., Gore, C. ve Figueiredo, J. B. (1995) "*Social Exclusion: Rhetoric, Reality, Responses*" International Institute for Labour Studies: UNDP.

Runciman, W.G. (1966) "*Relative Deprivation and Social Justice*", Routledge and Kegan Paul: London.

Selim, S. (2008) "Life Satisfaction and Happiness in Turkey" *Social Indicators Research*, 88(3):531-562.

Sen, A. (1985) "*Commodities and Capabilities",* Amsterdam: Elsevier.

Streeten, P., Burki, S. J., ul Haq, M., Hicks, N. ve Stewart, F. (1981) "First Things First: Meeting Basic Human Needs In The Developing Countries" Washington D. C.: OUP.

Tilak, J. B. (1989) "Education and Its Relation to Economic Growth, Poverty, and Income Distribution: Past Evidence and Further Analysis" World Bank Discussion Paper No.46.

Timur, B. ve Akay, E. Ç.(2017) "Kadın ve Erkeklerin Mutluluğunu Etkileyen Faktörlerin Genelleştirilmiş Sıralı Logit Modeli ile Analizi" *Social Sciences Res. Jou.*, 6(3): 88-105.

Townsend, P. (1979) Poverty in the United Kingdom, University of California Press.

TÜİK (2012) Gelir ve Yaşam Koşulları Araştırması. TÜİK: Ankara.

ul Haq, M. (1995) Reflections on Human Development. N.Y.: Oxford University Press.

UNDP (1990) Human Development Report 1990. New York: Oxford University Press.

van Praag, B.M.S. (1971) "The Welfare Function of Income in Belgium: An Empirical Investigation" *European Economic Review*, 2, 337-69.

van Praag, B.M.S. ve Kapteyn, A. (1973) "Further Evidence on the Individual Welfare Function of Income: An Empirical Investigation in the Netherlands" *European Economic Review*, 4: 33-62.

van Praag, B. M., Frijters, P. ve Ferrer-i-Carbonell, A. (2003) "The Anatomy of Subjective Well-being" *Journal of Economic Behavior & Organization*, 51(1):29-49.

Veenhoven, R. (1984) The Concept of Happiness, In R. Veenhoven (ed.), *Conditions of Happiness* (pp. 12-38). Springer: Dordrecht.

Veenhoven, R. (2000) "The Four Qualities of Life" *Journal of Happiness Studies*, 1(1), 1-39.

Veenhoven, R. (2008) "Sociological Theories of Subjective Wellbeing", In Michael Eid & Randy Larsen (Eds). *The Science of Subjective Well-being: A tribute to Ed Diener* (pp. 44-61), Guilford Publications, New York.

Walelign, S. Z., Charlery, L., Smith-Hall, C., Chhetri, B. B. K. ve Larsen, H. O. (2016) "Environmental Income Improves Household-Level Poverty Assessments and Dynamics" *Forest Policy and Economics*, 71:23-35.

World Bank (1990) World Development Report 1990: Poverty. New York: OUP.

EGE AKADEMİK BAKIŞ / *EGE ACADEMIC REVIEW*

Cilt 19 • Sayı 4 • Ekim 2019
SS. 411 / 424
Doi: 10.21121/eab.476419
Başvuru Tarihi: 30/10/2018 • Kabul Tarihi: 12/9/2019

A Cross-Cultural Validation of Consumer Arrogance Scale in Turkey and Romania

Tüketici Kibiri Ölçeğinin Türkiye ve Romanya'da Çapraz Kültürel Geçerliliği

Betül BALIKCIOGLU[1] *https://orcid.org/0000-0001-7043-2544*
Muzeyyen ARSLAN[2] *https://orcid.org/0000-0002-9852-7243*

ABSTRACT

Consumer arrogance (CA) is a new notion in consumer behavior, which is still unclear if CA has the same meaning in cross-cultures. The purpose of this study is to test the cross-cultural validation of the CA scale in Turkey and Romania, which are considered as collectivist cultures and developing countries. An empirical study was conducted on 145 Turkish and 147 Romanian university students. The cross-cultural validation tested via configural, metric, and covariance methods. A confirmatory factor analysis was used to test the validity of the measurement theory. The current research demonstrates the CA scale has an adequate fit to the data in each sample and across-cultural invariance between two countries. Also, the results indicate that the CA scale can be established as a second order construct and it is predicted by materialism.

Keywords: Arrogance, Consumer arrogance, Cross-cultural validity, Materialism.

ÖZET

Tüketici kibiri, tüketici davranışı alan yazınında yeni bir kavramdır. Bununla birlikte, tüketici kibirinin çapraz kültürlerde aynı anlama sahip olup olmadığı hala belirsizdir. Bu yazının amacı tüketici kibiri ölçeğinin, toplulukçu kültüre sahip olan ve gelişmekte olan ülkeler olarak kabul edilen Türkiye ve Romanya'da kültürler arası güvenilirliğini ve geçerliliğini test etmektir. Bunun için, 145 Türk ve 147 Romen üniversite öğrencisi üzerinde görgül bir çalışma yürütülmüştür. Ölçeğin kültürler arası geçerliliği, yapısal değişmezlik, metrik ve kovaryans yöntemleriyle test edilmiştir. Ölçüm teorisinin geçerliliğini test etmek için doğrulayıcı faktör analizi kullanılmıştır. Araştırma sonuçları, CA ölçeğinin her bir örneklemdeki verilerle istatistiksel olarak kabul edilebilir uyumu ve iki ülke arasındaki kültürel değişmezliği gösterdiğini ortaya çıkarmıştır. Ayrıca bulgular, CA ölçeğinin ikinci dereceden bir yapı olarak kurulabildiğini ve materyalizm tarafından tahmin edildiğini göstermektedir.

Anahtar Kelimler: Kibir, Tüketici kibiri, Çapraz-kültürel geçerlilik, Materyalizm.

Introduction

Most of attention in studying consumer behavior aspects is the consumption behaviors which serve to satisfy consumer's identity motives. For instance, consumers who need distinctiveness and self-esteem motives can satisfy these feelings by purchasing a rare, status-linked products (Shrum et al., 2013). In doing so, they make connections between positive self-identity motives and the symbolic meanings they attribute to the products (Ger & Belk, 1996). Due to view material possessions as an extension of themselves, they express their personal identity via consuming these products (Belk, 1988). Thus, understanding the personality variables and self-motives of individuals is prominent to explore the consumption related behaviors of individuals.

Although, consumption related behaviors of individuals to show their self-worth and superiority is intensely studied in the marketing literature, it is largely neglected how the trait of arrogance manifest itself as a consumer behavior. Psychologist Tim Kasser indicated that global consumer culture may evoke of

[1] **Hatay Mustafa Kemal University, Faculty of Economics and Administrative Sciences, Department of Business Administration, 31060 Hatay-Turkey, bbalikci@gmail.com**
[2] **Hatay Mustafa Kemal University, Kirikhan Vocational School, Hatay- Turkey, muzeyyenarslan82@gmail.com**

narcissistic personalities, which are generally acted with arrogance (Kasser, 2002). The underlying motives of goals that mentioned above are actually satisfying the needs of individuals (Vignoles, Golledge, Regalia, Manzi & Scabini, 2006).

Some researchers (Silverman, Johnson, McConnell & Carr, 2012) investigated arrogance as a trait and examined the situations in which people are perceived as arrogant. Also, people are perceived as arrogant when they communicate their quality and accentuate that it comes from their superiority and global self (Johnson et al., 2010). When considering arrogance as a personality trait which reflects the tendency to communicate one's qualities and self-worth to others (Lewis, 2000, 2016), it was extended to the marketing literature by addressing consumer arrogance notion. CA is conceptualized and defined by Ruvio & Shoham (2016) as people's proclivity for demonstrating their social superiority through the acquisition, utilization, or display of consumer goods (p. 3898). In this contribution, these two scholars developed a CA scale as a multi-dimensional trait. This new notion rooted from the symbolic meaning of consumption that suggesting consumers use products as symbols to create self-identity, to maintain their self-concept, to express their self, to convey personal and social achievements and to reflect their social status to others (Holman, 1981; Solomon, 1983; Belk, 1988; Hirschman & LaBarbera, 1990).

It seems that the scarcity of studies on arrogance stems from the lack of the measurement instrument of arrogance in marketing literature (Ruvio & Shoham; 2016). To fill this gap, Ruvio & Shoham (2016) developed a CA scale to measure it in the marketing context. Researchers have investigated CA in the USA and Israel which are western cultures also modern and developed countries. However, it is still uncertain that consumer arrogance has the same meaning across different cultures. The validity and reliability of the scale in non-Western societies are uncertain as well. There is a need to test the cross-cultural validity of CA construct in different cultures.

This study is designed to test reliability and validity of CA scale that have not been tested yet in Turkey and Romania, which are considered developing non-Western countries as well as collectivist. With this study, it is expected that the CA cross-culturally equivalent and valid scale. As mentioned above, CA is a new notion, investigating it in different cultures will facilitate to improve the comprehension of CA in cross cultural setting. In this manner, this study aims to establish measurement equivalence across cultures and to enrich the consumer behavior understanding by using materialism in predicting CA.

Literature Review

Arrogance is defined as a chronic belief of superiority and exaggerated self-importance that is demonstrated through excessive and presumptuous claims (Brown, 2012, p. 555) and used as to refer one of the seven deadly vices, which are strongly correlated with dark traits: machiavalism, psychopaty, narcissism (Veselka, Giammarco, Vernon, 2014). Paulhus & Williams (2002) considered dark trait an aspect of personality traits that are socially undesirable and resembles the seven deadly sins. Although the dark traits are undesirable traits in the interpersonal context, they are favorable on the personal point of view since they serve a purpose on the behalf of person instead of group (Kurt & Paulhus, 2008). Arrogance is grounded on interpersonal dynamics and sense of superiority by manifestation of overt and excessive behaviors over others (Johnson et al., 2010).

The early studies regarded arrogance as a psychoanalytic term which is a component of narcissism (Emmons, 1984). Although narcissism with the belief of superiority, shares the same meaning with arrogance they are different from each other. While arrogance emerges as open behaviors during interpersonal interactions, narcissism involves both open and confidential behavior (Johnson et al., 2010).

Although, arrogance is characterized by the words of 'hubris, contemptuous, vanity, and conceited', it differs from them. The conceptual origins rely on pride, which is a two-faced construct: authentic and hubristic. However, arrogance differs from pride at the attribution of success point. Pride arises from the attribution of a specific achievement to the efforts, prosocial behaviors whereas arrogance stems from the attribution of achievements to one's global self and abilities (Verbeke, Belschak & Bagozzi, 2004; Tracy & Robins, 2007).

Arrogance is embedded to social context and refers to exaggerate one's own superiority while underestimating others (Johnson et al., 2010). People feel better when they perceive themselves superior against to others (Locke & Nekich, 2000). Symbolically, possessions can be used as clues to show that people are superior to others (Hirschman & LaBarbera, 1990). Thus, Ruvio & Shoham (2016) suggested that arrogant consumers project their arrogant inclinations via

consumption. They constructed CA as a multi-dimensional trait and developed the original 21-items CA scale (see appendix A) consisting of four factors which are consumer superiority, consumer bragging, exhibitionism-based consumption, and image-based consumption. *Consumer superiority* refers to the individuals' sense of superiority. Individuals high on CA link their purchasing behaviors to their superior self. Arrogant consumers associate the superior quality of the products they bought with their global self, and they directly derive their superior qualities from the quality of products they purchased (Hayward & Hambrick, 1997; Verbeke et al., 2004). Furthermore, those who scored high on CA regard their purchasing behaviours as an achivement and tend to share these achievements with others vaingloriously (Verbeke et al., 2004). Any such intended behaviours are explained as *consumer bragging* dimension of CA.

Exhibitionism-based consumption refers to the social nature of CA. Arogance manifests in social interactions (Johnson et al.,2010) and arrogant people need social approval from others to enhance their self- views with feedbacks from those surrounding others (Aghababaei & Błachni, 2015; Egan, Chan, & Shorter, 2014; Johnson et al.,2010). Exhibisionistic and conspcious purchases signals their exceptional appearance to others and helps to inflate their ego (Ruvio & Shoham, 2016).

The desire to project arrogant consumers' belief of superiority over others leads them to purchase status products and brands that signal uniqueness, conspicious values (Lee, Ko & Megehee, 2015), which is called *image-based consumption* dimension of CA. Image-based and exhibitionism-based purchases are the way of representing ideal self-image, status and superior self to others.

In their study, Ruvio & Shoham (2016), used materialism as a predictor of CA and they found a positive relationship between the two. In the marketing literature one of the most common construct linking identity motives to consumption behaviors is materialism. Materialism is a construct emphasized the importance the individuals attached on possessions. On consumer perspective materialism, a construct that potently recognized with consumption, is explained as the importance that attached to the possessions. Individuals who are high on the materialism scores put these possessions on a central place in their life with the belief of providing the greatest source of happiness (Belk 1985; Richins & Dawson 1992). Also, materialism has been considered a kind of self-centered traits such as selfishness (Bauer et al. 2012). The most widely used materialism scale in consumer behavior literature is the Richins & Dawson's (1992) 18-items materialistic value scale that we used in this study (see appendix B), which involves three dimensions: (1) use of material possessions to make judgement about success of others and oneself, (2) acquisition centrality that is importance of possessions in one's life; and (3) acquisition of possession as the pursuits of happiness. Individuals scored high on materialism regard amount of possessions as an indicator of how success anyone in life, put material possessions in a prominent place in their life and believe material possessions lead them to happiness (Richins & Dawson, 1992).

The early studies regarded materialism as a western trait that is seen in developed and affluent countries since the marketing strategies are driving hedonism effectively in that countries. However with the effect of globalisation and Westernisation the consumers in developing countries or third world countries became consumer oriented (Campbell, 1987; Ger & Belk 1996). The level of materialism may vary according to the level of development of countries. For example, materialism was found higher in Turkey and Romania than in the USA and Europe (Ger & Belk, 1996).

On the dark side aspect of materialism the relationship between materialism and environmental and social concerns are examined. Materialistic values effect people negatively and leads them insensitivity to environmental and community issues (Clump, Brandel & Sharpe, 2002; Kilbourne & Pickett, 2008). Materialistic individuals generally consume huge amount of natural sources than necessary (Winter, 2004). Because people detest to think of themselves as selfish and environmentally insensitive. There may occur an incongruence between material self and ideal self that drives the feeling unhappiness and dissatisfaction (Kilbourne & Pickett, 2008). CA focuses on how consumption-related activities convey the superiority of individuals. Consumers with high levels of arrogance have a belief that they are superior to others in terms of the acquisition and use of possessions and they are linking the superior qualities of the products they purchase with their superior quality of their-selves which is related to materialism (Ruvivo & Shoham, 2016). Materialism and CA are related constructs as they reflect tendencies of individuals via consumption; therefore materialism was investigated as a predictor of CA in the current study.

Methodology

Sample

The data were gathered from two countries, Turkey and Romania. These two countries exhibit cultural similarities based on Hofstede's (1990) individualism and collectivism dimensions. The CA scale is developed and validated in Israel and the US. Although the American society has been considered as individualistic with a score of 91, in Hofstede's (2001) study, the Israeli society is considered as a blend of individualistic and collectivistic cultures with a score 54. Individualism emphasizes on the uniqueness of the person, self-reliance and competitiveness. On the other hand collectivism focuses on interpersonal connection and harmony, not bragging (Markus & Kitayama, 1991). According to Alicke & Sedikides (2009), people are self-promoting in individualistic cultures and when self-promotion suggests arrogance, it damages interpersonal relationships. In contrast, people are more modest in collectivist cultures and modesty avoids the risks of arrogance. Maheswaran & Shavitt (2000) provide useful literature that shows how culture, individualist versus collectivist, differentiated consumer behavior. Therefore, current study evaluate the construct validity of the CA within and across a collectivist cultural context in Turkey and Romania. These two countries were chosen because they are both less developed non-Western countries and collectivist with 37 (Turkey) and 30 (Romania) individualistic scores based on Hofstede's (1990) cultural dimensions. Thus, it is important to investigate the CA scale in a non-Western collectivist context.

An efficient way to make the samples comparable in two countries university students sample will be relatively homogeneous for this study. Therefore, we intend to reduce the random error that might occur by using heterogeneous sample (Calder, Lynn & Alice, 1981). The survey applied in the same way via self-report questionnaire administrated to convenience samples of university students in class in both country as voluntarily and without extra credit. The sample included a total of 368 young adults (192 Turkish and 176 Romanian). After deleting the outliers to reach the normally distributed data, the final sample revealed as 292 young adults (145 Turkish and 147 Romanian). Participants age range goes from 18-26 years old, and their average age is 20.76 *(Sd = 1.445)*. The majority of the participants were female (71%).

Measures

The questionnaire is consisted 39 items for scales and two questions for demographic information. CA was measured by using Ruvio & Shoham's (2016) scale, which is included 21 items across four dimensions. Materialism was measured by Richins & Dawson's (1992) scale, which is consisted of three dimensions a total of 18 items. The materialism scale has been tested on the Turkish sample and the Cronbach's alpha coefficient for whole scale is .80 (Karabati & Cemalcilar, 2010). Also, this scale has been tested on Romanian sample (Zait & Mihalache, 2014) and alpha coefficient was found 0.84. The Likert-type scale with a five point format was used for all items.

The five-step back-translation method of Brislin, Lonner & Thorndike (1973, p. 182) was used for translating the original English questionnaire to Turkish. For the Romanian sample, the English version of the questionnaire was administrated because the students were fluent in English. Because as indicated by Albaum, Erickson & Strandskov (1989) using the scale with its original language in cross-cultural research provides the same results as a translated one.

Data analysis procedure

To assess the construct validity and cross-cultural equivalence of the Ruvio & Shoham's (2016) four-factor CA (21 items) scale, the procedures that suggested by Hair, Black, Babin, Anderson & Tatham (2006, pp. 820-833) is followed. The analysis procedure began with performing the normality test in AMOS because of the confirmatory factor analysis (CFA) used with the maximum likelihood estimation. The univariate normality was assessed by using the proposed threshold of (Finch, West, & MacKinnon, 1997) for skewness $< \pm 2.0$ and for kurtosis $< \pm 7.0$ which reveals no significant problems in the data. Multivariate normality is determined by Mardia's multivariate kurtosis (Mardia, 1970).

Then, the internal consistency reliability of the original 21-items CA was assessed for an individual sample by using Cronbach's alpha coefficient (Nunnally, 1979). Because of the data was self-reported, we checked a common method variance (CMV) via Harmon's one factor test (Podsakoff & Organ, 1986) as loading all of the observable items on a single factor by using CFA.

The construct validity of CA assessed with convergent, discriminant, and nomological validity (Garver & Mentzer, 1999) for each samples separately as used by Ruvio & Shoham (2016). While convergent validity

confirms the relationship between the items and the factor that they belong, discriminant validity examines how the items different from other constructs. Following the procedures of convergent validity proposed by Fornell & Larcker (1981) were used in this study. These are the average variance extracted (AVE) and the composite reliability (CR) of each construct. Based on Hair et al. (2006) recommendation all CR values for the dimensions are expected to be a value of 0.70 and higher. Also, when AVE values equals or exceeds 0.50, convergent validity is assessed to be adequate. However, for discriminant validity, all AVE's should be higher than maximum squared variance (MSV) involving the constructs (Fornell & Larcker, 1981). MSV is the square of the highest correlation between the constructs. Nomological validity means that two or more theoretical and / or empirical concepts should also be correlated (Cronbach & Meehl, 1955).

In evaluating the model fit, we used five types of indicators, which are the Chi-Square (CMIN), the Root Mean Square Error of Approximation (RMSEA), the Normed Fit Index (NFI), the Tucker-Lewis Index (TLI), and the Comparative Fit Index (CFI). The CMIN value is the most basic measurement used to test the overall suitability of the model. An appropriate model is expected to give meaningless results at $p < 0.005$ (Barrett, 2007). RMSEA is a statistic that gives information about the compatibility of the hypothesized model parameters with the covariance matrix of the population (Byrne, 2011: 664). RMSEA can be within the range of 0.03 to 0.08 in the 95% confidence interval (Rigdon, 1996: 369-379). The NFI statistic compares the $\chi 2$ value of the model with the $\chi 2$ value of the zero model. The NFI value can be between 0 and 1, and a threshold value of 0.90 is considered to represent good fit (Hu & Bentler, 1999). TLI statistic were developed to eliminate the effect of sample size. In addition, the threshold value of TLI can be greater than 0.80 and high threshold values such as TLI> 0.95 are also found in the literature (Hu & Bentler, 1999; Byrne, 2011: 684). CFI also takes into account sample size and gives good results even in small samples and the threshold is generally considered greater than 0.90 (Byrne, 2011).

Since the purpose of this study is to test the cross-cultural invariance of the CA scale, the multi-group analysis of invariance used by assessing configural, metric, and factor covariance invariance as proposed by Hair et al. (2006). Configural invariance means that the construct is conceptualized with the same way across-samples without constraints imposed across samples. Metric invariance confirms that all factor loadings are equal across-samples by constraining the factor loadings across-groups. Factor covariance invariance indicates that the factor loadings and intercepts were equally constrained in both samples.

Findings

The results of the univariate normality test showed that the value of skewness ranged from -.337 to +1.138 and the value of kurtosis ranged from -.989 to +1.232 for the Turkish sample. For the Romanian sample, the value of skewness ranged from -.394 to +1.017 and the value of kurtosis ranged from -.925 to +.462. Although, these values are within the ranges of normal distribution on the multivariate kurtosis values revealed as 80.181 (c.r. = 17.873) for the Turkish sample and 53.593 (c.r. = 11.438) for the Romanian sample. According to Mardia (1974) since the c.r. is < 5, the data suggests multivariate normality distributed (Byrne, 2013). For lowering the multivariate kurtosis in this study, we deleted the outliers by using the measure of their Mahalanobis distances until the multivariate kurtosis index reaches the desired level. The second multivariate kurtosis resulted in a value of 14.916, with the 2.889 critical ratio (c.r.) for the Turkish sample and in a value of 11.364, with the 2.217 c.r. for the Romanian sample. Thus, the data supports multivariate normally distribution since the c.r. values < 5 (Byrne, 2013) for both samples. After deleting the outliers, which are fourty seven from Turkish sample and twenty nine from Romanian sample, the final sample included a total of 292 observations (145 Turkish and 147 Romanian) which is considered adequate sample size for SEM (at least one hundred observations) suggested by Hair et al. (2006).

After normality test, we first assessed the reliability of the overall CA for both samples, which are resulted in a good Cronbach's alpha coefficients ($\alpha = 0.92$ for the Turkish sample and $\alpha = 0.93$ for the Romanian sample. Then, the reliability based on four-dimensions of CA scale is addressed for the two samples. In the Turkish sample, the coefficients ranged from 0.84 (consumer superiority dimension) to 0.89 (image-based consumption dimension) and from 0.82 (consumer bragging dimension) to 0.89 (image-based consumption dimension) for the Romanian sample. Thus, overall CA and all of its dimensions have reliable Cronbach's alpha coefficients, which is considered to be good as suggested by Nunnally (1979).

Because the data were self-reported, we assessed a common method variance (CMV) via Harmon's one factor test (Podsakoff & Organ, 1986) as loading all of the observable items on a single factor using CFA. The results of Harmon's one factor test revealed a poor fit for both of countries (Turkey $\chi 2 = 998.848$, $df = 189$, $p \leq .000$; CFI = 0.56; NFI = 0.51; TLI = 0.51, RMSEA = 0.17 / Romania $\chi 2 = 755.749$, $df = 189$, $p \leq .000$; CFI = 0.66; NFI = 0.60, TLI = 0.66, RMSEA = 0.14). Thus, it can be inferred that there is no risk of CMV bias.

The CFA was carried out to determine the factor-structure of CA scale with 21 items for the Turkish sample and for the Romanian sample separately. The CFA results in both samples produced poor fit indexes. To improve fit indexes of both samples we followed the common practice by deleting five items that factor loadings lower than 0.50 (Hair et al., 2006, 133). Thus, the original scale reduced to 16 items for to maintain its validity. Table 1, reports the results of the CFA for the two samples.

Table 1: Factor Loadings and Correlations between Dimensions of CA – CFA Analysis

Factor loadings		**Turkey**	**Romania**
Image-based consumption (IBC)			
	IBC1	0.92	0.84
	IBC2	0.92	0.93
	IBC3	0.88	0.85
	IBC4	0.80	0.63
AVE		78%	67%
CR		0.93	0.89
Exhibitionism-based purchasing (EPB)			
	EPB3	0.81	0.80
	EPB4	0.89	0.81
	EPB5	0.72	0.74
	EPB6	0.60	0.85
AVE		58%	57%
CR		0.84	0.84
Consumer bragging (CB)			
	CB1	0.74	0.68
	CB2	0.94	0.96
	CB3	0.63	0.70
	CB4	0.54	0.57
AVE		56%	54%
CR		0.83	0.82
Consumer superiority (CS)			
	CS1	0.68	0.69
	CS2	0.86	0.87
	CS3	0.83	0.83
	CS4	0.66	0.74
AVE		58%	61%
CR		0.84	0.86
Correlations (variances)			
IBC	EPB	0.44 (0.19)	0.57 (0.32)
IBC	CB	0.38 (0.14)	0.51 (0.26)
IBC	CS	0.48 (0.23)	0.51 (0.26)
EPB	CB	0.56 (0.31)	0.72 (0.51)
EPB	CS	0.50 (0.25)	0.57 (0.32)
CB	CS	0.54 (0.29)	0.48 (0.23)

All loadings and correlations are significant at $p < .001$ level.

The modified measurement model of CA scale has a correlated four-factor structure with four items loading on the image-based consumption factor (IBC), four items loading on the consumer bragging factor (CB), four items loading on the exhibitionism-based purchasing (EBC), and four items loading on the consumer superiority factor (CS) (see Appendix A for retained and deleted items).

Although the re-specified four factor model has a significant chi-square (Turkey: $\chi 2(CMIN) = 202.589$, $df = 97$, $p \leq .000$, $\chi^2(CMIN)/df = 2.089$ / Romania: $\chi 2(CMIN) = 181.098$, $df = 97$, $p \leq .000$, $\chi^2(CMIN)/df = 1.867$) other fit indices demonstrate acceptable statistics based on Byrne (2011) recommendations (Turkey: CFI = 0.93; NFI = 0.89; TLI = 0.91, RMSEA = 0.08 / Romania: CFI = 0.94; NFI = 0.89, TLI = 0.93, RMSEA = 0.07).

All factor loadings are greater than 0.50, with dimensions' AVE values ranging from 54% to 78%, and CR values for the four factors above 0.80 are providing scale convergent validity. Additionally, discriminant validity was established based on Fornell & Larcker's (1981) procedure. All of the AVEs are greater than the shared variance between each pair of factors in both of the countries. Thus, this test doesn't suggest problems with discriminant validity. In addition, fit statistics of the measurement model of four-factor structure CA for both countries is good (Table 2). We then further examined whether the results from the sample of Turkish students can be cross-validated using the sample of Romania students. Table 2 summarizes the results of the measurement equivalence a cross-samples. Each progressive test supports the cross-validation.

The configural invariance test the same baseline measurement model on both countries simultaneously. The results demonstrated good fit statistics ($\chi 2 = 383.690$, $df = 194$; NFI = 0.89; TLI=.92; CFI = 0.94; RMSEA = 0.05) and supports the four factor measurement model in both countries. Constraining the loading estimates has not worsened fit of the unconstrained model and the change in chi-square in not significant ($\Delta\chi 2\ (12) = 10.72$; $p = .553$). A non-significant Chi-square supports that the factor loadings are equal across countries. Similarly, the test of covariance invariance provides a good fit ($\chi 2 = 402.910$, $df = 213$; NFI = 0.89; TLI=.93; CFI = 0.93; RMSEA = 0.05). Thus, statistically the model is not different from the baseline model ($\Delta\chi 2\ (17) = 8.5$; $p = .291$), and it means that the factor correlations are equal across countries.

Table 2: Cross-Cultural Validation Statistics for the CA Scale

Model	Chi *Square*	*df*	*p*	RMSEA	CFI	NFI	TLI	Δχ2	Δ*df*	*p*
Individual Groups:										
Turkey	202.59	97	.000	0.08	0.93	0.89	0.91			
Romania	181.10	97	.000	0.07	0.94	0.89	0.93			
TF-Factor Structure Equivalence	383.69	194	.000	0.05	0.94	0.89	0.92			
Factor Loading Equivalence	394.41	206	.000	0.05	0.93	0.89	0.92	10.72	12	0.553
Factor Covariance Equivalence	402.91	213	.000	0.05	0.93	0.89	0.93	8.5	7	0.291

To assess the nomological validity, materialism was used as an antecedent of CA, which is supported by Ruvio & Shoham's (2016) study that found a positive relationship. Before the nomological validity, we carried out CFA to confirm the three-factor structure of materialism for both countries. The CFA results in both samples indicate poor fit for the three dimensional materialism (Turkey: $\chi^2(CMIN)=327.019$, $df=132$, $\chi^2(CMIN)/df=2.477$; $p=.000$; CFI=.826; NFI=.743; TLI = .798; RMSEA=.101 / Romania: $\chi^2(CMIN)= 262.397$, $df=132$, $\chi^2(CMIN)/df=1.988$; $p=.000$; CFI=.886; NFI=.797; TLI = .868; RMSEA=.085). To improve the fits of the model we removed two indicators from success factor, four indicators from centrality, and one indicator from happiness that factor loadings lower than 0.50 as suggested by Hair et al. (2006,133) (see Appendix B for retained and deleted items). The modified measurement model of materialism demonstrates acceptable fit statistics (Turkey: $\chi^2(CMIN)=111.527$, $df=41$, $\chi^2(CMIN)/df=2.720$; $p=.000$; CFI=.90; NFI=.886; TLI =.904;

RMSEA=.080 / Romania: χ^2(CMIN)= 82.089, *df*=41, χ^2(CMIN)/*df*=2.002; *p*=.000; CFI=.942; NFI=.893; TLI = .923; RMSEA=.083). Next, CR and AVEs were computed for each dimension. The CR values of success were 0.82 and 0.84, for the Turkish and Romanian samples, respectively. The CR values of the happiness dimension were 0.84 for both of samples. However, the CR values of centrality factor with 0.70 and 0.65 values for the Turkish and Romanian samples, respectively, were failed to fit the recommended level of 0.80. In addition, the AVEs reached the required 0.50 values for success and happiness dimensions in both samples, except for the centrality dimension with 0.44 in the Turkish sample and 0.41 in the Romanian sample. Therefore, we omitted the centrality factor from the nomological validity analysis.

Before to test the nomological validity, CA subjected to a second order CFA to estimate of the primary construct on its sub-dimensions. The second order fit statistics of CA are good for both countries (unconstrained model: $\chi2 = 418.565$; d$f = 200$, p= 0.001; NFI=.865, TLI = 0.908; CFI = 0.924, RMSEA = 0.06 /the constrained model: $\chi2 = 433.378$; d$f = 215$, p = 0.001; NFI=0.865; TLI =0 .915; CFI = 0.924; RMSEA = 0.05) and the chi-square difference between models is not significant ($\Delta\chi2$ (15)=14.813; p =.465).

Then, we assessed the multi-group nomological validity with SEM (structural equation model) to understand the relationships between CA and materialism dimensions which are success and happiness. The results showed a good fit with the data (the baseline model: $\chi2 = 806.819$; d$f = 536$, p= 0.000; NFI=.830, TLI = 0.926; CFI = 0.934; RMSEA = 0.042). Then, the structural weights were constrained to be equal across the two samples. The constrained model was still acceptable and statistically not significant ($\chi2 = 827.142$; d$f = 560$, p= 0.000; NFI= 0.824; TLI = 0.930; CFI = 0.935; RMSEA = 0.041). By this constraint a change in Chi-square is non-significant, then the added constraints has not worsened the fit ($\Delta\chi2$ (24) =20.323; p =0.678). Therefore, the nomological validity of CA supported in both countries. The results of the SEM analysis partially confirmed the expected relationships. As expected, the results indicated that the coefficient of path from success dimension of materialism to CA for Turkey ($\beta = 0.55$, $p< 0.001$) and for Romania ($\beta = 0.85$, $p< 0.001$) were found statistically significant and positive. In addition, while the coefficient of path from happiness dimension of materialism to CA was found statistically significant and positive for Turkey ($\beta = 0.36$, $p< 0.001$), no significant relationships were found for Romania ($\beta = 0.12$, p=0.226).

DISCUSSION

This study investigates the configural, metric, and factor covariance invariance of the 21-items CA across two countries with a collectivist cultural context, Turkey and Romania, using multi-group analysis. In line with that purpose, the CFA analysis in each sample resulted with 16 items after deleting five items that factor loadings lower than 0.50 (see Appendix A). These items are IBC5 (The image of a product affects my purchase of it), IBC6 (I often buy products that emphasize my social status), EBP1 (I tend to buy products that attract attention), EBP2 (I tend to buy products that make me look meticulous), and CB5 (I tend to choose showy products). The reduced 16-item four-factor CA revealed evidence for full invariance across samples. Because, the three invariance tests indicate that the factor loading pattern, factor structure, and factor covariance revealed to be equivalent across the cultures. Thus, these findings suggest that the 16-item CA may use to make comparisons between samples from similar cultures. In addition, the results showed that CA might be used as a second order construct. While this result supports the findings of Ruvio & Shoham's (2016), it doesn't support Ruvio & Shoham's (2007) suggestion that CA is less desirable in a collectivist culture. This finding can be explained by Newman & Newman's (2014) argument that the globalization has moved cultures to individual values from collective values. Although Ruvio and Shoham (2016) confirmed the CA scale in Israel and the US, our study also contributes to the etic approach (Triandis & Marin, 1983) by supporting the cross-cultural measurement invariance of the CA in two countries with a collectivist cultural context.

The nomological validity findings show that materialism partially predicts CA. Our results show that the success dimension of materialism predicts CA in both samples and happiness dimension only predicts CA in the Turkish samples. Furthermore, the centrality dimension of materialism dropped from the analysis because of its low reliability. In this manner, our research make a significant contributions by displaying that materialism tendencies may understandable with the same way in collectivistic cultural context. This interesting findings may be interpreted in line of the problematic psychometric properties of Richnis & Dawson's (1992) materialism scale that found in some

studies (Griffin, Babin, & Christensen, 2004; Wong, Rindfleisch, & Burroughs, 2003; Watchravesringkan, 2012). As indicated by Webster & Beatty (1997) and supported by Watchravesringkan (2012), the mixed-worded scales particularly the reverse-worded items reduced the scale dimensionality in different cultural contexts. In his study Watchravesringkan (2012) found that the Cronbach's coefficient as well as the CR values of the centrality dimension of materialism below the acceptable ranges as recommended by Nunally (1979). Finally, Watchravesringkan (2012) showed that the reduced 10-items materialism scale, without reverse worded items (eight items), invariant across the US and the Thai samples. In addition, although our nomological findings do not fully supported that sub-dimensions of materialism as predictors of CA, it should be noted that Ruvio & Shoham (2016) had been used overall materialism score in their study. When materialism considered as one the important aspects of consumer behavior that influenced by culture (Ger & Belk, 1996), also it should be considered that the differences will be occured related to its dimensions. Thus, future research needs to replicate Richins & Dawson's materialism scale with other short versions in non-students samples.

In particular, future research ought to investigate the influence of culture at the individual level by classifying the respondents as independence and interdependence based on Singelis' (1994) self-concept scale. Nevertheless, while arrogance is conceptually distinct from self-esteem (Rosenberg, 1979), it is related with it. High self-esteem is associated with negative qualities such as arrogance and conceit (Gecas, 2009). Heatherton & Vohs (2000) found that when people with extremely high self-esteem they responded by acting more arrogantly than people with lower self-esteem. On the contrary with that Johnson et al. (2010) showed that arrogant behaviors may lead to low self-esteem. Although there is some inconsistency in results between self-esteem and arrogance in the literature, future research may also explore the relationship between self-esteem and CA.

Due to time and cost constraints the most important limitation of this study is that the CA scale is not translated into Romanian. The future research should established transnational equivalence through traditional translation-back translation procedures. Also, in this research student samples used, which is increased internal validity. However future research would use samples from non-student populations.

Acknowledgements

The authors would like to thank to Dr. Ayalla A. Ruvio and Dr. Aviv Shoham for their permission to use consumer arrogance scale, and to Prof. Dr. Mihaela Violeta GHICA for helping to collect the data in Romania for this study. Also, the authors warrant that this research has not received any grants from any funding agency.

REFERENCES

Ahuvia, A.C. & Wong, N.Y. (2002). Personality and values based materialism: their relationship and origins. *Journal of Consumer Psychology*, 12(4), 389-402.

Aghababaei, N., & Błachnio, A. (2015). Well-being and the Dark Triad. *Personality and individual differences*, *86*, 365-368.

Albaum, G., Erickson, R. & Strandskov, J. (1989). Questionnaire design in international and cross-cultural research: Is translation necessary? In Gardner M.P. (ed.), *proceedings of the society for consumer psychology*, American Psychological Association, Washington DC,10-15.

Alicke, M & Sedikides, C, (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, 20, 1-48.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual differences,* 42(5), 815-824.

Bauer, M. A., Wilkie, J. E. B., Kim, J. K., & Bodenhausen, G. B. (2012). Cuing consumerism: Situational materialism undermines personal and social well-being. *Psychological Science*, 23, 517–523. doi: 10.1177/0956797611429579

Belk, R.W. (1988). Possessions and the extended self, *Journal of Consumer Research,* 15(2), 139-168.

Brislin R. W., Lonner, W. J. & Thorndike, R. M. (1973). *Cross-cultural research methods,* John Wiley and Sons*,* New York, USA.

Burroughs, J. E., & Rindfleisch, A. (2002). Materialism and Well-Being: A Conflicting Values Perspective. *Journal of Consumer Research,* 29(3), 348-370.

Byrne, B. M. (2011). *Structural equation modeling with AMOS Basic concepts, applications, and programming (Multivariate Applications Series)*, Routledge, New York.

Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge.

Calder, B.J., Lynn, W.P., & Alice, M. (1981). Designing research for application. *Journal of Consumer Research*, 8, 197-207.

Campbell, Colin (1987). The Romantic Ethic and the Spirit of Modern Consumerism, New York: Basil Blackwell.

Clump, M. A., Brandel, J. M., & Sharpe, P. J. (2002). Differences in environmental responsibility between materialistic groups. Psychologia: An International Journal of Psychology in the Orient, 45(3), 155-161. http://dx.doi.org/10.2117/psysoc.2002.155.

Egan, V., Chan, S., & Shorter, G. W. (2014). The Dark Triad, happiness and subjective well-being. *Personality and Individual Differences*, *67*, 17-22.

Emmons, R. A. (1984). Factor analysis and construct validity of the narcissistic personality inventory. *Journal of Personality Assessment*, 48(3), 291-300.

Finch, J. F., West, S. G., & MacKinnon, D. P. (1997). Effects of sample size and nonnormality on the estimation of mediated effects in latent variable models. *Structural Equation Modeling: A Multidisciplinary Journal*, *4*(2), 87-107.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics.

Garver, M. S., & Mentzer, J. T. (1999). Logistics research methods: employing structural equation modeling to test for construct validity. *Journal of business logistics*, *20*(1), 33.

Gecas, V. (2009). The Self as a Social Force, available at: https://books.google.com.tr/books?hl=tr&lr=&id=2Nazn5H_cNYC&oi=fnd&pg=PA85&dq=Gecas,+Viktor+(2009).+The+Self+as+a+Social+-Force.&ots=u0SaFh2Ec4&sig=PPP1Bdpd5V2K6MfPfXSYh6pdBuA&redir_esc=y#v=onepage&q&f=-false (accessed 4 April 2018).

Ger, G. & Belk, R.W. (1996). Cross-cultural differences in materialism. *Journal of Economic Psychology*, 17, 55-77.

Griffin, M., Babin, B. J. & Christensen, F. (2004). A cross-cultural investigation of the materialism construct: Assessing the Richins and Dawon's materialism scale in Denmark, France, and Russia . *Journal of Business Research,* 57(8), 893 – 900 .

Hair, J. F., Jr., Black, W. C., Babin, B. J., Anderson, R. E. & Tatham, R. L. (2006). *Multivariate data analysis (6th Ed.)*, Pearson Prentice Hall, New Jersey.

Hayward, M. L., & Hambrick, D. C. (1997). Explaining the premiums paid for large acquisitions: Evidence of CEO hubris. *Administrative science quarterly*, 103-127.

Heatherton, T. F. & Vohs, K. D. (2000). Interpersonal evaluations following threats to self: Role of self-esteem. *Journal of Personality and Social Psychology*, 78, 725–736.

Hirschman, E. C. & LaBarbera, P. A. (1990). Dimensions of possession importance. *Psychology & Marketing*, 7(3), 215-233.

Hofstede, G. (1990). *Cultures and organizations: Software of the mind*. New York: McGraw-Hill.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Thousand Oaks, CA: Sage

Holman, R. H. (1981). Apparel as communication, avaliable at: http://acrwebsite.org/volumes/12220/volumes/sv04/SV-04 (accessed 4 April 2018).

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Johnson, R. E., Silverman, S. B., Shyamsunder, A., Swee, H. Y., Rodopman, O. B., Cho, E. & Bauer, J. (2010). Acting superior but actually inferior? Correlates and consequences of workplace arrogance. *Human Performance*, 23(5), 403-427.

Karabati, S. & Cemalcilar, Z. (2010). Values, materialism, and well-being: A Study with Turkish university students. *Journal of Economic Psychology*, 31(4), 624–633.

Kasser, T (2002). *The High Price of Materialism*. MIT Press, London.

Kastanakis, M. N. & Balabanis, G. (2012). Between the mass and the class: Antecedents of the bandwagon luxury consumption behavior. *Journal of Business Research*, 65(10), 1399-1407.

Kilbourne, W. & Pickett, G. (2008). How materialism affects environmental beliefs, concern, and environmentally responsible behavior. *Journal of Business Research*, 61, 885-893.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Kurt, A., & Paulhus, D. L. (2008). Moderators of the adaptiveness of self-enhancement: Operationalization, motivational domain, adjustment facet, and evaluator. Journal of Research in Personality, 42(4), 839-853.

Lee, J., Ko, E., & Megehee, C. M. (2015). Social benefits of brand logos in presentation of self in cross and same gender influence contexts. *Journal of Business Research*, *68*(6), 1341-1349.

Levy, S.J. (1959). Symbols for sale, *Harvard Business Review,* 37(4), 117-124.

Lewis, M. (2016). Self-conscious emotions: Embarrassment, pride, shame, guilt, and hubris in Lewis, B. & Jones H. (Ed.). *Handbook of emotions*, Guilford Press, New York, 792-814.

Lewis, M. (2000). Self-conscious emotions: Embarrassment, pride, shame, and guilt. In Lewis, M.& Haviland-Jones, J. M. (Ed.). *Handbook of emotions 2nd ed*., Guilford Press, New York, 623–636.

Locke, K. D. & Nekich, J.C. (2000). Agency and communion in naturalistic social comparison. *Personality and Social Psychology Bulletin*, 26(7), 864–874.

Maheswaran, D. & Shavitt, S. (2000). Issues and New Directions in Cultural Psychology. *Journal of Consumer Psychology,* 9(2), 59-66.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika, 57(3), 519-530.

Markus, H.R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224–253.

Newman, B. M. & Newman, P. R. (2014). Development through Life: A psychosocial approach, available at https://books.google.com.tr/books?id=_Vk8AwAAQBAJ&pg=PA61&hl=tr&source=gbs_toc_r&cad=4#v=onepage&q&f=false (accessed 05 April 2018).

Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric theory (3rd Ed.)*. McGraw-Hill, New York, NY.

Paulhus, D. L. & Williams, K. M. (2002). The dark triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of Research in Personality*, 36(6), 556 -563.

Podsakoff, P. M. & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12(4), 531–544.

Richins, M. & F.W. Rudmin, (1994). Materialism and economic psychology. *Journal of Economic Psychology,* 15, 217-232.

Richins, M. L. & Dawson S. (1992). Consumer values orientation for materialism and its measurement: Scale development and validation. *Journal of Consumer Research*, 19(3), 303-316.

Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(4), 369-379.

Ruvio, A., Shoham, A. & Shlomo, H. (2007). Consumers' arrogance: Construct conceptualization and preliminary validation evidence. Available at http://www.acrwebsite.org/volumes/v34 (accessed 4 April 2018).

Ruvio, A. A. & Shoham, A. (2016). Consumer arrogance: Scale development and validation. *Journal of Business Research*, 69(10), 3989-3997.

Shrum, L. J., Wong, N., Arif, F., Chugani, S. K., Gunz, A., Lowrey, T. M…Sundie, J. (2013). Reconceptualizing materialism as identity goal pursuits: functions, processes, and consequences. *Journal of Business Research*, 66(8), 1179-1185.

Silverman, S. B., Johnson, R. E., McConnell, N., & Carr, A. (2012). Arrogance: A formula for leadership failure. *The Industrial-Organizational Psychologist*, 50(1), 21-28.

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals, *Personality and Social Psychology Bulletin*, 20(5), 580-591.

Sirgy, M. Joseph (1998). Materialism and Quality of Life. *Social Indicators Research*, 43, 227-260.

Solomon, M. (1983). The role of products as social stimuli: a symbolic interactionism perspective. *Journal of Consumer Research,* 10, 319 –329.

Steenkamp, J.M. & Baumgartner, H. (1998). Assessing measurement invariance in crossnational consumer research. *Journal of Consumer Research*, 25, 78-90.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics (5th Ed.).* Pearson Allyn & Bacon, Upper Saddle River, NJ.

Tracy, J. L. & Robins, R. W. (2007). The psychological structure of pride: A tale of two facets. *Journal of Personality and Social Psychology*, 92(3), 506.

Triandis, H.C., & Marin, G. (1983). Etic plus emic versus pseudoetic: A test of basic assumption of contemporary cross-cultural psychology, *Journal of Cross Cultural Psychology*, 14(4), 489–500.

Verbeke, W., Belschak, F. & Bagozzi, R. P. (2004). The adaptive consequences of pride in personal selling. *Journal of the Academy of Marketing Science*, 32 (4), 386-402.

Veselka L, Giammarco, E. A., Vernon, P. A. (2014). The Dark Triad and the seven deadly sins. *Personality and Individual Differences*, 67, 75-80.

Vignoles, V. L., Golledge, J., Regalia, C., Manzi, C. & Scabini, E. (2006). Beyond self-esteem: Influence of multiple motives on identity construction. *Journal of Personality and Social psychology*, 90(2), 308–333.

Webster, C. & Beatty, R. C. (1997). *Nationality, materialism, and possession importance. Advances in Consumer Research*, 24 (1), 204 – 210.

Watchravesringkan, K. (2012). Cross-cultural equivalence of materialistic values scale (MVS): An exploratory study between the United States and Thailand. *Journal of Targeting, Measurement and Analysis for Marketing*, 20, 235 – 253.

Wong, N., Rindfleisch, A. & Burroughs, J. (2003). Do reverse-worded items confound measures in cross-cultural consumer research? The case of material values scale. *Journal of Consumer Research,* 30 (1), 72 – 91.

Zait, A. & Mihalache, C. M. (2014). How do we value what we see on tv? Television's consumption and the cultivation of materialism in Romania. *International Business and Economic Review,* 5, 24-43.

Appendix

A. The retained and deleted items from original 21-item CA scale (Ruvivo & Shoham, 2016) are shown in Table 3.

Table 3: A list of retained and deleted items from CA and materialism scales

CA scale
Image-Based Consumption
IBC 1. I prefer to buy only name brands.[a]
IBC 2. I look mostly for name brands when I shop.[a]
IBC 3. I tend to buy only in prestigious stores.[a]
IBC 4. I try to buy only expensive products.[a]
IBC 5. The image of a product affects my purchase of it.[b]
IBC 6. I often buy products that emphasize my social status.[b]
Exhibitionism-based purchasing
EBP 1. I tend to buy products that attract attention.[b]
EBP 2. I tend to buy products that make me look meticulous.[b]
EBP 3. I make sure to wear clothes that lead others to compliment me.[a]
EBP 4. I prefer to buy products that make others think that I am fashionable.[a]
EBP 5. I love it when people show interest in what I buy.[a]
EBP 6. It is important to me that others realize that I have the best things.[a]
Consumers' bragging
CB 1. I like to show others what I buy.[a]
CB 2. I frequently make sure that others know what I buy.[a]
CB 3. I always tell others how my purchases are the best.[a]
CB 4. I like to compare the things I have with others.[a]
CB 5. I tend to choose showy products.[b]
Consumers' superiority
CS 1. Compared to others, I usually know what the best buy is.[a]
CS 2. Not many people know the best buy as well as I do.[a]
CS 3. I tend to buy better products than most people I know.[a]
CS 4. I usually know where to get the best deals better than others.[a]

[a] Indicates retained items, [b] deleted items, and * reverse items.

B.The retained and deleted items from original 18-item Materialism scale (Rihcins & Dawson, 1992) are shown in Table 4.

Table 4: A list of retained and deleted items from materialism scale

Materialism scale
Success
MS 1. I admire people who own expensive homes, cars, and clothes.[a]
MS 2. Some of the most important achievements in life include acquiring material possessions.[a]
MS 3.I don't place much emphasis on the amount of material objects people own as a sign of success[b*]
MS 4. The things I own say a lot about how well I'm doing in life.[b]
MS 5. I like to own things that impress people.[b*]
MS 6. I don't pay much attention to the material objects other people own.[a]
Centrality
MC 1. I usually buy only the things I need.[b*]
MC 2. I try to keep my life simple, as far as possessions are concerned.[a*]
MC 3. The things I own aren't all that important to me.[b]
MC 4. I enjoy spending money on things that aren't practical.[a]
MC 5. Buying things gives me a lot of pleasure.[b*]
MC 6. I like a lot of luxury in my life.[a]
MC 7. I put less emphasis on material things than most people I know.[b*]
Happiness
MH 1. I have all the things I really need to enjoy life.[a*]
MH 2. My life would be better if I owned certain things I don't have.[a*]
MH 3. I wouldn't be any happier if I owned nicer things.[a]
MH 4. I'd be happier if I could afford to buy more things.[a]
MH 5. It sometimes bothers me quite a bit that I can't afford to buy all the things I'd like.[b]

[a] Indicates retained items, [b] deleted items, and [*] reverse items.

EGE AKADEMİK BAKIŞ / *EGE ACADEMIC REVIEW*

Cilt 19 • Sayı 4 • Ekim 2019
SS. 425 / 435
Doi: 10.21121/eab.527219
Başvuru Tarihi: 14/2/2019 • Kabul Tarihi: 12/9/2019

Customer Service Failure Evaluations in Diverse Airlines Business Models

Havayolu İş Modellerinde Müşterilerin Hizmet Hatası Değerlendirmeleri

Birce DOBRUCALI[1]
Bengü OFLAÇ[2]

ABSTRACT

Even though it is impossible to eliminate all failures in a service encounter, due to their negative outcomes, airlines strive at minimizing failures as much as possible. Customers may act differently after failures, but the important point for airlines is to understand the underlying psychological mechanisms for prevention. In this regard, this study reveals the dynamics by which expectancies of passengers affect stability attribution, word of mouth (WOM) and repurchase intentions after exposing service failures in diverse airlines business models. Structural equation modeling is conducted separately for two airlines business contexts: low cost carriers (LCCs) and flagship airlines, and differences in service failure evaluations are presented. Findings demonstrate that for LCC, passengers' expectations positively affect stability attributions. Additionally, stability attribution following flagship airlines service failure is found to have a negative effect on repurchase intentions. Finally, findings indicate that following both LCC and flagship service failures, stability attribution decreases WOM intention of passengers.

Keywords: Low cost carriers (LCC), Flagship airlines, Service failure, Attribution theory, Structural Equation Modeling (SEM)

JEL Classification: M31, L39

ÖZET

Hizmet sunumu sürecindeki tüm hataların ortadan kaldırılması imkânsız olsa da, havayolları, hizmet hatalarını, doğurdukları olumsuz sonuçlarından dolayı en aza indirgemeye çalışmaktadır. Müşteriler yaşadıkları hizmet hatalarına farklı tepkiler verseler de, havayolları için önemli olan nokta, müşterilerin verdiği bu tepkilerin altında yatan psikolojik mekanizmaları anlamaktır. Bu bağlamda, bu çalışma, farklı iş modellerinde hizmet veren havayolu şirketlerinin hizmet hatalarını takiben yolcuların beklentileri ile değişmezlik atfı, ağızdan ağıza pazarlama ve yeniden satın alma niyetleri arasındaki dinamikleri ortaya koymaktadır. Müşterilerin hizmet hatası değerlendirmeleri Yapısal Eşitlik Modellemesi ile düşük maliyetli taşıyıcılar (LCC'ler) ve bayrak taşıyıcı havayolları için ayrı ayrı incelenmiş ve değerlendirmeler arasındaki farklılıklar ortaya konulmuştur. Bulgular, düşük maliyetli havayolları için, yolcuların beklentilerinin değişmezlik atıflarını olumlu etkilediğini göstermektedir. Buna ek olarak, bayrak taşıyıcı havayollarının hizmet hatasını takip eden değişmezlik atfının, yeniden satın alma niyetleri üzerinde olumsuz bir etkisi olduğu bulunmuştur. Son olarak, bulgular hem düşük maliyetli hem de bayrak taşıyıcı havayollarının hizmet hatalarını takiben, değişmezlik atfının ağızdan ağıza pazarlama niyetini azalttığını göstermektedir.

Introduction

Airlines service providers strive at minimizing failures as much as possible due to its critical results such as decreased repurchase intentions (Folkes et al., 1987), and reduced profit ratios (Reichheld, 1996; Torres and Kline, 2006). Even though service failures

[1]**Research Assistant, Izmir University of Economics, Faculty of Business, Department of International Trade and Finance, birce.dobrucali@ieu.edu.tr**
[2]**Assoc. Prof. Dr., Izmir University of Economics, Faculty of Business, Department of Logistics Management, bengu.sevil@ieu.edu.tr**

cannot be eliminated, for taking either corrective or preventive actions, it is crucial for service providers to understand passengers and the altering causes that shape behavioral intentions. While digging into that issue, airlines should also be aware of the fact that customers using a particular type of service provider evaluate an event related to that type of service provider (Smith and Bolton, 1998).

With the global spread of deregulation and globalization of aviation market, the civil aviation market became more commercialized and number of airline service providers has started to increase resulting in the improvement of business models of low-cost carriers (LCC) and full network service airlines (FNSA) (Ozenen, 2003; Vidovic et. al., 2013). Core characteristics of FNSA include complex pricing structure, usage of primary airports, various class of seating (economic, business and first class), intensive usage of aircraft, free food and beverage delivery during flight, longer turnaround times and reliable customer service (O'Connell, 2005; Williams, 2001). In most cases, the flagship airlines, in other words national carriers (e.g. Turkish Airlines, British Airlines, Air France/KLM, and Austrian Airlines) carry the characteristics of FSNAs (Vidovic et al., 2013). Flagship airlines are defined as; airlines registered in a state, have privileges, are or were owned by the government, and hold their privileges even long after its privatization (Sull, 1999).

Low cost carriers (LCCs), budget or discount airlines, focus on cost leadership strategy and offer low-priced flight option to passengers by delivering no frills service, reducing expenses and using secondary airports with cheaper landing charges (Hunter, 2006). Main characteristics of LCCs can be listed as; shorter turnaround times, bundling food and beverage delivery during flight, single seating class option, high aircraft utilization, one-type aircraft fleets, minimum cabin crew with lower wage scales, point to point service and no connections (Mason, 2000; Doganis, 2001; Williams, 2001; Francis et al, 2004; O'Connell, 2005; Taumoepeau et al., 2017). Ryanair, EasyJet, and Pegasus Airlines, Debonair can be counted as successful examples of LCC.

Flying with a specific kind of airlines can alter and reshape behavioral intentions. More specifically, there is a possibility for passengers to act differently with the impact of their airlines' format based preset expectations. Parallel to these arguments, we aim to shed light on following questions: Do expectations impact stability attributions after airlines service failures? Does this relationship changes according to the type of airlines? Do stability attributions affect behavioral intentions (WOM and repurchase intentions) and do those relations change based on the airlines business models?

Even though airline service failures are widely studied in the literature, there exists a lack of understanding the role of expectations on their attribution behavior and post-failure intentions in relation with the airlines business models. By developing and examining the proposed research model (Figure 1) for both low cost and flagship airlines, this research is expected to provide more precise managerial implications for airlines industry.

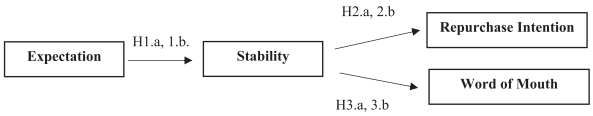


Figure 1. *Research Model*

Theoretical Framework

Expectancy Theories in Service Failures

Hess et al. (2003) explained service failure as the case where the provided service lacks the capability of meeting the expectation of customers. In airlines context, various situations ranging from flight cancellation due to weather conditions to check-in officer's misbehavior towards passenger may be interpreted as a service failure. With regard to the aforementioned definition, service failure perception and expectations are interrelated notions. Expectation is a belief or an estimation regarding the input's result in a certain level of performance (Teas, 1981) and also be defined as individualistic standards against which quality of received service is judged (McDougall and Levesque, 1998). Customers' expectations of a service is either determined before the first transaction with the company (through WOM or advertisements), by personal experience (Davis and Heineke, 1998) or service provider's image (Grönroos, 1984), and promises (Boulding et al., 1993). In airlines context, expectations have been taken into consideration in relation with LCC and FSNA business models. Wittman (2014) stated that LCC passengers hold lower expectations regarding quality of received service because they have paid less for tickets. In a similar vein, Bhadra (2009) has put forth the fact that expectations regarding level of service and ticket fares have a positive relationship, and FSNA passengers, who pay more for their tickets, have higher expectations. In accordance with the positive

relationship between ticket fares and level of expectations (Bhadra, 2009; Wittman, 2014), flagship airlines passengers are expected to have higher expectations for service experience.

Expectancy disconfirmation model states that the satisfaction of customers is dependent on their expectations, actualized performance, and disconfirmation of expectations and performance (Smith and Bolton, 2002). According to Bitner (1990), when initial service performance surpasses expectation, positive disconfirmation occurs and when performance falls behind expectation, negative disconfirmation occurs. When considered from expectancy disconfirmation framework, a failure may also be defined as a negative difference between expectation and actualized performance. Concordantly, negative disconfirmation caused by a failure experience may result in further negative evaluations (Rego et al., 2009).

Stability Attribution

Attribution theory focuses on individual's causal explanations in regard to their own behavior and others' actions, particularize underlying factors that stimulate them to examine causally relevant information, examines the way of processing information to relate the causes to the occurrences and the cognitive and behavioral consequences (Monson and Snyder, 1977; Weiner, 2000).

Weiner (1980) defines *stability* dimension as the evaluation of causes either as temporary or permanent over time. In other words, stability can be explained as the extent to which a cause is perceived as variable or enduring over time (Hess et al., 2003). Failures with enduring causes occur more frequently, and influence consumers' expectations regarding future performance of firms (Folkes 1984; Weiner, 2000). When compared with customers who attribute cause of a failure to unstable causes, customers who perceive the cause of a failure to be stable are more confident that the same failure will reoccur (Weiner, 1986). Concordantly, Folkes (1984) indicated that these inferences regarding the stability of a failure's cause impact customers' repurchase behaviors. For instance, if a flight delay is due to a stable reason, passengers are less willing to prefer the same airlines again, whereas if it is due to an unstable reason, passengers are more willing to repurchase service from the same airlines (Folkes et al., 1987).

Customers' level of expectation acts as a baseline for formation of judgments (Oliver, 1981) and affect the way they perceive the reality (Boulding et al., 1993). Starting from this point, it can be assumed that consumers' attributes regarding cause of a negative event, such as a service failure is being formed around their expectations. Consequently, it can be stated that expectations in relation with airlines' business model may have different type of relationship with stability dimension regarding service failure.

Oliver (1981) suggested that customers' level of expectation acts as a baseline for formation of judgements. Starting from this point, it can be assumed that consumers' attributes regarding the cause of a service failure is being formed around their expectations. When an airlines company falls behind expectation of passengers, and thus expectancy disconfirmation occurs, passengers are expected to show tendency to attribute failure to causes related with that airlines. Therefore, it is expected that following a service failure, expectations will have a positive influence on stability attributions.

H1.a: In flagship airlines service failures, expectation has a positive influence on stability attributions.

H1.b: In LCC service failures, expectation has a positive influence on stability attributions.

Repurchase Intentions

The post-failure process begins when customers evaluate the consumed service, and ends with completion of both behavioral and/or non-behavioral reactions to the failure (Day, 1980). Repurchase intention refers to customers' reaction to a service experience in terms of willingness of patronizing the firm in the future (Oliver and Swan, 1989).

According to Torres and Kline (2006), building long lasting relationships with customers by retaining them is more profitable than acquiring new customers. Furthermore, willingness of customers to keep on doing business with the same firm holds both economic and non-economic advantages for companies (Mostert et al., 2009). Economic advantages include ability of forecasting future sales, reduction in marketing and acquisition costs of customers, and willingness of customers to pay premium prices, while non-economic advantages include improvement of product and service offerings (Rosenberg and Czepiel, 1983; Payne, 1993; Reichheld, 1996; Ahmad and Buttle, 2001; Kassim and Souiden, 2007). For given reasons, repurchase intention holds important place in transportation sector including air travel service providers.

Perceived stability of a failure may influence consumers' intention to repurchase. When compared with customers who attribute cause of a failure to unstable causes, customers who perceive the cause of a failure to be stable are more confident that the same failure will reoccur (Weiner, 1986). Concordantly, Folkes (1984) indicated that these inferences regarding the stability of a failure impact customers' repurchase behaviors. For instance, if a flight delay is due to a stable reason, passengers are less willing to prefer the same airlines again, whereas if it is due to an unstable reason, passengers are more willing to repurchase service from the same airlines (Folkes et al., 1987). Hence, stability attribution is expected to have a negative influence on repurchase intentions both for LCCs and flagship airlines.

H2.a: In flagship airlines service failures, stability attribution has a negative influence on repurchase intentions.

H2.b: In LCC service failures, stability attribution has a negative influence on repurchase intentions.

Word-of-Mouth

Word of mouth (WOM) is defined as a person-to-person, informal conversation, regarding a certain product, service or brand, which does not contain any commercial bias (East et.al, 2008). WOM may either be positive and constructive, by encouraging usage of a specific product or service, or be negative and destructive complaint, by discouraging usage of a specific product or service.

According to Ziethaml et. al. (1996), WOM holds a special place for services due to their nature of being intangible which makes services riskier, harder to rely on, and difficult to standardize. It is found that, when compared with consumers of goods, consumers of services tend to have a greater confidence in personal sources of information (Murray, 1991). Additionally, negative effects of service failures spread as consumers share their dissatisfying experiences with other individuals (Kim et. al, 2010). Keaveney (1995) supported this fact by putting forward that 50% of service provider substitutions were done in this way.

Folkes (1984) and Folkes et al. (1987) analyzed complaining behavior from attributional perspective. They highlighted the fact that causal attribution dimensions; locus of control, controllability and stability of a failure influence the complaining frequency. Additionally, by analyzing the responses of airline passengers awaiting delayed flights, Folkes et al. (1987) found that anger acts as a mediator between causal attribution dimensions and complaining behavior or repurchase intentions. Therefore, stability attribution is expected to have a negative influence on WOM intentions both for LCCs and flagship airlines.

H3.a: In flagship airlines service failures, stability attribution has a negative influence on WOM intentions.

H3.b: In LCC service failures, stability attribution has a negative influence on WOM intentions.

Research Methodology

According to Churchill and Iacobucci (2005), conducting a pilot study before data collection plays a key role in ensuring questionnaire construction. Since the participation of at least ten individuals is suggested by Fink (1995), in this research, a pre-test was conducted by inviting 15 respondents to participate and share their insights. Academicians and doctoral, master's degree and bachelor students from different major areas of a well-known university participated to the pilot study. Valuable feedbacks regarding grammar, wording and design of the survey were provided. Accordingly, a definition and examples for "flagship airlines" were provided in order to make the term more understandable for respondents.

Firstly, a confirmatory factor analysis (CFA) has been carried out in order to test the measurement model's fit, which is built upon strong theoretical works. Thereafter, due to existence of various latent variables, structural equation modeling (SEM) is utilized, since it enables analyzing latent variables, their relationships, and holds an illustrative power of path diagrams (Nachtigall et al., 2003). By this way we provide a comparison for LCC and flagship airlines for the same model.

Sampling and Data Collection

The proposed research model is tested by surveying individuals living in Izmir. Two interviewers, who were trained for preventing any possible interviewer biases, were charged for gathering data from the field. Data was collected from four different provinces of Izmir which were selected based on their high population. Convenience sampling, which is a non-probabilistic sampling method, was employed to construct the sample handily. Since this study employs airline service failures as area of interest, the universe itself is bounded to the people who have experienced at

least one flight within the last year, which is the reason of existence for the question regarding the number of flights experienced within last one year.

In total 600 surveys were delivered in order to obtain accurate number of completed surveys. Survey was filled by 547 respondents, from which 512 valid surveys gathered. 35 of 547 surveys were invalid due to incomplete surveys. In order to provide construct validity, respondents with age of 18 and above were taken as the target of our study. Basing on different studies in Turkey (i.e. Eroğlu, 2016; Yenipazar and Turhan, 2017), we provide the summary of respondents' demographics.

Table 1. *Demographic Characteristics of the Sample*

		N=512	*%*
Gender	Female	290	56,6
	Male	222	43,4
Education	Primary School	13	2,5
	Secondary School	51	10
	High School	213	41,6
	Associate Degree	58	11,3
	Undergraduate Degree	142	27,7
	Master's Degree	24	4,7
	PhD Degree	11	2,1
Age	18-24	82	16
	25-31	139	27,1
	32-38	149	29,1
	39-45	87	17
	46-52	37	7,2
	53 and above	18	3,5
Monthly Income	0-1000	74	14,5
	1001-2000	141	27,5
	2001-3000	163	31,8
	3001-5000	104	20,3
	50001-7500	19	3,7
	7501 and above	11	2,1

Measures

Expectation measures were gathered from the study of Voss et al. (1998), who developed five-item expectation scale based on service quality dimensions identified by Parasuraman et al. (1988). We used the four-item stability attribution scale of Vázquez-Casielles et al. (2007). A four-item scale for repurchase intentions was obtained from the study of Nikbin et al. (2011). Finally, three-item WOM scale was taken from the study of Maxham III (2001). Multi-item scales were employed in order to increase the quality of the measurement. Ratings were made on a 5-point Likert scale with points ranging from 1= "strongly disagree" to 5= "strongly agree".

Confirmatory Factor Analysis (CFA) Results

We employed a confirmatory factor analysis (CFA) in order to test the measurement model's fit. Overall fit of the data to the model was tested for both low cost carriers and flagship carriers. Due to poor factor loadings, we deleted one item both from repurchase intention and WOM scales. As shown in Table 2, results of CFA for both LCCs and flagship airlines in terms of Chi-square Value, Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), and Root Mean Square Residual (RMR) indicates a good model fit. Solely the Normed Fit Index (NFI) values are to a minor extend below .90. Subsequently, CFA provides evidence of acceptable model fit and assures that the data fits to the proposed measurement model. Results are shown in Table 2.

Table 2. *Goodness of Fit Statistics Resulting from CFA*

Fit Index	LCC	FLG*	Criteria	Reference
χ 2/df	3.462	2.508	< 5	Bentler (1989)
P	0.000	.000	< 0.05	Awang (2012)
CFI	0.958	0.975	> 0.90	Hair et al. (2010), Awang (2012)
NFI	0.943	0.959	> 0.90	Awang (2012)
GFI	0.937	0.955	> 0.80	Forza & Filippini (1998),Greenspoon & Saklofske (1997)
AGFI	0.902	0.932	> 0.80	Cole (1987), Kim & Smith (2007)
TLI	0.944	0.967	> 0.90	Forza & Filippini (1998), Awang (2012)
RMSEA	0.069	0.054	< 0.08	Awang (2012)

Note: FLG: Flagship Airlines, LCC: Low Cost Carrier

Validity and Reliability of the Study

CFA was employed in order to assess the validity of the measures. For discriminant validity testing, Fornell–Larcker Test (Fornell and Larcker, 1981) was conducted. AVE values are taken as indicators of validity and since all the AVE values exceed 0.5 threshold level, convergent validity was confirmed (Hair et al., 2010). Besides, both maximum shared square variance (MSV) and average shared square variance (ASV) values are less than average variance extracted (AVE) values, and all of the AVE values exceed the square of

the correlation between constructs, and thus, discriminant validity was ensured (Fornell and Larcker, 1981; Hair et al., 2010). Table 3 shows all the validity related scores of repurchase intention, expectation, stability and WOM scales.

In order to test the reliability of scales, composite reliabilities were used and as shown in Table 3, they are all above 0.7 threshold (Fornell and Larcker, 1981; Hair et al., 2010).

Table 3. *Validity and Reliability Table for LCC and FLG*

LCC								
	CR	**AVE**	**MSV**	**ASV**	**RI**	**EXP**	**STAB**	**WOM**
RI	0,884	0,720	0,041	0,026	0,849			
EXP	0,928	0,722	0,041	0,026	0,202	0,849		
STAB	0,800	0,506	0,036	0,019	-0,010	0,191	0,711	
WOM	0,824	0,708	0,036	0,019	0,189	0,028	-0,143	0,841
FLAGSHIP								
RI	0,890	0,733	0,141	0,054	0,856			
EXP	0,906	0,662	0,007	0,002	0,086	0,814		
STAB	0,850	0,591	0,034	0,016	-0,121	-0,004	0,769	
WOM	0,810	0,687	0,141	0,058	0,375	-0,007	-0,185	0,829

Notes: CR: Composite Reliability, AVE: Average Variance Extracted, MSV: Maximum Shared Squared Variance, ASV: Average Shared Square Variance, EXP: Expectation, STAB: Stability Attribution, RI: Repurchase Intention

Findings

Results of SEM

As shown in Table 4, both for LCCs and flagship airlines, Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit (AGFI), Chi-square value (χ2/df) and Comparative Fit Index (CFI) are consistent with the required threshold levels. Additionally, both for LCCs and flagship airlines Root Mean Square Error of Approximation (RMSEA) values indicate a good fit by being close to 0.05 threshold level (MacCallum et al., 1996).

Table 4. *Goodness-of-Fit Statistics for LCC and Flagship Airlines*

Fit Index	**LCC**	**FLG**
χ 2/df	3,470	2,860
P	.000	.000
CFI	,950	,963
NFI	,932	,944
GFI	,930	,943
AGFI	,896	,917
RMSEA	,070	.,060

The findings indicate that expectations have a positive influence on stability attributions following low cost airlines service failures (β: 0.183; e: 0.036; p = 0.00). However, expectations do not have an influence on stability attributions following FLG service failures (β: -0.008; e: 0.06; p = 0.87). Therefore, H1.a. is not supported, whereas H1.b. is supported.

In a similar vein, stability attribution has no impact on repurchase intentions following LCC failures (β: 0.003; e: 0.06; p = 0.957). However, following flagship airlines service failures, stability negatively affects repurchase intention (β: -0,132; e: 0.052; p = 0.006). Hence, H2.a is supported, whereas H2.b is not supported.

Additionally, stability attribution has negative impacts on word of mouth in failures of both flagship (β: -0.117; e: 0.051; p= 0.000) and low cost airlines (β: -0.103; e: 0.059; p=0.007). Concordantly, both H3.a and H3.b are supported. Results of SEM are given in Table 5, and are schematized for flagship airlines and LCCs respectively in Figure 2 and Figure 3.

Table 5. *Results of SEM*

	Unstandardized Regression Weights	**Standardized Regression Weights**	**S.E.**	**C.R.**	**P**
Flagship Airlines					
STAB ← EXP	-0.01	-0.008	0.06	-0.163	0.87
RI ← STAB	-0.143	-0.132	0.052	-2.756	0.006
WOM ← STAB	-0.217	-0.117	0.051	-4.237	***
LCCs					
STAB ← EXP	0.135	0.183	0.036	3.774	***
RI ← STAB	0.003	0.003	0.06	0.053	0.957
WOM ← STAB	-0.161	-0.103	0.059	-2.708	0.007

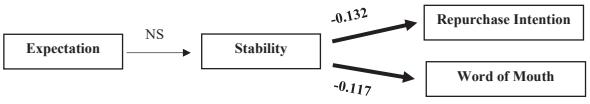


Figure 2. *Final Model with Hypotheses Test Results for Flagship Airlines*

Figure 3. *Final Model with Hypotheses Test Results for LCCs*

Discussion of the Results

This research focuses on attributions and reactions of passengers following failure incidences, and contributes to the service failure and aviation research by highlighting the relations among expectations, stability attributions and behavioral intentions. Furthermore, the current study broadens the extant body of knowledge in the field through examining the aforementioned relationships separately for two distinct passenger airlines business models, namely as flagship airlines and LCCs.

Results of this study indicate that following low cost airlines service failure, expectations and stability attribution have a positive relationship. Such relationship following a service failure may be explained by expectation disconfirmation paradigm, which states that when initial service performance exceeds pretrial expectation, positive disconfirmation occurs and in cases where performance falls behind pretrial expectation, negative disconfirmation occurs (Bitner, 1990). Concordantly, negative disconfirmation caused by a failure experience may lead further negative evaluations (Rego et al., 2009). In line with this line of thought, it may be noted that, following LCC service failures, passengers experience a negative expectation disconfirmation that, in turn, leads attributing cause of failure to be stable over time. On the other hand, surprisingly, relationship between expectation and stability is found to be insignificant for flagship airlines.

Related literature states that customers who perceive the cause of a failure to be stable are more confident that the same failure will reoccur (Weiner, 1986), and concordantly, stability impacts customers' repurchase behaviors (Folkes, 1984). For flagship airlines, findings of the study are congruent with the previous literature. When the customers perceive high stability in failures, they lower repurchase intentions for the flagship airlines. On the other hand, for LCC service encounter, a relationship between stability and repurchase intentions was not observed. The reason lying behind the non-existence of a statistically significant relationship between stability attribution and repurchase intentions following LCC service failures may be the nature of air transport industry. Passengers may be patronizing airlines in accordance with convenience of flights, availability of destinations, price deals and/or connected flight availability and options. Based on the literature (e.g. Vlachos and Lin, 2014; Calisir et al., 2016; Rajaguru, 2016) the underlying reason may be linked to the fact that passengers make decisions and airlines choice case-by-case for each flight and take buying action in accordance with facts regarding flight options and prices without considering past failure experiences.

Related literature indicates that when customers attribute cause of a failure to be relevant to various diverse situations, they make negative evaluations about the provider (e.g. Hess, 1999; Hess et al., 2007). Moreover, as mentioned previously, Folkes (1984) and Folkes et al. (1987) highlighted the fact that causal attribution dimensions of locus of control, controllability and stability effect the complaining behavior. In congruence with extant body of knowledge, findings of this study point out a negative relationship between stability attribution and WOM intention both for flagship airlines and LCCs. To put it another way, when passengers attribute the cause of a service failure to be stable, or permanent, over time, their intention to engage in positive WOM intention decreases. Since failures with enduring causes occur more frequently, and thus concordantly influence consumers' expectations regarding future performance of firms (Folkes 1984; Weiner, 2000), negative relationship between stability attribution and WOM intention is legitimate.

This study contributes to the practitioners in aviation industry by providing managerial implications separately for LCCs and flagship carriers. Findings reveal that airlines' business model creates noteworthy differences among dynamics by which passengers evaluate service failures. The practitioners in low cost airlines should be aware of the link among expectations and stability attributions. Besides, in flagship airlines, if the customers perceive the failures as enduring and stable, their intention to repurchase and engage in positive WOM decrease. Since positive WOM plays a significant role in marketing, practitio-

ners in LCC and flagship airlines should be the ones that convince the customers about the instability of failures experienced. Information provision with explanation and communication during and after failures gain importance in this persuasion process.

Limitations and Recommendation for Further Research

This study has its own limitations, which may also limit findings of the study. First, the severity effect of the failure has not been investigated in this study. For a passenger going to an important business meeting, the effect of a 30-minutes-delay may conclude in different attributional responses. Thus, failure severity in airlines context is worthy for further investigation.

Secondly, service failures may occur at different levels; at employee level or at organizational level. For different cases involving different level of service failures, behavioral intentions and attributions following a service failure may vary. For example, a passenger may be more propitious in terms of repurchase intentions in case of having a negative experience with an impolite stewardess, when compared with the case of cancellation of a flight without giving a notice. In accordance, failures may be categorized as employee-level and organization-level for further investigation of the model.

Furthermore, this study examines the data obtained solely from Turkish passengers. The study may be broadened by obtaining data from international passengers from different cultures. Consumers from different cultures are found to have dissimilar levels of service expectations, and behavioral and attributional tendencies (Patterson and Smith, 2001; Mattila and Paterson, 2004; Carpenter, 2000). Accordingly, in further researches another set of data may be obtained from citizens of a western country and a comparison may be made between individuals belonging to eastern and western culture. The comparison may supply significant managerial implications in terms of behavioral and attributional patterns of customers.

References

Ahmad, R.; Buttle, F., 2001. *Customer Retention: A Potentially Potent Marketing Management Strategy*. Journal of Strategic Marketing, 9(1): 29-45.

Awang, Z., 2012. Structural equation modeling using AMOS graphic. Penerbit Universiti Teknologi MARA.

Bentler, P.M., 1989. EQS Structural Equations Program Manual, BMDP Statistical Software, Los Angeles.

Bhadra, D. (2009). You (expect to) get what you pay for: A system approach to delay, fare, and complaints. *Transportation Research Part A: Policy and Practice*, *43*(9), 829-843.

Bitner, M. J. (1990). Evaluating service encounters: the effects of physical surroundings and employee responses. *the Journal of Marketing*, 69-82.

Boulding, W., Kalra, A., Staelin, R., & Zeithaml, V. A., 1993. A dynamic process model of service quality: from expectations to behavioral intentions. Journal of marketing research, 30(1), 7.

Calisir, N., Basak, E., & Calisir, F., 2016. Key drivers of passenger loyalty: A case of Frankfurt–Istanbul flights. Journal of Air Transport Management, 53, 211-217.

Carpenter, S., 2000. Effects of cultural tightness and collectivism on self-concept and causal attributions. Cross-Cultural Research, 34(1), 38-56.

Churchill, G. A., and Iacobucci, D., 2005. Marketing Research: Methodological Foundations (9th ed.). Mason, Ohio: Thomson South-Western.

Cole, D. A., 1987. Utility of confirmatory factor analysis in test validation research. Journal of consulting and clinical psychology, 55(4), 584.

Davis, M. M.; Heineke, J., 1998. How disconfirmation, perception and actual waiting times impact customer satisfaction. International Journal of Service Industry Management, 9(1): 64 – 73.

Day, R., 1980. "Research perspectives on consumer complaining behavior", in Theoretical Developments in Marketing, Lamb, C.W. and Dunne, P.M. (Eds), American Marketing Association, Chicago, IL, pp. 211-5.

Doganis, R., 2001. The Airline Business in the 21st Century. Routledge, London.

East, R. and et al. (2008). *Measuring the Impact of Positive and Negative Word of Mouth on Brand Purchase Probability.* International Journal of Research in Marketing, 25(3):215-224.

Eroğlu, F. (2016). *Kompulsif Satin Alma: Kişisel Faktörler, Postmodern Tüketim Şekilleri Ve Reklama Karşi Tutumun Rolüne Yönelik Bir Araştirma. Balikesir University Journal of Social Sciences Institute, 19(35).*

Fink, A., 1995. The survey handbook (Vol. 1). Thousands Oaks, CA: Sage.

Folkes, V. S. (1984). Consumer reactions to product failure: An attributional approach. Journal of consumer research, 398-409.

Folkes, V. S. (1984). Consumer reactions to product failure: An attributional approach. *Journal of consumer research*, *10*(4), 398-409.

Folkes, V. S., Koletsky, S., and Graham, J. L., 1987. A field study of causal inferences and consumer reaction: the view from the airport. Journal of consumer research, 534-539.

Fornell, C., and Larcker, D. F., 1981. Evaluating structural equation models with unobservable variables and measurement error. Journal of marketing research, 39-50.

Forza, C., & Filippini, R., 1998. TQM impact on quality conformance and customer satisfaction: a causal model. International journal of production economics, 55(1), 1-20.

Francis, G.; Humphreys, I.; Ison, S., 2004. Airports' Perspectives on the Growth of Low-Cost Airlines and the Remodeling of the Airport-Airline Relationship. Tourism Management, 25(4): 507-514.

Greenspoon, P. J., & Saklofske, D. H., 1997. Validity and reliability of the multidimensional students' life satisfaction scale with Canadian children. Journal of Psychoeducational Assessment, 15(2), 138-155.

Grönroos, C., 1984. A Service Quality Model and Its Market Applications. European Journal of Marketing, 18(4): 36-44.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L., 2010. Multivariate data analysis. Pearson.

Harris, J.; Uncles, M., 2007. Modeling the Repatronage Behavior of Business Airline Travellers. Journal of Service Research, 9(4): 297-311.

Hess, R. L., Ganesan, S., and Klein, N. M., 2003. Service failure and recovery: the impact of relationship factors on customer satisfaction. Journal of the Academy of Marketing Science, 31(2), 127-145.

Hess, R. L., Ganesan, S., and Klein, N. M., 2007. Interactional service failures in a pseudorelationship:

The role of organizational attributions. Journal of Retailing, 83(1), 79-95.

Hunter, L., 2006. Low Cost Airlines: Business Model and Employment Relations. European Management Journal, 24(5): 315-321.

Kassim, N.M.; Souiden, N., 2007. Customer retention measurement in the UAE banking sector. Journal of Financial Services Marketing, 11(3): 217–228.

Keaveney, S. M. (1995). *Customer switching behavior in service industries: An exploratory study*. The Journal of Marketing, 71-82.

Kim, W., and Ok, C. (2010). Customer orientation of service employees and rapport: Influences on service-outcome variables in full-service restaurants. *Journal of Hospitality and Tourism Research*, *34*(1), 34-55.

Kim, Y. K., and Smith, A. K., 2007. Providing a critical service today for tomorrow's consumers: A relational model of customer evaluations and responses in the child care industry. Journal of Retailing and Consumer Services, 14(3), 232-245.

Mason, K.J., 2000. The Propensity of Business Travelers to Use Low Cost Airlines. Journal of Transport Geography, 8(2): 107-119.

Mattila, A. S., and Patterson, P. G., 2004. The impact of culture on consumers' perceptions of service recovery efforts. Journal of Retailing, 80(3), 196-206.

Maxham III, J. G. (2001). Service recovery's influence on consumer satisfaction, positive word-of-mouth, and purchase intentions. *Journal of business research*, *54*(1), 11-24.

McDougall, G.; Levesque, T., 1998. The Effectiveness of Recovery Strategies After Service Failure: An Experiment in the Hospitality Industry. Journal of Hospitality and Leisure Marketing, 5 (2/3): 27–39.

Monson, T.; Snyder, M., 1977. Actors, observers and the attribution process: Toward a reconceptualisation. Journal of Experimental Social Psychology, 43, 89-11.

Mostert, P.G.; De Meyer, C.F.; Van Rensburg, 2009. The Influence of Service Failure and Service Recovery on Airline Passengers' Relationships with Domestic Airlines: An Exploratory Study. Southern African Business Review, 13(2): 118-140.

Murray, K. B. (1991). *A test of services marketing theory: consumer information acquisition activities*. The journal of marketing, 10-25.

Nachtigall, C., Kroehne, U., Funke, F., and Steyer, R., 2003. Pros and Cons of Structural Equation Modeling. Methods of Psychological Research Online, 8(2), 1-22.

Nikbin, D., Ismail, I., Marimuthu, M., & Younis Abu-Jarad, I., 2011. The impact of firm reputation on customers' responses to service failure: the role of failure attributions. Business Strategy Series, 12(1), 19-29.

O'Connell, J. F.; Williams, G., 2005. Passengers' Perceptions of Low Cost Airlines and Full Service Network Carriers: A Case Study Involving Ryanair, Aer Lingus, Air Asia and Malaysia Airlines. Journal of Air Transport Management, 11(4): 259-272.

Oliver, R .L., Swan, J. E., 1989. Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach. Journal of Marketing, 53: 21-35.

Oliver, R. L., 1981. Effect of Satisfaction and Its Antecedents on Consumer Preference and Intention. Advances in Consumer Research, 8: 89-93.

Ozenen C.G., 2003. Havaalanı Yatırımlarında Özelleştirme Dünyadaki Uygulamalar ve Türkiye İçin Öneriler. Ankara: Devlet Planlama Teşkilatı.

Parasuraman, A., Zeithaml, Valerie A. and Berry, Leonard L., 1988. "SERVQUAL: A Multiple Item Scale for Measuring Consumer Perceptions of Service Quality," Journal of Retailing, 64 (Spring), 12-40.

Patterson, P. G., and Smith, T., 2001. Relationship benefits in service industries: a replication in a Southeast Asian context. Journal of Services Marketing,15(6), 425-443.

Payne, J., 1993. The Adaptive Decision Maker. Cambridge, UK: Cambridge Press.

Rajaguru, R., 2016. Role of value for money and service quality on behavioural intention: A study of full service and low cost airlines. Journal of Air Transport Management, 53, 114-122.

Rego, L.L., Billett, M.T., & Morgan, N.A. (2009). Consumer-based brand equity and firm risk. Journal of Marketing, 73, 47–60.

Reichheld, F., 1996. The Loyalty Effect. Boston: Harvard Business School Press.

Rosenberg, L. J., Czepiel, J. A., & Cohen, L. C. (1977). Consumer Affairs Audits: Evaluation and Analysis. *California Management Review*, *19*(3), 12-20.

Smith, A. K., and Bolton, R. N., 1998. An experimental investigation of customer reactions to service failure and recovery encounters paradox or peril? Journal of service research, 1(1), 65-81.

Sull, D., 1999. EasyJet's $500 million gamble. European Management Journal, 17(1), 20-32.

Taumoepeau, S., Towner, N., & Losekoot, E., 2017. Low-Cost Carriers in Oceania, Pacific: Challenges and opportunities. Journal of Air Transport Management, 65, 40-42.

Teas, R. K., 1981. An Empirical Test of Models of Salespersons, Job Expectancy and Instrumentality Perceptions. Journal of Marketing Research, 18(2): 209-226.

Torres, E.N. and Kline, S., 2006. From Satisfaction to Delight: A Model for the Hotel Industry. International Journal of Contemporary Hospitality Management., 18(4):290 – 301.

Vázquez-Casielles, R. et al., 2007. Quality of Past Performance: Impact on Customers' Responses to Service Failure. Marketing Letters, 18(4): 249-264.

Vidović, A.; Štimac, I.; Vince, D., 2013. Development of Business Models of Low-cost Airlines. International Journal for Traffic and Transport Engineering, 3(1): 69-81.

Vlachos, I., & Lin, Z. (2014). Drivers of airline loyalty: Evidence from the business travelers in China. *Transportation Research Part E: Logistics and Transportation Review*, *71*, 1-17.

Voss, G. B., Parasuraman, A., & Grewal, D., 1998. The roles of price, performance, and expectations in determining satisfaction in service exchanges. The Journal of Marketing, 46-61.

Weiner, B., 1980. A cognitive (attribution)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. Journal of Personality and Social psychology, 39(2), 186.

Weiner, B., 1986. Attribution, emotion, and action. Handbook of motivation and cognition: Foundations of social behavior, 1, 281-312.

Weiner, B., 2000. Attributional thoughts about consumer behavior. Journal of Consumer research, 27(3), 382-387.

Williams, G., 2001. Will Europe's Charter Carriers Be Replaced by "No-Frills" Scheduled Airlines? Journal of Air Transport Management, 7(5): 277-286.

Wittman, M. D., 2014. Are low-cost carrier passengers less likely to complain about service quality? Journal of Air Transport Management, 35, 64-71.

Yenipazar, E., & Turan, A. H. (2017). *Son Kullanici Bilgi Teknolojileri Tatmini Sakarya Üniversitesi Sabis Örneği*. Suleyman Demirel University Journal of Faculty of Economics & Administrative Sciences, 22(2).

Zeithaml, V. A., Berry, L. L., and Parasuraman, A. (1996). *The behavioral consequences of service quality*. The Journal of Marketing, 31-46.

EGE AKADEMİK BAKIŞ / *EGE ACADEMIC REVIEW*

Cilt 19 • Sayı 4 • Ekim 2019
SS. 437 / 455
Doi: 10.21121/eab.486249
Başvuru Tarihi: 21/11/2018 • Kabul Tarihi: 08/10/2019

Does Trade Credit Channel Operate in Turkey? An Analysis with CBRT Sector Statistics

Türkiye'de Ticari Kredi Kanalı İşliyor mu? TCMB Sektör İstatistikleri ile Bir Analiz

Arzu ŞAHİN[1]

ABSTRACT

This paper aims to test trade credit channel in Turkey over the period from 2008 to 2016 using two data sets at sectoral level obtained from the aggregated accounts of Central Bank of the Republic of Turkey (CBRT) statistics. First sample consists of 7 main sectors and other sample involves 12 manufacturing-sub sectors. Dynamic panel data methods were applied to test whether monetary and nonmonetary factors of trade credit had an impact on both gross and net trade credit status. According to analysis results about trade credit received (gross), monetary policy restrictions has increased the trade credit usage of SME sized firm classes. Findings for net trade credit (net) showed that main-sectors medium sized firm groups became net trade credit receiver. These findings confirming substitution hypothesis conclude that trade credit channel has operated with the strongest effects on SME firms.

Key words: trade credit channel, trade credit, credit channel

JEL Classification: E44, E52, G32.

ÖZET

Bu çalışma, Türkiye Cumhuriyet Merkez Bankası (TCMB) istatistiklerinin toplam hesaplarından elde edilen sektör düzeyinde iki veri setini kullanarak, 2008'den 2016'ya kadar olan süreçte Türkiye'de ticari kredi kanalını test etmeyi amaçlamaktadır. İlk örneklem 7 ana sektörden ve diğer örneklem 12 imalat alt sektöründen oluşmaktadır. Ticari kredilerin parasal ve parasal olmayan faktörlerinin hem brüt hem de net ticari kredi durumu üzerinde bir etkisi olup olmadığını test etmek için dinamik panel veri yöntemleri uygulanmıştır. Alınan ticaret kredisi (brüt) ile ilgili analiz sonuçlarına göre, para politikası kısıtlamaları KOBİ ölçekli firma sınıflarının ticari kredi kullanımını artırmıştır. Net ticaret kredisi (net) için bulgular, ana sektör orta ölçekli firma gruplarının net ticari kredi alıcısı olduğunu göstermiştir. İkame hipotezini doğrulayan bu bulgular, ticari kredi kanalının, en güçlü etkisi KOBİ firmaları üzerinde olmak üzere, faaliyet gösterdiği sonucuna varmıştır.

Anahtar Kelimeler: ticari kredi kanalı, ticari kredi, kredi kanalı

To explain the way how monetary policy stance affects spending, borrowings and consequently the entire economy, two basic approaches known as money channel and credit channel have been suggested. In money view (or interest rate view), monetary authorities strive to control aggregate demand and production level by basically changing money market interest rates to alter the cost of capital (Bernanke & Gertler, 1995: 27; Hong, 2017: 40). Money view disregards the financial sector and the role of bank assets (bank loans), while the credit view accepts the importance of credit market imperfections and bank loans in the monetary policy transmission (Nilsen, 2002: 226). Financial market imperfections that lie under the credit view can magnify the influence of monetary policy impulses via two channels known as balance sheet channel and bank-lending channel. The balance sheet channel emphasises the importance of borrower's financial strength as the basic factor to be affected from the monetary policy innovations, whereas the bank-lending channel stress the way how monetary policy shocks influence the bank loan supply (Bernanke & Gertler, 1995: 35; Atanasova & Wilson, 2003: 505; Hong, 2017: 4-5) and then borrowers separately from

[1]asahin@atu.edu.tr

the cost of capital. Bank-lending channel posits that contractions in bank loans in the periods of monetary restraint force the bank-dependent firms (firms that are dependent on bank loan as most preferable external finance) to cut back spending regardless of interest rate (Nilsen, 2002: 248; Bernanke & Gertler, 1995: 35-40; Hong, 2017: 32-37). Therefore, bank-lending channel have usually the most severe effect on bank dependent borrowers (Atanasova & Wilson, 2003: 505). Trade credit as a substitute fund source may provide a solution to the liquidity shortages of bank dependent firms especially under the conditions of the intensive credit rationing or high premium on external finance[1] (Guariglia & Mateut, 2006: 2836) during economic downturns.

Trade credit received (TCR) means payables to suppliers; trade credit extended (TCE) shows the receivables from customers (Brechling & Lipsey, 1963: 620). TCR, TCE, and their changes are expressed as gross trade credits (GTC) whereas the net difference between them is called net trade credits (NTC). Movements in both gross and net trade credits may serve as a tool to frustrate the workings of monetary policy (Brechling & Lipsey, 1963; Chiplin & Wright, 1985: 221). Meltzer (1960) highlighted the trade credit role in the credit channel and firstly proposed that trade credit may work as a mechanism to offset bank-lending channel by serving as a substitute for bank loan (De Blasio, 2003; Nilsen, 2002; Guariglia & Mateut, 2006; Gama & Auken; 2015). According to the Meltzer's (1960) suggestion, small firms can make greater use of trade credit to overcome bank credit rationing during monetary contraction periods and this part of his suggestion may more relate to hypothesis known as substitution theory. Besides substitution theory, Meltzer (1960) also suggests the redistribution hypothesis meaning that trade credit should flow from larger firms to small firms. Larger firms are more liquid, less credit constrained and high access to credit market, whereas smaller firms are less liquid, more credit constrained and low access to credit market (Meltzer, 1960). Redistribution theory suggests that unconstrained firms that can readily access to bank credit will offer loans received, as trade credit, to credit rationed firms (Hong, 2017: 216) without ability to access to bank credit and with greater default risk, in order to maintain their own sales volume by helping customers with relaxed accounts receivable terms. With these features, trade credit may play a buffer role for credit channel and more generally may frustrate the monetary policy implications (Meltzer, 1960). Literature developed and used different terms that are based on Melzer's (1960) idea such as trade credit channel (TCC), trade credit offsetting effect, substitution hypothesis, redistribution or reallocation hypothesis, complementarity hypothesis and, helper theory.

Bond, equity and commercial paper markets as an alternative financing source are weak while bank loans and trade credits are dominant in financial structure of firms in Turkey. According to 2016 company accounts of CBRT, bank loans make up 32.2%, TCR from suppliers (trade debt) accounts for 13.8%, financial leasing takes percentage of 2.6% and other financial liabilities including bond, equities and commercial paper constitute only 1.5% of all sources. TCE (trade receivables) constitutes 16% of Turkish firms' assets. In spite of the critical role of inter-firm trade credit as the best alternative to bank loan, it is noticed that existence and operation of trade credit offsetting channel has not been adequately researched with Turkish data so far. Özlü and Yalçın (2012) directly investigated the TCC at the manufacturing firm level with gross trade credit terms using static panel data between 1996 and 2008. Without directly dealing with TCC, two papers (Demirgunes, 2016; Şahin, 2018) examined the trade credit determinants and theories on the data sets from Borsa Istanbul. The purpose of this paper is to examine TCC in Turkish real sector in terms of both GTC and NTC terms using dynamic panel data method on sectoral level samples. Because TCC operates mostly through bank-dependent (financially constrained) firms and small firms are most likely to be accepted as severely reliant on bank loans and as financially constrained (Gertler & Gilchrist, 1994: Atanasova & Wilson, 2003; Özlü & Yalçın, 2012:106) small firms (sector aggregations of small sized firms) are the centre of this study. Therefore, this paper aims to find out whether bank-dependent small firms have increased their trade credit usage in the wake of monetary contraction.

Trade Credit Channel and Literature Review

Trade credit provided by vendors as a short-term fund to finance purchases of customers (Nilsen, 2002: 228) is an inevitable source of finance (Chiplin & Wright, 1985: 221) and a key component of most company's activities (Gama & Auken, 2015: 888) allowing firms to borrow from each others. Suppliers are willing to grant trade credit as a relatively cheaper way of promoting sales when credit rationing decreases the demand for products during tight money periods (Meltzer, 1960: 429 - 431). Meltzer's (1960) argues that financially constrained firms substitute bank credit with trade

credit by lengthening TCR or squeezing TCE, following a tight period (De Blasio, 2003: 85). Therefore, when bank loans become rationed, expensive and not readily accessible to firms that are reliant on bank credit, these bank-dependent firms can maintain their real activities up to a level to which they can obtain trade credit. Due to the presence of TCC, impact of a monetary restriction may be weakened (Guariglia & Mateut, 2006: 2837) and a single firm can frustrate monetary policy by using NTC to finance its current activity level that would have been constrained with a restrictive monetary policy (Brechling & Lipsey, 1963: 627).

The Meltzer's (1960) intuition also suggests that trade credit imperfectly substitutes bank loans and plays a complementary sources of financing (Gama & Auken, 2015: 889) when loan constraints are imposed by financial institutions (Carbo-Valverde, Mansilla-Fernandez & Rodriguez-Fernandez, 2017: 25). Trade credit can be engaging but it also has some repellent sides. Prompt payment discounts and strict penalties for defaults cause trade credit to be more expensive than bank borrowings (Petersen & Rajan, 1997: 688). Despite the higher trade credit cost that makes it a less desirable external finance alternative to bank loan, corporations would still be willing to use this more expensive short-term financing and to make more usage when they face restrictions (Petersen & Rajan, 1997; Atanasova and Wilson, 2003; Mateut, Bougheas & Mizen, 2006; Guariglia & Mateut, 2006: 2837; Hong, 2017).

We might list three underlying factors behind the trade credit becoming attractive. First of all, monetary contractions do not increase the cost of trade credit as bank credit (Kohler, Britton & Yates, 2000: 10), delinquencies in accounts payables may be less costly (Chiplin & Wright, 1985: 223) and trade credit terms are likely to maintain the same level over time (Ng, Smith & Smith, 1999: 1128). Therefore, trade credit as the best option for bank credit (Fisman & Love, 2003: 373) may remain relatively cheaper for some firms (Atanasova & Wilson, 2003: 505). Secondly, trade credit appears to alleviate the information problems (Gama & Auken; 2015: 887) and lower the risks of trade credit suppliers than risk incurred by financial institutions (Guariglia & Mateut, 2006: 2837). Sellers have some advantages in lending over financial institutions through business relationships. Trade credit extenders have greater ability to collect current information, to monitor repayment and early-payment discount acceptance behaviours, and so to evaluate financial strength quickly, to enforce repayment, to apply precautions for defaults and to liquidate the goods sold which is a kind of collateral for vendors (Petersen & Rajan, 1997: 688; Kohler et al., 2000: 9-10). Furthermore, and as the third factor, from the profit maximizing perspective, suppliers' interests extend beyond financial institutions because they benefit from the survival of their customer relationships (Gama & Auken, 2015: 899). Future business potential in case of ongoing transaction with customer leads suppliers to finance suspect but growing firms via granting trade credit (Petersen & Rajan, 1997: 688). To conclude, owing to these costs and information advantageous over banks, vendors can adjust credit terms both for themselves and borrowers and therefore trade credit becomes the optimal external financial choice for financially constrained firms (Mateut et al., 2006, 622-627).

The structures of industries and corporations are important factors that shape their responses to monetary impulses and it is obvious that wide variations exist across different industries (Arslan & Ergeç, 2011: 90). The economy-wide total volume of NTC may remain stable because total lenders should equal total borrowers in the absence of foreign flows. However, from the sectoral viewpoint, NTC varies per sector. Some can be net receivers whereas the others can be net extenders. That is why reactions of industries to monetary policy shocks may differ from each other (Chiplin & Wright; 1985: 221). Ng et al. (1999) concluded from the survey applied to 950 US firms, that credit terms and policies differ greatly across industries while tend to be similar within the same industry. Various buyer characteristics and non-salvageable (sunk) investment requirements in industries contribute to these differences (Ng et al., 1999: 1127-1128). Dai and Yang (2015), included industry dummies while Chiplin and Wright (1985) presented regression results run for each industry in their firm-level analyses. Herbst (1974) studied on one sector and emphasized the need for empirical reports at firm or industry level.

Reviewed literature that analyse monetary policy and trade credit relations is presented according to the economies where study is conducted and a summary of them is shown at the Table 1. Meltzer (1960) investigated the trade credit behaviour of US manufacturing companies as response to the tightened monetary conditions of 1955-1957. By directly relating the money market policy to liquidity position of different size groups, Meltzer (1960) finds a reduction in liquidity level of sample firms and an increase in the bank

credit level. Large firms with easy access to bank and non-bank funds and with greater liquid assets reallocate assets towards accounts and notes receivables. According to the author, the main purpose of trade credit extension at tight periods is to avoid losses from price reduction and to overcome sales decreases. As a usual result of tight monetary controls, discrimination through credit rationing against small firms can be limited by lengthening trade credit terms (Meltzer, 1960). For the US market Meltzer's (1960) hypothesis was confirmed by Laffer (1970), Schwartz (1974), Nilsen (2002) and Choi and Kim (2003), partly supported by Herbst (1974) and Petersan and Rajan (1997), while totally rejected by Nadiri (1969) and Walker (1985).

The findings of Laffer (1970) who theoretically and empirically examined the unutilized trade credit available as a component of the total money stock, signalled the substitution between bank money and trade credit money. Laffer (1970) stated that being a close substitute for bank money and not being subject to regulations made trade credit a tool to largely avoid the policy implications to change bank money. Schwartz (1974) who mainly investigated the factors that explain differing trade credits terms between firms, sectors and time, presented results for monetary policy effect on trade credit. Schwartz (1974) indicated that small firms and unrated large firms both suffering from credit constraints met their loan demand by trade credit as the propitious but costly and unfavourable substitute credit during tight monetary episodes. These findings revealed that large firms with the advantage of financing motive were able to ease the effects of tight policies by extending trade credit terms. Nilsen (2002) revealed the irrelevance of firm size by providing evidence about increased trade credit usage of non-credible large firms that have no bond rating. Comparing quoted large firms with non-quoted small firms, Choi & Kim (2003) confirmed substitution effect of trade credit following higher interest rate and supported smoothing effect of trade credit on credit contraction. In their study, inter-firm liquidity flowed more actively but as a contrary to redistribution theory, liquidity mainly flowed to larger firms from smaller firms.

Focusing on trade credit determinants of the U.S. lumber and wood products industry, Herbst (1974) found no direct monetary influences on accounts payables but reported that higher current obligation of long-term bank loans may force firms to rely more on trade credit. Petersan and Rajan (1997) studied the determinants of trade payables and receivables with a broad SME sample for U.S. Although authors didn't directly relate monetary indicators with accounts payable, their result suggesting that small firms made a higher usage of vendor financing when financial institutions constraint credits might provide a partially support for the Meltzer's (1960) hypothesis. Nadiri (1969) found no evidence that accepts the reaction of manufacturing sector to the inflationary pressures by supplying trade credit to other sectors. In a profit maximization model while gross and net credits react to changes in their user costs, NTC seems to be insensitive to monetary policy stance (Nadiri, 1969: 421). Walker (1985), who surveyed the role of trade credit supply in SME financing with 27 firms' questionnaires, stated that besides other findings, trade credit had been relatively insensitive to interest rate changes, recession and recovery periods in related years.

Apart from Chiplin and Wright (1985) almost all trade credit researchers, involved in this study (Brechling & Lipsey, 1963; Kohler et al., 2000; Mateut & Mizen, 2003; Atanasova & Wilson, 2003, Mateut et al., 2006; Guariglia & Mateut, 2006; Hong, 2017), investigating the existence and magnitude of TCC for the UK concluded that monetary restrictions boosted the trade credit level of UK firms. With analysis of 18 industries, Chiplin and Wright (1985) showed that while inter-industries differences existed, monetary policy generally seemed to neutralize them and eliminated the balancing effect of NTC on tight politics. Brechling and Lipsey (1963) suggested that NTC rather than gross credit had a very strong potential to frustrate monetary policy and to be a source of inflationary finance. In addition, NTC signalled the redistribution of cash balances from those firms having them to those firms in need of them. Kohler et al., (2000) confirmed the argument of TCC reporting that, after tight times, publicly held UK firms became a net extender and helped bank-dependent firms to whom capital market funds are not readily available. The findings of Mateut and Mizen (2003) in absolute and relative terms, in the wake of stringent money conditions, supported indirectly that bank-lending channel operated and trade credit could offset the monetary policy implications to some degree in the UK. They indicated that vendor financing was an important short-term external source alternative of credit constrained firms (typically small, young and risky firms).

Atanasova and Wilson (2003) confirmed strongly smoothing effect of TCC showing that during mone-

tary restraints, rationed companies in UK requested trade credit, although it is an unattractive/expensive substitute, far greater than un-constrained firms did. Mateut et al. (2006) observed that, by modelling and empirically analysing UK manufacturing firms, tightened monetary policy had a reducing effect on bank loans, but mid-wealthy (mid credit rating) small firms with few assets, financed their project by trade credits. They suggested that any increase in interest rate would raise threshold wealth level for both bank loan and inter-firm credits, but bank-lending volume decreased more than trade credit because firms suffering from the decreasing bank loan resorted to trade credit. Based on investment equation Guariglia & Mateut (2006) verified both credit channel and TCC offsetting effect. More recently, using UK's SMEs, Hong (2017) analysed credit rationing and interactions between monetary implications, financing mix and trade credit usage for borrowing constrained firms. The author tested substitution hypothesis from the perspective of trade credit users, and tested redistribution theory from the side of trade credit grantor and supported both theories. Estimates revealed that credit-constrained firms substituted greatly bank loan with trade credit whereas credit-unconstrained firms were more willing to redistribute bank loan by supplying trade credit to constrained firms.

Using data from Italy (De Blasio, 2003; Marotta,1997), from Turkey (Özlü and Yalçın, 2012), from China (Dai &Yang, 2015), from Portugal (Gama & Auken, 2015), from Spain (Carbo-Valverde et al., 2017), and from multi-country samples (Fisman & Love, 2002; Saiz, Azofra, Olmo & Gutierrez, 2017), 8 articles analysing the real sector trade credit reaction to economic downturns have come to the similar conclusion that TCC operates in other countries the same way as in the US and the UK. For Italian case, examining the Italian manufacturing firms' inventory behaviour De Blasio (2003) approved TCC with findings suggesting that the restrictive effect, with a modest magnitude, of the trade credit availability on inventory investments was as twice as powerful in the periods of monetary contractions. Besides, substitution effect was more likely relevant to small firms and firms having less collateralizable assets. Evidences of Marotta (1997) documented redistribution hypothesis worked but in the opposite direction in which trade credit flows to larger firms, meaning that trade credit did not behave as a shield for small firms to protect them from monetary restrictions.

Gama & Auken (2015) examined the bank loan and trade credit interrelationship focusing on the holdup problems and credit constraints caused from the monopolistic power of a single main bank on Portuguese SME. Their findings suggested that some advantages of suppliers over banks allowed them to provide interest rate more competitive than main banks and trade credit became a solution to customers lacking liquidity. Findings of Dai & Yang (2015) for China showed that the positive relationship of accounting conservatism with trade credit was more powerful under strict monetary conditions when the need for conservative accounting to decrease information asymmetries raised. Using Spanish firm level data, Carbo-Valverde at al. (2017), concluded that in spite of increasing cost of trade credit after restrictive monetary policy, firms were still likely to become more trade credit users. While rising interest rates motive large firms to being trade credit lenders, financial crisis turns their behaviour towards being trade borrower. Crisis effect is more pronounced for SMEs that are strongly reliant on the trade credit especially during crisis in which cuts in bank lending is seen. According to their findings, restrictive policy narrows the cost distance between bank borrowing and trade credit and this result, the opposite of substitution effect, is expressed as complementary hypothesis. Fisman and Love (2002) provided country and industry level evidence for Meltzer's suggestion on a sample of 43 countries and 37 industries. They argued that in countries with poorly developed financial markets, firms financed growth by trade credit and industries heavily reliant on trade credit grew more rapidly. Saiz et al. (2017) who investigated the relation of sovereign risk with trade credit by taking the crisis and monetary contractions into consideration for non-financial firms from 12 Eurozone countries, has documented restrictive monetary stance raised the trade credit payables in lower sovereign risk countries.

Özlü and Yalçın (2012) offered supporting evidence for the trade credit offsetting effect for Turkey with a manufacturing firm level data set over 1996-2008. They documented that small firms were more prone to be credit constrained and these constraints were stronger at tight periods. Thus financially constrained firms, defined as small firms with low export sales, substituted trade credit for bank loan during hard times. Moreover, because trade credit flowed from large manufacturing firms (not bank loan constrained) to small ones, findings also confirmed the redistribution hypothesis. In addition, for Turkey trade credit markets,

Demirgunes (2016) investigated the determinants of TCR (accounts payable) for 10 publicly held SMEs between 2008-2015 and Şahin (2018) studied determinants of TCE (accounts receivable) and analysed the validity of three trade credit theories on the 23 firms from Borsa Istanbul with 2016 data.

Table 1: Trade Credit Channel, Literature Review Summary

	Paper	Country, Model, Data, Period	OE
1	Meltzer, 1960	US, man. industry, 1955-1957	Yes
2	Brencling & Lipsey, 1963	UK, Model + 75 firms, 1950-1959	Yes
3	Nadiri, 1969	US, man. industry, 1949-1964	No
4	Laffer, 1970	US, Model +1946-1966	Yes
5	Scwartz, 1974	US, Model	Yes
6	Herbst, 1974	US, one industry, 1956-1966	Partly
7	Walker, 1985	US, 27 firms, survey	No
8	Chiplin & Wright, 1985	UK, 915 firms, 1970-1977	No
9	Petersen & Rajan, 1997	US, 3404 SME 1987	Partly
10	Marotta, 1997	Italy, 1982-1993	Partly
11	Kohler et al., 2000	UK, 2000 firms, 1983-1995	Yes
12	Nilsen, 2002	US, aggregate + man. firms, 1959-1992	Yes
13	De Blasio, 2003	Italy, 3862 man. firms, 1982-1999	Yes
14	Mateut & Mizen, 2003	UK, 16,000 man. firms, 1990-1999	Yes
15	Atanasova & Wilson, 2003	UK, 639 firms, 1989-1999	Yes
16	Choi & Kim, 2003	US, 1975-1997 659 S&P + 689 non-S&P firms	Yes
17	Fisman & Love, 2003	43 countries, 37 industries	Yes
18	Mateut at al., 2006	UK, Model +16,000 man. firms, 1990-1999	Yes
19	Guariglia & Mateut, 2006	UK, 609 man. firms, 1980-2000	Yes
20	Özlü & Yalçın, 2012	Turkey, 5,655 man. firms, 1996-2000	Yes
21	Dai & Yang, 2015	China, 1880 firms, 2003-2012.	Yes
22	Gama & Auken, 2015	Portuguese, 468 SME, 1998-2006	Yes
23	Carbo-Valverde at al., 2017	Spain, Model +, 13,364 firms, 1998-2009	Yes
24	Hong, 2017	UK, 700,591 SME, 1991-2010	Yes
25	Saiz et al., 2017	12 Eurozone countries, 45,864 firms, 2005-2012	Yes

Note. Among the 25 examined TCC studies, 19 studies provided supporting evidence for Meltzer's (1960) TCC theory, whereas three of them confirmed partially and three ones found no evidence for trade credit offsetting effect. Nine of these studies were carried out in the US, eight in the UK, six in the other nations' markets (one in Turkey) and two in the cross-country samples. OE implies the existence of offsetting effect of trade credit on bank credit channel or more generally on restrictive monetary policy. Man. denotes manufacturing. In the third column Model refers to hypothetical model.

Methodology

Some researchers demonstrated that bank-lending channel and trade credit channel worked through restricted bank loan supply to bank-dependent (credit or financially constrained) firms. Since size is inversely related to the information opaqueness and inability to access credit market, small size firms are more likely to be accepted as bank-dependent firms (Gama & Auken; 2015: 892-893; Guariglia & Mateut, 2006: 2836). There has been consensus on that tight-money periods further enhances large and small firm diversity (Bernanke & Gertler, 1995: 39). Therefore, testing TCC through trade credit usage of small firms when monetary policy tightens (Nilsen, 2002: 228) seems appropriate. Following the previous literature, we tested TCC by examining two relationships, firstly between short-term interest rates and trade credit usages and the second relationship is among bank loan and trade credit usages. TCR and NTC, proxy for trade credit usages and constitute dependent variables. We regressed the trade credit variables on lags of themselves, financial variables (monetary policy indicators and bank loan) and other trade credit control variables. The dynamic panel model was applied on the 2 sector-level panel data set, derived from the CBRT databases, 7 sectors (main sectors sample) and 12 manufacturing sub-sectors (sub manufacturing sample), in Turkey over the period of 2008-2016. This section explains sample and variable set constructions, provides rationales for the methodology followed and presents the analysis results.

Sample

Industry level aggregated figures of financial tables and ratios were obtained from real sector statistics of CBRT we think as the only available and most comprehensive data source for particularly non-publicly held companies in Turkey. Since 1992, CBRT have been reporting company accounts and selected financial ratios of real sector firms set for three-year periods, aggregated based on main sector, sub-sector, scale and years. From the reporting year 2011 (including 2008-2009-2010 years), a different economic activity classification for sectors have been adopted. In CBRT database, the size classes are determined by "net sales" and "asset size" criteria[2]. Adhering with the data methodology of CBRT statistics, to ensure the data continuity, periods during which a different economic activity classification had adopted were excluded, and sectors for which aggregated data for size groups are not available were dropped. In this way, we identified

7 main sectors and 12 sub-manufacturing sectors that provide periodic data from 2008 to 2016. The analysis based on the CBRT company accounts 2011 (covering 2008-2009-2010), CBRT company accounts 2014 (covering 2011-2012-2013) and CBRT company accounts 2017 (covering 2014-2015-2016). Taking averages of the number of companies included company accounts release periods, we developed two tables shown in Appendix to clarify the sample structure. Concerning the number of companies, selected sector sets comprise to approximately 85% of all CBRT company accounts. Samples made up of relatively small and non-quoted firms. Main sector sample consists of small firms accounted for 49% and medium firms made up 35%, publicly-held company percentage[3] is %2. Similar size pattern exists for manufacturing sub-samples with small and medium sized firm that accounted for about 82 percent (www. tcmb.gov.tr).

Variable Construction

Almost all variables are in ratio formation[4] and scaled by total assets as done in Petersen and Rajan (1997), Dai and Yang (2015), Hong (2017) and Gama & Auken (2015). Ratios directly taken from the CBRT ratio reports reflect aggregated accounts rather than arithmetic mean of the individual ratios. Remaining ratios were calculated from related balance sheet and income statement accounts. Table 2 shows the names, abbreviations and calculations for variables of the analysis.

Table 2: Analysis Variables

Variable	Abbreviation	Formula
Trade Credit Received	TCR	(Accounts Payable/Total Assets) (%)
Net Trade Credit	NTC	(Accounts Payable- Accounts Receivables)/Assets %
Lagged Trade Credit Received	LTCR	One lagged value of TCR (%)
Lagged Net Trade Credit	LNTC	One lagged value of NTC (%)
Growth in Purchases	GROWP	$[(COGS_t - COGS_{t-1})/COGS_{t-1}]$ (%)
Inventory Level	STOCK	Inventories/Total Assets (%)
Internal Reserves	INTRES	Reserves from Retained Earnings / Total Assets (%)
Bank Loan	BANK	Bank Loans / Total Assets (%)
Policy Interest Rates	POIR	Average Annual Policy Interest Rates (%)
Overnight Interest Rates	ONIR	Average Annual Overnight Interest Rates (%)

Note. COGS imply cost of goods sold.

Trade credit variables

Trade credit is a comprehensive term used to express delayed payments for purchases among firms[5]. Dependent variables of this analysis are TCR and NTC. TCR is determined as the accounts payable as a percentage of total assets of industries. This ratio may represent the firms' demand for trade credit funds as in Petersen and Rajan (1997) and Hong (2017) or the credit extended to the firm by its suppliers (if there is any supplier information) (Petersen & Rajan, 1997: 667-668). Fisman and Love (2003), Nilsen (2002), Atanasova and Wilson (2003) use TCR to test conventional bank lending channel and trade credit channel too. The difference between TCR and TCE as a percentage of assets equals to NTC. Positive NTC means that the sector is a net trade credit receiver and negative NTC reflects net trade credit extender. NTC produce insight about joint or complete effect of TCR and TCE to ease the monetary pressure and verify findings[6]. Meltzer (1960), Brechling and Lipsey (1963), De Blasio (2003), Petersen and Rajan (1997), Kohler et al. (2000), Guariglia and Mateut (2006), Carbo-Valverde et al. (2017) define and measure net trade credit position.

Financial variables

Monetary policy indicator (MPI) and bank loan are classified as financial or monetary variables. However, in the empirical analyses determining a suitable MPI may be a whopping difficulty (Chiplin & Wright, 1985: 225) and different indicators proxy tighter monetary policy in the recent literature (De Blasio, 2003: 89), many researchers have used interest rate or some weighted measure of it as MPI (Nadiri, 1969: 413). Reilly and Sarte (2010) showed that for most interest rate time series over 1991 and 2009, common movements in interest rates highly reflected the monetary policy settings determined by federal funds rate. Many paper examined have utilized several type of short-term money market rates[7] directly or combining them with other monetary policy factors in a monetary policy stance index and some of them rely on predetermined tight dates[8]. Changing short-term interest rates as the monetary policy tool have been growing in importance. During transition year of 2005 and the following implementation periods, short-term interest rates have been used as a main policy instrument in Turkey (www.tcmb.gov.tr; CBRT 2005 annual report, 2006: 73-74; Eroğlu, 2009: 25; Aklan & Nargeleçekenler, 2008; 111). In this paper following the literature and CBRT monetary policy regimes in recent years[9], two short term interest rates, as policy rates and overnight rates,

were used in order to proxy tight and loose money periods. First MPI is the annual average of lending policy rates set by the CBRT as the one-week repo rate (www.tcmb.gov.tr). Annual average of lending overnight interest rates that have a 97.8% correlation[10] with policy rate is alternative MPI to verify robustness. These two MPI entered in the regressions with one period lag[11].

The ratio of bank loan over assets (BANK) is second financial variable as proxy for reduced bank loan accessibility during credit constrictions. A negative relationship between bank loan and TCR or NTC, especially for small firms, indicate that substitution hypothesis hold or trade credit channel operate. The bank loan scaled by either assets or short or long-term liabilities have been analysed in literature[12]. Dai & Yang (2015) use bank ratio to test financing substitution theory (or alternative financing theory) stating that lower bank loan raise trade credits. Guariglia & Mateut (2006) define bank loan ratio to represent bank loan dependency (level of being financially constrained). In the analysis of Gama & Auken (2015) trade credit to total debt ratio represents the bank credit or trade credit dependency. Mateut et al., (2006) and Saiz et al. (2017) expect a decline in bank loan while expecting an increase in trade credit received during tight periods, focusing on stronger effect for small firms that lack bank credit access. In consistent with the intuition of substitution hypothesis, bank loan that can be substituted with trade credit, is an influencing factor for gross or net trade credit.

Control variables

To capture the non-financial (non-monetary) factors that have ability to influence trade credit, dependent variables are regressed on the ratio of lagged trade credit variable (LTCR and LNTC), inventories (STOCK), retained earnings reserves (INTRES) as a proportion to total assets and on purchases growth rate (GROWP).

As in some other researches (Chiplin & Wright, 1985; Kohler et al., 2000), lagged ratio of trade credit (gross or net) is added to the regressions. The relationship between TCR (NTC) and LTCR (LNTC) intends to control for the industry specific intensity in trade credit usage. Trade credit ratio indicates the degree of reliance on trade credit by showing the asset proportion financed by trade credit (Gama & Auken, 2015: 892). Industries with intensive trade credit utilization are more subject to the consequences of bank loan constraints (Chiplin & Wright, 1985: 225). The higher the usage of trade credit historically in an industry, the stronger the positive relationship will be for that variables. As done in Chiplin and Wright (1985), the high correlations between TCR (NTC) and its one-period lagged amounts[13] prove that lagged values has an explanatory power but are not just critical factors.

As one of the fundamental financial relationship, TCR (TCE) is the product of sales (purchases) and credit period (Brechling & Lipsey, 1963: 620-622; Herbst, 1974: 379). Sales increases raise the firms' demand for credit (Petersen & Rajan, 1997: 683), thus, sales level (Nilsen, 2002; Kohler et al., 2000) or growth in sales level (Saiz et al., 2017) may explain the trade credit motive. From the borrower side, higher growing firms will need more TCR to finance growth (Marotta, 1997; Atanasova & Wilson, 2003; Gama & Auken, 2015; Saiz et al., 2017) and from the supplier side, growth opportunities will be more attractive to vendor financing (Dai & Yang, 2015). Growth rate in purchases of industries between two consecutive years serve as control variable to reflect this relationship. Based on the similar assumptions made by mentioned researchers, we expect *GROWP* will motive TCR.

STOCK variable is the ratio among inventory over total assets. Firms with high current assets demand significantly more trade credit (Petersen & Rajan, 1997: 684). The more a firm carry inventories, the greater are the requirements of short-term funds and trade credit. In addition, while inventory provides relatively liquid collateral in bankruptcy, greater inventory stock will attract suppliers to grant trade credit (Guariglia & Mateut, 2006: 2838; Saiz et al., 2017: 44-45). Therefore, stock level proportioned to assets (Saiz et al., 2017) or sales (Nilsen, 2002) is one of the trade credit determinants that is expected to relate dependent variables positively.

INTRES, the ratio of reserves from retained earnings to total assets is also included in the analysis to control for internal funds availability of industries. As an indicator of the strength to produce funds internally, literature[14] has used different ratio formations in which cash flows or profits became numerator. Meltzer (1960) states that firms generate finance from their liquid assets as a response to a tight monetary policy (Nadiri, 1969: 413). Petersen and Rajan (1997), Dai and Yang (2015), Gama and Auken (2015) and Saiz et al. (2017) expect a negative relation between internal reserves and TCR in consistency with the pecking order theory. Pecking order theory suggest that low-cost and low risk internal funds should take

the first order in the hierarchy of financing sources therefore firms with strong internal reserve (or having high liquidity) demand less trade credit (Kohler et al., 2000). By expecting the negative relation, Gama & Auken (2015) focus on the greater funds needs of liquidity-constrained firms. However, according to the Nilsen (2002) who relate trade credit to cash position positively, cash-rich but credit-constrained firms that are obligated to hold cash by precautionary needs, react to inventory shock with increasing trade credit usage. Kohler et al. (2000) also argue that the liquidity may attract more trade credits. Following the pecking order view, we predict that lower internal reserves may raise the usage of trade credit.

Related category, brief descriptions of all variables and their expected relationship with dependent variables are summarized in Table 3.

Table 3: Variable Groups and Expected Signs

Variable Group	Variable	Brief Description	Expected Sign
Dependent Variables	TCR	Gross trade credit	
	NTC	Net trade credit	
Financial Variables	BANK	Bank credit accessibility	-
	POIR	Monetary policy indicator	+
	ONIR	Monetary policy indicator	+
Control Variables	LTCR	Sectoral reliance on TCR	+
	LNTC	Sectoral reliance on NTC	+
	GROWP	Growth in purchases	+
	STOCK	Stock level to finance	+
	INTRES	Internal funds availability	-

Method

Since one period lagged dependent variable enter in the regressions as one of the regressors, analysis relationship characterized as dynamic and autoregressive model can be illustrated as in the first equation (Baltagi, 2005: 135; Tataoğlu, 2012: 65-66). When we apply this simple dynamic/autoregressive model to our variables, in broad terms, 2nd equation for TCR and 3rd one for NTC are set. 4th and 5th equations expanded with explanatory variables and replicated with a different money market interest rates, are derived for the first dependent variable, TCR. After repeating the same process for the second dependent variable NTC, the equations take the forms as seen in 6th and 7th equations.

$$Y_{it} = \delta Y_{it-1} + \beta X'_{it} + v_{it}, \text{ (X'= vector of regressors, } v_{it} = u_i + u_{it}), \quad (1)$$

$$TCR_{it} = \delta TCR_{it-1} + \beta X_{it} + v_{it} \quad (2)$$

$$NTC_{it} = \delta NTC_{it-1} + \beta X_{it} + v_{it} \quad (3)$$

$$TCR_{it} = \delta TCR_{it-1} + \beta(LTCR) + \beta(GROWP) + \beta(STOCK) + \beta(INTRES) + \beta(BANK) + \beta(POIR) + v_{it}, \quad (4)$$

$$TCR_{it} = \delta TCR_{it-1} + \beta(LTCR) + \beta(GROWP) + \beta(STOCK) + \beta(INTRES) + \beta(BANK) + \beta(ONIR) + v_{it}, \quad (5)$$

$$NTC_{it} = \delta NTC_{it-1} + \beta(LNTC) + \beta(GROWP) + \beta(STOCK) + \beta(INTRES) + \beta(BANK) + \beta(POIR) + v_{it}, \quad (6)$$

$$NTC_{it} = \delta NTC_{it-1} + \beta(LNTC) + \beta(GROWP) + \beta_0 (STOCK) + \beta(INTRES) + \beta(BANK) + \beta(ONIR) + v_{it}, \quad (7)$$

As indicator of bank-dependence criteria, CBTR small, medium and large size classifications were used. The last four regression equations (4th to 7th equations) were estimated separately for small, medium and large size groups. Indeed, this process created 3 sub-sample split by size, generating 6 samples. To avoid misunderstandings, main samples will be referred by adding related size sub-sample, e.g. main sector-small (or only main-small) or manufacturing sub sectors-medium (or only manufacturing-medium). 6 regressions (3 for TCR and remaining 3 for NTC) for main sectors sample (Table 5 and Table 6) and another 6 regressions for manufacturing sub-sectors sample (Table 7 and Table 8) were run. In the above regression equations, Y_{it-1} correlated with error term that violates strict exogeneity principle is the most basic problem with autoregressive panel. Two estimators that suggest to use instrumental variables to solve this autocorrelation problem are standard (or first-difference) generalized method of moments (GMM) and system generalized method of moments (system GMM), and they differ in the type of valid instruments used and in the way of determining them. GMM renders greater data loss particularly on un-balanced panel data set whereas system GMM minimizes the data loss. On the other hand, system GMM generates efficient estimates with observations having shorter time (T) compared to unit (N) (Tataoğlu, 2012: 65-104; Baltagi, 2005:135-148). Standard GMM estimates main-sectors sample with longer T (T=9, N=7) and system GMM fits for the manufacturing sub-sectors sample with small T (T=9, N=12)[15].

Findings and Discussions

This section begins with descriptive statistics of variables (Table 4 and Table 5) and continues with estimation results from GMM regressions[16] (Table 6, Table 7, Table 8 and Table 9) and discussions of findings.

Table 4: Descriptive Statistics of Main Sectors Sample

Size		TCR	NTC	GROWP	STOCK	INTRES	BANK	PR	OR
	Obs.	63	63	56	63	63	63	63	63
Small	Mean	15.08	-0.18	0.11	13.13	4.22	35.39	9.85	12.41
	Std.dev.	8.09	8.54	0.36	9.20	2.84	11.62	5.33	4.42
	Min	3.44	-10.08	-0.78	0.19	0.53	7.58	4.75	7.6
	Max	54.13	45.55	1.17	31.41	10.78	64.270	20.13	21.13
Medium	Mean	15.27	-2.93	8.37	11.68	4.96	32.85	9.85	12.41
	Std.dev.	5.28	6.25	20.42	8.69	3.11	13.31	5.33	4.42
	Min	6.45	-14.61	-39.18	1.46	0.57	4.95	4.75	7.6
	Max	29.31	11.73	83.23	26.33	13.18	67.050	20.13	21.13
Large	Mean	15.51	-3.01	15.33	11.04	7.45	24.07	9.85	12.41
	Std.dev.	8.64	6.08	29.64	8.56	4.26	13.38	5.33	4.42
	Min	4.82	-18.34	-36.31	1.48	0.29	2.12	4.75	7.6
	Max	39.73	11.69	152.51	26.72	17.79	65.600	20.13	21.13

Table 5: Descriptive Statistics of Manufacturing Sub-Sectors Sample

Size		TCR	NTC	GROWP	STOCK	INTRES	BANK	PR	OR
	Obs.	108	108	96	108	108	108	108	108
Small	Mean	15.69	-6.02	6.77	19.57	5.18	32.24	9.85	12.41
	Std.dev.	3.86	3.58	19.22	4.78	4.16	10.06	5.31	4.41
	Min	6.1	-14.91	-30.4	10.02	1.31	14.39	4.75	7.6
	Max	23.93	3.17	105.36	34.74	25.79	69.970	20.13	21.13
Medium	Mean	17.28	-8.07	7.42	20.00	6.19	30.37	9.85	12.41
	Std.dev.	3.65	4.27	16.62	4.28	2.23	6.98	5.31	4.41
	Min	11.25	-17.17	-29.72	12.78	2.74	16.85	4.75	7.6
	Max	28.46	2.52	94.82	31.16	11.18	49.550	20.13	21.13
Large	Mean	16.15	-7.56	6.77	18.92	11.27	28.04	9.85	12.41
	Std.dev.	5.63	5.82	19.22	5.64	3.85	8.51	5.31	4.41
	Min	7.52	-20.79	-30.4	9.52	4.12	9.05	4.75	7.6
	Max	32.58	7.33	105.36	44.36	20.93	52.140	20.13	21.13

In Table 4, means of TCR indicate that trade credit utilization into asset financing is approximately 15% for all size but deviate largely for small firms. Negative NTCs means that during analysis period, 7 different sectors extend more trade credit on average with stronger mean value for medium and large groups but with greatest standard deviation for small group. Means and standard deviations for the remaining variables vary between size groups, as may be expected. Main-small sample's growth performance and internal fund capacity is lower while investment need in inventory and reliance on bank loan are greater than medium and large size averages. Looking at Table 4 and Table 5 together as well as the separate evaluation reveals the differences in trade credit stance between a main-sectors sample and manufacturing industries sample. As expected within the same industry, manufacturing sector has lower standard deviations for TCR, NTC, STOCK, INTRES and BANK variables. Medium and large size groups use more trade credit with TCR means of 17.3 and 16.2 respectively. Compared to main-sectors sample, manufacturing industry extend more trade credit, make greater inventory investment, hold more retained reserves while using slightly less bank loan. Small manufacturing firms can grow faster, at a rate of 6.8% on average than their main-sector counterparts, but with a substantially high standard deviation of 19.2%. In both samples, size inversely related to the bank dependency or in other words, small companies are likely to have the highest bank dependency.

Table 6: Estimation Results of 7 Main Sectors, Dependent Variable is TCR

Variable	Main-Small		Main-Medium		Main-Large	
LTCR	0.159*	0.176*	0.706**	0.661**	-0.046	0.089
	(2.20)	(2.01)	(2.89)	(3.32)	(-0.40)	(1.03)
GROWP	11.587***	11.474***	0.583***	0.057***	0.254*	0.024**
	(3.68)	(3.70)	(3.63)	(3.66)	(2.30)	(2.38)
STOCK	0.072	0.031	0.234	0.230	-0.007	0.013
	(0.33)	(0.17)	(1.07)	(1.06)	(-0.02)	(0.05)
INTRES	-1.639	-1.573	-0.092	-0.044	-0.822***	-0.808**
	(-1.75)	(-1.71)	(-0.49)	(-0.24)	(-3.59)	(-3.44)
BANK	-0.344***	-0.350***	-0.098	-0.100	-0.304	-0.252
	(-3.65)	(-3.60)	(-1.01)	(-0.97)	(-1.62)	(-1.54)
POIR	0.137		0.001		0.483	
	(0.80)		(0.02)		(1.25)	
ONIR		0.079		-0.041		0.436
		(0.44)		(-0.48)		(1.31)
Observations	49	49	49	49	49	49
Groups/Instruments	7/7	7/7	7/7	7/7	7/7	7/7
P_AR(2)	0.518	0.510	0.164	0.160	0.893	0.621
P_Hansen	0.126	0.079	0.891	0.789	0.243	0.178

Note. Table 6 presents statistics for 6 regressions applied difference GMM for the panel data of 7 main sectors. However, observations were derived beginning from 2008 to 2016 covering 9 years, growth rate calculation induce a year loss and after the GMM differencing procedure we are left an observation of 49. Size categories represent CBTR size classifications. The regressand is TCR. First regressors are the one-lagged values of TCR. First regressions of each size include policy interest rates while next ones use overnight interest rates. T- statistics corrected for robust standard errors are presented in parentheses under coefficients. *, ** and *** denotes 10%, 5% and 1% significance level. The figures in the lowest 4 row report probability statistics (p) for autocorrelation and endogeneity tests. Ho hypothesis (no autocorrelation) of Arellano-Bond tests (AR2) exhibit no autocorrelation problem for all estimations. Ho hypothesis of Hansen test (instrument variables or over identifying restrictions are valid) are accepted for all policy rate regressions but weakly rejected at 10% significance level for one of the overnight rate regressions.

In Table 6 that shows findings for selected 7 main sectors, the responses to tight money implications are significant for small firm group however, medium and large firm groups seem irrelevant to financial variables (short-term interest rates and bank loan), hence small firms seem more influenced by changes in money policy. TCR and BANK relate negatively at 1% confidence degree meaning lower bank loan availability raise TCR. Small firms receive more trade credit to substitute bank loan during tight periods that force financially constrained firms to resort an unfavourable trade credit alternative to overcome credit constraints. The reaction of small firms to a decrease in bank loan supply as taking more trade credit or being slower to make trade credit payables and smooth the adverse effects of credit channel fits the substitution hypothesis. Substitution hypothesis predicts an increase in TCR of credit constrained firms to substitute for bank loan after strict policy. The explanatory powers of GROWP control variable are strong for all main samples. On the TCR of main-large group, INTRES has significant effect too. LTCR is positive and significant at 5% or 10% level for main-small and main-medium subgroups. Combining these results, main small and medium samples tend to take more trade credit when they depend heavily on trade credit and grow rapidly; however, for main large group lower internal fund is also a significant trade credit determinant. Non-financial factors are likely to behave as predicted by literature.

Table 7: Estimation Results of 7 Main Sectors, Dependent Variable is NTC

Variable	Main-Small		Main-Medium		Main-Large	
LNTC	0.191	0.196	-0.691**	-0.698***	-0.144	-0.146
	(0.81)	(0.85)	(-3.28)	(-3.76)	(-0.30)	(-0.31)
GROWP	8.121*	7.843*	-0.027*	-0.343*	0.017	0.019
	(2.08)	(2.02)	(-2.10)	(-2.29)	(0.72)	(0.75)
STOCK	-0.369	-0.416	0.166	0.197	0.137	0.132
	(-1.17)	(-1.37)	(0.33)	(0.40)	(0.27)	(0.26)
INTRES	-1.328	-1.282	-0.631	-0.672*	-0.690	-0.683
	(-1.16)	(-1.16)	(-1.78)	(-1.99)	(-1.80)	(-1.82)
BANK	-0.136	-0.150	-0.287*	-0.281*	0.200	0.199
	(-1.41)	(-1.50)	(-2.08)	(-2.08)	(0.94)	(0.93)
POIR	0.077		0.332*		0.035	
	((0.41)		(2.25)		(0.07)	
ONIR		-0.029		0.465*		0.137
		(-0.15)		(2.19)		(0.26)
Observations	49	49	49	49	49	49
Groups/Instruments	7/7	7/7	7/7	7/7	7/7	7/7
P_AR(2)	0.972	0.964	0.235	0.211	0.690	0.703
P_Hansen	0.603	0.704	0.296	0.262	0.624	0.619

Note. Table 7 presents the coefficients of difference GMM estimations for each size group of 7 main-sectors sample. The dependent variable is NTC. Arellano-Bond tests and Hansen test reject the presence of autocorrelation and endogeneity problems. Robust t- statistics are presented in parentheses under coefficients. *, ** and *** represents 10%, 5% and 1% significance level.

When we replaced NTC as dependent variable, we gather additional information to test existence of TCC. NTC regressions on 7 main-sectors sample presented in Table 7 indicate that main-medium group tends to be net trade credit user, meaning that even taking into account for trade credit given they still keep being a trade credit receiver. Since negative impact of BANK and positive impact of POIR and ONIR should mean net receiver, according to both negative significant (10%) sign of bank loan financial variable and positive significant (10%) sign of short-term interest rates, only main-medium is accepted as net receiver. In NTC replications, apart from bank variable, further support is given by short-term interest rates for main sectors-medium group at 10% importance level. Short-term interest rates provide confirmation of medium firms being a trade credit receiver implying that they request greater trade credit than they offer during monetary contractions.

Table 8 shows the statistics of system GMM regressions run for our second sample covering 12 sub-sector aggregations of only manufacturing industry. Consequence derived from this sample is similar with results found in the main-sectors sample but provide stronger relationship for manufacturing-medium group. Small and medium manufacturing firms experienced a significant rise in TCR during tight money periods and BANK (1%, 5% and 10%) variable confirms this finding. Large manufacturing sub-sector again does not react to monetary and credit restrictions by changing their gross trade credit usage. Signs of the control variables indicators are similar to the regressions of main sector sample.

Table 8: Estimation Results of 12 Manufacturing Sub-Sectors, Dependent Variable is TCR

Variable	Manufacturing-Small		Manufacturing-Medium		Manufacturing-Large	
LTCR	0.424**	0.412*	1.140***	1.104***	1.059***	1.073***
	(2.22)	(2.02)	(9.90)	(9.65)	(3.39)	(3.48)
GROWP	0.034***	0.038***	0.055***	0.051***	0.013	0.01
	(3.21)	(4.12)	(4.35)	(3.98)	(0.67)	(0.38)
STOCK	0.427***	0.405***	-0.049	-0.013	-0.047	-0.05
	(3.76)	(4.01)	(-0.53)	(-0.14)	(-0.44)	(-0.53)
INTRES	-0.002	-0.019	-0.004	0.093*	-0.017	-0.007
	(-0.02)	(-0.30)	(-0.07	(1.98)	(-0.22)	(-0.16)
BANK	-0.033***	-0.044**	-0.059**	-0.045*	-0.001	-0.002
	(-3.10)	(-2.90)	(-2.45)	(-1.94)	(-0.03)	(-0.08)
POIR	0.143		-0.030		0.039	
	(1.41)		(-0.83)		(0.39)	
ONIR		0.198		-0.120*		0.011
		(1.38)		(-1.90)		(0.06)
Observations	96	96	96	96	96	96
Groups/Instruments	12/8	12/8	12/8	12/8	12/8	12/8
P_AR(2)	0.132	0.098	0.382	0.232	0.644	0.635
P_Hansen	0.397	0.319	0.335	0.341	0.397	0.428
P_Diff. in Hansen	0.128	0.111	0.293	0.353	0.174	0.204

Note. Table 8 exhibits statistics for 6 regressions applied system GMM estimator for 12 manufacturing sub-industries. Panel data consists of aggregated amounts for 12 manufacturing sub-industries for small, medium and large size firms separately over 2008 and 2016. Observation number of 96 equals 12 unit times 8 periods. Dependent variable is TCR. The figures in the last 5 row report probability statistics (p) for autocorrelation and endogeneity tests. Second level Arellano-Bond tests (AR2) reject autocorrelation problem. Hansen test confirms exogeneity of instruments. t-statistics corrected for robust standard errors are presented in parentheses under coefficients. *, ** and *** implies 10%, 5% and 1% confidence level.

Table 9: Estimation Results of 12 Manufacturing Sub-Sectors, Dependent Variable is NTC

Variable	Manufacturing-Small		Manufacturing-Medium		Manufacturing-Large	
LNTC	0.901***	0.796**	0.573***	0.601***	0.611	0.0617
	(3.75)	(2.98)	(5.43)	(6.41)	(0.94)	(0.90)
GROWP	0.004	-0.004	0.027	0.028	0.006	-0.002
	(0.07)	(-0.07)	(0.95)	(1.04)	(0.20)	(-0.07)
STOCK	-0.040	-0.019	0.085	0.085	0.058	0.068
	(-0.83)	(-0.33)	(1.05)	(1.13)	(0.60)	(0.59)
INTRES	-0.014	-0.000	-0.397***	-0.313***	-0.134	-0.101
	(-0.28)	(-0.00)	(-3.90)	(-3.19)	(-0.66)	(-0.52)
BANK	0.019	0.011	-0.115	-0.101	-0.069	-0.060
	(0.41)	(0.24)	(-1.73)	(1.58)	(-0.44)	(-0.35)
POIR	-0.049		0.031		-0.030	
	(-0.46)		(1.23)		(-0.31)	
ONIR		-0.108		-0.042		-0.081
		(-0.92)		(-1.06)		(-0.80)
Observations	96	96	96	96	96	96
Groups/Instruments	12/9	12/9	12/8	12/8	12/8	12/8
P_AR(2)	0.402	0.374	0.661	0.823	0.811	0.747
P_Hansen	0.167	0.196	0.202	0.265	0.088	0.084
P_Diff. in Hansen	0.088	0.107	0.111	0.137	0.131	0.109

Note. Table 9 presents the coefficients of system GMM estimator for 12 manufacturing sub-industries when the dependent variable is NTC. Ho hypothesis of autocorrelation and endogeneity tests are mostly accepted. Robust t-statistics are presented in parentheses under coefficients. *, ** and *** refers 10%, 5% and 1% significance level.

However, coefficients of bank variable in the SME manufacturing groups in the Table 8 support the substitution hypothesis, negative TCR reaction of manufacturing-medium group to the increases on the ONIR at %10 significance level mitigates the power of substitution hypothesis. Negative sign of ONIR means that medium manufacturing firms may receive less trade credit during hard times, and we can accept this behaviour as consistent with the redistribution or helper theory. Manufacturing-medium group assumed as less financially unconstrained compared to small ones, may help their small sized suppliers by receiving less trade credits from them or making early payments to them. Despite, when we combine this weak evidence of redistribution theory at Table 8 (negative gross trade credit response of medium manufacturing group) with the net trade credit status at Table 9, we cannot provide further support for redistribution theory. The financial variables in the regressions examining the net effect of tightened periods on trade credit flowing within manufacturing industry, displayed in Table 9, lost their significances. Thus, we have no significant evidence to interpret whether manufacturing sample is net trade credit extender or net credit receiver.

Size and industry effects are obvious in the analysis. Small and medium sized samples appear to be more sensitive to monetary policy changes. In terms of gross ratios, small firm groups made more trade credit usage in both two samples, according to net terms, there is no significant evidence about their net trade credit status (receiver or extender). Medium sized group were sensitive to monetary shocks with gross figures in the manufacturing sample and engaged in trade credit flow by becoming net receiver in the main-sectors sample. Large firms group in all regressions seemed to be irrelevant. Analysis results across main sectors (multi-sectors or economy wide) and among manufacturing industry support substitution theory. Despite the trade credit user behaviour of main-medium group, net trade credit results of other sized main groups (main-small and main-large) and net trade credit findings for all manufacturing sample do not provide sufficient evidences to interpret redistribution hypothesis that can make clear the route to which trade credit flows.

Most regressions reveal that financing mix of small and medium firms shifted to trade credit from bank credit during tight periods in terms of both gross and net terms. Consequently, analysis findings support trade credit offsetting channel with the findings showing that SMEs tend to overcome credit constraints by using more trade credit (despite its disadvantages) from vendors. Trade credit can serve as a tool to dampen the effects of credit channel and to frustrate the workings of monetary policy.

Conclusion

The bank-lending channel explains the fall in the small firm's bank loan usage by the credit supply constraints rather than firms' reduced fund demand. Firms typically strive to avoid the expensive trade credit in normal times. But at times of tight policy or recessions, when lending channel begin to operate, and cut back the credits, especially bank-dependent small firms are forced to use trade credit as their most common non-bank debt alternative. (Nilsen, 2002: 228-234). Trade credit channel initiated by Meltzer (1960), suggests that wider usage of trade credit as a substitute for bank loans dampens the credit channel at tight monetary periods during which bank loan accessibility becomes more difficult for bank-dependent firms due to both severe credit rationing and flight to quality implications of financial institutions and deteriorated balance sheet structure of them. For financially constrained firms, trade credit act as a supplement source of fund to finance their activities, hence have a moderating effect on the credit channel, more broadly on monetary policy implications. Trade credit channel of monetary policy is more prone to emerge in informational opaque small firms that are in the great need of overcoming bank loan constraints with trade credit. To our knowledge, there has been no sufficient works investigating the reaction of Turkish real sector to monetary restrictions in both gross and net trade credit terms. This paper examined whether the trade credit channel weakened the bank credit channel for Turkish economy at sector level. To see whether gross and net trade credit variables have any response to tight monetary policy and bank loan constraint, dynamic panel models extrapolated data sets involving 7 sectors and 12 manufacturing sub-sectors in Turkey from 2008 to 2016.

Considering the reaction of trade credit received to strict monetary conditions, small and medium firms from main-sectors and manufacturing sub-sectors raised their trade credit usage when bank loan decreased. Bank loan level became the main determinant and short-term interest rates provide support for manufacturing-medium sample. Trade credit taken by large size firm groups appears irrelevant to monetary

contractions. If we base our assessment of substitution hypothesis on gross trade credit results, we can say substitution theory holds. Trade credit as the best substitute of bank loan for small firms alleviates credit channel effect of monetary policy that aims to restrict the bank loans available for particularly SMEs. According to net trade credit position, sector wide medium group (main-medium) acted as a net trade credit receiver that is supported both with the bank loan-trade credit relationships and with policy rates. Other findings of net trade credit status are not sufficient to support redistribution theory. Combining these results, we can conclude that trade credit channel operated and mitigated the effects of credit channel in Turkey during 2008-2016. Trade credit has helped credit-constrained small firms smooth the effects of credit channel implications.

Some more possible extensions of this paper are to analyse the bank-loan and monetary policy relationships, testing the trade credit channel by using different monetary policy indicators and if possible updating the analysis with firm level data.

REFERENCES

Aklan, N.A., & Nargeleçekenler, M. (2008). Para politikalarının banka kredi kanalı üzerindeki etkileri. İstanbul Üniversitesi *Siyasal Bilgiler Fakültesi Dergisi, (39),* 109-132.

Arslan, B.G., & Ergeç, E. H. (2011). Türkiye ekonomisinde para politikasının sektörel etkileri. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, *11* (1), 89-103.

Atanasova, C. V., & Wilson, N. (2003). Bank borrowing constraints and the demand for trade credit: Evidence from panel data. *Managerial and Decision Economics,* 24, 503–514.

Baltagi, B. H. (2005). *Econometric analysis of panel data* (3rd edition). Chichester: Wiley. ISBN: 13-978-0-470-01456-1.

Bernanke, B.S., & Gertler, M. (1995). Inside the black box: The credit channel of monetary policy transmission. *Journal of Economic Perspective, 9(4),* 27-48.

Brechling, F. P. R., & Lipsey, R. G. (1963). Trade credit and monetary policy. *The Economic Journal, 73(292),* 618-641.

Carbo-Valverde, S., Mansilla-Fernandez, J. M., & Rodriguez-Fernandez, F. (2017). Monetary policy, implicit interest rate, and relative net trade credit. *Revista De Economia Aplicada, 25(73),* 21-54.

Chiplin, B., & Wright, M. (1985). Inter-industry differences in the response of trade credit to changes in monetary policy. *Journal of Business Finance & Accounting*, *12(2),* 221-248.

Choi, W.G., & Kim, Y. (2003). Trade credit and the effects of macro-financial shocks: Evidence from US panel data. *International Monetary Fund Working Paper No. 03/127*.

Dai, B., & Yang, F. (2015). Monetary policy, accounting conservatism and trade credit. *China Journal of Accounting Research, 8,* 295-313.

De Blasio (2003). Does trade credit substitute for bank credit? Evidence from firm-level data. *International Monetary Fund Working Paper No. 03/166.*

Demirgunes, K. (2016).Determinants of access to trade credit for small and medium sized enterprises in Turkey after the 2008 global crisis. *Journal of Economics, Finance and Accounting, 4(3)*, 266-284.

Eroğlu, N. (2009). Türkiye'de 1990'lı ve 2000'li yıllarda para politikaları. Marmara Üniversitesi, İ.İ.B.F Dergisi, 26(1), 23-42.

Fisman, R., & Love, I. (2003). Trade credit, financial intermediary development and industry growth. *Journal of Finance 58,* 353–374.

Gaiotti, E., & Generale, A. (2002). Does monetary policy have asymmetric effects? A Look at the investment decision of Italian firms. *Giornale degli Economisti e Annali di Economia, 61 (1),* 29-59.

Gama, M., Paula, A., & Van Auken, H. (2015). The interdependence between trade credit and bank lending: Commitment in intermediary firm relationships. *Journal of Small Business Management, 53(4),* 886–904.

Gertler, M.& Gilchrist, S. (1993). The role of credit market imperfections in the monetary transmission mechanism: Arguments and evidence. *The Scandinavian Journal of Economics, (1),* 43-64.

Gertler, M.& Gilchrist, S. (1994). Monetary policy, business cycles, and the behaviour of small manufacturing firms. *The Quarterly Journal of Economics, (2),* 309-340.

Guariglia, A., & Mateut, S. (2006). Credit channel, trade credit channel, and inventory investment: Evidence from a panel of UK firms. *Journal of Banking & Finance, 30(10),* 2835–2856.

Herbst, F. (1974). Some empirical evidence on the determinants of trade credit at the industry level of aggregation, *Journal of Financial and Quantitative Analysis,* 377–394.

Hong, B.P. (2017). The impact of monetary policy on bank credit and trade credit for the UK's SMEs: A disequilibrium model of credit rationing. PhD Dissertation. The University of Leeds.

Kohler, M., Britton, E., & Yates, T. (2000). Trade credit and the monetary transmission mechanism. *Discussion Paper, Bank of England.*

Laffer, A. B. (1970). Trade credit and the money market. *Journal of Political Economy, (2),* 239-267.

Marotta (1997), Does trade credit redistribution thwart monetary policy? Evidence from Italy, *Applied Economics, 29(12),* 1619–1629.

Mateut, S., & Mizen, P. (2003). Trade credit and bank lending: An investigation into the determinants of UK manufacturing firms' access to trade credit. *EUI Working Paper ECO, No. 2003/3.*

Mateut, S., Bougheas, S., & Mizen, P. (2006). Trade credit, bank lending and monetary policy transmission. *European Economic Review*, *50 (3),* 603-629.

Meltzer, A.H. (1960). Mercantile credit, monetary policy, and size of firms. *The Review of Economics and Statistics, (4),* 429-437.

Nadiri, M. I. (1969). The determinants of trade credit in the U.S. Total manufacturing sector. *Econometrica, (3),* 408-423.

Ng, C.K., Smith, J.K & Smith, R.L. (1999). Evidence on the determinants of credit terms used in interfirm trade, *The Journal of Finance, 54 (3),* 1109-1129.

Nilsen, J. H. (2002). Trade credit and the bank lending channel. *Journal of Money, Credit and Banking, 34(1),* 226–253.

Petersen, M. A., & Rajan, R.G. (1997). Trade credit: Theories and evidence. *The Review of Financial Studies, (3),* 661-691.

Reilly, D., & Sarte, P. G. (2010). Changes in monetary policy and the variation in interest rate changes across credit markets. *Economic Quarterly (10697225), 96(2),* 201-229.

Romer, C.D. & Romer, D.H. (1993). Credit channel or credit actions? An Interpretation of the Postwar Transmission Mechanism. *NBER Working Paper No. 4485*.

Roodman, D. (2009a). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics 71(1)*, 135–58.

Roodman, D. (2009b). How to do Xtabond2: An Introduction to Difference and System GMM in Stata. *The Stata Journal, 9(1),* 86-136.

Saiz, C.M., Azofra, S.S., Olmo, B.T, & Gutierrez, C.L. (2017). Trade credit, sovereign risk and monetary policy in Europe. *International Review of Economics & Finance, 52,* 39-54.

Şahin A. (2018). Firmalar arası ticari alacak kredisini açıklayan teoriler: Borsa İstanbul'da bir uygulama, *Maliye ve Finans Yazıları, 110,* 37-70.

Schwartz, Robert A. (1974). An economic model of trade credit. *The Journal of Financial and Quantitative Analysis, (4),* 643-657.

Tatoğlu, F. Y. (2012). *İleri panel veri analizi (1. basım).* İstanbul: Beta Basım A.Ş. ISBN: 978-605-377-721-2.

Özlü, P. & Yalçın, C. (2012). The trade credit channel of monetary policy transmission: Evidence from non-financial manufacturing firms in Turkey. *Emerging Markets Finance & Trade, 48(4),* 102–117.

Walker, D. A. (1985). Trade credit supply for small businesses. *American Journal of Small Business*, 9(3), 30-40.

http://www.tcmb.gov.tr (04.04.2018 - 10.11.2018)

http://www3.tcmb.gov.tr (04.11.2018- 17.11.2018)

Appendix

Table A: The average number of companies in CBRT for 7 main sectors

Main Sectors	Code	Total	Small	Medium	Large
Manufacturing	C	3322	1548	1205	569
Electricity, gas, steam and air conditioning supply	D	252	163	44	45
Construction	F	936	438	317	181
Wholesale and retail trade, repair of motor vehicles*	G	2305	1099	888	319
Transportation and storage	H	360	211	107	42
Information and communication	J	88	39	29	20
Administrative and support service activities	N	244	167	57	20
Sample Total		7507	3664	2647	1196
Size/Sample %			49	35	16
Database Total		8804			
Sample/Database %		85.3			

Note. Sector name is "Wholesale and retail trade, repair of motor vehicles and motorcycles". Table contents were derived from the company accounts statistics of CBRT, from website of www.tcmb.gov.tr.

Table B: The average number of firms in CBRT for 12 manufacturing sub-sectors

Manufacturing Sub-Sectors	Code	Total	Small	Medium	Large
Manufacture of food products	C10	483	168	205	110
Manufacture of textiles	C13	435	195	187	53
Manufacture of wearing apparel	C14	262	140	100	22
Manufacture of paper and paper products	C17	90	34	38	18
Manufacture of chemicals and chemical products	C20	145	65	45	35
Manufacture of rubber and plastic products	C22	204	95	79	30
Manufacture of other non-metallic mineral products	C23	288	147	93	48
Manufacture of basic metals	C24	194	62	64	68
Manufacture of fabricated metal products*	C25	235	118	92	25
Manufacture of electrical equipment	C27	119	44	43	33
Manufacture of machinery and equipment n.e.c.	C28	206	121	68	17
Manufacture of motor vehicles, trailers and semi-trailers	C29	138	56	46	36
Sample Total		2799	1244	1060	494
Size/Sample %			44	38	18
Database Total		3322			
Sample/Database %		84.2			

Note. Sector name is "Manufacture of fabricated metal products, except machinery and equipment". All information was derived from the website of www.tcmb.gov.tr.

(Endnotes)

1 External finance premium is the cost premium paid to raise external funds over the opportunity cost of internal funds (Bernanke & Gertler, 1995: 28-35).

2 As of the reporting year, the data of those firms with continuous data for the preceding three years are involved. From 2008, sectors have been classified based on economic activity classification of NACE Rev.2, before then, NACE Rev.1.1 had been in use. However, net

sales are the classification criteria for sizing almost all sectors except Construction Sector and Holding Companies Sector for which assets reflect the true size. Net sales (assets) less than EUR 10 million is threshold for small companies. Net sales between EUR 10-50 million (assets EUR 10-43 million) determine medium size firms. Firms with net sales greater than EUR 50 million (assets 43 million) are large ones. In determining the number of companies to be included in the CBRT reports, ensuring the highest possible level of those firms having credit balance in banks and those involved in the GDP calculations in the last years is taken account (www. tcmb.gov.tr).

3 As stated in the two available evaluation report of sector balance sheets of CBRT (2011 and 2014 reports) an average of 178 firms traded in Borsa Istanbul (account for 2%), and 727 companies identified as among the first 1000 largest industrial organization by Istanbul Chamber of Industry (make up 8.3%) were covered (www. tcmb.gov.tr).

4 Ratio formation eliminates potential demand effects controlling size level (Mateut et al., 2006) and scaling by total assets provides control over activity level variations (Gama & Auken, 2015: 892). As scaling variable Nadiri (1969), Chiplin and Wright (1985) and Saiz et al., (2017) used sales or purchases, Mateut et al., (2006) utilized liabilities but expressed that they repeat analysis by asset scaled trade credit variable and obtained the same results.

5 In examined literature, accounts payable often has similar meaning with the phrases of trade payable, trade credit taken, received, demanded or requested. For those demanding trade credit; trade credit user, taker, receiver, borrower, demander, requestor or debtor phrases have generally been used. Following the same logic, accounts receivable generally means trade receivables, trade credit given, extended, made, offered, granted or supplied. Trade creditor, trade credit giver, supplier, extender, lender or vendor expresses those supplying trade credits.

6 NTC represent extend to which trade payable finance the trade receivables. Brechling and Lipsey (1963) express this effect as passing on effect. TCR is expressed as passing on part of TCE because the difference between gross credit given and net credit is passed on to other firm (Brechling & Lipsey, 1963: 636).

7 Overnight rates, minimum bank lending rates, official interest rates, rates on three or six-month Treasury Bill, Fed fund rates, bank prime rate, LIBOR, EURIBOR, CONSOL yield, bank interest rate, the rate bankers' acceptances, four-to-six-month corporate/commercial paper rates are among the short-term interest rates to proxy for MPI.

8 Among the papers examined Choi and Kim (2003), Mateut et al. (2006), Özlü and Yalçın (2012) used both indicators by using a money market rate and adding a dummy whether determined previously or determined by themselves based on the level of interest rates. Choi and Kim (2003) used dummy for U.S. restrictive periods identified respectively by Romer and Romer (1993). De Blasio (2003) utilized only a chronology generated by Gaiotti and Generale (2002) for Italy. On the other side, Meltzer (1960), Bernanke and Gertler (1995), Kohler et al. (2000), Aklan and Nargeleçekenler, (2008), Carbo-Valverde et al. (2017), Hong, (2017), Saiz et al. (2017) all used different type of short term money market interest rates as monetary policy indicator. Kohler et al. (2000) used crisis dummy and Saiz et al. (2017) used year dummy but not for tight periods, Hong, (2017) applied quantitative easing periods. Difficulties on clearly assigning years as tight or loose after 2008 and on finding a commonly used predetermined dates led us to focus on money market rates.

9 The Central Bank manages total demand and inflation expectations by using policy interest rates and other monetary policy instruments in order to meet inflation targets. Following the crisis in February 2001, CBRT started to launch inflation targeting regime, up to 2006 implementing a transition process under implicit inflation targeting and as the beginning of 2006 fully switching to the inflation targeting. During transition year of 2005 and the following implementation periods, short-term interest rates have been used as a main policy instrument. Since 2005, monetary policy decisions including policy rates are taken by the Monetary Policy Committee at pre-scheduled meetings and announced to the public (www.tcmb.gov.tr; CBRT 2005 annual report, 2006: 73-74; Eroğlu, 2009).

10 Short-term interest rates often move closely even though they belong to different markets (Eroğlu, 2009: 27).

11 Interest rates generally affect the other variables with a one period lag (Gertler & Gilchrist, 1993: 53).

12 Herbst, (1974), Marotta (1997), Guariglia and Mateut (2006), Mateut et al. (2006), *Özlü* and Yalçın (2012), Dai and Yang (2015), Gama and Auken (2015), Saiz et al. (2017).

13 Within the main sector sample, TCR and LTCR correlations for small, medium and large size groups are respectively 0.89, 0.84 and 0.89. NTC and LNTC correlations are 0.64, 0.84 and 0.64. For the manufacturing sub-sectors sample these are 0.82, 0.82 and 0.92 for TCR and 0.70, 0.78 and 0.87 for NTC.

14 Ratios among cash flows to total assets (Dai &Yang, 2015; Saiz et al., 2017; Carbo-Valverde et al., 2017), cash to sales ratio (Nilsen, 2002), profits before taxes (Özlü & Yalçın, 2012) and earnings before interest and taxes over total assets (Gama & Auken, 2015) have been used in papers.

15 Some authors used GMM methods in studies of trade credit channel (Guariglia & Mateut, 2006, Gama & Auken, 2015; Saiz et al., 2017) and in papers about credit channel existence (Aklan & Nargeleçekenler, 2008).

16 Regressions are run through Stata "xtabond2" commands for difference GMM and system GMM (Tataoğlu, 2012; Roodman, 2009a; Roodman 2009b).

EGE AKADEMİK BAKIŞ / *EGE ACADEMIC REVIEW*

Cilt 19 • Sayı 4 • Ekim 2019
SS. 457 / 467
Doi: 10.21121/eab.476531
Başvuru Tarihi: 30.10.2018 • Kabul Tarihi: 08.10.2019

Development of A Disaster Attitude Scale and Assessment of University Students' Attitudes Towards Disasters

Afet Tutum Ölçeği Geliştirilmesi ve Üniversite Öğrencilerinin Afetlere Karşı Tutumlarının Değerlendirilmesi

Ayça Hatice TÜRKAN[1] *https://orcid.org/0000-0002-4375-9733*
İbrahim KILIÇ[2] *https://orcid.org/0000-0003-0595-8771*
Murad TİRYAKİOĞLU[3] *https://orcid.org/0000-0002-1874-0678*

ABSTRACT

Humanity is exposed to numerous disasters all around the world. Both natural and man-made disasters might lead to very important economic and social consequences and might pose a significant threat to public health and its future. Therefore, the attitudes of individuals before, during and after a possible disaster are worth considering. This study aims to develop a scale to measure attitudes towards disasters by considering three dimensions of attitude: cognitive, affective and behavioural. A survey-based study was carried out on a sample of 787 undergraduates at Afyon Kocatepe University in Turkey. A five-point Likert scale including 23 items was used to quantify students' attitudes. The results obtained from the survey revealed a high level of affective attitudes and a low level of behavioural attitudes among the university students. Moreover, affective and behavioural attitudes differed based on the students' gender and their study programme. Attitudes towards disasters are not at the required level according to these measurements made at the local level. It is necessary to produce and implement policies in the scope of preparation for a possible crisis.

Keywords: Attitude scale, disaster, earthquake, public health.

JEL Codes: C10, C83, O13

ÖZET

İnsanoğlu, dünyanın dört bir yanında birçok afete maruz kalmaktadır. Hem doğal hem de insan kaynaklı afetler, önemli ekonomik ve sosyal sonuçlara yol açabilmekte; halk sağlığı ve geleceği için önemli bir tehdit oluşturabilmektedir. Bu yüzden bireylerin olası bir afet öncesi, sırası ve sonrası tutumları incelenmeye değer bir konudur. Bu çalışma, tutumun bilişsel, duyuşsal ve davranışsal olan üç boyutunu dikkate alarak, afetlere yönelik tutumları ölçmek için bir ölçek geliştirmeyi amaçlamaktadır. Afyon Kocatepe Üniversitesi'nde 787 lisans öğrencisi üzerinde ankete dayalı bir çalışma yapılmıştır. Öğrencilerin tutumlarını ölçmek için 23 maddeden oluşan beşli Likert ölçeği kullanılmıştır. Araştırmadan elde edilen sonuçlar, üniversite öğrencileri arasında yüksek düzeyde duyuşsal tutum ve düşük düzeyde davranışsal tutum olduğunu ortaya koymuştur. Ayrıca, duyuşsal ve davranışsal tutumlar, öğrencilerin cinsiyetleri ve çalışma programlarına göre farklılaşmaktadır. Afetlere yönelik tutumlar, yerel düzeyde yapılan bu ölçümlere göre arzu edilen düzeyde değildir. Olası bir kriz için hazırlık kapsamında politikalar üretmek ve uygulamak gerekmektedir.

Anahtar Kelimeler: Tutum ölçeği, afet, deprem, halk sağlığı.

JEL Kodları: C10, C83, O13

[1]**Asst. Prof., Afyon Kocatepe University, Faculty of Science and Literature, Department of Statistics, Afyonkarahisar, aturkan@aku.edu.tr**
[2]**Assoc. Prof., Afyon Kocatepe University, Faculty of Veterinary Medicine, Department of Biostatistics, Afyonkarahisar, kilicibrahim@hotmail.com**
[3]**Assoc. Prof., Afyon Kocatepe University, Faculty of Economics and Administrative Sciences, Department of Economics, Afyonkarahisar, murad@tiryakioglu.org**

INTRODUCTION

Natural and man-made disasters result in significant economic and social costs throughout the world. Moreover, they affect directly public health and increase fragility depending on their severity. Together with global climate change and industrial accidents, natural disasters threaten particularly women, disabled people, immigrants, minorities, children, the elderly and youth. As a natural result of this impact, the severity of physical and economic damage caused by natural disasters is higher in less-developed countries than in developed countries. Rottach (n.d.) underlined that poverty is both the cause and the result of disaster vulnerability for both nations and individual households. Therefore, the first step should be taken on the individual level in order to foster disaster awareness at the national level. However, the critical point to be considered is that poor people and underdeveloped countries have to focus on meeting the basic needs before they take measures in disaster awareness, the benefits of which would be seen in the long term.

107 million people were affected by a total of 317 natural disasters in 2014 (IFRC, 2015). 1,753 floods, 223 droughts, 1,254 earthquakes, 988 storms and 236 high-temperature events occured between 2005–2014, affecting billions of people and causing billions of dollars in economic damage. When examining the geographical distribution of the economic and social impacts of the disasters, one can observe that the less-developed and developing countries are more likely to be affected by natural disasters, and the normalisation of both economic and social activities takes a much longer time with higher costs than developed countries.

Especially being highly fragile to natural disasters, Turkey has experienced numerous disaster experiences. The damage to the environment depending on the growing industrialization increases the frequency and severity of man-made disasters in Turkey. Erzincan earthquake occurred in 1939 was a major disaster in the history of the Republic of Turkey. It caused more than 32,000 people to lose their lives according to official records. The Marmara (Kocaeli) earthquake killed nearly 17,000 people and led to an economic loss of $12-$19 billion, according to official records. In recent history, the Van and Simav earthquakes once again revealed the high disaster risk of Turkey, affecting significant numbers of people. Recent experiences have shown that Turkey has developed its capacity and capability for disaster emergency response and search and rescue since the 1999 Marmara earthquake. Nonetheless, disaster awareness remains inadequate among the Turkish people, except among those who have experienced disasters or whose relatives have suffered disasters.

The common attitude of all communities after the disaster is that awareness reaches the highest level immediately after the disaster but after a while this level of awareness decreases. Likewise, disaster awareness in Turkey is very high in the post-disaster period, but after a certain period of awareness, the level of awareness decreases. The underlying reason for this is that the mechanisms to keep the level of awareness at a certain level and the disaster culture to ensure the sustainability of these mechanisms are not sufficiently developed. This finding suggests that the level of individual and social preparedness and awareness must develop together.

In countries with high vulnerability to disasters such as Turkey, high levels of knowledge, awareness and preparedness of individuals about disasters ensure relatively less impact of disasters. Pre-and post-disaster behaviours and attitudes vary depending on the capacity of preparing for disasters, and the educational activities in the relevant field. In this context, it is important to examine in depth the concept of attitude. In the literature, there are numerous definitions of the concept of attitude. One of the definitions of attitude is that it comprises feelings, thoughts and behaviours towards something (Rosenberg & Hovland, 1960; Smith, 1968; Zanna & Rempel, 1988; Olson & Zanna, 1993). These components are known as the ABCs of attitudes (affective, behavioural and cognitive). The affective component refers to feelings and emotions, the behavioural component to behaviours and the cognitive component to thoughts and beliefs.

Previous studies generally handled the disasters together with risk perception (Lindell & Perry, 2003; Baytiyeh & Öcal, 2016), preparedness (Dooley et al., 1992; Junn & Guerin, 1996; Tierney et al., 2001) and attitudes (Honda et al., 2014; Ahayalimudin & Osman, 2016). Examining the studies about the attitudes of the individuals related to disasters, a limited literature is discovered. Qi et al. (2003) designed a questionnaire that measures disaster perception from three aspects: disaster knowledge, disaster attitude and responding behaviour. Honda et al. (2014) developed a six-point scale to assess individuals' multidimensional attitudes about the Fukushima nuclear accident. Ahayalimudin

and Osman (2016) explored emergency medical personnel's knowledge, attitude and practice towards disaster management using a questionnaire. Tavan et al. (2016) developed a questionnaire on the nurses' knowledge, attitude and practice of disaster preparedness. Bhat et al. (2017) designed a questionnaire to assess the level of awareness, attitudes and practices about the disaster preparedness among college students in district Ganderbal of Kashmir Valley. Pekez-Pavliško et al. (2018) designed a questionnaire including items related to attitudes (risk awareness, response to disaster), self-assessed preparedness and previous experience in order to respond or assist in mass casualty incidents in Croatia, and conducted to family physicians. Certainly, there are studies in which affective, cognitive and behavioural responses to disasters are examined in other theoretical frameworks or in the context of other constructs (Terpstra, 2011; Grimm et al., 2014; Lindholm et al., 2015). However, this study was mainly aimed at developing a scale based on the ABCs of attitudes to measure the attitudes towards disasters in Turkey using a sample of students who were considered to be representatives of future policy makers and the next generation. In the scale, the attitudes of the individuals were evaluated by considering stages of the disaster: before, during and after a disaster. Before a possible disaster, the proper attitudes had been considered regarding to the measures to prevent or minimize the damages caused by disaster. During a disaster, attitudes towards dealing with a disaster in an appropriate manner and after a disaster attitudes regarding the actions to be taken were considered. The developed scale was applied to students at Afyon Kocatepe University in Afyonkarahisar province where the earthquake risk is high.

The paper is organised as follows. First, the scale's construct validity and reliability were assessed. Second, the dimensions of the scale were compared. Third, the effects of gender and study programme on each dimension were examined. Finally, the disaster attitude level of undergraduates was determined.

METHOD

PARTICIPANTS

The participants of the study comprised 787 undergraduates pursuing different programmes at Afyon Kocatepe University in Turkey. In this framework, stratified sampling was used and the students' study programmes (natural sciences, health sciences and social sciences) were defined as strata. The sample size was calculated by the formula, $n = s2Z2/d2$, as recommended by the NEA (1965) for quantitative research, assuming an infinite universe ($N > 10000$). Depending on the formula, the minimum acceptable sample size was found to be 650 ($s = 1.30$, $Z = 1.96$, $d = 0.10$; s: standard deviation, Z: critical or theoretical table value for a significance level of 0.05, d: effect size). In the determination of the unknown values in the formula, the results of the pilot study of 35 participants were used. After the pilot study, the questionnaires were distributed to 820 students. From the collected questionnaires, 33 of them were incomplete and were therefore excluded, leaving 787 questionnaires for the analyzes given in Data Analysis section.

INSTRUMENT

The questionnaire used in this study was in Turkish language, including three subscales to measure the students' attitudes towards disasters and their responses to additional demographic questions (gender and undergraduate study programmes). The scale items (i.e. statements) were written according to the structure of the attitude scale (Edwards, 1983). The affective dimension comprised items measuring fear, anxiety and worry of individuals in the event of a possible disaster. The behavioural dimension comprised items related to behavioural responses to a possible disaster and disaster measures. The cognitive dimension contained items that express ideas of individuals and also included some basic knowledge.

In the development of our scale, Delphi method was utilised by which experts are subjected to a series of questionnaires with controlled opinion feedback to collect data from a group of experts and to obtain the most reliable opinion consensus (Dalkey & Helmer, 1963). But a two-round Delphi technique was used instead of the three-round Delphi technique, which was widely used in the literature, because the experts reached consensus after the second round. Primarily, experts were detected from official institutions operating on disasters in Turkey, relevant non-governmental organizations, engaged in research in this area and academicians working in universities in the same area. Some of the experts have been working on disasters officially, some of them voluntarily and the rest both officially and voluntarily. A total of 7 experts were contacted via email and their approvals were received for the expertise. By email, the experts were asked with open-ended questions what survey items could be asked to participants in order to determine their affective, behavioural, cognitive attitudes towards

disasters. An item pool of 67 statements was created with the first feedback from experts (round 1). Then, the number of items was reduced to 34 with the evaluation of the information obtained by a comprehensive literature review about the disasters and by combining statements similar to each other in the item pool. The completed 34-item draft scale was sent to the experts and they were asked to rate whether or not each of the item is appropriate and to make additional relevant suggestions, if there are any (round 2). Thus, assessments were made by all experts whether the items in the scale were relevant, and whether they represented the structure to be used to measure. Consensus was established with absolute majority of agreement by experts for an item. After the second consultation round, 5 items were removed and the number of items in the item pool was reduced to 29. Then, a focus group consisting of 6 students and a pilot study of 35 participants were employed. Respondents reported that 26 out of 29 items were clear and understandable. After the two-round Delphi which was followed by a focus group discussion and a pilot study were completed, thorough discussions with Turkish experts were held for cultural and linguistic assessments of items. In the end the final draft of the scale consisting of 26 items was obtained.

This scale was applied face to face to 787 undergraduate students between May and June 2016. All students were provided a questionnaire that included the statement of informed consent, personal questions, and related scale. Participants were informed that the survey aimed to assess attitudes towards disasters, their participation was voluntary and the data would be maintained confidential. Each item of the scale was answered on a five-point Likert scale, ranging from 1 ('totally disagree') to 5 ('totally agree'). It should be noted that the questionnaire was in Turkish language, but to be included in the paper, the items were translated into English.

DATA ANALYSIS

First, the reliability and internal consistency statistics for the subscales were measured. To investigate patterns in the participants' responses to the attitude items, exploratory factor analysis (EFA) was performed using varimax rotation. Following EFA, confirmatory factor analysis (CFA) with a different sample was used to test the factor structure. A repeated measure ANOVA was performed for comparing the factors obtained from the EFA and Bonferroni test was used for paired comparisons. The independent samples t test and one-way ANOVA with Tukey's honest significant difference post-hoc test were conducted to investigate whether the attitude dimensions differed depending on the gender and the study programmes of the participants.

FINDINGS

Male and female participants composed 39.4% (310) and 60.6% (477) of the sample, respectively. Their ages ranged from 18 to 26. Of all participants, 31.8% were students of natural sciences programmes, 33.5% health sciences programmes, and 34.7% social sciences programmes.

The above-mentioned draft of the scale contained 26 items. To select items for the final version, we performed a reliability analysis. Table 1 presents the corrected item-total correlation in the first column and the Cronbach's alpha if item deleted in the second column. The corrected item-total correlation values smaller than 0.3 suggest the removal of the items from the scale (Maltby et al., 2010). Removal of items 4, 25 and 26 from the scale according to this criterion also leads to an increase in Cronbach's alpha value. The results of this procedure led to the removal of three items and in the final scale, a set of 23 items was offered to measure attitudes towards disasters. The means and standard deviations for these items are also presented in Table 1.

Initially, an EFA was conducted on the scale comprising 23 items, yielding a three-factor solution. The Bartlett's test for sphericity and Kaiser-Meyer-Olkin measure of sampling adequacy verify the factorability of data (Barlett's test of sphericity gives (approx.) χ^2 = 7345.292, df = 253 and sig. 0.000; Kaiser-Meyer-Olkin measure of sampling adequacy is 0.888). Table 2 displays items, factor loadings, eigenvalues and total variance explained. Factor analysis was performed employing orthogonal varimax rotation, resulted in 55.273% of variance explained. The number of factors was established on the basis of those with eigenvalues greater than 1 as the selection criteria. Affective attitude made the most significant contribution to the scale, accounting for 21.004% of the variance. The second and third most significant factors were determined to be cognitive and behavioural attitudes, respectively. The factor loadings confirmed the scale's accurate assessment of the three types of attitude: affective, cognitive and behavioural.

Table 1: The results of the reliability analysis for the subscales.

Factors	Items	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	$\overline{X}$	SD
FACTOR2: COGNITIVE	I1	0.674	0.815	3.10	1.08
	I2	0.670	0.817	2.94	1.00
	I3	0.585	0.826	2.58	1.10
	I4	**0.246**	**0.868**	3.60	1.21
	I5	0.624	0.821	2.63	1.18
	I6	0.662	0.816	2.69	1.20
	I7	0.714	0.809	2.82	1.17
	I8	0.518	0.834	3.57	1.15
FACTOR1: AFFECTIVE	I9	0.545	0.820	3.39	1.33
	I10	0.409	0.860	3.70	1.25
	I11	0.556	0.819	3.79	1.20
	I12	0.530	0.822	3.68	1.25
	I13	0.681	0.807	3.70	1.19
	I14	0.734	0.803	3.83	1.14
	I15	0.711	0.805	3.74	1.18
	I16	0.600	0.815	3.77	1.23
	I17	0.453	0.829	3.84	1.18
FACTOR3: BEHAVIOURAL	I18	0.500	0.789	2.53	1.17
	I19	0.575	0.780	2.28	1.23
	I20	0.709	0.764	1.84	1.16
	I21	0.651	0.770	1.80	1.24
	I22	0.560	0.781	1.98	1.29
	I23	0.602	0.776	2.31	1.26
	I24	0.557	0.783	2.98	1.14
	I25	**0.168**	**0.825**	3.98	1.10
	I26	**0.285**	**0.824**	3.27	1.56

$\overline{X}$=Mean; SD=Standard Deviation

Table 2: The results of the exploratory factor analysis.

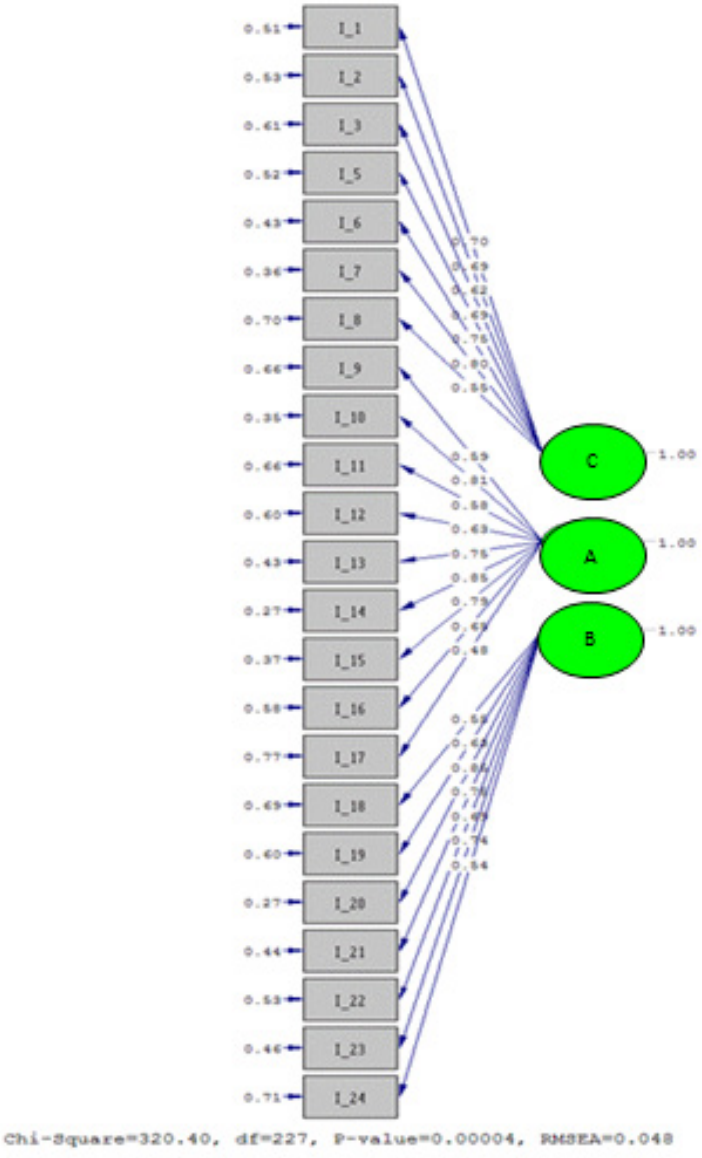
No	Items	Factor Loadings		
		Factor1	Factor2	Factor3
1	I have a basic knowledge about disasters.	0.072	**0.778**	0.079
2	I know how to reduce and/or eliminate the risk factors related to disasters.	0.076	**0.757**	0.119
3	I have enough information about the family disaster plan.	0.051	**0.633**	0.284
5	I have enough information about the safety of the house/dorm I live.	-0.034	**0.718**	0.166
6	I know what the non-structural risk factors are at home/dorm I live.	-0.053	**0.770**	0.147
7	I know what I need to do to reduce the non-structural risk factors of the house/dorm I live.	-0.015	**0.804**	0.155
8	I know how to behave during a disaster in crowded places (shopping centers, schools, public transport, social activity areas, etc.).	0.100	**0.619**	0.150
9	The risk of experiencing a disaster in the city I live scares me.	**0.669**	-0.042	0.145
10	The risk of experiencing a disaster in the country I live scares me.	**0.816**	0.035	-0.049
11	It makes me uneasy that necessary measures are not taken against a potential disaster.	**0.663**	0.013	0.004
12	Being in crowded places (shopping centers, schools, public transport, social activity areas, etc.) worries me during a disaster.	**0.708**	-0.025	-0.059
13	I am worried that I may not be accessed within a short time during a possible disaster.	**0.764**	0.064	-0.105
14	The possibility of not being able to accessed by search and rescue teams after a possible disaster worries me.	**0.832**	0.054	-0.114
15	The thought of not being able to get enough support (physical, psychological, housing) after a possible disaster worries me.	**0.808**	0.011	-0.082
16	I am afraid of experiencing communication problems with relatives after a potential disaster.	**0.704**	0.012	-0.120
17	The fact that social awareness increases only during disaster times is worrisome.	**0.543**	0.059	-0.032
18	I think I am prepared for a potential disaster.	-0.071	0.301	**0.561**
19	I back up my personal information and documents in case of exposure to a disaster.	0.021	0.213	**0.678**
20	We prepared family disaster plan against a possible disaster.	-0.084	0.092	**0.865**
21	I have a disaster and emergency bag.	-0.077	0.058	**0.812**
22	We took individual measures, such as fire extinguishers, at home where I live with my familiy.	-0.045	0.056	**0.761**
23	I have the required knowledge and training to protect myself during disasters.	-0.095	0.275	**0.688**
24	I can communicate correctly and accurately in case of an emergency.	-0.047	0.351	**0.545**
Eigenvalues		4.831	4.080	3.802
% of Variance		21.004	17.738	16.530
Cumulative %		21.004	38.743	55.273

To test the factor structure obtained by the EFA, CFA was conducted. The fit indices for construct validity in the CFA are shown in Table 3. The root mean square error of approximation, Non-normed fit index and comparative fit index indicate a good fit, whereas the other indices indicate an acceptable fit. Further, Graph 1 shows the factor structure: $\chi^2 = 320.40$ and df = 227. The value of chi-squared divided by degrees of freedom (Chi-square/df) is 1.41, where any chi-squared/df ratio less than 2 indicates an excellent fit (Tabachnick & Fidell, 2007). According to the fit indices, a three-factor model fits the data adequately.

Table 3: Goodness of fit indices for the data set (Schermelleh-Engel et al., 2003).

Fit Measure	Good Fit	Acceptable Fit	Goodness of Fit Statistics
RMSEA	0 ≤ RMSEA ≤ 0.05	0.05 < RMSEA ≤ 0.08	0.048
NFI	0.95 ≤ NFI ≤ 1	0.90 ≤ NFI < 0.95	0.920
NNFI	0.97 ≤ NNFI ≤ 1	0.95 ≤ NNFI < 0.97	0.980
CFI	0.97 ≤ CFI ≤ 1	0.95 ≤ CFI < 0.97	0.980
SRMR	0 ≤ SRMR ≤ 0.05	0.05 < SRMR ≤ 0.10	0.061
GFI	0.95 ≤ GFI ≤ 1	0.90 ≤ GFI < 0.95	0.910
AGFI	0.90 ≤ AGFI ≤ 1	0.85 ≤ AGFI < 0.90	0.880

RMSEA=Root Mean-Square Error of Approximation; NFI=Normed Fit Index; NNFI=Nonnormed Fit Index; CFI=Comparative Fit Index; SRMR=Standardized Root Mean Square Residual; GFI=Goodness-of-Fit Index; AGFI=Adjusted Goodness-of-Fit Index.



Graph 1. *Confirmatory factor analysis for students' attitudes towards disasters (A = Affective, B = Behavioural, C = Cognitive).*

Table 4 shows the item numbers, means, standard deviations, and Cronbach's alpha values for each subscale, as well as the results of a repeated measure ANOVA. The affective component is at a high level ($\overline{X}$ = 3.71 and SD = 0.89), when the cognitive and behavioural dimensions are under the moderate level ($\overline{X}$= 2.91 and SD = 0.83; $\overline{X}$= 2.25 and SD = 0.89). Cronbach's alpha values for the subscales are 0.868 for cognitive, 0.836 for affective and 0.852 for behavioural dimensions, indicating high-reliability statistics.

To test whether the factors demonstrate a significant difference due to gender or the study programme, an independent samples t test and a one-way ANOVA were conducted, respectively. The test results are given in Table 5.

According to the independent samples t test, there is a significant difference in the affective and behavioural factors between male and female participants. However, there is no significant difference in the cognitive factor. The level of affective attitude is higher in females ($\overline{X}$ = 3.89) than in males ($\overline{X}$ = 3.43). Conversely, the level of behavioural attitude is higher in males ($\overline{X}$ = 2.44) than in females ($\overline{X}$ = 2.12).

Results from a comparison of the dimensions of attitude of the students related to their study programmes are shown in Table 5. The students' study programmes influence both affective and behavioural dimensions. The level of affective attitude is higher for students of health sciences ($\overline{X}$ = 3.83) than for those of natural sciences ($\overline{X}$ = 3.61). The level of behavioural attitude is higher for students of natural sciences ($\overline{X}$ = 2.39) compared with those of the other programmes.

Table 4. *The item numbers, means, standard deviations and reliability statistics for each factor and the results of repeated measure ANOVA.*

Factors	Item Number	Mean	SD	F	p	Cronbach's Alpha
Cognitive	7	2.91[b]	0.83	639.962	<0.001	0.868
Affective	9	3.71[a]	0.89			0.836
Behavioural	7	2.25[c]	0.89			0.852
Total	23	3.02	0.56			0.810

[a, b, c]: The difference between means having different letters is significant (p<0.05).

Table 5. *The results of independent samples t test and one-way ANOVA.*

Variables	Groups	Cognitive		Affective		Behavioural	
		Mean±SD	p	Mean±SD	p	Mean±SD	p
Gender	Male	2.96±0.86	0.136	3.43±0.89	<0.001	2.44±0.85	<0.001
	Female	2.87±0.80		3.89±0.83		2.12±0.89	
Type of program	Natural Science	2.96±0.82	0.279	3.61±0.89[b]	<0.049	2.39±0.84[a]	<0.01
	Social Science	2.84±0.87		3.69±0.86[ab]		2.24±0.92[b]	
	Health Science	2.91±0.77		3.83±0.88[a]		2.12±0.87[b]	

[a, b]: The difference between groups having different letters is significant (p<0.05).

DISCUSSION

The main purpose of this study is to develop a new disaster attitude scale. The common feature of scale development studies is to produce a tool that will measure a certain phenomenon in terms of distinct dimensions and items. Likert type scales (Likert, 1932) are generally used in scale development studies. In this framework, unlike the few relevant studies in literature, this study presents a disaster attitude scale consisting of three dimensions (affective, behavioural and cognitive) and 23 items developed to determine attitudes towards disasters. By using this new disaster attitude scale, which has been found to be both reliable and valid, the attitudes of university students living in Afyonkarahisar, a high-risk earthquake area in Turkey, towards disasters have been determined. Some important results obtained from the study are as follows: University students' attitudes in the affective dimension were higher than cognitive and behavioural dimensions. This means that despite the high levels

of fear and anxiety of the individuals, the scores for the cognitive dimension that includes the thoughts and the beliefs, and the behavioural dimension that includes readiness and preventive behaviours are not at the required level. Additionally the cognitive dimension, representing information and beliefs about disasters, did not differ based on gender or study programme. However, there were differences in attitudes in the affective and behavioural dimensions associated with gender and study programme.

The scores for the affective dimension showed that female students had more fear, anxiety and worry about disaster and less disaster preparedness than male students, as reported previously by a study on different communities of Lovekamp (2006). Despite a high level of anxiety and fear among students, there was a lack of action, such as taking measures towards disasters. It should be noted that this was more apparent in female students than in males. Previous studies that investigated both gender and disasters demonstrated important results regarding women: both during and after disasters, women are more likely to be affected than men (Croson & Gneezy, 2009; Neumayer & Plümper, 2007).

When attitudes of students of natural sciences were compared with attitudes of students of health sciences, a significant difference in the affective dimension was detected. It can be said that the levels of anxiety and fear of students in the health sciences were higher than those in students of other programme. Moreover, there was a difference between the attitudes of students of natural sciences and the attitudes of students of other programmes in terms of the behavioural dimension: students of natural sciences had relatively higher scores in behavioural attitude. In fact, since students studying health sciences in order to provide services to public health are thought to have certain qualities such as managing emergency situations, they were expected to score lower in the affective dimension compared to those studying in other programs (natural and social sciences). Also, they were expected to score higher in the behavioural dimension because they were expected to be superior in terms of sensitiveness and responsibility. This contradiction can be attributed to several factors such as the deficiencies and shortcomings in terms of creating an awareness on community health, which result from the fact that the issues of protecting and promoting health are approached in a rather individualized manner within the process of education in health sciences.

The results of the analysis show that although disasters have destructive and agonizing impacts on the environment as well as on people's lives, attitudes towards disasters are not at the required level according to these measurements made at the local level. In a study, Ozkazanc and Yuksel (2015) aimed to determine the level of disaster awareness and disaster sensitivity of students in Turkey; their study revealed similar results. To cope with this problem, it is necessary to produce and implement policies in the scope of preparation for a possible crisis.

As a result, this particular study, which provides a valuable insight and contribution into the attitudes of undergraduate students, suggests the use of popular social media tools among young people to increase disaster awareness of university students. Short videos and animations posted on social media channels can make a significant contribution to attitude change. These videos and animations should take into consideration the short attention span of the target audience. The drills to be held during the time periods covering the course and library times, and the trainings to be organized afterwards, will raise awareness, help develop disaster culture and a reflex.

In addition to contributing to literature, the results obtained from this study are important data sources for institutions and organizations concerned with disasters. Besides, using the present scale, it will be possible for different researchers to determine the attitudes of people in different sample groups (in different geographies, countries, cultures, individuals with different socio demographic characteristics) towards disasters. The scale developed in this study is also an important first step in the development of individuals' awareness on and readiness to disasters.

ETHICAL STATEMENT

Permission for the study was granted by the Scientific Research and Publication Ethics Committee of Afyon Kocatepe University in May 2016; the data were collected between May and June 2016. Participants were informed that the survey aimed to collect data for a scientific study and volunteers willingly participated in the study.

REFERENCES

Ahayalimudin, N., & Osman, N. N. S. (2016). Disaster management: Emergency nursing and medical personnel's knowledge, attitude and practices of the East Coast region hospitals of Malaysia. *Australasian Emergency Nursing Journal, 19*(4), 203-209.

Baytiyeh, H., & Öcal, A. (2016). High school students' perceptions of earthquake disaster: A comparative study of Lebanon and Turkey. *International Journal of Disaster Risk Reduction, 18*, 56-63.

Bhat, B. A., Anees, S-U-M, Geelani, S. N. Z., Nusrat, Jan, I., & Zargar, B. A. (2017). A study on disaster awareness and preparedness among college students in district Ganderbal of Kashmir valley. *International Journal for Science and Advance Research in Technology*, *3*(12), 156-159.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature, 47*(2), 448-474.

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*, *9*(3), 458-467.

Dooley, D., Catalano, R., Mishra, S., & Serxner, S. (1992). Earthquake preparedness: predictors in a community survey. *Journal of Applied Social Psychology, 22*(6), 451-470.

Edwards, A. L. (1983). *Techniques of attitude scale construction*. Ardent Media.

Grimm, A., Hulse, L., Preiss, M., & Schmidt, S. (2014). Behavioural, emotional, and cognitive responses in European disasters: results of survivor interviews. *Disasters, 38*(1), 62-83.

Honda, A., Wiwattanapantuwong, J., & Abe, T. (2014). Japanese university students' attitudes toward the Fukushima nuclear disaster. *Journal of Environmental Psychology, 40*, 147-156.

IFRC (2015). World Disasters Report 2015. International Federation of Red Cross and Red Crescent Societies, Geneva, Switzerland. Retrieved from https://ifrc-media.org/interactive/wp-content/uploads/2015/09/1293600-World-Disasters-Report-2015_en.pdf

Junn, E. N., & Guerin, D. (1996). Factors related to earthquake preparedness among child care professionals: Theory and policy implications. *International Journal of Mass Emergencies and Disasters, 14*(3), 343-359.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22*, 5-55.

Lindell, M. K., & Perry, R. W. (2003). *Communicating environmental risk in multiethnic communities*. Sage Publications.

Lindholm, J., Carlson, T., Djupsund, G., Gögväg, J., & Strandberg, K. (2015). Citizens' emotional and cognitive responses to focusing events--an Experimental study. *International Journal of Mass Emergencies and Disasters, 33*(3), 407-427.

Lovekamp, W. E. (2006). *Gender, race/ethnicity and social class differences in disaster preparedness, risk and recovery in three earthquake-stricken communities* (Doctoral dissertation). Southern Illinois University Carbondale, United States – Illinois.

Maltby, J., Day, L., & Macaskill, A. (2010). *Personality, individual differences and intelligence*. Harlow: Pearson Education.

NEA (National Education Association) (1965). *Sampling and statistics handbook for surveys in education*. Washington, DC: National Education Association Press.

Neumayer, E., & Plümper, T. (2007). The gendered nature of natural disasters: The impact of catastrophic events on the gender gap in life expectancy, 1981–2002. *Annals of the Association of American Geographers, 97*(3), 551-566.

Olson, J. M., & Zanna, M. P. (1993). Attitudes and attitude change. *Annual Review of Psychology, 44*(1), 117-154.

Ozkazanc, S., & Yuksel, U. D. (2015). Evaluation of disaster awareness and sensitivity level of higher education students. *Procedia-Social and Behavioral Sciences, 197*, 745-753.

Pekez-Pavliško, T., Račić, M., & Jurišić, D. (2018). A questionnaire study on the attitudes and previous experience of Croatian family physicians toward their preparedness for disaster management. *Bulletin of Emergency and Trauma*, *6*(2), 162-168.

Qi, Z., Jing-yi, L., & Jing-bo, Z. (2003). Study on index system of assessment of public disaster perception in the western China. *Chinese Geographical Science*, *13*(3), 284-288.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. *Attitude organization and change: An analysis of consistency among attitude components, 3*, 1-14.

Rottach, P. (n.d.) *Background and Components of Disaster Risk Reduction*. Retrieved from http://

www.preventionweb.net/files/24122_24122backgroundcomponentsofdrractdi.pdf

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23-74.

Smith, M. B. (1968). *Attitude change. International encyclopedia of the social sciences*. Crowell Collierand Mac Millan.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson/Allyn & Bacon.

Tavan, H., Menati, W., Azadi, A., Sayehmiri, K., & Sahebi, A. (2016). Development and validation of a questionnaire to measure Iranian nurses' knowledge, attitude and practice regarding disaster preparedness. *Journal of Clinical and Diagnostic Research*, *10*(8), IC06-IC09.

Terpstra, T. (2011). Emotions, trust, and perceived risk: Affective and cognitive routes to flood preparedness behavior. *Risk Analysis, 31*(10), 1658-1675.

Tierney, K. J., Lindell, M. K., & Perry, R. W. (2001). *Facing the unexpected: Disaster preparedness and response in the United States*. Joseph Henry Press.

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. *The Social Psychology of Knowledge, 8*, 315-334.

EGE AKADEMİK BAKIŞ / *EGE ACADEMIC REVIEW*

Cilt 19 • Sayı 4 • Ekim 2019
SS. 469 / 485
Doi: 10.21121/eab.638575
Başvuru Tarihi: 28/1/2018 • Kabul Tarihi: 17/9/2019

Türkiye'de Vadeli İşlem ve Opsiyon Piyasası'nın Etkinliği ve Sözleşmelerin Karşılaştırmalı Fiyat Öngörüsü*

Effectiveness of Turkish Derivatives Market and Forecasting Comparative Prices for the Contracts

Taner TAŞ[1]
Sibel SELİM[2]

http://orcid.org/0000-0002-2861-5467
http://orcid.org/0000-0002-8464-588X

ÖZET

Finansal piyasalarda oluşan belirsizliğin ve riskin giderilmesi amacıyla geliştirilmiş olan türev piyasalarda, piyasaya duyulan güven, doğru bilginin piyasaya dahil olan tüm unsurlara aynı anda ulaşması sayesinde piyasanın etkin olarak işlemesi durumunda gerçekleşebilmektedir. Böylece, geçmiş dönem fiyat hareketlerinden yararlanarak gelecek döneme ilişkin öngörüler yapmak mümkün olmamaktadır. Bu bağlamda çalışmada öncelikle, Türkiye'de faaliyet gösteren Vadeli İşlem ve Opsiyon Piyasası'nın etkinliği; Genişletilmiş (Augmented) Dickey-Fuller (ADF), Phillips-Perron (PP) ve Kwiatkowski vd. (KPSS) doğrusal birim kök testleri ve Kapetanios vd. (KSS) doğrusal olmayan birim kök testi uygulanarak sınanmıştır. Rassal yürüyüş sergilemediğine karar verilen seriler sebebiyle piyasanın etkin olmadığı sonucuna ulaşılmıştır. Ardından, Vadeli İşlem ve Opsiyon Piyasası'nda işlem gören TL/Dolar ve Bist-30 sözleşmelerinin gün sonu uzlaşma fiyatının öngörüsünde en yüksek performansı gösteren yöntemin belirlenmesi amaçlanmıştır. Bu amaçla, Borsa İstanbul A.Ş'den temin edilen ve 04.02.2005 – 31.12.2015 tarihleri arasını kapsayan veriler kullanılmıştır. Analiz bulgularına göre, TL/Dolar sözleşme serisi için ARMA(4,4) modeli, RBF-1-B-L yapay sinir ağı modeli ve ARCH(1) modeline kıyasla daha yüksek öngörü performansı gösterirken, Bist-30 sözleşme serisi için ise TDNN-1-B-L yapay sinir ağı modeli, ARMA(4,5) ve ARCH(1) modeline kıyasla daha yüksek öngörü performansı gösteren model olmuştur.

Anahtar Kelimeler: Türev Piyasalar, Etkin Piyasalar Hipotezi, Birim Kök Testi, ARIMA Modelleri, ARCH/GARCH Modelleri, Yapay Sinir Ağları

Jel Kodu: C45, C22, D53

ABSTRACT

Derivative markets developed for eliminating uncertainty and risk arising from financial markets can make predictions about the future by using past price movements in case the market is not effective. In this context, in this study, firstly, the effectiveness of the Turkish Derivatives Market was tested by applying the Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) and Kwiatkowski et al. (KPSS) linear unit root tests and Kapetanios et al. (KSS) nonlinear unit root test. As a result of all unit root tests, it was concluded that the series did not show random walk, so that the market was not effective. Then, the method that shows the highest performance is tried to be determined when forecasting the end of day settlement price of the TL/Dollar and Bist-30 contracts which is traded in the Derivatives Market. For this purpose, the forecasting results produced by the time series analysis methods are compared with the results of the artificial neural network model which has the best performance by employing different architectures, layer numbers, cell numbers in layers, activation functions and learning methods using the data which is provided from Borsa Istanbul Inc. and covering the dates between 04.02.2005 and 31.12.2015.According to the results of analysis, ARMA (4,4) model performed better than RBF-1-BL artificial neural network model and ARCH (1) model for TL/Dollar contract series. For the Bist-30 contract series, TDNN-1-B-L artificial neural network model has higher predictive performance than ARMA (4.5) and ARCH (1) models.

Keywords: Derivatives Markets, Efficient Market Hypothesis, Unit Root Test, ARIMA Models, ARCH/GARCH Models, Artificial Neural Networks

Jel Codes: C45, C22, D53

***"Türkiye'de Vadeli İşlem ve Opsiyon Piyasası'nın Etkinliği ve Sözleşmelerin Karşılaştırmalı Fiyat Öngörümlemesi" adlı doktora tezinden üretilmiştir. Ayrıca bu çalışma Manisa Celal Bayar Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi tarafından 2015-175 no'lu proje ile 04.01.2016-29.12.2016 tarihleri arasında desteklenmiştir.**

[1]Dr. Öğr. Üyesi, Manisa Celal Bayar Üniversitesi, İşletme Fakültesi, Ekonomi ve Finans Bölümü, taner.tas@gmail.com
[2]Prof. Dr., Manisa Celal Bayar Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ekonometri Bölümü, sibel.selim@cbu.edu.tr

Giriş

Belirsizliğin giderilmesi ve risk yönetimi konusunda önemli bir yere sahip olan türev piyasalarda, herhangi bir varlığın bugün belirlenen fiyat ve miktar üzerinden alım satımı yapılmakta ancak teslimat ileri bir tarihte gerçekleşmektedir. BrettonWoods sabit kur sisteminin yerine dalgalı kur sisteminin kullanılmaya başlanması ile döviz kurlarında ve faiz oranlarında meydana gelen oynaklıklardan en az düzeyde etkilenmek için forward sözleşmeleri (alivre işlem sözleşmeleri), futures sözleşmeleri (vadeli işlem sözleşmeleri), opsiyon sözleşmeleri ve swap sözleşmeleri (takas sözleşmeleri) gibi çeşitli türev ürünler kullanılmaya başlanmıştır.

Tüm finansal piyasalarda olduğu gibi, risk yönetimi ve geleceğe dönük fiyat oluşumu gibi iki temel fonksiyonu bulunan türev piyasalar açısından da güvenilirlik kavramı işlevselliği önemli ölçüde etkilemektedir. Piyasaların güvenirliği de, piyasaya dahil olan tüm unsurların doğru bilgiye aynı anda ulaşabilmeleri ve bunun sonucunda doğru kararlar alabilmelerine bağlıdır ve bu durum ancak piyasaların etkin olarak işlemesi ile gerçekleşebilmektedir. Piyasa etkinliği kavramı, bilgi (enformasyon) etkinliği, kaynak dağılımı etkinliği ve faaliyet etkinliği olmak üzere üçe ayrılmaktadır. Bilgi (fiyatlama) etkinliği kavramı ile ilgili olan Etkin Piyasalar Hipotezi (EPH), Fama (1965) tarafından geliştirilmiş ve o tarihten günümüze kadar geçen sürede iktisadi anlamda ilgi çeken konular arasındaki yerini korumuştur. Hipotezin temelinde yatan düşünce, piyasada bulunan tüm bilgilerin finansal varlıkların fiyatlarına tam ve doğru olarak yansımış olması sebebiyle, herhangi bir yöntem kullanılarak uzun dönemde piyasa ortalama getirisinin üzerinde getiri sağlamak mümkün değildir. Fama (1970), bilgi etkinliği olarak da tanımlanabilen ve temel hipotezi, fiyatların piyasaya gelen yeni bilgiler doğrultusunda ve rassal olarak belirlenmesi üzerine kurulu olan piyasa etkinlik türlerini bilgi girişinin çeşidine göre, Zayıf Formda, Yarı Güçlü Formda ve Güçlü Formda Piyasa etkinliği olmak üzere üç gruba ayırmaktadır. Zayıf formda piyasa etkinliğinde, gelecek dönem fiyatlarının geçmiş dönem fiyatları kullanılarak öngörülemeyeceği, yarı güçlü formda piyasa etkinliğinde, kamuya açıklanan tüm bilgilerin finansal varlık fiyatlarına çok hızlı bir şekilde yansıdığı, güçlü formda piyasa etkinliğinde ise finansal varlık fiyatlarının kamu ile henüz paylaşılmamış özel bilgileri de içerdiği varsayılmaktadır.

Etkin piyasalar hipotezi genel olarak rassal yürüyüş kavramı ile ilişkilendirilmektedir. Bu durumun en önemli nedeni, etkin piyasalar hipotezi gereği rassal olarak oluşan bir bilginin finansal piyasalardaki fiyatlara yansımasının da aynı şekilde rassal olduğu varsayımıdır. Bu varsayımın arkasında, birbiri ardından gelen fiyat değişimlerinin veya getirilerin birbirinden bağımsız olması ve bu fiyat değişimi veya getirilerin benzer dağılım göstermesi şeklinde rassal yürüyüş hipotezinin temelini oluşturan iki varsayım bulunmaktadır. Rassal yürüyüş hipotezini ve dolayısıyla piyasaların zayıf formda etkinliğini sınayabilmek için kullanılan birim kök testlerinde, rassal yürüyüş boş hipotezi durağanlık alternatifine karşı test edilmekte ve birim kökün varlığı durumunda, piyasanın zayıf formda etkin olduğu sonucuna ulaşılmaktadır. Eğer seride birim kök yoksa yani seri I(0) durağan ise geçmiş dönem fiyatlar kullanılarak gelecek dönemdeki fiyatları öngörebilmenin mümkün olabileceği ihtimali kabul edilmekte ve piyasanın zayıf formda etkin olmadığı sonucuna ulaşılmaktadır. Zayıf formda etkin olmayan piyasalarda yer alan ekonomik ve finansal değişkenlerin gelecek dönem değerlerinin öngörülebilmesi için geliştirilen modeller ve bu modellerin gelecek dönem öngörü performanslarının kıyaslandığı çalışmalara ekonomi ve finans literatüründe sıkça karşılaşılmaktadır. Ancak ekonomik ve finansal zaman serileri kullanılarak oluşturulacak öngörü modelleri, bu alanda çok sayıda bağımlı ve bağımsız değişkenin mevcut olması sebebiyle oldukça karmaşık ve zor bir süreç olarak karşımıza çıkmaktadır.

Bu çalışmanın literatüre katkısı, Türkiye'de faaliyet gösteren Vadeli İşlem ve Opsiyon Piyasası'nın zayıf formda etkinliğinin birim kök testleri ile sınanması ve sözleşme verilerinin öngörüsü gerçekleştirilirken kullanılan yapay sinir ağları ve zaman serileri modelleri karşılaştırılarak en iyi yöntemin belirlenmesidir. Türkiye ekonomisi için Vadeli İşlem ve Opsiyon Piyasası'nda işlem gören sözleşmelerin gün sonu uzlaşma fiyatları için üç ayrı yöntemin öngörü performanslarını karşılaştırarak öngörü işleminde en iyi yöntemi ortaya çıkaran bu çalışma, literatürde bulunan ve finansal değişkenlerin gelecek dönem değerini öngören diğer çalışmalardan farklılaşmaktadır. Bu çalışma ile zaman serisi analizi yöntemlerine karşı yapay sinir ağları yönteminin öngörü performansı başarısının, zayıf formda etkin olmadığı belirlenen Vadeli İşlem ve Opsiyon Piyasası'nda işlem gören TL/Dolar ve Bist-30 endeks sözleşmelerinde de gerçekleşip gerçekleşmediği belirlenerek literatüre katkı sağlanmaya çalışılmıştır.

Bu çalışmanın izleyen bölümleri şu şekilde düzenlenmiştir. İkinci bölümde konuyla ilgili literatür araştırması ele alınmıştır. Üçüncü bölümde çalışmada

kullanılan veriler ve yöntem sunulmuştur. Çalışmanın dördüncü bölümünde ise, analizlerden elde edilen bulgulara ve modellerin performanslarının karşılaştırılmasına yer verilmiştir. Son olarak, sonuç bölümünde model tahminlerinden elde edilen bulgular literatürdeki çalışmalarla desteklenerek yorumlanmıştır.

Literatür Araştırması

Bu çalışmanın literatür araştırması iki ayrı kısımdan oluşmaktadır. İlk kısımda zayıf formda etkinlik çalışmaları incelenmiş, ikinci kısımda ise finansal ve iktisadi değişkenlerin gelecek dönem değerlerinin öngörüldüğü çalışmalara yer verilmiştir. Zayıf formda piyasa etkinliğini birim kök testleri yardımıyla inceleyen çalışmalar genel olarak ele alındığında, piyasalara ilişkin verilerin birim kök içermediği yani piyasaların zayıf formda etkin olmadığı görüşü daha kuvvetli olsa da, bu konuda net bir fikir birliği söz konusu değildir. Fakat ülkelerin gelişmişlik düzeyleri ile birlikte finansal piyasalarının derinliğinin artması ile orantılı olarak piyasa etkinliğinin de arttığı görülmektedir.

İstanbul Menkul Kıymetler Borsası için zayıf formda piyasa etkinliğinin analiz edildiği çalışmalardan, Buguk ve Brorsen (2003), Taş ve Dursunoğlu (2005), Özcan ve Yılancı (2009)'da, elde edilen birim kök testi sonuçlarına göre rassal yürüyüş hipotezinin desteklendiği ve sonuç olarak İMKB'nin zayıf formda etkin olmadığı kanıtlanmaktadır. Ancak aksi yönde kanıtlar elde eden çalışmalardan, Zengin ve Kurt (2004), Pınar (2009)'da ise İMKB'ye ilişkin serilerin birim kök içerdiği bu sebepten piyasanın zayıf formda etkin olduğu sonucu ortaya koyulmaktadır. İzmir Vadeli İşlem ve Opsyion Borsası'nda işlem gören İMKB 30, İMKB 100, Dolar ve Euro vadeli işlem (future) sözleşmelerinin zayıf formda etkinliğini test eden Durmuşkaya (2011) ise çalışmasında sözleşme serilerinin rassal yürüyüş göstermediği ve İzmir Vadeli İşlem ve Opsiyon Borsası'nın zayıf formda etkin olmadığı bulgusuna ulaşmıştır.

Aralarında Türkiye'nin de bulunduğu gelişmekte olan ülke borsalarının piyasa etkinliğine ilişkin olarak yapılan çalışmalardan, Çelik (2007), Marashdeh ve Wilson (2005), Erdem (2011) inceledikleri serilerin birim kök içerdiğini tespit etmişler ve piyasaların zayıf formda etkin olduğunu kanıtlamışlardır. Ancak benzer şekilde, Worthington ve Higgs (2003b), Chang vd. (2004), Calomfir (2015), Shaikh (2016), Stakicvd (2016), gelişmekte olan ülkeler şeklinde nitelendirilebilecek ülkelerin borsalarını incelemişler ve piyasaların zayıf formda etkin olmadığını ispatlamışlardır. Son olarak, gelişmiş ülkelerin borsalarına ait verilerle gerçekleştirdikleri zayıf formda etkinlik çalışmalarında Worthington ve Higgs (2003a), Demireli vd. (2010) piyasaların rassal yürüyüş hipotezini gerçekleştirdiğini belirlemişler ve bundan dolayı zayıf formda etkin piyasalar olduğu sonucuna ulaşılmışlardır.

Finansal ve iktisadi değişkenlerin gelecek dönem değerlerinin öngörülmesi ile ilgili literatür incelendiğinde, herhangi bir yöntem kullanılarak örneklem içi ve örneklem dışı öngörüler üretilen çalışmalar olduğu gibi, öngörü işleminde birden fazla yöntem kullanan ve bu yöntemler içerisinde en yüksek performans göstereni belirleyen çalışmaların çoğunluğu oluşturduğu göze çarpmaktadır. Öngörü yöntemi olarak yalnızca otoregresif hareketli ortalama (ARIMA) modelini kullanan çalışmalardan olan Subaşı (2005), Erdoğan (2006), Özkan (2006), Türkiye ekonomisine ilişkin olarak enflasyon, ithalat, dış borç gibi değişkenlerin farklı dönemler için öngörülerini gerçekleştirmişlerdir. Çörtük (2007) ve Adlığ (2009) döviz kurlarına ilişkin olarak, Çolak (2013) ise İstanbul Menkul Kıymetler Borsası (İMKB) verileri ile otoregresif koşullu değişen varyans modellerini (ARCH, GARCH…) kullanarak öngörü çalışmalarında bulunmuşlardır. Öngörü işleminde yalnızca yapay sinir ağı modellerinin kullanıldığı çalışmalardan, Avcı (2007), Dumlu (2011) ve Akcan ve Kartal (2011), İMKB endeksine, Qi ve Zhang (2008), ABD ekonomisine, Zahedi ve Rounaghi (2015) ise İran ekonomisine ilişkin verileri kullanarak öngörü çalışmaları gerçekleştirmişlerdir.

Birden çok öngörü yöntemi kullanarak performans karşılaştırmasının yapıldığı çalışmalardan Kaynar ve Taştan (2009) ve Polat (2009) haricindeki çalışmalarda yapay sinir ağlarının üstünlüğü göze çarpmaktadır. Eğrioğlu ve Aladağ (2005), Ataseven (2007), Özdemir (2008), İnsel ve diğerleri (2010), Uslu (2011), Bozdağ (2011), Akdağ (2010), Türkiye ekonomisine ilişkin veriler kullanarak yapay sinir ağları ile ARIMA modellerinin öngörü sonuçlarını karşılaştırmış ve öngörü işleminde yapay sinir ağlarının daha yüksek performans gösterdiğini ortaya koymuşlardır. Benzer şekilde, dünya ekonomisine ilişkin farklı veriler kullanarak gerçekleştirilen çalışmalardan, Kihoro vd. (2004), Panda ve Narasimhan (2007), Zou vd. (2007), Adebiyi vd. (2014), Lasheras vd. (2015)'de, yapay sinir ağlarının üstünlüğü kanıtlanmıştır.

Aydın ve Cavdar (2015), Usta (2007) ve Yıldız (2009) gelecek dönem öngörü işleminde vektör otoregresif (VAR), ARIMA ve yapay sinir ağları modellerini kullanmışlar ve Moralı (2011), Aygören vd. (2012)'nin ARIMA, nümerik arama modelleri ve yapay sinir ağlarını kul-

lanarak üç ayrı yöntemi karşılaştırdıkları çalışmalarına benzer şekilde, yapay sinir ağları yönteminin daha başarılı olduğunu göstermişlerdir. Yine, Kadılar vd. (2009), yapay sinir ağlarına alternatif olarak kullanılan ARIMA ve ARCH yöntem karşılaştırması sonucunda, doğrusal olmayan zaman serisi yöntemlerine karşı yapay sinir ağlarının performansını, son olarak Arabacı (2007) ve Godarzi vd. (2014)'de aynı şekilde yapay sinir ağları yöntemini daha başarılı bulmuşlardır.

Veri ve Yöntem

Veri

Bu çalışmada kullanılan veriler Borsa İstanbul A.Ş'den elde edilmiştir. Veriler 4 Şubat 2005 ile 31 Aralık 2015 tarihleri arasını kapsamaktadır. Analizde kullanılan TL/Dolar sözleşme serisi ile BİST-30 sözleşme serisi, her gün itibariyle en yakın vadeye sahip vadeli işlem sözleşmesinin gün sonu uzlaşma fiyatı verisinden oluşmaktadır. Uygulamada serilerin logaritmik birinci farkları alınarak elde edilen getiri serileriyle çalışılmıştır. Getirilerin hesaplanmasında kullanılan formül aşağıdaki şekildedir;

$$r_t = \ln(d_t/d_{t-1}) r_t = lnd_t - lnd_{t-1} \quad (1)$$

Bu formülde r_t, ilgili sözleşme serisinin t dönemindeki getirisini, d_t, t dönemindeki sözleşmenin gün sonu uzlaşma fiyatını, d_{t-1} ise t-1 dönemindeki gün sonu uzlaşma fiyatını göstermektedir. Bu çalışmada getiri serileri için kullanılan kısaltmalarda *rusd*; VİOP – TL/DOLAR sözleşmesini, *rbist30*; VİOP – BİST-30 sözleşmesini ifade etmektedir.

Yöntem

Zaman serisi analizinin ilk aşaması, analizde kullanılacak serilerin durağanlığının belirlenmesidir. Çünkü herhangi bir zaman serisine ait geçmiş değerler kullanılmak suretiyle serinin sabit parametreli modeline ulaşabilmek için, bu zaman serisi modelinden elde edilen stokastik sürecin niteliğinin zaman boyunca değişmiyor olması gerekmektedir. Üç çeşit durağanlık tipi vardır, ilki zayıf durağanlık veya diğer adı ile kovaryans durağanlıktır.

Ortalamanın ve varyansın zaman içinde değişmemesi kovaryansın gecikmeye bağlı olup zamana bağlı olmamasıdır. Özetle ortalama, varyans ve kovaryans zamana bağlı değildir (Sevüktekin ve Nargeleçekenler, 2005). Literatürde zaman serilerinin doğrusallığını test eden çeşitli testler bulunsa da, seriler için doğrusal olmamaya sebebiyet veren şeylerin farklı olmasından dolayı testler arasında birbirlerine karşı belirli bir üstünlük söz konusu değildir (Guhathakurta, 2011). Bu çalışmada parametrik olmayan bir yöntem olan BDS (Brock-Dechert-Scheinkman) testi kullanılmış olup bu yöntem, modelin hatalarının birbirinden bağımsızlığını ve hatalara ilişkin dağılımın aynı olup olmadığını incelemektedir. Yöntemin boş hipotezi, modelin hatalarının birbirinden bağımsız ve dağılımlarının aynı olması durumudur (Brock vd. 1987).

Otoregresif Hareketli Ortalama (ARMA/ ARIMA) Modelleri

Bir zaman serisi kendi gecikmeli değerlerinin bir fonksiyonu olarak tanımlanıyorsa buna otoregresif (AR- Autoregressive) süreç, t dönemindeki değerinin, hata teriminin (e_t) cari ve gecikmeli değerleri tarafından belirlendiği sürece de hareketli ortalama (MA-Moving Average) süreci adı verilmektedir. Durağan rassal süreçlerin sadece otoregresif veya sadece hareketli ortalama ile modellenememesi durumunda ise bu iki süreç aynı anda dikkate alınmaktadır. Bu durumda otoregresif hareketli ortalamalar modelinin (ARMA) kullanılması önerilmektedir (Tsay, 2010). Böylece zaman serisi modeli p ve q'uncu dereceden AR ve MA bileşenleri ile birlikte ARMA(p,q) olarak tanımlanmaktadır. ARMA(p,q) modeli Denklem (2)'teki gibidir.

$$Y_t = \delta + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \quad (2)$$

Burada kesme terimi δ , Y_t' nin ortalamasını ilgilendirmekte iken, hataların e_t, $E(e_t)=0$ ve varyansın $Var(e_t)= \sigma_e^2$ birbiri ile korelasyonu olmayan rassal değişkenler olduğu varsayılmaktadır. Zaman serisinin durağan olmadığı durumlarda ise ARIMA(p,d,q) modeli kullanılmaktadır. Bu modelde d, zaman serisinin durağanlaştırılması için alınması gereken fark sayısını göstermektedir. Bu modele aynı zamanda d. dereceden entegre bir ARIMA(p,d,q) modeli denir.

Koşullu Değişen Varyans Modelleri

Hata terimlerinin varyansının zaman içinde değişmediği varsayımı klasik doğrusal regresyon modelinin temelini oluşturan birkaç unsurdan bir tanesidir. Engle (1982) tarafından geliştirilen model sayesinde, hata terimlerinin sabit varyansa sahip olduğu yönündeki ortak görüşün geçerli olmadığı kanıtlanmıştır. Buradaki temel mantık, bir sonraki döneme ait varyans tahmininin önceki dönemler ile ilişkili olmasıdır. Engel (1982) koşullu değişen varyansı basit bir AR(1) modelinden yola çıkarak tanımlamıştır.

$$y_t = b_0 + b_1 y_{t-1} + \varepsilon_t \quad (3)$$

$$\varepsilon_t = z_t h_t^{1/2} z_t \text{ , iid, } E(z_t)=0, \quad Var(z_t)=1 \quad (4)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \quad (5)$$

Denklem (3)'deki modelde y_{t-1}, bağımlı değişkeni açıklayan gecikme değerini, ε_t hata terimini, b_0 ile b_1 bilinmeyen parametreleri ve Denklem (4)'deki z_t ortalaması 0, varyansı 1 olan rassal hata sürecini temsil etmektedir. Bu sayede finansal verilerin modellenmesinde görülen değişen varyans ve oynaklık kümelenmesi bu model içerisinde yer almaktadır.

ARCH modeli oynaklık sürecini açıklarken gecikme sayısının fazla olduğu durumlarda çok fazla parametreye ihtiyaç duyar ve bu durum tahmincilerin etkinliğini düşürmektedir. Ayrıca gecikmelerin uzun olduğu durumlarda koşullu varyans denklemindeki parametrelerin negatif olmama kısıtı da sağlanamamaktadır. Bundan dolayı Bollerslev (1986: 308), ARCH modelinin genişletilmiş hali olan, içerisinde daha fazla geçmiş bilgi barındıran ve gecikme yapısı daha esnek olan bir model geliştirmiştir. Bu modele genelleştirilmiş ARCH veya GARCH modeli denilmektedir. Bu model ARCH modelinin bir alternatifi değildir sadece ARCH modelinin eksikliklerini gidermeyi amaçlamaktadır.

Varyansın etkisinin simetrik olduğu kabul edilen ARCH ve GARCH modellerinde, pozitif veya negatif olduğuna bakılmaksızın büyüklüğü mutlak değerce eşit olan şokların oynaklık üzerinde aynı etkiye sahip olduğu varsayılmaktadır. Ayrıca bu modeller oynaklığın sadece büyüklüğünü ele almaktadır. Oysa finansal piyasaların algılama şekillerine göre, olumlu ve olumsuz bir haberin oynaklık üzerindeki etkisi farklı olmakta, azalış yönündeki dalgalanmalar artış yönündeki dalgalanmalardan daha fazla oynaklığa sebep olmaktadır (Nelson, 1991). Bu sebeple, asimetrik ilişkinin varlığının yanında kaldıraç etkisi olarak bilinen etkinin de belirlenmesine imkan tanıyan EGARCH modeli geliştirilmiştir.

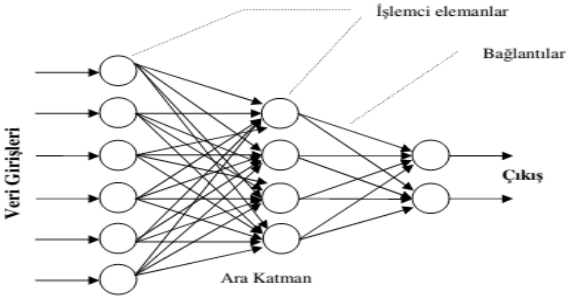
Volatiliteyi açıklamada asimetrik etkiyi dikkate alan bir diğer modelde eşik değerli ARCH (TARCH) modelidir (Zakoian, 1994). Bu çalışmada kullanılan ve koşullu varyans üzerindeki asimetriklik etkisini hesaba katan son model ise, Ding vd. (1993) tarafından geliştirilen asimetrik üslü ARCH (Asymmetry Power ARCH-APARCH) modelidir.

Yapay Sinir Ağları

Yapay sinir ağları, insan beynine ait biyolojik sinir ağlarından esinlenerek geliştirilen, birbirlerine belirli ağırlıklar ile bağlı öğrenme algoritmalarından ve transfer fonksiyonlarından oluşmaktadır (Elmas, 2003). Daha açık bir ifadeyle yapay sinir ağlarının yapısı, girişi yapılan verilerden çıktılar üretebilen, basit bir matematiksel fonksiyonu ifade eden bir kara kutu olarak tanımlanmaktadır. Yapay sinir ağları işlem süreci Şekil 1'de gösterilmiştir. Sistemin ilk adımı olan veri girişlerinin ardından işlemci elemanlar (nöronlar) ve bağlantı elemanları (aksonlar) işlemi çıkışa doğru devam ettirmektedir.

Yapay sinir hücresi, biyolojik sinir hücresinin gösterdiği davranışlar temel alınarak geliştirilen matematiksel kalıbı çalıştıran bir algoritma şeklinde tanımlanabilir. X_1, X_2, …,X_n gibi girdilerin ağırlıklarıyla çarpılıp toplanmasının ardından doğrusal veya doğrusal olmayan bir aktivasyon (transfer) fonksiyonundan f(net) geçilmektedir. Bu durum, W_{ni} ve W_{n0}'ın ağırlıkları ifade ettiği Denklem (6)'da görülmektedir (İnan ve Aras, 2005).

$$y_n = f(net_n) = f(\sum_{i=1}^{n} w_{ni} x_i + w_{n0}) \quad (6)$$



Şekil 1: *Yapay Sinir Ağları İşlem Süreci (Kaynak: Öztemel, 2012:30)*

Transfer (aktivasyon) fonksiyonu, hücreye ulaşan net girdiyi işleme tabi tutarak bu girdiden elde edilecek çıktıyı oluşturmaktadır. Bunun için farklı formüller kullanılmaktadır. Transfer fonksiyonu seçiminde deneme-yanılma yöntemi kullanılır. Çünkü uygun fonksiyonu gösteren bir formül yoktur. Yapay sinir ağları, ileri beslemeli (feedforward) ve geri beslemeli (feedback veya recurrent) ağlar olmak üzere ikiye ayrılmaktadır. İleri beslemeli ağlarda katmanlar halinde bulunan işlem elemanları tüm katmanlardaki işlem elemanları ile bağlantı kurarlar. İleri beslemeli ağlarda sinyallerin iletimi girdi katmanından çıktı katmanına doğru olarak yalnızca tek yönde gerçekleşmektedir. Ayrıca gizli katmandaki nöronların doğrusal olmayan davranışları, ileri beslemeli ağlarında doğrusal olmayan bir yapı göstermesine neden olmaktadır (Efe ve Kaynak, 2000).

Çok Katmanlı Algılayıcı (Multilayer Perceptrons) ve Doğrusal Vektör Parçalama Modeli (Learning Vector Quantization) ileri beslemeli ağ örneklerindendir (Sağıroğlu vd. 2003). Zaman serilere ait yapılan öngörülerde oldukça başarı sağlayan geri beslemeli ağların, ileri beslemeli ağlardan temel farkı hücrelerin ürettiği çıktılarının girdi katmanına gönderilerek yeniden girdi olarak kullanılabilmesidir (Zhang, 2003). Bu duruma geri besleme denmektedir ve katman içinde bulunan hücreler arasında olabildiği gibi, katmanlar arasında bulunan hücreler arasında da olabilmektedir.

Yapay sinir ağlarının eğitimi konusunda kullanımı en yaygın olan öğrenme algoritması geri yayılım (backpropagation) algoritmasıdır. Ancak standart geri yayılım algoritmasının hesaplama zamanının çok uzun olması ve yerel minimumda yakınsama riski gibi bazı zayıf noktaları vardır. Bu çalışmada bahsedilen sebeplerden dolayı, çeşitli optimizasyon yöntemleri ile geliştirilen daha yüksek performanslı Levenberg-Marquardt yöntemi kullanılmıştır (Aşkın vd. 2011).

BULGULAR

Bu çalışmada ilk olarak, doğrusal bir yapı sergilemediği Brock vd. (BDS) testiyle belirlenen Vadeli İşlem ve Opsiyon Piyasası'nın etkinliği, çeşitli birim kök testleri uygulanarak sınanmıştır. Ardından serilere, tek değişkenli zaman serisi yöntemi olan Box-Jenkins (1976) (ARMA – Otoregresif Hareketli Ortalama), otoregresif koşullu değişen varyans modelleri (ARCH, GARCH, EGARCH…) ve son olarak yapay sinir ağları yöntemi uygulanarak gelecek dönem öngörü işlemi gerçekleştirilmiştir.

Vadeli İşlemler ve Opsiyon Piyasası'nın Zayıf Formda Etkinliğinin Test Edilmesi

BDS testinin boş hipotezi olan serinin bağımsız ve özdeş dağılıma sahip olması durumunun reddedilmesi halinde seride doğrusal bağımlılığın olduğu, serinin durağan olmadığı ve son olarak seride doğrusal olmayan bağımlılık olduğu anlamlarından herhangi birisi oluşabilmektedir. En genel haliyle doğrusal olmayan yapının sorgulandığı bu testte, Tablo 1'de de görüldüğü gibi, getiri serilerinin tüm boyutlarında doğrusal olmayan bir yapının var olduğu ortaya çıkmaktadır.

Getiri serisinin zayıf formda etkinliğinin sınanabilmesi serinin birim kök içerip içermediğini araştırabilmek amacıyla, Genişletilmiş (Augmented) Dickey-Fuller (ADF), Phillips-Perron (PP), Kwiatkowski vd. (KPSS) doğrusal birim kök testleri ve Kapetanios, Shin ve Snell (KSS) doğrusal olmayan birim kök testi uygulanmıştır. ADF, PP ve KSS testleri serinin birim kök içerdiğini savunan boş hipotezi test ederken, KPSS testi ise serinin durağan olduğunu savunan boş hipotezi test etmektedir.

Tablo 1: *BDS Testi Sonuçları*

	Boyut	BDS istatistiği	Standart Hata	Z-istatistiği
rusd	2	0.0214	0.0017	12.1029
	3	0.0395	0.0028	14.0281
	4	0.0530	0.0033	15.7551
	5	0.0598	0.0035	17.0427
	6	0.0619	0.0033	18.2748
rbist30	2	0.0119	0.0016	7.1266
	3	0.0280	0.0026	10.5254
	4	0.0406	0.0031	12.8214
	5	0.0475	0.0032	14.4335
	6	0.0500	0.0031	15.7664

Tablo 2: *Getiri Serilerinin Durağanlık Testi Sonuçları*

	ADF test istatistiği		
	Sabitli	**Sabitli, Trendli**	**Sabitsiz**
rusd	-51.44	-51.46	-51.40
rbist30	-51.04	-51.03	-51.02
	PP test istatistiği		
	Sabitli	**Sabitli, Trendli**	**Sabitsiz**
rusd	-51.44	-51.46	-51.40
rbist30	-51.03	-51.02	-51.02
	KPSS test istatistiği		
	Sabitli		**Sabitli, Trendli**
rusd	0.18		0.02
rbist30	0.04		0.03
	KSS Birim Kök Testi		
rusd	-8.244		
rbist30	-4.085		

Not: KPSS testi için kritik değerler, %1, %5 ve %10 için sabitli ve trendli modeller için sırasıyla, 0.73, 0.46, 0.34 ve 0.21, 0.14, 0.11' dir. KPSS testi için ise, -3.48, -2.93 ve -2.66' dır.

Tablo 2'deki birim kök testleri sonuçlarına göre, getiri serisi düzeyde durağandır yani birim kök içermemektedir. Bu sonuçlara dayanarak Türkiye'de faaliyet gösteren Vadeli İşlem ve Opsiyon Piyasası'nın zayıf formda etkin bir piyasa olmadığı sonucuna ulaşılmakta ve geçmiş dönemlerde oluşan fiyatlar kullanılarak gelecek dönemlerde oluşacak olan fiyatları öngörme ihtimali söz konusu olmaktadır. Bundan dolayı, tek değişkenli zaman serisi yöntemi olan Box-Jenkins (ARMA – Otoregresif Hareketli Ortalama), otoregresif koşullu değişen varyans modelleri (ARCH, GARCH, EGARCH…) ve yapay sinir ağları yöntemi, getiri serisinin gelecek

dönem değerini öngörmek için kullanılacak ve bu yöntemler arasından öngörü performansı en yüksek olan yöntem belirlenebilecektir.

Getiri Serilerinin Fiyat Öngörüsü

Getiri serilerinin tanımlayıcı istatistiklerine ilişkin bilgiler Tablo 3'de görülmektedir.

Tablo 3: *Getiri Serilerinin Tanımlayıcı İstatistikleri*

	Rusd	Rbist30
Ortalama	0.000126	0.000167
Medyan	-0.000146	0.000195
Maksimum	0.035784	0.041940
Minimum	-0.027607	-0.043563
Standart Sapma	0.003810	0.008066
Çarpıklık	0.806074	-0.165997
Basıklık	11.09327	6.109282
Jarque-Bera	7706.663	1106.527
Olasılık	0.0000	0.0000

Ortalama değerleri düşük olan getiri serilerinin basıklık değerlerinin normal dağılım için kritik değer olan 3'den oldukça büyük olduğu görülmektedir. Getiri serilerinin sıfırdan farklı olan çarpıklık değerleri Jarque-Bera istatistiğini büyütmekte ve bu durum normal dağılım hipotezinin reddedilmesi anlamına gelmektedir. Ayrıca Jarque-Bera testine göre, getiri serilerinin normal dağılması üzerine kurulu olan boş hipotez tüm getiri serileri için %1 anlamlılık düzeyinde reddedilmektedir. Çarpıklık değerinin negatif olması serinin soldan çarpık dağılıma sahip olduğu anlamına geldiği için, Tablo 3'deki değerlere göre, *Rusd* serisi sağa çarpık, *Rbist30* serisi ise sola çarpıktır. Basıklık değeri ise normal dağılımda 3 değerini alırken getiri serilerinde 3'den oldukça büyük olduğu için, seriler normal dağılıma göre daha sivri bir tepeye (leptokurtik) ve daha kalın kuyruk (thick-tail) yapısına sahip olduğu görülmektedir.

En uygun otoregresif hareketli ortalama modelinin bulunabilmesi için Box-Jenkins metodolojisi izlenerek model parametrelerinin anlamlılığı, belirlilik katsayısı, bilgi kriterleri, hata karelerinin toplamı (SSE), modelin F istatistiği ve olabilirlik oranı gibi değerler dikkate alınmış ve uygulanan modellerin korelogramlarından otokorelasyon ve kısmi korelasyon katsayıları incelenmiştir. Öncelikle getiri serileri düzey değerlerinde durağan olduğu için, ARIMA modelinde bütünleşme derecesini gösteren kısım olan "I" sıfır değerini alır ve model ARIMA(p,0,q) yani ARMA(p,q) olur. AR(p), MA(q) ve ARMA(p,q) modellerine ait p ve q terimlerinin belirlenebilmesi için p,q=1,2,3,4,5 modelleri tahmin edilmiştir. Öncelikle farklı p ve q kombinasyonları ile kurulan modellerde katsayıların istatistiksel olarak anlamlılığı %1 anlamlılık düzeyinde test edilmiştir. Sonraki aşamada katsayıların istatistiksel olarak anlamlı olan ARMA(p,q) modellerinin kontrolü F testi yardımıyla yapılmıştır. Son olarak, diğer iki koşulun sağlanması şartıyla en düşük Akaike Bilgi Kriteri (AIC) değerine sahip olan model uygun model olarak seçilmiştir. Tablo 4' de yer alan sonuçlara göre, rusd için ARMA(4,4), rbist30 için ise ARMA(4,5) uygun model olarak belirlenmiştir.

Tablo 4: *rusd ve rbist30 Getiri Serileri İçin Bulunan Uygun ARMA Modellerinin Öngörü Performansı*

	rusd – ARMA(4,4)	rbist30 – ARMA(4,5)
RMSE	0.003786	0.008001
MAE	0.002615	0.005795
MAPE	110.0826	112.2808
TIC	0.886945	0.875445

Getiri serilerini koşullu varyans modelleri ile modelleyebilmek ve bu modellerle öngörü yapabilmek için öncelikle bulunan ortalama modellerinin hata terimlerinin karesinde ARCH etkisinin olup olmadığı araştırılmalıdır. Bu nedenle serilere önce ARCH-LM testi uygulanmış ardından hata terimlerinin korelogramlarından ARCH etkisinin varlığı araştırılmıştır.

Tablo 5 incelendiğinde, rusd ve rbist30 getiri serilerinin olasılık değerleri %1'den küçük olduğu için bu serilerde ARCH etkisinin olmadığını ifade eden sıfır hipotezi reddedilmiştir. Yani ARCH etkisinin varlığı tespit edilmiş olup bu etkinin giderilmesi gerekmektedir.

Tablo 5: *Getiri Serisine Uygulanan ARCH-LM Testi Sonuçları*

	rusd ARMA(4,4)	rbist30 ARMA(4,5)
F-İstatistiği	13.2028	49.5573
Olasılık F	0.0000	0.0000
Gözlem sayısı*R^2	537.1757	48.7030
Olasılık Ki-Kare(1)	0.0000	0.0000

Bu aşamadan sonra varlığı tespit edilen ARCH etkisini giderecek en uygun modeli belirleyebilmek için, normal dağılım, student-t dağılımı ve GED dağılımları için ayrı ayrı olarak, ARCH(1,2,3,4), GARCH(1,1), GARCH(1,2), GARCH(2,1), GARCH(2,2), EGARCH(1,1), TGARCH(1,1), APARCH(1,1) modelleri kurulmuştur. Model sonuçlarına göre, rusd getiri serisi için yalnızca ARCH(1), ARCH(2), ARCH(3), ARCH(4), GARCH(2,1) ve

GARCH(2,2) modellerinde, rbist30 getiri serisi için ise tüm modellerde ARCH etkisinin giderildiği görülmüştür. Bu çalışmada, üç ayrı dağılım için hesaplanan 11 ayrı koşullu değişen varyans modelinin öngörü işlemi gerçekleştirilmiş ve en iyi öngörü performansı gösteren modellere ilişkin sonuçlar Tablo 6'da sunulmuştur. Elde edilen sonuçlara göre, ortalama hata karesinin kökü (root mean squared error – RMSE), kriterine göre yapılan sıralamada, normal dağılım ARCH(1) modeli her iki getiri serisi için de en yüksek performansı gösteren model olmuştur.

Bu çalışmada kullanılan TL/Dolar ve Bist30 endeksi veri setinden elde edilen getiri serilerinin (-1,1) aralığında yer alması ve durağan olması sebebiyle serilere herhangi bir normalizasyon işlemi uygulanmamıştır. Veri seti, %70-15-15, %70-20-10 ve %80-10-10 olmak üzere üç farklı veri ayırma yüzdesi kullanılarak, eğitim, geçerlilik ve test kısımlarına ayrılmış ve farklı girdi katman nöron sayılarına göre performansları ölçülmüştür. Farklı yüzdelik kısımlara ayrılan veri setleri, 1,2,3,4,5 ve 10 girdi katmanı nöron sayıları ile çok katmanlı algılayıcılar (Multilayer perceptrons - MLPs), genelleştirilmiş ileri beslemeli ağlar (Generalized feed forward networks – GFF), temel bileşenler analizi ağları (Principal component analysis networks – PCAs), radyal tabanlı fonksiyon (Radial basis function – RBF), zaman gecikmeli geri beslemeli ağlar (Time lagged recurrent networks – TLRNs), geri beslemeli ağlar (Recurrent network) türündeki yapay sinir ağı modelleri kullanılarak analiz edilmiştir. Bu işlemlerin amacı, kullanılacak olan en iyi yapay sinir ağına ilişkin veri setini ve bu veri seti kullanılarak oluşturulan ağdaki en düşük RMSE değerine sahip girdi katmanı nöron sayısını belirleyebilmektir.

Tablo 7'de yer alan sonuçlara göre, rusd getirisi serisi için, %70 eğitim, %15 geçerlilik ve %15 test veri seti ve girdi katmanında 2 adet nöron bulunan, gizli katmanında Linear Sigmoid Axon, çıktı katmanında LinearTanhAxon aktivasyon fonksiyonlarının kullanıldığı RBF-1-B-L modeli en düşük RMSE değerine sahip model olarak belirlenmiştir.

Tablo 6: *rusd ve rbist30 Getiri Serileri İçin En İyi Öngörü Performansı Gösteren Koşullu Değişen Varyans Modelleri*

rusd	Model	Dağılım	RMSE	MAE	MAPE	TIC
1	**ARCH(1)**	Normal	0.003795	0.002622	110.0102	0.892563
2	**ARCH(1)**	Student-t	0.003806	0.002601	102.7311	0.925485
3	**GARCH(1,2)**	GED	0.003811	0.002599	101.1398	0.932783
rbist30	**Model**	**Dağılım**	**RMSE**	**MAE**	**MAPE**	**TIC**
1	**ARCH(1)**	Normal	0.008027	0.005823	112.2579	0.882351
2	**ARCH(4)**	GED	0.008032	0.005796	109.9203	0.891834
3	**ARCH(1)**	Student-t	0.008041	0.005799	107.0116	0.905711

Tablo 7: *Getiri Serileri i*çin En iyi Performans Gösteren *Yapay Sinir Ağı Modelleri*

		Rusd	Rbist30
Ağın Türü		RBF-1-B-L	TDNN-1-B-L
Veri Sayısı (Eğitim-Test-Doğrulama)		%70 (1932) - %15 (407) - %15 (407)	%70 (1932) - %15 (407) - %15 (407)
Ağdaki Toplam Katman Sayısı		3	3
Girdi Katmanındaki Nöron Sayısı		2	2
Gizli Katmandaki Nöron Sayısı		18	18
Çıktı Katmanındaki Nöron Sayısı		1	1
Gizli Katmandaki Aktivasyon Fonksiyonu		Linear Sigmoid Axon	LinearTanhAxon
Çıktı Katmanındaki Aktivasyon Fonksiyonu		LinearTanhAxon	Sigmoid Axon
Öğrenme Algoritması		LevenbergMarquardt	LevenbergMarquardt
Eğitim Türü		Batch	Batch
Döngü (epoch) Sayısı		1000	1000
RMSE	**Eğitim**	0.004094	0.008564
	Doğrulama	0.012020	0.007799
	Test	0.003138	0.005505
MAE	**Eğitim**	0.002832	0.006245
	Doğrulama	0.002841	0.005621
	Test	0.002346	0.004189

Rbist30 getirisi serisi için ise benzer şekilde, %70 eğitim, %15 geçerlilik ve %15 test veri seti ve girdi katmanında 2 adet nöron bulunan, gizli katmanında LinearTanhAxon, çıktı katmanında Sigmoid Axon aktivasyon fonksiyonlarının kullanıldığı TDNN-1-B-L modeli en düşük RMSE değerine sahip model olarak belirlenmiştir.

Modellerin Performanslarının Karşılaştırılması

Bu çalışmada kullanılan modellere ilişkin bir performans karşılaştırması yapabilmek amacıyla modellerin ortalama hata karesinin kökü (RMSE) kullanılarak sıralama yapılmıştır. Yapay sinir ağı modelleri için kullanılan RMSE değeri eğitim, doğrulama ve test setine ait değerlerin ortalamasından oluşmaktadır. Tablo 8'de sonuçları görülmekte olan bu sıralama sonucunda, rusd getiri serisi için, ARMA(4,4) modeli bu çalışmada ele alınan tüm modeller içerisinde en yüksek performansı göstermiştir. Bu modeli ARCH(1) modeli ile yapay sinir ağlarına ilişkin RBF-1-B-L modeli izlemiştir. Benzer şekilde rbist30 getiri serisi için yapılan sıralamada ise yapay sinir ağlarına ilişkin TDNN-1-B-L modeli bu çalışmada kullanılan tüm modeller içerisinde en yüksek performansı göstermiş ve bu modeli ise ARMA(4,5) ve ARCH(1) modeli izlemiştir. Böylece yapay sinir ağları modellerinin doğrusal olmayan serilerdeki öngörü başarısı (Sterbaand ve Hilovska, 2010) rbist30 getiri serisi için geçerli olurken, rusd getiri serisi için ise ARIMA modellerinin yapay sinir ağı modellerine kıyasla daha yüksek performans gösterdiği ortaya çıkmıştır (Lee vd. 2007; Merh vd. 2010). Bu sonuçlar çalışmayı geleneksel zaman serileri yöntemlerine karşı öngörü işleminde yapay sinir ağlarının daha üstün olduğu sonucuna ulaşan çalışmalardan (Moralı, 2011; Adebiyi vd. 2014; Kadılar vd. 2009; Panda ve Narasimhan, 2007; Akdağ, 2010) kısmen farklılaşmaktadır.

Tablo 8: *rusd ve rbist30 Getiri Serileri* için *Tüm Modeller Açısından Performans Karşılaştırması*

		rusd	
	Model	**RMSE**	**MAE**
1	**ARMA(4,4)**	0.003786	0.002615
2	**ARCH(1)**	0.003795	0.002622
3	**RBF-1-B-L**	0.006417	0.002673
		rbist30	
	Model	**RMSE**	**MAE**
1	**TDNN-1-B-L**	0.007289	0.005351
2	**ARMA(4.5)**	0.008001	0.005795
3	**ARCH(1)**	0.008027	0.005796

SONUÇ

Bu çalışmada amaç, Türkiye'de faaliyet gösteren Vadeli İşlem ve Opsiyon Piyasası'nın zayıf formda etkinliğinin hem doğrusal hem de doğrusal olmayan birim kök testleri ile sınanması ve sözleşme verilerinin öngörüsü gerçekleştirilirken yapay sinir ağları ve zaman serileri modellerinin karşılaştırılarak en iyi yöntemin belirlenmesidir.

Bu sebepten, öncelikle Türkiye'de faaliyet gösteren Vadeli İşlem ve Opsiyon Piyasası'nın etkinliği; Genişletilmiş (Augmented) Dickey-Fuller (ADF), Phillips-Perron (PP) ve Kwiatkowski vd. (KPSS) doğrusal birim kök testleri ve Kapetanios vd. (KSS) doğrusal olmayan birim kök testi uygulanarak sınanmıştır. Uygulanan tüm birim kök testleri sonucunda serilerin birim kök içermediğine yani rassal yürüyüş sergilemediğine karar verilmiş, böylece piyasanın etkin olmadığı sonucuna ulaşılmıştır. Ardından, zayıf formda etkin olmadığı belirlenen Vadeli İşlem ve Opsiyon Piyasası'nda işlem gören TL/Dolar ve Bist-30 sözleşmelerinin gün sonu uzlaşma fiyatının öngörülmesinde en yüksek performansı gösteren yöntemin belirlenmesi amaçlanmıştır. Bu amaçla, Borsa İstanbul A.Ş'den temin edilen ve 04.02.2005 – 31.12.2015 tarihleri arasını kapsayan veriler kullanılmış ve farklı öngörü sonuçları elde edilmiştir. Son olarak, zaman serileri analizi yöntemlerinin ürettiği öngörü sonuçları ile farklı mimariler, katman sayıları, katmanlardaki hücre sayıları, aktivasyon fonksiyonları ve öğrenme yöntemleri denenerek ulaşılan yapay sinir ağı modeli sonuçları karşılaştırılmıştır. Bu çalışmadan elde edilen bulgulara göre, TL/Dolar sözleşme serisi için, ARMA(4,4) modeli, RBF-1-B-L yapay sinir ağı modeli ve ARCH(1) modeline kıyasla daha yüksek öngörü performansı göstermiştir. Bist-30 sözleşme serisi için ise TDNN-1-B-L yapay sinir ağı modeli, ARMA(4,5) ve ARCH(1) modeline kıyasla daha yüksek öngörü performansı gösteren bir model olmuştur.

Türkiye'de faaliyet gösteren Vadeli İşlem ve Opsiyon Piyasası'nın zayıf formda etkin olmadığını kanıtlayıp, sözleşme serilerinin öngörü işlemini gerçekleştirerek literatüre katkı sağlayan bu çalışmanın sonuçlarına göre, türev piyasalara ilişkin finansal verilerin gelecek dönem değer öngörüsü yapılırken yapay sinir ağları, zaman serileri analizine bir alternatif olarak kullanılmıştır. Finansal varlık fiyatlarının rassal olarak hareket etmediği yani zayıf formda etkin olmadığı kanıtlanan piyasalarda geleceğe ilişkin risk ve belirsizliği ortadan kaldırmak için finansal varlıkların fiyatlarına ilişkin öngörülerde bulunmak isteyen karar alma birimlerinin

ARIMA modelleri ve yapay sinir ağları modellerini kullanmasının daha uygun olduğu görülmektedir. Ancak bu çalışmanın sonuçlarına göre öngörü işleminde iki yöntemin birbirine karşı kesin bir üstünlüğü olmamakla birlikte modellerin öngörü performansı da analizde kullanılan seriye göre değişmektedir.

KAYNAKÇA

Adebiyi, A. A., Adewumi, A., O. ve Ayo, C. K. (2014). Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction. *Journal of Applied Mathematics*. http://dx.doi.org/10.1155/2014/614342.

Akcan, A. ve Kartal, C. (2011). İMKB Sigorta Endeksini Oluşturan Şirketlerin Hisse Senedi Fiyatlarının Yapay Sinir Ağları ile Tahmini. *Muhasebe ve Finansman Dergisi*, Temmuz, 27-40.

Akdağ, Y. (2010). *Vadeli İşlem ve Opsiyon Borsası'nda TL/Dolar Vadeli İşlem Sözleşmelerinin Gün Sonu Uzlaşma Fiyatının Yapay Sinir Ağları ile Tahmini*. (Yayınlanmamış Yüksek Lisans Tezi). Marmara Üniversitesi/ Bankacılık ve Sigortacılık Enstitüsü, İstanbul.

Arabacı, Ö. (2007). *Makroekonomik Zaman Serisi Analizi ve Yapay Sinir Ağı Uygulamaları*. (Yayınlanmamış Doktora Tezi). Uludağ Üniversitesi/ Sosyal Bilimler Enstitüsü, Bursa

Aşkın, D., İskender, İ. ve Mamizadeh, A. (2011). Farklı Yapay Sinir Ağları Yöntemlerini Kullanarak Kuru Tip Transformatör Sargısının Termal Analizi. *Gazi Üniv. Müh. Mim. Fak. Dergisi*, *26*(4), 905-913.

Avcı, E. (2007). Forecasting Daily and Sessional Returns of the ISE-100 Index with Neural Network Models. *Doğuş Üniversitesi Dergisi*, *8*(2), 128-142.

Aydın, A. D. ve Çavdar, S. Ç. (2015). Comparison of Prediction Performances of Artificial Neural Network (ANN) and Vector Autoregressive (VAR) Models by Using the Macroeconomic Variables of Gold Prices, Borsa İstanbul (BIST) 100 Index and US Dollar-Turkish Lira (USD/TRY) Exchange Rates. *Procedia Economics and Finance*, *30*, 3-14.

Aygören, H., Sarıtaş, H. ve Moralı, T. (2012). İMKB 100 Endeksinin Yapay Sinir Ağları ve Newton Nümerik Arama Modelleri ile Tahmini. *International Journal of Alanya Faculty of Business*, *4*(1), 73-88.

Box, G. ve Jenkins G. (1976). *Time Series Analysis, Forecasting and Control*. San Francisco: Holden Day.

Çelik, B. (2008). *Yapay Sinir Ağları Metodolojisi İle Zaman Serisi Analizi: Teori ve Uygulama.* (Yayınlanmamış Yüksek Lisans Tezi). Marmara Üniversitesi/Sosyal Bilimler Enstitüsü, İstanbul.

Dickey, D. A. ve Fuller, W. A. (1979). Distribution of Estimators for Time Series Regressions with a Unit Root. *Journal of the American Statistical Association*, *74*, 427-431.

Efe, M. Ö. ve Kaynak O. (2000). *Yapay Sinir Ağları ve Uygulamaları*. İstanbul: Boğaziçi Üniversitesi.

Elmas, Ç. (2003). *Yapay Sinir Ağları*. Ankara: Seçkin.

Engel, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K Inflation. *Econometrica*, *50*, 987-1008.

Kadılar, C., Şimşek, M. ve Aladağ Ç., H. (2009). Forecasting the Exchange Rate Series with ANN: The Case of Turkey. *Ekonometri ve İstatistik Dergisi*, *9*, 17-29.

Kapetanios, G., Shin, Y. ve Snell, A. (2003). Testing for a Unit Root in the Nonlinear STAR framework. *Journal of Econometrics*, *112*(2), 359-379.

Kaynar, O. ve Taştan, S. (2009). Zaman Serisi Analizinde MLP Yapay Sinir Ağları ve ARIMA Modelinin Karşılaştırılması. *Erciyes Üniversitesi İktisadi İdari Bilimler Dergisi*, *33*, 161-172.

Kwiathowski, D., Phillips, P. C. B., Schmidt, P. ve Shin, Y. (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root, How Sure Are We That Economic Time Series Have a Unit Root?. *Journal of Econometrics*, *54*, 159-178.

Lee, C. K., Sehwan, Y. ve Jongdae, J. (2007). Neural Network Model Versus SARIMA Model in Forecasting Korean Stock Price Index (KOSPI). *Issues in Information System*, *8*(2), 372–378.

Merh, N., Saxena, V. P. ve Pardasani, K. R. (2010). A Comparison Between Hybrid Approaches of ANN and ARIMA for Indian Stock Trend Forecasting. *Journal of Business Intelligence, 3*(2), 23–43.

Moralı, T. (2011). İMKB *100 Endeksinin yapay Sinir Ağları ve Newton Nümerik Arama Modelleri ile Tahmini ve Sonuçlarının Karşılaştırılması*. (Yayınlanmamış Yüksek Lisans Tezi). Pamukkale Üniversitesi/ Sosyal Bilimler Enstitüsü, Denizli.

Özdemir, Ö. (2008). *Zaman Serisi Modellemesinde Yapay Sinir Ağlarının Kullanımı ve Bir Uygulama*. (Yayınlanmamış Yüksek Lisans Tezi). Anadolu Üniversitesi/ Fen Bilimleri Enstitüsü, Eskişehir.

Öztemel, E. (2012). *Yapay Sinir Ağları*. İstanbul: Papatya.

Panda, C. ve Narasimhan, V. (2007). Forecasting Exchange Rate Better with Artificial Neural Network. *Journal of Policy Modelling*, *29*, 227-236.

Phillips, P. C. B. ve Perron, P. (1988). Testing for a Unit Root in Time Series Regressions. *Biometrica*, 75, 335-346.

Qi, M. ve Zhang, G.,P. (2008). Trend Time-Series Modelling and Forecasting with Neural Networks. *Ieee Transactions on Neural Networks*, *19*(5), 808 – 816.

Sağıroğlu, Ş., Beşdok, E. ve Erler, M. (2003). *Mühendislikte Yapay Zeka Uygulamaları-1: Yapay Sinir Ağları*. Kayseri: Ufuk.

Servüktekin, M. ve Nargeleçekenler, M. (2005). *Zaman Serileri Analizi*. Ankara: Nobel.

Shazly, M. R. ve Shazly, H. E. (1999). Forecasting Currency Prices Using a Genetically Evolved Neural Network Architecture. *International Review of Financial Analysis*, *8*(1), 67-82.

Sterbaand, J. ve Hilovska, K. (2010). The Implementation of Hybrid ARIMA Neural Network Prediction Model for Aggregate Water. *Aplimat Journal of Applied Mathematics*, *3*(3), 123-131.

Tsay, R. ve S. (2010). *Analysis of Financial Time Series*. New York: Wiley Series in Probability and Statistics.

Yavuz, N. Ç. (2014). *Finansal Ekonometri*. Der.

Zahedi, J. ve Rounaghi, M. M. (2015). Application of Artificial Neural Network Models and Principal Component Analysis Method in Predicting Stock Prices on Tehran Stock Exchange. *Physica A*, *438*, 178-187.

Zhang, P. G. (2003). *Business Forecasting with Artificial Neural Networks: An Overview, Neural Networks in Business Forecasting*. (Ed.: Zhang, G. P.) Idea Group Publishing.

EKLER:

EK 1: *Rusd için Normal Dağılıma Göre Öngörü Performansı*

Model	RMSE	MAE	MAPE	TIC
ARCH(1)	0.003795	0.002622	110.0102	0.892563
GARCH(2.1)	0.003799	0.002613	103.7832	0.924174
EGARCH(1.1)	0.003799	0.002617	106.1735	0.91882
GARCH(1.2)	0.003802	0.00262	103.4306	0.920226
GARCH(2.2)	0.003802	0.00262	103.0366	0.921508
ARCH(4)	0.003806	0.002617	102.1267	0.941749
GARCH(1.1)	0.003808	0.002622	104.07	0.919991
ARCH(2)	0.003809	0.002623	102.5257	0.939764
ARCH(3)	0.003809	0.002616	100.2623	0.094826
TGARCH(1.1)	0.003809	0.002624	104.822	0.934761
APARCH(1.1)	0.003809	0.002619	101.6738	0.942717

EK 2: *Rusd için Student-t Dağılımına Göre Öngörü Performansı*

Model	RMSE	MAE	MAPE	TIC
ARCH(1)	0.003806	0.002601	102.7311	0.925485
ARCH(2)	0.003811	0.002603	102.3651	0.937056
GARCH(2.1)	0.003811	0.0026	100.1716	0.939353
APARCH(1.1)	0.003812	0.002606	99.22791	0.956989
GARCH(1.1)	0.003813	0.002599	100.255	0.935775
GARCH(1.2)	0.003813	0.002611	98.47579	0.09724
ARCH(3)	0.003814	0.002605	100.453	0.953414
GARCH(2.2)	0.003814	0.002601	98.84675	0.094983
EGARCH(1.1)	0.003814	0.002605	98.65923	0.959763
TGARCH(1.1)	0.003814	0.002604	99.07964	0.957599
ARCH(4)	0.003819	0.002602	101.4962	0.092836

EK 3: *Rusd için GED Dağılımına Göre Öngörü Performansı*

Model	RMSE	MAE	MAPE	TIC
GARCH(1.2)	0.003811	0.002599	101.1398	0.932783
GARCH(2.2)	0.003813	0.002598	101.765	0.92453
APARCH(1.1)	0.003813	0.002599	100.9585	0.92889
EGARCH(1.1)	0.003815	0.002605	100.7931	0.948106
ARCH(1)	0.003817	0.002605	99.30509	0.955436
GARCH(1.1)	0.003818	0.002605	99.89437	0.953384
TGARCH(1.1)	0.003818	0.002605	99.70433	0.954711
ARCH(2)	0.003819	0.002605	99.757	0.952182
ARCH(3)	0.003819	0.002605	99.93614	0.951113
GARCH(2.1)	0.003819	0.002599	100.0343	0.942208
ARCH(4)	0.003821	0.002605	100.4317	0.947654

EK 4: *Rbist30 için Normal Dağılıma Göre Öngörü Performansı*

Model	RMSE	MAE	MAPE	TIC
ARCH(1)	0.008027	0.005823	112.2579	0.882351
ARCH(2)	0.00803	0.005824	112.4953	0.883779
ARCH(3)	0.008041	0.005822	111.5436	0.897498
ARCH(4)	0.008046	0.005827	114.6397	0.882846
GARCH(1,1)	0.008063	0.005836	108.3833	0.915173
GARCH(2,1)	0.008063	0.005836	108.5376	0.91464
APARCH(1,1)	0.008063	0.005838	107.2363	0.91569
TGARCH(1,1)	0.008071	0.005846	107.2359	0.916984
GARCH(1,2)	0.008073	0.005849	109.4991	0.910945
GARCH(2,2)	0.008073	0.005849	109.5196	0.910893
EGARCH(1,1)	0.008077	0.005849	105.4851	0.926817

EK 5: *Rbist30 için Student-t Dağılımına Göre Öngörü Performansı*

Model	RMSE	MAE	MAPE	TIC
ARCH(1)	0.008041	0.005799	107.0116	0.905711
ARCH(2)	0.008043	0.005799	108.8764	0.907338
APARCH(1,1)	0.008048	0.005811	106.7911	0.907693
GARCH(2,1)	0.008054	0.005816	106.048	0.925921
EGARCH(1,1)	0.008058	0.005823	104.2293	0.933836
GARCH(1,1)	0.008059	0.005822	105.454	0.929945
GARCH(1,2)	0.008059	0.005822	105.5028	0.929686
GARCH(2,2)	0.00806	0.005822	105.5692	0.929318
ARCH(3)	0.008062	0.005804	106.363	0.913994
ARCH(4)	0.008078	0.00584	105.7233	0.932666
TGARCH(1,1)	0.008082	0.005833	108.6424	0.90705

EK 6: *Rbist30 için GED Dağılımına Göre Öngörü Performansı*

Model	RMSE	MAE	MAPE	TIC
ARCH(4)	0.008032	0.005796	109.9203	0.891834
GARCH(1,2)	0.008041	0.005801	108.3312	0.910859
TGARCH(1,1)	0.008047	0.005811	104.8247	0.921901
EGARCH(1,1)	0.008049	0.005806	107.5018	0.913988
ARCH(1)	0.008054	0.005797	105.3116	0.903912
GARCH(1,1)	0.008059	0.005819	105.359	0.927399
GARCH(2,1)	0.008059	0.005819	105.5089	0.9267
GARCH(2,2)	0.008059	0.005819	105.5128	0.926665
APARCH(1,1)	0.00806	0.005825	106.5053	0.912698
ARCH(2)	0.008061	0.005804	104.9472	0.915003
ARCH(3)	0.008063	0.005803	105.795	0.909505

EK 7: *Rusd Getiri Serisi İçin %70-%15-%15 Modele Ait Sonuçlar*

Model Adı	Girdi Katmanındaki Nöron Sayısı	Eğitim		Geçerlilik		Test	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
RBF-1-B-L (RadialBasisFunction)	2	0.004077	0.002798	0.012024	0.002795	0.003133	0.002313
RBF-1-B-L (RadialBasisFunction)	3	0.004079	0.002805	0.01201	0.002791	0.003147	0.002342
RBF-1-B-L (RadialBasisFunction)	1	0.00408	0.002795	0.012017	0.002806	0.003149	0.002345
MLPR-1-B-L (Regression MLP)	1	0.004065	0.002818	0.012356	0.00299	0.003152	0.002357
RBF-1-B-L (RadialBasisFunction)	4	0.004078	0.002797	0.012006	0.00279	0.003156	0.002337
MLPR-1-B-L (Regression MLP)	2	0.004095	0.002833	0.012013	0.002833	0.00316	0.002365
GFFR-1-B-R (Reg Gen Feedforward)	2	0.004092	0.002815	0.012037	0.002848	0.00316	0.002348
MLPCPC-1-B-L (Class MLP with PCA)	4	0.004088	0.002827	0.012016	0.002814	0.003161	0.002361
LogR-0-B-R (LogisticRegression)	4	0.004093	0.002808	0.012047	0.002866	0.003162	0.002351
MLPCPC-1-B-R (Class MLP with PCA)	5	0.004091	0.002815	0.012011	0.0028	0.003163	0.002353
GFFC-1-O-M (Class Gen Feedforward)	1	0.004095	0.002813	0.012032	0.002845	0.003163	0.002352
MLPC-1-O-M (Classification MLP)	5	0.004091	0.002813	0.012004	0.002806	0.003163	0.002356
MLPR-2-B-L (Regression MLP)	3	0.004087	0.002808	0.012012	0.002797	0.003164	0.002349
MLPRPC-1-B-R (Reg MLP with PCA)	3	0.004084	0.002804	0.012014	0.00281	0.003164	0.002348
MLPC-2-O-M (Classification MLP)	5	0.004091	0.002817	0.012012	0.002801	0.003166	0.002355
MLPC-2-O-M (Classification MLP)	10	0.00409	0.002814	0.012014	0.002798	0.003167	0.002353
MLPCPC-1-B-L (Class MLP with PCA)	10	0.004085	0.002823	0.012018	0.002821	0.003168	0.00236
MLPC-1-B-R (Classification MLP)	10	0.004088	0.002806	0.012022	0.00281	0.003169	0.002349

EK 8: *Rbist30 Getiri Serisi İçin %70-%15-%15 Modele Ait Sonuçlar*

Model Adı	Girdi Katmanındaki Nöron Sayısı	Eğitim		Geçerlilik		Test	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
MLPR-2-B-R (Regression MLP)	5	0.008445	0.006182	0.007754	0.005627	0.005486	0.004165
TLRN-1-O-M (Time-Lag Recurrent Network)	3	0.008452	0.006202	0.007947	0.00573	0.005489	0.004158
TDNN-1-B-L (Time-Delay Network)	2	0.008338	0.006121	0.007762	0.005547	0.005513	0.004179
MLPCPC-1-B-L (Class MLP with PCA)	4	0.008504	0.006192	0.007818	0.005657	0.005516	0.004201
RBF-1-B-R (Radial Basis Function)	4	0.008533	0.006226	0.007868	0.005648	0.005517	0.004209
MLPR-2-B-R (Regression MLP)	10	0.008375	0.006141	0.007732	0.005561	0.005517	0.004165
RBF-1-B-R (Radial Basis Function)	10	0.008508	0.006204	0.007924	0.00572	0.005517	0.004193
GFFR-1-B-R (Reg Gen Feedforward)	4	0.008482	0.006198	0.007845	0.005642	0.005519	0.004193
MLPR-1-B-R (Regression MLP)	5	0.008361	0.006171	0.007693	0.005558	0.005522	0.004197
GFFR-1-B-L (Reg Gen Feedforward)	5	0.008349	0.006152	0.007799	0.005602	0.005523	0.004199
MLPC-1-B-R (Classification MLP)	10	0.008556	0.006228	0.007854	0.005652	0.005523	0.004193
LinR-0-B-R (Linear Regression)	3	0.008549	0.006223	0.007868	0.00565	0.005531	0.004195
MLPR-1-B-L (Regression MLP)	1	0.008545	0.006218	0.007869	0.005662	0.005533	0.004203
MLPRPC-1-B-L (Reg MLP with PCA)	3	0.008541	0.006213	0.007839	0.00565	0.005533	0.004204
GFFC-1-O-M (Class Gen Feedforward)	2	0.008548	0.006224	0.007873	0.005659	0.005537	0.004209
MLPRPC-1-B-L (Reg MLP with PCA)	1	0.008544	0.006221	0.007876	0.005669	0.005537	0.004211
MLPRPC-1-B-R (Reg MLP with PCA)	1	0.008547	0.006225	0.00788	0.005671	0.005537	0.004211
MLPC-1-B-R (Classification MLP)	2	0.008555	0.006222	0.007851	0.005638	0.005537	0.004206
MLPR-2-B-R (Regression MLP)	5	0.008445	0.006182	0.007754	0.005627	0.005486	0.004165
TLRN-1-O-M (Time-Lag Recurrent Network)	3	0.008452	0.006202	0.007947	0.00573	0.005489	0.004158
TDNN-1-B-L (Time-Delay Network)	2	0.008338	0.006121	0.007762	0.005547	0.005513	0.004179

EK 9: *Rusd Getiri Serisi İçin %70-%20-%10 Modele Ait Sonuçlar*

Model Adı	Girdi Katmanındaki Nöron Sayısı	Eğitim		Geçerlilik		Test	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
RBF-1-B-R (RadialBasisFunction)	2	0.004073	0.002798	0.01052	0.002593	0.00323	0.002477
RBF-1-B-L (RadialBasisFunction)	2	0.004082	0.002801	0.010492	0.002594	0.003269	0.002496
MLPRPC-1-B-L (Reg MLP with PCA)	4	0.004081	0.002853	0.010905	0.002901	0.003274	0.002537
RBF-1-B-L (RadialBasisFunction)	1	0.004081	0.002798	0.010497	0.002609	0.003276	0.002514
RBF-1-B-L (RadialBasisFunction)	4	0.004074	0.002798	0.010495	0.002588	0.003279	0.00251
RBF-1-B-R (RadialBasisFunction)	3	0.004078	0.002806	0.010482	0.0026	0.003283	0.002531
MLPRPC-1-B-R (Reg MLP with PCA)	3	0.004093	0.002826	0.010498	0.002606	0.003286	0.002534
TDNN-1-B-R (Time-Delay Network)	3	0.004007	0.002795	0.010578	0.002741	0.003286	0.002512
RBF-1-O-M (RadialBasisFunction)	2	0.004115	0.002794	0.010535	0.002625	0.00329	0.0025
RBF-1-O-M (RadialBasisFunction)	4	0.004145	0.002827	0.010455	0.002623	0.00329	0.002475
RBF-1-B-R (RadialBasisFunction)	1	0.004077	0.002794	0.010477	0.002592	0.003297	0.002535
MLPCPC-1-B-R (Class MLP with PCA)	5	0.004091	0.002816	0.010498	0.002594	0.003299	0.002536
MLPR-2-B-R (Regression MLP)	1	0.004088	0.002806	0.010498	0.002593	0.0033	0.002533
MLPC-1-O-M (Classification MLP)	5	0.004093	0.002811	0.010489	0.002592	0.003302	0.002532
RBF-1-B-R (RadialBasisFunction)	5	0.004069	0.002798	0.010498	0.002593	0.003303	0.002528
MLPC-2-O-M (Classification MLP)	10	0.004091	0.002817	0.010495	0.002594	0.003305	0.002539
MLPC-2-B-R (Classification MLP)	10	0.004092	0.002809	0.010499	0.002589	0.003305	0.002536
MLPCPC-1-O-M (Class MLP with PCA)	10	0.004094	0.002822	0.010494	0.002598	0.003307	0.002544

EK 10: *Rbist30 Getiri Serisi İçin %70-%20-%10 Modele Ait Sonuçlar*

Model Adı	Girdi Katmanındaki Nöron Sayısı	Eğitim		Geçerlilik		Test	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
TDNN-1-B-L (Time-Delay Network)	5	0.008254	0.006099	0.007255	0.005235	0.005613	0.004257
TDNN-1-B-L (Time-Delay Network)	3	0.008408	0.006159	0.007258	0.005162	0.005622	0.004235
RBF-1-B-R (Radial Basis Function)	5	0.008496	0.006213	0.007309	0.005257	0.005638	0.004253
MLPRPC-1-B-L (Reg MLP with PCA)	10	0.008465	0.006211	0.007319	0.005278	0.005638	0.004261
TDNN-1-B-R (Time-Delay Network)	4	0.008346	0.00614	0.007247	0.005199	0.005638	0.004231
GFFR-1-B-R (Reg Gen Feedforward)	10	0.008332	0.00615	0.007316	0.005203	0.005644	0.004262
GFFC-1-B-L (Class Gen Feedforward)	10	0.008472	0.006194	0.007329	0.005237	0.005646	0.004247
TLRN-1-B-R (Time-Lag Recurrent Network)	3	0.008397	0.006149	0.007281	0.005256	0.005651	0.004253
MLPR-2-B-L (Regression MLP)	5	0.008428	0.006189	0.007225	0.005266	0.005653	0.00424
TDNN-1-B-R (Time-Delay Network)	3	0.008463	0.006181	0.007296	0.005261	0.005659	0.004249
MLPRPC-1-B-L (Reg MLP with PCA)	1	0.00855	0.006226	0.007298	0.005269	0.005671	0.00426
TLRN-1-B-R (Time-Lag Recurrent Network)	1	0.008549	0.006224	0.007299	0.005262	0.005671	0.004258
GFFR-1-B-R (Reg Gen Feedforward)	4	0.008494	0.006204	0.007311	0.005254	0.005672	0.004297
MLPR-2-O-M (Regression MLP)	4	0.008479	0.006203	0.007274	0.00525	0.005673	0.004275
RBF-1-B-L (Radial Basis Function)	1	0.008536	0.006225	0.007323	0.005277	0.005678	0.004297
TDNN-1-B-L (Time-Delay Network)	2	0.008364	0.006146	0.00718	0.005171	0.005684	0.004297
GFFC-1-B-R (Class Gen Feedforward)	2	0.008549	0.006226	0.007298	0.005259	0.005684	0.004276
LinR-0-B-L (Linear Regression)	2	0.008545	0.006224	0.007311	0.00527	0.005685	0.004283

EK 11: *Rusd Getiri Serisi İçin %80-%10-%10 Modele Ait Sonuçlar*

Model Adı	Girdi Katmanındaki Nöron Sayısı	Eğitim		Geçerlilik		Test	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
RBF-1-B-L (RadialBasisFunction)	2	0.003933	0.002666	0.014599	0.003448	0.003258	0.002488
GFFC-1-O-M (Class Gen Feedforward)	5	0.003985	0.00272	0.014689	0.003719	0.003269	0.002513
GFFC-1-B-L (Class Gen Feedforward)	4	0.003959	0.00269	0.014669	0.003628	0.003277	0.002527
TDNN-1-B-L (Time-Delay Network)	4	0.003974	0.002747	0.014576	0.003659	0.00328	0.002514
RBF-1-B-L (RadialBasisFunction)	3	0.003937	0.00268	0.014585	0.003444	0.003281	0.002518
RBF-1-O-M (RadialBasisFunction)	2	0.003996	0.002789	0.014639	0.003546	0.003281	0.002522
MLPRPC-1-B-R (Reg MLP with PCA)	3	0.003957	0.002689	0.014586	0.003498	0.003283	0.002529
GFFC-1-O-M (Class Gen Feedforward)	2	0.003952	0.00269	0.014634	0.003566	0.003287	0.002526
MLPRPC-1-B-L (Reg MLP with PCA)	5	0.003918	0.002661	0.01634	0.00442	0.003291	0.002532
RBF-1-O-M (RadialBasisFunction)	4	0.004	0.002814	0.014594	0.003558	0.003295	0.002557
MLPC-2-B-L (Classification MLP)	3	0.003947	0.002701	0.01458	0.003448	0.003298	0.002545
MLPC-1-B-R (Classification MLP)	5	0.003944	0.002677	0.014594	0.003478	0.003301	0.002539
MLPC-1-B-R (Classification MLP)	10	0.00395	0.002676	0.014574	0.003516	0.003303	0.002537
MLPC-2-B-R (Classification MLP)	10	0.003945	0.002683	0.014582	0.003459	0.003304	0.00254
MLPC-2-O-M (Classification MLP)	10	0.003944	0.00268	0.014585	0.003469	0.003306	0.002541
MLPR-2-B-R (Regression MLP)	1	0.003945	0.002688	0.014467	0.003378	0.003346	0.002571
TDNN-1-B-R (Time-Delay Network)	1	0.003985	0.002738	0.014528	0.003604	0.003352	0.002573
GFFR-1-B-L (Reg Gen Feedforward)	1	0.003938	0.002684	0.014548	0.003493	0.003355	0.002576

EK 12: *Rbist30 Getiri Serisi İçin %80-%10-%10 Modele Ait Sonuçlar*

Model Adı	Girdi Katmanındaki Nöron Sayısı	Eğitim		Geçerlilik		Test	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
TDNN-1-B-R (Time-Delay Network)	3	0.008257	0.006022	0.006839	0.005145	0.005609	0.004208
GFFR-1-B-R (Reg Gen Feedforward)	10	0.008452	0.006117	0.007009	0.005319	0.00563	0.004246
TLRN-1-O-M (Time-Lag Recurrent Network)	3	0.008351	0.006065	0.007	0.005278	0.005632	0.004258
TDNN-1-B-L (Time-Delay Network)	4	0.008164	0.005995	0.007059	0.005259	0.005634	0.004244
TLRN-1-O-M (Time-Lag Recurrent Network)	2	0.008375	0.006071	0.006956	0.005191	0.00564	0.004254
GFFR-1-B-L (Reg Gen Feedforward)	5	0.008287	0.006041	0.006918	0.005183	0.005648	0.00426
RBF-1-O-M (Radial Basis Function)	5	0.008402	0.0061	0.006987	0.005218	0.005649	0.004276
TDNN-1-B-L (Time-Delay Network)	10	0.008281	0.00602	0.007016	0.005253	0.005651	0.004251
MLPR-2-B-L (Regression MLP)	5	0.008317	0.006047	0.006927	0.005171	0.005655	0.00427
TLRN-1-B-R (Time-Lag Recurrent Network)	4	0.008428	0.006103	0.006936	0.005249	0.005656	0.00424
MLPR-2-B-R (Regression MLP)	4	0.008303	0.00605	0.00688	0.005167	0.005658	0.004261
RBF-1-B-R (Radial Basis Function)	10	0.008459	0.006116	0.006914	0.005227	0.005659	0.004266
RBF-1-O-M (Radial Basis Function)	1	0.008434	0.006113	0.006986	0.005254	0.005661	0.004294
RBF-1-O-M (Radial Basis Function)	3	0.008418	0.006098	0.006984	0.005229	0.005662	0.004269
MLPCPC-1-B-L (Class MLP with PCA)	2	0.008435	0.006111	0.006986	0.00523	0.005671	0.004268
MLPR-1-B-L (Regression MLP)	1	0.008436	0.006107	0.006975	0.00522	0.005672	0.004266
MLPRPC-1-B-L (Reg MLP with PCA)	1	0.008432	0.006107	0.006987	0.005232	0.005676	0.004269
TDNN-1-O-M (Time-Delay Network)	2	0.008436	0.006099	0.006977	0.00521	0.00568	0.004278

MAKALE GÖNDERME VE YAZIM KURALLARI

Dergide yayınlanmak üzere gönderilen yazıların daha önce başka bir yerde yayınlanmamış veya yayınlanmak üzere gönderilmemiş olması gerekir. Daha önce konferanslarda sunulmuş ve özeti yayınlanmış çalışmalar, bu durum belirtilmek üzere kabul edilebilir. Yayın için gönderilmiş çalışmalarını gecikme veya diğer bir nedenle dergiden çekmek isteyenlerin bir yazı ile başvurmaları gerekir. Yayın kurulu Ege Akademik Bakış Dergisi için gönderilmiş yazılarda makale sahiplerinin bu koşullara uymayı kabul ettiklerini varsayar. Dergiye gönderilen yazılara telif hakkı ödenmez.

Yayın kurulu, yayın koşullarına uymayan yazıları yayınlamamak, düzeltmek üzere yazarına geri vermek, biçimce düzenlemek yetkisine sahiptir. Yayınlanmak üzere gönderilen yazılar, yayın kurulunun uygun gördüğü en az iki hakem tarafından değerlendirildikten sonra yayınlanması uygun görülürse dergide basılır.

Derginin yayın dili Türkçe ve İngilizce'dir. Yazıların Türk Dil Kurumu'nun Türkçe Sözlüğü'ne ve Yeni Yazım Kılavuzu'na uygun olması gerekir.

Yazışma Adresi: Doç. Dr. Nazif Çatık, Ege Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ege Akademik Bakış Dergisi Editörü, e-posta: eab@mail.ege.edu.tr

Elektronik (Online) Makale Başvurusu

Ege Akademik Bakış dergisine makale başvuruları www.onlinedergi.com/eab adresinden ulaşabileceğiniz çevrimiçi makale gönderme sistemi üzerinden yapılır. Elektronik ("online") başvuru sırasında karşılaşacağınız herhangi bir sorunla ilgili olarak dergi sekreteryası ile iletişime geçebilirsiniz.

Makalenin editör kurulu tarafından yapılacak ön değerlendirmede reddedilmesi durumunda yazarlardan başvuru ücreti alınmamaktadır. Makale, editör kurulu tarafından yapılacak ön değerlendirmeyi geçtiği takdirde dergiye yönelik maliyetlerin karşılanması amacıyla 140 TL tutarındaki değerlendirme ücreti dergi sekreteryası tarafından bildirilecek banka hesabına yatırılmalıdır. Ödemenin yapıldığına dair banka dekontunun dergi sekretaryasına e-mail ile (eab@mail.ege.edu.tr) ulaştırılması durumunda makale hakem sürecine sokulacaktır. Makale başvuru ücreti geri ödenmemektedir.

İletişim:

Gül Huyugüzel Kışla-Fatma Demircan Keskin

Ege Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Bornova, 35100 İzmir, Türkiye Tele-Faks: +90 232 342 38 21 E-posta: eab@mail.ege.edu.tr

Yazım Kuralları Lütfen makalenizi dergi yazımı ile ilgili kuralları ve aşağıdaki örnek makaleyi dikkate alarak dergimize gönderiniz.

1. Özet

Özet 250 kelimeden az olmalı çalışmanın kapsamını ve temel bulguları içermelidir. Özetten sonra en az 3 anahtar sözcüğe yer verilmelidir. Anahtar sözcük listesinden sonra, varsa en az 3 JEL kodu eklenmelidir.

2. Ana Metin

Ana metin tablo, şekil ve referanslar dahil 35 sayfayı aşmamalıdır. Makaleler çift satır aralıklı ve A4 boyutunda yazılmalıdır. Makale gönderiminden önce yazı ve referanslarda hata olup olmadığı kontrol edilmelidir.

3. Tablo ve Şekiller

Tablo ve şekiller metnin uygun yerlerinde ardışık numaralandırılmış bir şekilde gösterilmelidir. Her şekil ve tabloya bir başlık verilmelidir. Tablo ve şekillere başka bir yerden alındığı takdirde kaynak gösterilmelidir.

4. Kaynaklar

4.1. Kaynaklar metin içinde şöyle gösterilmelidir: Arellano ve Bond (1991), kaynak gösterilen makale birden fazla yazara aitse Arreallano ve diğerleri (1995). Eğer bir yazarın aynı yıla ait birden fazla çalışması var ise Arellano (1997a) ve Arellano (1997b) gibi her yıl için alfabetik olarak sıralanmalıdır.

4.2. Kaynaklar makalenin sonunda yer almalıdır ve aşağıdaki örneklerde de gösterildiği gibi yazarın soyadına göre sıralanmalıdır.

Dergiler

İki yazar için: Arellano, M. ve Bond, S. (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations" *Review of Economic Studies*, 58: 277-297.

İkiden fazla yazar için:

Arellano, M., Bover, O. ve McLaugh, D. (1995) "Another Look at the Instrumental Variable Estimation of Error-Component Models" *Journal of Econometrics*, 68(1): 29-52.

Kitaplar

İki yazar için:

Cooke, P. (2004) "Regional Innovation Systems – An Evolutionary Approach" Cooke et al. (eds.) *Regional Innovation Systems*, 2nd Edition, London, Routledge.

İkiden fazla yazar için:

Cooke, P., Stephen, R. ve Wylie, P. (2003) *Northern Ireland's Evolving Regional Innovation System*, 3rd Edition, London, Routledge

İnternet Kaynakları

TÜBITAK (2006), http://www.tubitak.gov.tr/ hakkimizda/2004/ek7/EK_7.pdf, (05.05.2006) (Kaynağa ulaşım tarihi parantez içinde verilmelidir.)

SON NOTLAR

Lütfen dipnot yerine son notları kullanınız. Son notlar metnin ilgili yerinde üst simgeler kullanılarak verilmeli ve makalenin sonunda yer almalıdır.

SUBMISSION GUIDELINES

All submitted works must be original, must not have appeared elsewhere, and must not be submitted for publication elsewhere while under consideration by Ege Academic Review. Those works which have been presented in conferences and whose abstracts have already been published can be accepted for consideration on the condition that the presentation and publication dates and places are stated to the editor. Authors who wish to withdraw their submitted works for delays or any other reason should write to the editor. Submission of a manuscript for review is assumed by the editorial board the acceptance of these conditions by the authors. There will be no royalty payment to the authors.

The editorial board reserves the right not to accept manuscripts for consideration if they do not follow submission guidelines, to return them for revisions, or to organize them in the house style. All submitted works for publication undergo peer review, based on initial screening by editors and subsequent doubleblind refereeing by at least two reviewers. They are published only if they are found proper for publication.

The working languages of Ege Academic Review are Turkish and English. Articles in Turkish should follow the Dictionary and Spelling Guide of the Turkish Linguistic Society (http://www.tdk.org.tr).

Manuscripts and all editorial correspondence should be addressed to: The Editor, Assoc. Prof. Dr. Nazif Çatık, Ege University, Faculty of Economics and Administrative Sciences, The Journal of Ege Academic Review, e-mail: eab@mail.ege.edu.tr.

Online Manuscript Submission

Manuscript submissions to the Journal of Ege Academic Review must be made using the online manuscript submission system at www.onlinedergi.com/eab. Users should register when accessing the online system for the first time. If you experience any problem with the online submission please contact with the journal secretariat.

If the paper is rejected by the Editorial board, authors do not have to pay a fee. If the paper passes the preliminary assessment process by the Editorial board, of $50 should be paid to the bank account provided by the Journal Secratariat as the submission fee to cover processing costs. The manuscript will be sent to referees when the scanned copy of the bank receipt of payment is e-mailed to the Journal Secratariat. Submission fee is non-refundable.

Contact

Gül Huyugüzel Kışla-Fatma Demircan Keskin

Ege Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Bornova, 35100 İzmir, Türkiye Tele-Fax: +90 232 342 38 21

E-mail: eab@mail.ege.edu.tr

The Journal Writing Style

Please submit your papers to our journal by considering the style guidelines and the sample paper given below.

1. Abstract

The abstract should be fewer than 250 words and should cover the scope of the work and summarize the main findings. At least 3 keywords should be listed below the abstract. If available, at least 3 Journal of Economic Literature (JEL) classification codes should be added after the keywords list.

2. The manuscript

Manuscripts should not exceed 35 pages of text including tables, figures and references. Manuscripts should be typed with a double spacing on one side of A4 paper only. The text and references should be checked for the errors before submission.

3. Tables and Figures

Tables and Figures should be embedded in the convenient place of the text and numbered consecutively in Arabic numerals. Each figure and table should have a caption. The source of tables and figures should be given if they are taken from elsewhere.

4. References

4.1. References should be cited in the text as follows: Arellano and Bond (1991), if the paper is written by more than two authors Arellano et al., (1995). If there is more than one work by an author in a given year, they should be labeled alphabetically within each year, such as Arellano (1997a) Arellano (1997b).

4.2. The references should be placed at the end of the paper and arranged in alphabetical order by authors' last names, as shown in examples below.

Periodicals / Journals

For two authors:

Arellano, M. ve Bond, S. (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations" Review of Economic Studies, 58: 277-297.

More than two authors:

Arellano, M., Bover, O. ve McLaugh, D. (1995) "Another Look at the Instrumental Variable Estimation of Error-Component Models" Journal of Econometrics, 68(1): 29-52.

Books

Cooke, P. (2004) "Regional Innovation Systems – An Evolutionary Approach" Cooke et al. (eds.) Regional Innovation Systems, 2nd Edition, London, Routledge.

More than two authors:

Cooke, P., Stephen, R. ve Wylie, P. (2003) Northern Ireland's Evolving Regional Innovation System, 3rd Edition, London, Routledge

Internet sources:

TUBITAK (2006), http://www.tubitak.gov.tr/ hakkimizda/2004/ek7/EK_7.pdf, (05.05.2006) (Access date should be written in parenthesis)

END NOTES

Please use endnotes rather than footnotes. Endnotes should be marked by consecutive superscript numbers in the related part of the text, and listed at the end of the article.