Journal of
Accounting and Finance

North American Business Press
Atlanta – Seattle – South Florida - Toronto

Journal of Accounting and Finance

Volume 20(8)
ISSN 2155-3625

North American Business Press, Inc.
Atlanta – Seattle – South Florida - Toronto

For submission, subscription or copyright information, contact the editor at:

customerservice@nabpress.com

Subscription Price: US$ 310/yr

Our journals are indexed by UMI-Proquest-ABI Inform, EBSCOhost, GoogleScholar, and listed with Cabell's Directory, Ulrich's Listing of Periodicals, Bowkers Publishing Resources, the Library of Congress, the National Library of Canada. Our journals have been accepted through precedent as scholarly research outlets by the following business school accrediting bodies: AACSB, ACBSP, & IACBE.

This Issue

This study examines several determinants of student performance in Managerial Accounting at a small private college and contrasts its results with those of a published study at a larger public university. Intended grade has a significant association with student performance at both schools but intention to take the CPA exam is significant only at the small college. Students' work hours and course loads do not impact student performance negatively at the small college but they do at the larger university. Math and Writing abilities, Financial Accounting grade, and GPA are strong predictors of student performance at both institutions. Based on these results suggestions are provided for faculty.

In this paper, I study insider trading behavior following analyst earnings forecast revisions. Contrary to the notion that insiders trade consistent with analyst earnings forecast revisions in order to take advantage of the post forecast revision drift, I find that insiders react opposite to analyst earnings forecast revisions and are more likely to sell stock after upwards earnings forecasts revisions and are more likely to purchase stock after downwards earnings forecast revisions. My findings are consistent with insiders trading opportunistically following analyst earnings forecasts revisions.

This article examines what macroeconomic factors affect the NPL ratio and how to better measure this ratio to reduce the valuation bias caused by credit growth over the reporting period. The study is based on current publications in the context of the topic and the macroeconomic indicators available in the Eurostat database for regression analysis during the period 2001-2018. As a result of the study, it is proposed to adjust the traditionally used NPL ratio with credit growth during the reporting period. Statistical tests that were carried out with 623 regression models provide strong evidence to conclude that the adjusted NPL ratio leads to higher explanatory power of macroeconomic indicators and thus increase the level of confidence of NPL ratio predictions.

Estimates show that NPL ratio is mostly determined by some key macroeconomic variables, such as loans to GDP ratio, loan growth, unemployment, foreign investment growth, household income growth, inflation rates, and others. Macroeconomic development tendencies must be carefully considered when formulating policies in order to reduce credit risk.

In Germany and Russia, the majority of the population shows significant deficits in financial literacy. To overcome this situation both countries need people who are able to share and teach their knowledge. Economic students could be such multipliers. Therefore, it is important to understand the qualification of these students as ambassadors of financial literacy.

We examine the financial competence of German and Russian economic students. Based on a survey with international-comparable questions we find that most students perform indeed better than the international average. But only few students are able to answer all questions correctly with gender and age specific divergences. Women perform better at numeracy but worse at financial questions and younger students perform worse than older. Finally, we show that students with better understanding and education in finance are more likely to assign a probability to future crisis and that Russian students are more pessimistic than German students.

GUIDELINES FOR SUBMISSION

Journal of Accounting and Finance
(JAF)

Domain Statement

The *Journal of Accounting and Finance* (JAF) is dedicated to the advancement and dissemination of research across all the leading fields of financial inquiry by publishing, through a blind, refereed process, ongoing results of research in accordance with international scientific or scholarly standards. Articles are written by business leaders, policy analysts and active researchers for an audience of specialists, practitioners and students in all areas related to financial and accounting in business and education. Studies reflecting issues concerning budgeting, taxation, process, investments, regulatory procedures, and business financial analysis are suitable themes. JAF also covers theoretical and empirical analysis relating to financial reporting, asset pricing, financial markets and institutions, corporate finance, and corporate governance. Articles of regional interest are welcome, especially those dealing with lessons that may be applied in other regions around the world.

Submission Format

Articles should be submitted following the American Psychological Association format. Articles should not be more than 30 double-spaced, typed pages in length including all figures, graphs, references, and appendices. Submit two hard copies of manuscript along with a disk typed in MS-Word.

Make main sections and subsections easily identifiable by inserting appropriate headings and sub-headings. Type all first-level headings flush with the left margin, bold and capitalized. Second-level headings are also typed flush with the left margin but should only be bold. Third-level headings, if any, should also be flush with the left margin and italicized.

Include a title page with manuscript which includes the full names, affiliations, address, phone, fax, and e-mail addresses of all authors and identifies one person as the Primary Contact. Put the submission date on the bottom of the title page. On a separate sheet, include the title and an abstract of 100 words or less. Do not include authors' names on this sheet. A final page, "About the Authors," should include a brief biographical sketch of 100 words or less on each author. Include current place of employment and degrees held.

References must be written in APA style. It is the responsibility of the author(s) to ensure that the paper is thoroughly and accurately reviewed for spelling, grammar and referencing.

Review Procedure

Authors will receive an acknowledgement by e-mail including a reference number shortly after receipt of the manuscript. All manuscripts within the general domain of the journal will be sent for at least two reviews, using a double blind format, from members of our Editorial Board or their designated reviewers. In the majority of cases, authors will be notified within 45 days of the result of the review. If reviewers recommend changes, authors will receive a copy of the reviews and a timetable for submitting revisions. Papers and disks will not be returned to authors.

Accepted Manuscripts

When a manuscript is accepted for publication, author(s) must provide format-ready copy of the manuscripts including all graphs, charts, and tables. Specific formatting instructions will be provided to accepted authors along with copyright information. Each author will receive two copies of the issue in which his or her article is published without charge. All articles printed by *JAF* are copyrighted by the *Journal*. Permission requests for reprints should be addressed to the Editor. Questions and submissions should be addressed to:

North American Business Press
5720 Old Ocean Blvd
Boynton Beach, FL 33435
jaf@nabpress.com
866-624-2458

Student Performance in Managerial Accounting: An Empirical Study at a U.S. Small Private College

Mostafa M. Maksy
Kutztown University of Pennsylvania

Janis K. Zaima
Menlo College

This study examines several determinants of student performance in Managerial Accounting at a small private college and contrasts its results with those of a published study at a larger public university. Intended grade has a significant association with student performance at both schools but intention to take the CPA exam is significant only at the small college. Students' work hours and course loads do not impact student performance negatively at the small college but they do at the larger university. Math and Writing abilities, Financial Accounting grade, and GPA are strong predictors of student performance at both institutions. Based on these results suggestions are provided for faculty.

Keywords: accounting program, accounting education, student success

INTRODUCTION

Several prior studies have explored various factors to explain student performance in college-level accounting courses. While many past studies on introductory accounting courses were conducted in the late 1980s-1990s at large public institutions, few studies were directed at smaller residential private colleges (e.g. Doran et. al 1991, Eskew & Faley 1988, Gist et. al 1996, Gul & Fong 1993, Tho 1994; Wooten 1998; Ibrahim 1989). Moreover, recent studies have included several new variables such as distraction, self-perceived ability, prior ability, as well as other factors that provide insights into student performance in upper level accounting courses (e.g. Maksy & Rodrieguez 2017). This study investigates the associations between these recently applied variables and student performance in a lower-level Managerial Accounting course at a small residential private college in the U.S.

A study of Managerial Accounting is an essential bridge to student success in the business field because it introduces students to the various methods to measure cost for a company as well as examines its impact on the company's financial positions. Typically, students are required to take Introductory Financial Accounting as a prerequisite to Managerial Accounting. Hence, both courses are required for all business students at most universities and colleges, and are considered to be difficult courses for most business students. However, we found that Managerial Accounting is generally more difficult than Financial Accounting. At the small private college, after numerous years of administering the AACSB assessment tests, student average score continues to hover around 62 -72% compared to Financial Accounting with average scores of 85-88%. For this reason, we believe a better understanding of student performance in

Managerial Accounting is worthwhile to explore. This study examines the association of the new variables, motivation, distraction, self-perceived abilities, and prior ability, and student performance. Additionally, we contrast the results with that of a study conducted at a larger public institution at which this course is also required, allowing us to explore whether a closer student-faculty learning environment at a small residential college changes the impact of these variables to student performance. Murdoch & Guy (2002), for example, find that students in small section of an introductory accounting class perform better than in a large section of the same class.

The remaining parts of the paper present a review of prior research followed by a discussion of the study objectives. Next, we discuss the variables and hypotheses used in the study followed by research methodology and its results. The paper ends with conclusions, recommendations, study limitations, and some suggestions for further research.

REVIEW OF PRIOR RESEARCH

Several researchers, using data from various universities, find evidence supporting that overall Grade Point Average (OGPA) is a significant predictor of performance in accounting courses (Ingram and Peterson 1987; Eskew and Faley 1988; Doran et al. 1991). Others find evidence that both OGPA and the grade(s) earned in the prerequisite course(s) as significant predictors of performance in the course under investigation (Maksy & Rodriguez 2017 and 2018; Hao and Maksy 2019, and Maksy & Yoon 2019). In Saudi Arabia, Al-Twaijry (2010) finds that performance in high school, achievement in pre-university mathematics, and the grade in the Financial Accounting course as strong predictors of performance in Managerial Accounting. In contrast, Gist et al. (1996) find no significant association between academic performance and performance in principles of accounting courses at the university level.

Accounting is a subject area that requires quantitative skills and several studies find that strong mathematical abilities help students succeed in accounting classes. Eskew and Faley (1988), Gul and Fong (1993), and Al-Twaijry (2010) suggest that students with mathematical backgrounds outperform students with weaker mathematical backgrounds. Yet Gist (1996) observe no significant association between mathematics and performance in the accounting course.

Prior studies examine the influence of motivation and effort on student performance. For example, Wooten (1998) finds that motivation significantly affects effort which in turn significantly affects performance in an introductory accounting course. Other studies focused on upper-level accounting courses use “the grade the student intends to earn in the course” as a proxy for motivation and find it to be significantly associated with student performance in various upper-level accounting courses (Maksy & Rodriguez 2017; Maksy & Yoon 2019; Hao & Maksy 2019.)

In recent years, there has been increased interest in studying the influence of intervening variables on student performance. Lynn and Robinson-Backmon (2005) find a significant adverse association between employment status and learning outcomes in upper-division accounting courses. They also indicate that a student’s self-assessment of course learning objectives is significantly and directly related to grade performance. In contrast, recent studies find no significant negative associations between job hours, job type (if it is not related to accounting or business in general) or course load and student performance in various accounting courses (Maksy & Rodriguez 2017 and 2018; Maksy & Yoon 2019.) Al-Twaijry (2010) finds that students carrying more than 15 hours course load per semester perform better than others in a Managerial Accounting course. In contrast, Hao and Maksy (2019) find a significant negative association between course load and student performance in Advanced Accounting, an upper-level course.

These new variables and conflicting results motivate a second look at an introductory Managerial Accounting course along with an ability to contrast it with the results of a study conducted at a larger, public university offering the same course.

STUDY OBJECTIVES AND HYPOTHESES

Since motivation and effort has generally been positively associated with student performance, this study includes selected motivation factors to determine whether it affects student performance in the Managerial Accounting course. The authors also look at several factors which are commonly viewed as possibly distracting student's performance such as number of hours worked. Moreover, the study investigates the impact of four self-perceived abilities factors and student performance and whether students make accurate assessment of those abilities. Finally, the study investigates the impact of two specific measures of prior abilities on student performance, and also uses them as control variables while testing for the association between motivation, distraction, self-perceived abilities, and student performance in the Managerial Accounting course.

We use hypotheses H_1 to H_3 to test for the association between motivation and performance described in Appendix A. Hypotheses, H_4 to H_6 test the effects of distraction on student performance while H_7 to H_{10} examine the students self-perceived abilities and its impact on performance. Finally, H_{11} to H_{12} examine how prior abilities of the students impact their outcome in the Managerial Accounting course.

STUDY DEPENDENT VARIABLES

In addition to the 12 independent variables described above, the study uses two dependent variables. We use the course letter grade (A, B, C, etc.) as a measure of student performance (dependent variable, Grade). However, if the instructor does not use pluses and minuses for the letter grade, a student with a total percentage points of 80% and another with a total percentage points of 89% would earn a B and be considered having equal performance, even though the first student is one percentage point away from a C grade and the other student is one percentage point away from an A grade. As a result, we utilize an overall percentage as an alternative dependent variable (Points) defined as the overall total points earned by a student divided by the total possible points for the course (before any curving by the instructor).

RESEARCH METHODOLOGY

Survey Instrument and Sample

Besides the study variables, the survey instrument includes some demographic variables and other information. The survey was administered, in Fall 2018 and Spring 2019, to 46 of 49 students enrolled in two sections of the Managerial Accounting course at a U.S. small residential private college. The enrollment in each semester averaged 23 per section. The college enrolls approximately 745 students, and is considered to be one of the smallest AACSB accredited school. The instructor teaching the two sections provided data representing the two dependent variables (the 'letter grade,' and 'overall points' before any curving). For confidentiality purposes, performance data were matched with survey responses using student IDs only. We compare our results to the results in Maksy & Rodriguez (2018) study which was conducted at a larger public university that serves almost 12,000 students with a business school that enrolls about 2,100, or about 3 times the size of the small private college at which this study was conducted.

Table 1 presents descriptive statistics (minimum, maximum, mean, and standard deviation) of all variables used in the study. It is interesting to note that the mean of Intended Grade (IG) of 3.58 is higher than the mean of the Letter Grade of 2.78 earned in the course. It is also higher than the mean of each of the prior ability factors (the grade in Financial Accounting and overall GPA) – 3.38 and 3.27 respectively. It is even higher than the mean of each of the self-perceived ability factors (Writing, Math, Reading, and Listening) that have means of 2.76, 2.80, 2.67, and 3.00 respectively. This indicates that the students were overly optimistic about the grades they intended to earn in the Managerial Accounting course, which suggests students' overconfidence. In comparison, Maksy and Rodriguez (2018) study of performance determinants in a Managerial Accounting course at a New England public university report a Managerial Accounting course grade of 2.81 (which is just slightly higher than the average grade in this study, 2.78), GPA in the prerequisite Financial Accounting course of 3.40 (which is just slightly higher than in this study,

3.38), and overall GPA of 3.02 (which is lower than the OGPA in this study, 3.27). It is noteworthy that the negative difference between the average course letter grade and the average Financial Accounting prerequisite course grade for our sample is .60 and very close to the negative difference of 0.59 reported by Maksy and Rodriguez (2018). However, the negative difference of 0.49 between the average course letter grade and overall GPA is much higher than the negative difference of 0.21 reported by Maksy and Rodriguez (2018). This may imply that the students at a large public university likely has a lower average OGPA as compared to the private residential college. That is, assuming the same grade outcome for the Managerial Accounting course (2.78), the OGPA on average is 3.27 for the small private residential college as compared to OGPA of 2.99 for a large public school in New England. Though we cannot disaggregate whether the difference in OGPA is due to higher grade inflation at the small private college or students working harder (or getting more faculty help), a significant benefit exists for students who attend a smaller private school.

Typically, Managerial Accounting is well-known for being a difficult required business class, especially for those not majoring in accounting or finance. This may explain the large difference between the larger public institution and small private school when it comes to the difference between the average course letter grade and the average Intended Grade for the Managerial Accounting course. The *negative* difference of .80 between the average course letter grade and the average Intended Grade in this study is significantly higher (by 1.07) than the *positive* difference of 0.27 reported by Maksy and Rodriguez (2018). Perhaps, student expectations at a small residential college is much higher than at a larger public university. This is another indication that the students in this study were overly optimistic about the grades they intended to earn in the Managerial Accounting course, which again suggests students' overconfidence. Also in a closer student-faculty environment, students are better acquainted with faculty, thereby conflating acquaintance with greater optimism about their performance. Moreover, it may imply that students at a larger public university perceive their success in Managerial Accounting to be lower given that most students consider it to be a difficult course for non-accounting majors.

Hence, the overconfidence could be related to the environment at a small residential college, where students are encouraged to be successful in all their classes where they study together and feel comfortable seeking faculty help.

Data Analysis

To test the formulated hypotheses we use standard statistical analysis such as one-way analysis of variance (ANOVA), Pearson and Spearman's correlation coefficients and ordinary least square linear (OLS) regressions.

STUDY RESULTS

The analysis of the statistical results of the association between student performance and the four categories of independent factors; motivation, distraction, self-perceived abilities, and prior abilities are presented in the following five sections.

Motivation Factors Associated With Student Performance

As Tables 2 through 5 indicate, of the three motivation variables discussed in H_1 to H_3, Intended Grade (IG) is significantly associated with student performance at varying levels of significance depending on how student performance is measured. When student performance is measured by Letter Grade, the significance level is .10 under the ANOVA test (Table 2) and the regression test (Table 5) and .05 under the Pearson and Spearman correlations (Table 3). However, when student performance is measured as Points, which is a finer measure than Grade, the significance level becomes stronger: .05 under ANOVA and regression tests, and .01 under Pearson and Spearman correlations. As Table 4 indicates, after controlling for prior ability, as measured by the grade earned in the prerequisite Financial Accounting course (ACC 201 Grade) and OGPA, this significant association continues, but at a lower level (almost .10 when performance is measured as Grade and .05 when performance is measured as Points). These results

differ from the study conducted at a larger public institution where the IG effect disappears completely with no statistical significance when controlled for ACC201 Grade and OGPA effects while it continues to exist for the small private college. It appears that ACC201 Grade and OGPA drives student performance in the subsequent Managerial Accounting course at a larger public university, rather than motivated by IG.

Intention to take the CPA exam (ICPA) is also significantly associated at the .05 level with student performance under the Spearman correlation and at almost the same level (.05) under the Pearson statistics (Table 3) and at the .10 level under the regression test using both performance measures, Grade and Points. The private school shows a statistical association between ICPA and performance while the larger public university does not exhibit any relationship to ICPA. These results provide some evidence that smaller school may be advising students more closely regarding their future endeavors. However, the ANOVA test (Table 2) does not show any significant association between ICPA and student performance, however defined. As Table 4 indicates, after controlling for prior ability factors (ACC 201 Grade and OGPA) the significant associations under the correlations tests disappear, indicating that strong prior abilities override the motivation to take the CPA exam when related to student performance. Another confounding factor is whether the students are intending to be accounting majors. Only if they intend to go into accounting would they consider taking the CPA exam.

The third motivation variable, intention to attend graduate school (IGS) is not significantly associated with student performance (however defined) under any test. Since students typically take Managerial Accounting as sophomores or possibly first semester juniors it may be too early for them to consider their future beyond the undergraduate program. As expected, this explanation would apply to both a small private school as well as a larger, public university.

The results of this study, which exhibit significant association between IG and student performance and no significant association between IGS and student performance, are in agreement with several prior studies (e.g., Maksy & Rodriguez 2017 and 2018, and Hao & Maksy 2019.) It implies that the motivation factor at a small residential U.S. college displays similar effects as a larger, public university for IG and IGS. However, the results for the relationship between ICPA and student performance differ. It appears that students at a smaller school may become informed earlier in their academic career about the importance of taking the CPA exams to succeed in the accounting profession, thereby leading to a statistical relationship between the intention to take the CPA exams and performance.

Distraction Factors Associated With Student Performance

As Tables 2 through 5 indicate, none of the three distraction factors discussed in H_4 to H_6, has any significant negative association with student performance (however defined) under any of the four statistical tests used in the study. In fact, Table 3 shows some significant *positive* association between course load and student performance, but only when it is defined as Points, and only under the Pearson correlation, and at the lowest level of significance of .10. Furthermore, when controlled for the prior ability factors (ACC 201 Grade and OGPA), as Table 4 indicates, the significant *positive* association between course load and student performance disappears. This disappearance implies that the grade in Financial Accounting and overall GPA are the determining factors of the student performance earned in Managerial Accounting, i.e., regardless of how many courses per semester the students are taking, those who have high grades in the prerequisite Financial Accounting (ACC201 Grade) and high GPAs earn high grades in Managerial Accounting. The lack of negative associations between each of the three distraction factors and student performance, are in agreement with several recent studies (e.g., Maksy & Rodriguez 2017 and 2018, and Maksy & Yoon 2019.) Also, working while in college is not detrimental to student performance in the Managerial Accounting course at the private college. However, in contrast, it is significantly detrimental to student performance at a larger, public school. Also, at the large public school, job Type is significant at the .05 level relative to Grade while it is statistically significant at the .10 level relative to Points. Once controlled for the Financial Accounting class and OGPA, Job Hours becomes significant at the .10 level. These results imply that there is a negative association for students working at larger, public institutions while there is no effect for the private school. It may be related to better student advising by academic counselors and faculty at the smaller college. For example, the smaller college provides mid-term grades to

students that is recorded on the student academic record (though not on their transcripts). It provides a formalized method of alerting students about their academic progress, thereby, encouraging students to reassess their work hours should they earn low mid-term grades or to drop the course. Additionally, students with a mid-term grade lower than a C must meet with their advisors, who are given written suggestions by faculty as to how the student might improve their performance.

Self-Perceived Ability Factors Associated With Student Performance

The four self-perceived ability factors discussed in H_7 through H_{10}, have varying associations with student performance depending on how performance is defined and what statistical test is used. As Table 2 indicates, none of the four self-perceived abilities are significantly associated with student performance, however defined under the ANOVA test. As Table 3 indicates, Spearman correlations show significant association between Writing and student performance, but only when it is measured by Grade, and at the lowest level of significance of .10. Furthermore, when we control for the prior ability factors (ACC 201 Grade and OGPA), Table 4 indicates that the significant association between Writing and student performance disappears. Table 5 regression test demonstrates a significant association between Writing and student performance, however defined, at the .10 significance level. When examining the association between Math and student performance, Pearson and Spearman correlations display significance, at the .10 and .05 levels, respectively using both measures, Grade and Points. However, when controlled for prior ability factors (ACC 201 Grade and OGPA), Table 4 indicates that the significant association between Math and student performance disappears at the small private school, exhibiting similar behavior to that of a larger public university.

As shown in Tables 2 and 3, the ANOVA test and the correlation tests did not show any significant association between Reading and student performance, however defined for this study. But, interestingly, the regression test in Table 5 shows a significant *negative* association between Reading and student performance. The significance level of this association is .10 when performance is defined as Grade and .05 when defined as Points. It implies that students who reported that their reading abilities are Very Good or Good earned lower grades whereas students who reported that their reading abilities are Average (or even Poor) secured relatively higher grades. This seems to be a clear indication that students (particularly those with poor performance) severely overestimated their reading abilities.

As Tables 2, 3 and 5 indicate, none of the statistical tests used in the study showed any significant associations between Listening and student performance, however defined. In contrast, the larger school shows a .05 significance between Listening and student performance using both dependent variable measures. The Pearson correlation for the public university shows a .01 significance level between Listening and performance. Perhaps, students at larger institution believe their listening skills are better than writing or reading skills. We would need to conduct a more detailed evaluation as to the reasons for this interesting result.

In summary, the results from the former study at a larger public institution show a different association between perceived abilities and student performance on Writing, Reading and Listening. Writing showed a .10 significance for the small college while the larger institution showed no correlation using both Pearson and Spearman tests. The differences are also exhibited with Reading and Listening. Reading showed no relationship to performance for the small college using both performance measures, while the larger school showed a .05 significance using Points with the Spearman test, and both measures showed a .05 significance level for the Pearson test. Similarly, perceived Listening skills also exhibit .05 significance level for the larger university for both performance measures, while the relationship is nonexistent for the small private college. Moreover, the relationship did not disappear (for the larger school) after controlling for ACC201 Grade and OGPA, still indicating a .10 level using Grade and .05 using Points. Students at a larger, public school do not display the same confidence that exists with the private school counterparts.

Prior Actual Ability (Control) Factors Associated With Student Performance

The ANOVA test (in Table 2) and the regression test (in Table 5) do not show any significant association between ACC 201 Grade and student performance, however defined, but both Pearson and

Spearman correlations tests (in Table 3) show significant association (at .01) between ACC 201 Grade and student performance, however defined. The ANOVA test (in Table 2) shows significant association between OGPA and student performance, (at the .10 significance level when performance is defined as Grade and at the .05 when performance is defined as Points). The Pearson correlations test (in Table 3) shows significant association between OGPA and student performance, (at the .01 significance level when performance is defined as Grade and at the .05 level when performance is defined as Points). The Spearman correlations test (in Table 3) shows significant association (at .01) between OGPA and student performance, however defined. The regression test (in Table 5) does not show any significant association between OGPA and student performance, however defined.

The results of this study showing significant associations between the prerequisite course (ACC201) and student performance, are in agreement with the larger, public institutions in prior studies (e.g., Al-Twaijry 2010, Maksy & Rodriguez 2017 and 2018, Hao and Maksy 2019, and Maksy & Yoon 2019). The results of this study showing significant associations between OGPA and student performance, are in agreement with almost all prior studies mentioned in this paper. Therefore, we can generalize the relationship between the prerequisite course grade and student performance in the second accounting course, and similar association exist for overall GPA and student performance in Managerial Accounting.

CONCLUSIONS AND RECOMMENDATIONS

Similar to the larger public institution, the study finds that student performance at a small private school are impacted primarily by prior abilities, measured by Financial Accounting Grade (ACC201 Grade) and overall GPA (OGPA). As in past studies, students who are academically strong do well in Managerial Accounting. However, motivation has some bearing on their outcome in the course, and the results differ somewhat between larger and small institutions. Using Intended Grade (IG) as a measure of motivation, both schools show a strong correlation to student performance (using Grade the significance is .05 and using Points it equals .10). It differs when we control for prior abilities which indicates that IG response from the larger school has no association to student performance while the smaller school in this sample continues to exhibit statistical significance at the .10 and .05 for the dependent variable, Grade and Points, respectively. This implies that motivated students at the private school perform significantly better in the Managerial Accounting course than non-motivated students notwithstanding their prior abilities. Other differences between the two types of schools exist. Students appear more confident with their perceived abilities in writing and math at the small college, and they are unaffected by distractions as compared to students at a larger, public university.

In light of the above general conclusion, we recommend that accounting faculty should encourage their students to measure the student intention at the beginning of the semester, raising their awareness that being motivated by their intention are correlated to doing well in the class. We recommend giving this short survey to identify the students who have intentions to do well in Managerial Accounting (or any course). If the intent to excel exists, a faculty might be able to use methods to engage these students to keep them motivated. While a faculty at a small college might be able to target the entire class of 24 students, a professor at a larger institution might use the survey to identify the students who are interested in exceling in the course. These results are consistent with Batra & Klein (2020) who find that undergraduate business courses with class size of 27 or less benefit from various pedagogy that fits the students' learning style.

The results suggest that students are not distracted from working too many hours or taking too many classes at a small private school. This may be a result of more time advising students to ensure they do not overcommit. In comparison, students at a larger, public institution show some negative effects from these factors, implying they are left to make their own decisions regarding the number of hours they work or courses taken.

The study provides evidence that there is a strong significant association between students' self-perceived math abilities and their performance in the Managerial Accounting course. A small residential college allows faculty to interact more closely with students with lower abilities or lack confidence in math. This is possible at any institution, large or small, but the ability to work with students in a class of 24 is far

greater than one with 45 students. There is also some moderate significant association between students' self-perceived writing abilities and their performance in the Managerial Accounting course at the small college while none exists at a larger school.

Because of the strong significant association that exists between students' grades in the Financial Accounting course (ACC201 Grade) and OGPA and their performance in the Managerial Accounting course, accounting faculty might target students who earned lower grades in ACC201 and mentor them, knowing a priori, they will need more help, and especially if they are motivated to do well. This is certainly more doable at a smaller private school, than a larger public one. However, with the survey response faculty at a larger institution may be able to identify students who need help sooner rather than later.

The findings of the study are helpful in designing the Managerial Accounting course which is known to be difficult for most business students. A faculty could use some of the factors mentioned in the study to help design a course that will enhance learning among all levels of students.

STUDY LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

This study is subject to some limitations. One limitation is that the study examines one small private college contrasted to a larger public institution examined in past studies. A more comprehensive study of several small, private schools might be helpful to determine whether other private schools exhibit similar characteristics. It will allow us to see if all private schools are alike or not. Another limitation is the small sample size relative to the number of independent variables analyzed and, hence, the results may not be as robust as they would have been if the sample were larger. Thus, another suggestion for further research is to replicate the study using a somewhat larger sample at the same school.

REFERENCES

Almunals, T., Alfraih, M., & Alharbi, F. (2014). Determinants of accounting students' performance. *Business Education and Accreditation*, *6*(2), 1-9.

Al-Twaijry, A.A., & Abdulrahman, A. (2010). Student academic performance in undergraduate Managerial Accounting Courses. *The Journal of Education for Business*, *85*(6), 311-322.

Baldwin, B.A., & Howe, K.R. (1982). Secondary-level study of accounting and subsequent performance in the first college course. *The Accounting Review*, *57*(3), 619-626.

Batra, M.M., & Klein, A. (2020, June 1). Effective pedagogy for small undergraduate classes. *Journal of Education for Business*. https://doi.org/10.1080/08832323.2020.1770160

Bergin, L.J. (1983). The effect of previous accounting study on student performance in the first college-level financial accounting course. *Issues in Accounting Education*, *1*, 19-28.

Buckless, F.A., Lipe, M.G., & Ravenscroft, S.P. (1991). Do gender effects on accounting course performance persist after controlling for general academic aptitude? *Issues in Accounting Education*, *6*(2), 248-261.

Doran, B., Bouillon, M.L., & Smith, C.G. (1991). Determinants of student performance in accounting principles I and II. *Issues in Accounting Education*, *6*(1), 74-84.

Eskew, R.K., & Faley, R.H. (1988). Some determinants of student performance in the first college-level financial accounting course. *The Accounting Review*, *63*(1), 137-147.

Gist, W.E., Goedde, H., & Ward, B.H. (1996). The influence of mathematical skills and other factors on minority student performance in principles of accounting. *Issues in Accounting Education*, *11*(1), 49-60.

Gul, F.A., & Fong, S.C. (1993). Predicting success for introductory accounting students: Some further Hong Kong evidence. *Accounting Education: An International Journal*, *2*(1), 33-42.

Hao, Q., & Maksy, M.M. (2019). Factors Associated with Student Performance in Advanced Accounting: An Empirical Study at a U.S. Residential Public University. *Journal of Accounting and Finance*, *19*(4), 62-80.

Ibrahim, M.E. (1989). Effort-expectation and academic performance in managerial cost accounting. *Journal of Accounting Education*, *7*(1), 57-68.

Ingram, R.W., & Peterson, R.J. (1987). An evaluation of AICPA tests for predicting the performance of accounting majors. *The Accounting Review*, *62*(1), 215-223.

Lynn, S.A., & Robinson-Backmon, I. (2005). Course-level outcomes assessment: an investigation of an upper-division undergraduate accounting course and the factors that influence learning outcomes. *Journal of Accounting and Finance Research*, *13*(4), 133-140.

Maksy, M.M., & Rodriguez, C. (2017). Motivation and distraction factors associated with student performance in auditing: An empirical study at a residential university. *Journal of Accounting and Finance*, *17*(7), 173-193.

Maksy, M.M., & Rodriguez, C. (2018). Factors associated with student performance in managerial accounting: An empirical study at a New England public university. *Journal of Applied Business and Economics*, *20*(7), 102-125.

Maksy, M.M., & Yoon, M. (2019). Factors associated with student performance in cost accounting II: An empirical study at a US commuter public university. *Journal of Applied Business and Economics*, *21*(1), 71-88.

Murdoch, B., & Guy, P.W. (2002). Active learning in small and large classes. *Accounting Education*, *11*(3), 271-282. https://doi.org/10.1080/0963928021000031448

Tyson, T. (1989). Grade performance in introductory accounting courses: Why female students outperform males. *Issues in Accounting Education*, *4*(1), 153-160.

Wooten, T. (1998). Factors influencing student learning in introductory accounting classes: a comparison of traditional and nontraditional students. *Issues in Accounting Education*, *13*(2), 357-373.

APPENDIX A
STUDY FORMAL STATEMENTS OF HYPOTHESES

Motivation Factors

H_1: *There is a significant positive association between the grade the student intends to earn in the Managerial Accounting course and student performance in that course.*

H_2: *There is a significant positive association between the student's intention to take the CPA exam and student performance in the Managerial Accounting course.*

H_3: *There is a significant positive association between the student's intention to attend graduate school and student performance in the Managerial Accounting course.*

Distraction Factors

H_4: *There is a significant negative association between the student's average number of hours of work per week and student performance in the Managerial Accounting course.*

H_5: *There is a significant negative association between the student's job type (if it is not related accounting, or business in general) and student performance in the Managerial Accounting course.*

H_6: *There is a significant negative association between the number of semester courses a student is taking and that student's performance in the Managerial Accounting course.*

Self-Perceived Ability Factors

H_7: *There is a significant positive association between the student's self-reported writing ability and student performance in the Managerial Accounting course.*

H_8: *There is a significant positive association between the student's self-reported math ability and student performance in the Managerial Accounting course.*

H_9: *There is a significant positive association between the student's self-reported reading ability and student performance in the Managerial Accounting course.*

H_{10}: *There is a significant positive association between the student's self-reported listening ability and student performance in the Managerial Accounting course.*

Prior Ability Factors

H_{11}: *There is a significant positive association between the grade the student earned in the Financial Accounting course and student performance in the Managerial Accounting course.*

H_{12}: *There is a significant positive association between the student's overall GPA and student performance in the Managerial Accounting course.*

APPENDIX B

TABLE 1
DESCRIPTIVE STATISTICS OF THE STUDY VARIABLES

	N	Minimum	Maximum	Mean	Std. Deviation
Letter Grade[1]	46	0.00	4.00	2.78	0.891
Overall Points (in %)	46	19.00	93.00	75.72	12.968
Intended Grade[2]	46	2	4	3.58	0.621
ICPA[3]	31	0	3	1.87	0.885
IGS[3]	46	1	3	2.00	0.943
Job Hours	46	0	56	11.87	13.562
Job Type[4]	46	1	4	2.13	0.934
Course Load	46	3	7	5.38	0.936
Writing Ability[5]	46	2	4	2.76	0.766
Math Ability[5]	46	1	4	2.80	0.885
Reading Ability[5]	46	1	4	2.67	0.818
Listening Ability[5]	46	1	4	3.00	0.943
ACC 201 Grade[1]	45	1	4	3.38	0.658
OGPA (out of 4.0)	44	2.00	4	3.27	0.514

[1]A = 4.00; A- = 3.67; B+ = 3.33; B = 3.00; B- = 2.67; C = 2.00; D = 1.00; F = 0.00.
[2]An A 4.00; At least a B = 3.00; C is fine with me = 2.00
[3]No = 1; Maybe = 2; Yes = 3
[4]Do not work = 1; Other = 2; Business Related (but not accounting) = 3; Accounting related = 4.
[5]Very Good =4; Good =3; Average =2; Poor =1.

TABLE 2
ONE-WAY ANALYSIS OF VARIANCE
(All numbers are for Between Groups Only)

		Dependent Variables			
		Letter Grade (Grade)		**Overall Points % (Points)**	
Independent Variables	**DF**	**F Value**	**Sig.**	**F Value**	**Sig.**
Intended Grade	2/44	2.795	.072*	3.983	.026**
CPA	3/30	1.355	.278	1.298	.295
Grad School	2/45	0.820	.447	0.167	.847
Job Hours	21/45	0.666	.825	0.599	.880
Job Type	5/45	0.663	.654	0.968	.449
Course Load	5/45	1.080	.386	0.881	.503
Write	2/45	1.509	.233	0.696	.504
Math	3/45	1.678	.186	2.012	.127
Read	3/45	1.271	.297	1.094	.362
Listen	3/45	0.981	.411	0.617	.608
ACC 201 Gr	9/44	1.741	.116	1.092	.393
OGPA	28/43	2.051	.073*	2.539	.031**

*Significant at 10% level of significance using two tails test
**Significant at 5% level of significance using two tails test
***Significant at 1% level of significance using two tails test

TABLE 3
PEARSON/SPEARMAN CORRELATION COEFFICIENTS[a]

	Letter Grade	Points	IG	ICPA	Grad Sch IGS	Job Hours	Job Type	Course Load	Write	Math	Read	Listen	ACC 201 Gr.	OGPA
Letter Gr		.941***	.343**	.354*	.089	-.004	-.240	.211	.234	.262*	.141	.114	.441***	.444***
Points	.957***		.389***	.353*	-.011	-.022	-.236	.260*	.105	.266*	.023	-.020	.380***	.329**
IG	.313**	.402***		.071	.217	-.070	-.072	.061	-.125	.191	.030	-.217	.581***	.256*
ICPA	.379**	.431**	.041		.060	.363**	-.212	.365**	.145	.464***	.175	.154	.240	.340*
IGS	.128	.084	.231	.044		.071		.126	.123	.133	.259*	.150	.253*	.199
Job HRS	-.002	-.034	-.085	.314*	.140		.330**	.240	.220	.124	.275*	.226	-.028	.069
Job Type	-.201	-.243	-.081	-.080	.022	.499***		-.289*	-.022	-.119	-.173	-.064	-.092	-.222
C Load	.097	.173	.127	.327*	.158	.287*	-.017		.119	.091	.263*	.136	.051	.333**
Write	.249*	.154	-.082	.137	.130	.206	.059	.096		-.071	.724***	.554***	.007	.331**
Math	.319**	.363**	.192	.456***	.139	.178	-.094	.063	-.087		.125	.080	.509***	.507***
Read	.191	.160	.074	.173	.239	.247*	-.122	.216	.742***	.143		.461***	.154	.360**
Listen	.143	.033	-.237	.125	.154	.246*	.088	.041	.553***	.103	.451***		.017	.248
ACC 201 Gr.	.518***	.582***	.502***	.351*	.216	.145	-.086	.090	.063	.515***	.196	.001		.644***
OGPA	.508***	.523***	.320**	.353*	.227	.173	-.151	.248	.305**	.510***	.334**	.191	.693***	

a Pearson correlations are above the diagonal and Spearmen correlations are below the diagonal.
*Significant at 10% level of significance using two tails test
**Significant at 5% level of significance using two tails test
***Significant at 1% level of significance using two tails test

TABLE 4
PEARSON PARTIAL CORRELATION COEFFICIENTS
(CONTROLLING FOR ACC 201 Grade AND OGPA)

	Letter Grade	Points	IG	ICPA	IGS	Job Hours	Job Type	Course Load	Write	Math	Read	Listen
Letter Gr	1											
Points	.941***	1										
IG	.321*	.418**	1									
ICPA	.294	.321	-.099	1								
IGS	-.137	-.165	.182	.089	1							
Job HRS	-.100	-.087	-.067	.392**	.057	1						
Job Type	-.181	-.172	.078	-.171	-.112	.366*	1					
C Load	.224	.272	.144	-.348*	.275	.226	-.122	1				
Write	.139	.006	-.072	.114	.049	.100	-.035	.002	1			
Math	-.031	.106	-.167	.378*	.012	.094	-.105	.187	-.352*	1		
Read	-.135	-.261	-.034	.180	.320	.207	.001	.095	.794***	-.277	1	
Listen	.061	-.073	-.407**	.163	.038	.116	.019	.260	.464**	-.101	.341*	1

*Significant at 10% level of significance using two tails test
**Significant at 5% level of significance using two tails test
***Significant at 1% level of significance using two tails test

TABLE 5
REGRESSION ANALYSIS
(All numbers are for 46 Observations)

Independent Variables	Dependent Variables			
	Letter Grade		Overall Points %	
	t Coeff.	Sig.	t Coeff.	Sig.
Constant	-1.813	0.089*	-.510	0.617
IG	1.785	0.093*	2.493	0.024**
ICPA	1.871	0.080*	2.082	0.054*
IGS	-.438	0.667	-.717	0.484
Job Hours	-.602	0.555	-.547	0.592
Job Type	-.397	0.697	-.510	0.617
Course Load	.670	0.512	1.018	0.324
Write	1.778	0.094*	2.014	0.061*
Math	-0.515	0.613	0.046	0.964
Read	-2.063	0.056*	-2.546	0.022**
Listen	0.349	0.732	0.048	0.962
ACC 201 Gr	1.162	0.262	1.192	0.251
OGPA	0.481	0.637	-.364	0.721
Adj. R^2	0.309		0.434	
F	2.044	0.091*	2.789	0.029**

*Significant at 10% level of significance using two tails test
**Significant at 5% level of significance using two tails test
***Significant at 1% level of significance using two tails test

Insider Trading Behavior Following Analyst Earnings Forecast Revisions

Benjamin C. Anderson
San José State University

In this paper, I study insider trading behavior following analyst earnings forecast revisions. Contrary to the notion that insiders trade consistent with analyst earnings forecast revisions in order to take advantage of the post forecast revision drift, I find that insiders react opposite to analyst earnings forecast revisions and are more likely to sell stock after upwards earnings forecasts revisions and are more likely to purchase stock after downwards earnings forecast revisions. My findings are consistent with insiders trading opportunistically following analyst earnings forecasts revisions.

Keywords: Insider Trading, Analyst Earnings Forecast Revisions, Financial Analysts

INTRODUCTION

Corporate insiders are generally perceived by regulators, academics, and the business press to have superior information about the value of their own firms. Accordingly, prior research documents that trades by insiders of their own companies' stock (hereafter, "insider trades") are useful for predicting future stock returns and firm performance (Keown & Pinkerton, 1981; Piotroski & Roulstone, 2005; Cohen et al., 2012). Nonetheless, while insiders are generally assumed to have the highest quality and quantity of firm-specific information useful for determining their firms' value, prior empirical research also suggests that outsiders such as financial analysts provide incremental information about firm value and researchers speculate that this is due to broader knowledge of how industry-wide trends impact firm operations (Hutton et al., 2013). Thus, when analyst forecasts revise their estimates of future earnings, they provide incremental information useful for both capital markets and insiders. In this study I empirically examine insider trades that are preceded by analyst earnings forecast revisions (hereafter, "post forecast revision trades").

Insiders could agree with or learn from analysts and trade in a manner that is consistent with the information signaled by analyst earnings forecast revisions. Purchasing after an upward earnings forecast revision or selling after a downward earnings forecast revision signals to external capital market participants that their information is consistent with the information contained in analyst earnings forecast revisions and also allows insiders to take advantage of the subsequent drift that follows analyst earnings forecast revisions documented by Givoly and Lakonishok (1980). Alternatively, insiders could disagree with analysts and trade in the opposite direction of prior analyst forecast revisions due to private knowledge about firm value that differs from the information known by financial analysts and which is revealed through their earnings forecasts revisions. Since both insiders and analysts hold useful incremental information about firm value, I believe market participants and academic researchers will be interested in examining whether insider trading behavior is related to analyst earnings forecast revisions. Specifically, my study opens up a broad avenue of research allowing for the examination of circumstances in which the information in post forecast revision trades aligns with or differs from the information in analyst earnings forecast revisions.

Using a sample of 244,511 firm trade-days, I study the information content of insider trades and analyst forecast revisions. I obtain insider trades for all insiders required to file their trading activity with the SEC, including executives, directors, and large shareholders, and aggregate insider purchases and sales separately to single trade-days while removing days which have both purchases and sales. I obtain the mean analyst earnings forecast from the Institutional Brokers' Estimate System (I/B/E/S) summary file and measure the consensus analyst forecast revision as the difference between the most recent consensus analyst forecast before the trade date and the consensus analyst forecast one month prior. I find that insiders are more likely to trade in a manner that contradicts the information contained in analyst earnings forecast revisions. Specifically, I find strong evidence that insiders have a greater propensity to purchase their own companies' own stock following downwards earnings forecast revisions and sell their own companies' stock following upwards earnings forecast revisions. My finding is inconsistent with the notion that insiders seek to benefit from the post forecast revision drift that follows analyst earnings forecast revisions. My finding is consistent with insiders opportunistically trading on private information that differs from the information known by financial analysts and revealed by their earnings forecast revisions.

The remainder of the paper is organized as follows. Section 2 summarizes the literature and presents my formal hypothesis. Section 3 describes the data and empirical design. Section 4 describes the empirical results. Section 5 concludes the paper.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Insider Trading

Despite regulations such as the Securities and Exchange Act of 1934 that prohibit corporate insiders from benefiting from private information by trading their company's own stock, prior research documents that insider trades contain relevant information about the value of insiders' firms. Specifically, insider trades are associated with future earnings, the firms' market-to-book ratio, and are often contrary to prior changes in the firm's stock price (Piotroski & Roulstone, 2005). Insider trades are also associated with positive abnormal returns, particularly when they are diverge from a normal trading routine (Cohen et al., 2012) or when information asymmetry is high (Aboody & Lev, 2000; Huddart & Ke, 2007). There are two primary explanations as to why insider trades have information about firm value. First, despite the apparent illegality of trading on private information, insiders could still benefit from trading in their company's own stock because they believe they will be able to shield themselves from future litigation by carefully timing their trades using their superior information (Ke et al., 2003). Second, the information in insider trades could be caused by insiders attempting to credibly signal information about firm value (Fidrmuc et al., 2006).

Despite the opportunity for mangers to use their control over voluntary disclosure to benefit from insider trading, the empirical evidence is mixed regarding whether managers actually seize such opportunities. Noe (1999) finds that insiders do not appear to use their influence over discretionary disclosure in order to benefit from insider trades. Cheng and Lo (2006) further investigate the finding by Noe (1999) and find that managers tend to increase purchases following management forecasts that are negative in nature. However, they find no evidence that managers adjust the amount of selling they undertake following similar news disclosures. Their finding provides support for the notion that sales by insiders are given much greater scrutiny by regulatory authorities and so are less likely to be driven by managerial opportunism. In addition, capital market participants are able to assess the truthfulness of management guidance as it relates to insider trades and identify that management's willingness to misrepresent information varies with the capital market participants' ability to assess that bias (Rogers and Stocken, 2005).

Financial Analysts and Earnings Forecasts

Financial analysts discover and analyze information about the firms they follow in order to create financial reports used by capital market participants in order to make investing decisions. Lang and Lundholm (1996) find evidence most consistent with the notion that analysts act as intermediaries of information released by firm management. This notion is supported by other prior research which indicates

that analysts' strongest source of information is information given to them by managers at the firm they cover (Lees, 1981; Knutson, 1992) and that an analysts' decisions to follow firms are strongly related to the amount of information generated by a firm's management (Byrd et al., 1993; Lang & Lundholm, 1996). Other research further documents that analysts incorporate the information in management guidance into their forecasts (Jennings, 1987; Baginski & Hassell, 1990), particularly when management has established itself as having a strong forecasting reputation (Williams, 1996). However, analysts also play an important role as information discoverers (Livnat & Zhang, 2012). This is particularly true immediately prior to releases of information by firms and when firm disclosure is scarce (Chen et al., 2010). Analysts also appear to have an advantage in incorporating the effect of industry-wide information compared to managers (Hutton et al., 2012). Analyst earnings forecasts in particular are demonstrated to be useful to market participants and are superior to alternative time series forecasts for predicting earnings (Fried & Givoly, 1982; Brown & Hagerman, 1987), particularly in the short term (Bradshaw et al., 2012) and that the market responds strongly to whether companies are able to meet expectations formed by these forecasts (Kasznik & McNichols, 2002). The response of capital market participants to analyst forecast revisions us ultimately consistent with the notion that investors use analyst earnings forecasts to predict companies' future cash flows and, thus, fundamentally incorporate the information in them into their pricing of firms' common stock (Francis & Soffer, 1997; Gleason & Lee, 2003).

Hypothesis Development

There is considerable prior literature that studies the information content of insider trades and the motivations of insiders to purchase or sell stock in their own firms. There is also considerable prior literature that studies the information content of analyst forecasts and how the market processes the information in forecast revisions. When examining insider trading behavior following analyst earnings forecasts revisions, insiders could trade in a manner consistent with the information provided by the earnings forecast revision. In this case, insiders have information consistent with the information that they wish to profit from or alternatively seek to benefit from the post forecast revision drift that follows analyst earnings forecast revisions (Givoly & Lakonishok, 1980). Alternatively, insiders could trade in a manner inconsistent with the information provided by analyst earnings forecast revisions. If this is the case, insiders have information opposite to the information analysts have and reveal through their earnings forecast revisions and seek either to profit from this information, to credibly signal their private information to the market (Fidrmuc et al., 2006) or to mislead the market regarding future prospect for their firms (Benabou & Laroque, 1992). Since it is initially unclear how insider trading behavior change following analyst forecast revisions, I pose my hypothesis in the null form:

H1: *The likelihood of an insider to trade in the stock in their own firm is unrelated to analyst forecast revisions.*

DATA AND SAMPLE

Data

I gather data on insider trades from the Thomson Reuters Insiders Data Table 1 and consensus analyst forecast data from the I/B/E/S Summary File.[1] I gather financial statement information from Compustat and daily stock price data from the Center for Research in Securities Prices (CRSP). I restrict my analysis to only trades occurring from 2003 onward because of the substantial change in the regulatory environment caused by passage of the Sarbanes-Oxley Act of 2002 and Regulation FD. Finally, I winsorize all data at the 1 and 99% levels to mitigate the influence of outliers on my results.

I aggregate insider purchases and sales separately across all insiders by firm-day (hereafter, 'firm trade-day'). I do not net sales and purchases on a given trade-day in order to examine the asymmetric behavior that insider purchases and sales represent. Sales reflect an immediate payout or loss to the insider but purchases result in exposure to future changes in the company's stock price. Thus, I believe that there are considerable differences in the incentives between buying and selling and so I do not treat purchases as

'negative sales' and vice-versa. To control for any potentially confounding effects of days which have both a sale and a purchase, I delete any firm trade-days in which there are both purchases and sales.[2]

I measure analyst earnings forecast revisions ($Revision_{i,t}$) as the change in the consensus analyst forecast, measured as the change in the median analyst forecast as of each firm (i) trade-day (t) from 30 days prior to the firm trade-day until the trade-day, scaled by stock price (P) on the trade day, such that $Revision_{i,t} = (ConsenFC_{i,t} - ConsenFC_{i,t-30})/P_{i,t}$. This generates a signed change in the analyst forecast preceding each insider trade-day. Consistent with prior literature, I eliminate all firms with a stock price of one dollar in order to avoid the influence of penny stocks and in order to avoid the confounding effects of fractional denominators in the revision calculation.

To build my sample I first start with all insider purchases and sales available in Thomson Reuters beginning in January 2003.[3] Aggregating total insider trades (2,711,266 trades) to firm trade-days results in a sample of 642,263 firm trade-days. In order to create my analyst earnings forecast revision measure, I gather consensus analyst forecasts from the I/B/E/S Summary file.[4] Since my consensus analyst forecast revision variables are scaled by stock price, I eliminate all firms with a stock price of less than one dollar. This results in a sample of 338,594 total firm trade-day observations which I can associate with a prior consensus analyst forecast revision. I gather annual financial statement data from Compustat to create my control variables and calculate abnormal returns using daily stock price data from CRSP. This results in a final sample of 244,511 firm trade-day observations for my primary empirical tests. I present the details of my sample selection in Table 1:

TABLE 1
SAMPLE SELECTION

All insider purchases and sales from Thomson Reuters from 2003-2012 with transaction value data	2,711,266
Aggregated to firm trade-day	642,263
Less: Observations without analyst earnings forecast data	(269,740)
Less: Missing stock price data from CRSP	(30,056)
Less: Firms with stock price less than one dollar	(3,873)
Total firm trade-day observations for which measure can be created:	338,594
Data missing for control variables:	
Less: Missing returns data from CRSP	(41,280)
Less: Missing financial statement data from Compustat	(49,506)
Less: Trade-days with both purchase and sale	(3,297)
Final Sample (firm trade-days)	244,511

EMPIRICAL ANALYSIS

Univariate Analysis

I first examine how insider trading behavior changes following analyst forecasts using univariate analysis. In Table 2 below I sort insider trade-days into six different classifications according to the analyst forecast revision and whether insiders purchase or sell stock on that day. In order to simplify visual examination of univariate differences across each group, I present the table in two ways: in Panel A each group is presented in percentage terms of total insider trade-days, whereas in Panel B each group is presented in percentage terms of days in which insiders only purchase or sell stock.

TABLE 2
UNIVARIATE ANALYSIS OF POST FORECAST REVISION TRADING

PANEL A - PERCENT OF TOTAL INSIDER TRADES

	Purchase	Sale	Total
Upward Revision	4.78%	32.07%	36.85%
No Revision	6.93%	31.95%	38.89%
Downward Revision	6.76%	17.51%	24.27%
Total	18.47%	81.53%	100.00%
Difference between Upward Revision and Downward Revision	1.98%*** (<.0001)	14.56%*** (<.0001)	

PANEL B - PERCENT OF TOTAL INSIDER PURCHASES OR SALES

	% of Purchases	% of Sales
Upward Revision	25.88%	39.33%
No Revision	37.52%	39.20%
Downward Revision	36.60%	21.47%
Total	100.00%	100.00%
Difference between Upward Revision and Downward Revision	10.72%*** (<.0001)	17.86%*** (<.0001)

P-values are presented in parentheses and are based on two-tailed tests. ***, **, and * represents significance at the 1%, 5%, and 10% levels, respectively. Purchases Only is the subset of trade-days in which one or more insider purchases occur. Sales Only is the subset of trade-days in which one or more insider sales occur. Upward Revision equals 1 if the consensus analyst forecast rose prior to the insider trade. No Revision equals 1 if the consensus analyst forecast stayed the same prior to the insider trade. Downward Revision equals 1 if the consensus analyst forecast fell prior to the insider trade.

Insider purchases following downward analyst forecast revisions account for 6.76 percent of total insider trade-days and 36.60 percent of total purchase-days. Insider sales following upward revisions account for 32.07 percent of total insider trade-days and 39.33 percent of total sale-days. I find that insiders are significantly more likely to purchase following downwards analyst forecast revisions than after upwards analyst forecast revisions (p-value <0.0001) and that insiders are significantly more likely to sell following upwards analyst forecast revisions compared to after downwards forecast revisions (p-value <0.0001). Thus, I find univariate evidence that insiders change their trading behavior following analyst forecast revisions. Nonetheless, there are other factors that influence both insider trading behavior and analyst forecast revisions and so in the next section I adopt a multivariate regression approach to examine how insider trading behavior changes following analyst forecast revisions.

Multivariate Regression Design

In order to examine the impact of analyst forecast revisions on insider trading behavior, I estimate the following logistic regression with standard errors clustered by firm:

$$\begin{aligned}\text{Purchase}_{i,t} = {} & \alpha_1\text{UpwardRevision}_{i,t} + \alpha_2\text{NoRevision}_{i,t} + \alpha_3\text{DownwardRevision}_{i,t} \\ & + \alpha_4\text{Earnings}_{i,t} + \alpha_5\text{Loss}_{i,t} + \alpha_6(\text{Earnings}_{i,t}*\text{Loss}_{i,t}) + \alpha_7\text{Market Value}_{i,t} + \alpha_8\text{Book to Market}_{i,t} \\ & + \alpha_9\text{Dividends}_{i,t} + \text{Industry Fixed Effects} + \text{Year Fixed Effects} + \varepsilon_{i,t} \end{aligned} \quad (1)$$

I present variable definitions for each variable in equation (1) in Table 3. Note that all firm-specific variables are formed as of firm (i) trade-day (t):

TABLE 3
VARIABLE DEFINITIONS

Variable Name:	Definition:
$\text{Purchase}_{i,t}$	Dummy variable which equals 1 if there is one or more purchases by insiders, and 0 if there is one or more sales by insiders.
$\text{Upward Revision}_{i,t}$	Dummy variable which equals 1 if the consensus analyst forecast increased, and equals 0 otherwise.
$\text{No Revision}_{i,t}$	Dummy variable which equals 1 if the consensus analyst forecast does not change, and equals 0 otherwise.
$\text{Downward Revision}_{i,t}$	Dummy variable which equals 1 if the consensus analyst forecast decreased, and equals 0 otherwise.
$\text{Earnings}_{i,t}$	Earnings for the prior fiscal year (Compustat item OIADP)
$\text{Loss}_{i,t}$	Dummy variable which equals 1 if earnings is negative, and 0 otherwise.
$\text{Market Value}_{i,t}$	Market value, measured as common shared outstanding (Compustat item CSHO) times per share closing price at the end of the prior fiscal year (Compustat item PRCC_F).
$\text{Book to Market}_{i,t}$	Book to maket ratio, measured as the book value of equity (Compustat item AT minus compustat item LT) divided by market value (Compustat item CSHO times Compustat item PRCC_F).
$\text{Dividends}_{i,t}$	Dividends paid in the prior fiscal year (Compustat item DVC)
Industry Fixed Effects	Industry dummy variables, based on Fama-French 48 industries.
Year Fixed Effects	Year dummy variables, based on calendar years.

Descriptive Statistics

Table 4 below lists the variables used in my univariate and multivariate regression analysis and presents descriptive statistics for each.

Sales by insiders are substantially greater in magnitude and frequency compared to purchases by insiders for the firms in my sample. Specifically, 81.53 percent of the trade-days in my sample are due to sales whereas 18.47 percent of the trade-days in my sample are due to purchases. This is due to most companies offering employees stock in the companies they manage in order to align their incentives with shareholders. Thus, by definition, insiders will need to sell their own companies' stock much more often and in much greater amounts than how much they purchase. Nonetheless, the mean insider purchase is 13,440 shares of stock, representing a substantial increase in the investment in their own firms when insiders decide to purchase their own companies' stock. Next, I present multivariate regression results.

TABLE 4
DESCRIPTIVE STATISTICS

Variable	N	Mean	1Q	Median	3Q	Std. Dev.
Earnings	244,511	0.0668	0.0280	0.0730	0.1184	0.1143
Loss (Dummy)	244,511	0.1539	0.0000	0.0000	0.0000	0.3608
Size	244,511	5.3259	0.3102	0.8539	2.8112	15.5970
Market to Book	244,511	0.5006	0.2466	0.4116	0.6497	0.3850
Dividends	244,511	58.0181	0.0000	0.0000	14.4860	207.2793
Purchase Days (Dummy)	244,511	0.1847	0.0000	0.0000	0.0000	38.8100
Shares Purchased (in thousands)	244,511	13.4398	0.0000	0.0000	0.0000	1200.2372
Sales (Dummy)	244,511	0.8153	1.0000	1.0000	1.0000	0.3881
Shares Sold (in thousands)	244,511	70.6300	1.0000	6.3460	22.8500	1402.3488
Upward Forecast Revision (Dummy)	244,511	0.3685	0.0000	0.0000	100.0000	0.04824
No Forecast Revision (Dummy)	244,511	0.3889	0.0000	0.0000	100.0000	0.4875
Downward Forecast Revision (Dummy)	244,511	0.2427	0.0000	0.0000	0.0000	0.4287

Regression Results

I present results of estimating equation (1) using logistic regression with robust standard errors clustered by firm below in Table 5:

TABLE 5
MULTIVARIATE ANALYSIS OF POST FORECAST REVISION TRADES

Dependent Variable = 1 if Insider Purchase, 0 if Insider Sale	
Revision Variables:	
Upward Revision	-4.3720***
	(<.0001)
No Revision	-4.0073***
	(<.0001)
Downward Revision	-3.6285***
	(<.0001)
F-Test of Coefficients:	
Upward Revision - Downward Revision	-0.7435***
	(<.0001)
Control Variables:	
Earnings	0.7568
	(0.1131)
Loss	0.6267***
	(<.0001)
Earnings*Loss	-2.6076***
	(0.0003)
Market Value	-0.0254***
	(<.0001)
Book to Market	0.8578***
	(<.0001)
Dividends	0.0007***
	(0.0066)
Industry Fixed Effects	Included

Year Fixed Effects	Included
Observations	244511
Wald Chi-Squared Statistic	3565.53***

P-values are presented in parentheses and are based on two-tailed tests based on robust standard errors clustered by firm. *** represents statistical significance at the 1 percent level. Upward Revision is a dummy variable which equals 1 if the consensus analyst forecast rose prior to the insider trade, and equals 0 otherwise. No Revision is a dummy variable which equals 1 if the consensus analyst forecast stayed the same prior to the insider trade, and equals 0 otherwise. Downward Revision is a dummy variable which equals 1 if the consensus analyst forecast fell prior to the insider trade, and equals 0 otherwise. Control variables are as previously defined.

Consistent with results from my univariate analysis of the relation between analyst earnings forecast revisions and insider trading behavior, I find using multivariate regression results that insider trading behavior is contrary to the information provided by analyst earnings' forecasts. Specifically, I find that insiders are substantially less likely to sell their companies' own stock (and, thus, more likely to buy their companies' own stock) following downward earnings forecast revisions compared to upwards earnings forecasts revisions. My findings are consistent with the notion that insiders behave opportunistically following analyst earnings forecasts revision and insider trades reflect private information about the value of the firm and these trades can be perceived as a signal to the capital markets regarding this private information.

CONCLUSION

Insider trading poses a curious problem to academic research in accounting and finance. Specifically, insider trading is a heavily regulated activity with potentially massive consequences to insiders who are identified as taking advantage of private information to further their own personal wealth at the cost of outside investors. Insider trading is a fundamentally necessary activity due to many corporations choosing to compensate their managers using their own companies' stock in order to align their incentives with shareholders. Nonetheless, even though insider trading is highly regulated, insider trades are consistently documented to have incremental information about firm value. My paper seeks to examine the problem in a new light by examining insider trading behavior following major information events, analyst earnings forecast revisions. I find that insider trading behavior differs according to whether analyst earnings forecasts are revised upwards or downwards. My findings are inconsistent with the notion that insider trades are purely random or rhythmic. Instead, insiders are systematically responding to these external information events. My findings also provides evidence that insider trades are a valuable source of information to capital market participants and that insider trades are particularly useful information events when considered in the context of other events such as analyst earnings forecast revisions. Accordingly, my paper should be of interest to capital market participants who seek to identify information events relevant to determining firm value.

My paper should also be of interest to regulators and academics who are interested in identifying the information content of insider trades. My empirical findings are consistent with the notion that insider trades are responding to an important information event, analyst earnings forecast revisions, in a systematic way. Specifically, I find that insiders are more likely to sell following upward earnings forecast revisions and more likely to buy following downward earnings forecast revisions. Surprisingly, my findings are inconsistent with insiders seeking to benefit from the post forecast revision drift documented by Givoly and Lakonishok (1980). These insider trades will, in expectation, be unprofitable unless there is information known by insiders that is inconsistent with the changes in analysts' earnings forecasts. Accordingly, post forecast revision trades may be a means by which insiders signal to the capital markets beliefs that their firms are undervalued following downward earnings forecast revisions or overvalued following upward earnings forecast revisions while also serving to benefit themselves from this private contrarian information. My paper opens an avenue of research examining post forecast revision insider trades and

whether insider trades that are either consistent or inconsistent with prior earnings forecast changes provide unique information to the capital markets.

ENDNOTES

1. My findings remain similar if I use data from the I/B/E/S Analyst Detail File to generate consensus analyst forecasts.
2. My findings remain similar if I keep firm trade-days in which I allow both purchases and sales to occur on a given trade-day.
3. I begin my sample in 2003 due to the passage of the Sarbanes Oxley Act and Regulation Fair Disclosure in 2002. Both regulations substantially changed the role of managerial insiders in corporations and provided stricter limitations on the way that financial analysts can interact with corporate insiders.
4. I use the median analyst forecast in my reported results. My results remain similar when I use the mean analyst forecast instead.

REFERENCES

Aboody, D., & Lev, B.I. (2000). Information asymmetry, R&D, and insider gains. *Journal of Finance*, *55*(6), 315-346.

Baginski, S.P., & Hassell, J.M. (1990). The market interpretation of management earnings forecasts as a predictor of subsequent financial analyst forecast revision. *The Accounting Review*, *65*(1), 175-190.

Benabou, R., & Laroque, G.R. (1992). Using privileged information to manipulate markets; Insiders, Gurus, and Credibility. *The Quarterly Journal of Economics*, *107*(3), 921-958.

Bradshaw, M.T., Drake, M.S., Myers, J.N., & Myers, L.A. (2012). A re-examination of analysts' superiority over time-series forecasts of annual earnings. *Review of Accounting Studies*, *17*(4), 944-968.

Brown, L.D., Hagerman, R.L., Griffin, P.A., & Zmijewski, M.E. (1987). Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting and Economics*, *9*(1), 61-87.

Byrd, J.W., Johnson, M., & Johnson, M.S. (1993). *Investor relations and the cost of capital*. Working paper, University of Michigan, Ann Arbor, MI.

Chen, X., Cheng, Q., & Lo, K. (2010). On the relationship between analyst reports and corporate disclosures: Exploring the roles of information discovery and interpretation. *Journal of Accounting and Economics*, *49*(3), 206-226.

Cheng, Q., & Lo, K. (2006). Insider trading and voluntary disclosures. *Journal of Accounting Research*, *44*(5), 815-848.

Cohen, L.H., Malloy, C.J., & Pomorski, L. (2012). Decoding insider information. *Journal of Finance*, *67*(3), 1009-1043.

Fidrmuc, J.P., Goergen, M., & Renneboog, L. (2006). Insider trading, news releases, and ownership concentration. *Journal of Finance*, *61*(6), 2931-2973.

Francis, J.R., & Soffer, L.C. (1997). The relative informativeness of analysts' stock recommendations and earnings forecast revisions. *Journal of Accounting Research*, *35*(2), 193-211.

Fried, D., & Givoly, D. (1982). Financial analysts' forecasts of earnings: A better surrogate for market expectations. *Journal of Accounting and Economics*, *4*(2), 85-107.

Gleason, C.A., & Lee, C.M.C. (2003). Analyst forecast revisions and market price discovery. *The Accounting Review*, *78*(1), 193-225.

Givoly, D., & Lakonishok, J. (1980). Financial analysts' forecasts of earnings. *Journal of Banking and Finance*, *4*(2), 221-233.

Huddart, S.J., & Ke, B. (2007). Information asymmetry and cross-sectional variation in insider trading. *Contemporary Accounting Research*, *24*(1), 195-232.

Hutton, A.P., Lee, L.F., & Shu, S.Z. (2012). Do managers always know better? The relative accuracy of management and analyst forecasts. *Journal of Accounting Research*, *50*(5), 1217-1244.
Jennings, R.H. (1987). Unsystematic security price movements, managerial earnings forecasts, and revisions in analyst earnings forecasts. *Journal of Accounting Research*, *25*(1), 90-110.
Ke, B., Huddart, S.J., & Petroni, K.R. (2003). What insiders know about future earnings and how they use it: Evidence from insider trades. *Journal of Accounting and Economics*, *35*(3), 315-346.
Kasznik, R., & McNichols, M.F. (2002). Does meeting earnings expectations matter? Evidence from analyst forecast revisions and share prices. *Journal of Accounting Research*, *40*(3),727-759.
Keown, A.J., & Pinkerton, J.M. (1981). Merger announcements and insider trading activity: An empirical investigation. *Journal of Finance*, *36*(4), 855-869.
Knutson, P.H. (1992). *Financial reporting in the 1990s and beyond: A position paper of the Association for Investment Management and Research*. Working paper, University of Pennsylvania, Philadelphia, PA.
Lang, M.H., & Lundholm, R.J. (1996). Corporate disclosure policy and analyst behavior. *The Accounting Review*, *71*(4), 467-492.
Lees, F.A. (1981). *Public Disclosure of Corporate Earnings Forecasts*. New York, NY: The Conference Board.
Livnat, J., & Zhang, Y. (2012). Information interpretation or information discovery: Which role of analysts do investors value more? *Review of Accounting Studies*, *17*, 612-641.
Noe, C.F. (1999). Voluntary disclosures and insider transactions. *Journal of Accounting and Economics*, *27*(3), 305-326.
Piotroski, J.D., & Roulstone, D.T. (2005). Do insider trades reflect both contrarian beliefs and superior knowledge about future cash flow realizations? *Journal of Accounting and Economics*, *39*(1), 55-81.
Rogers, J.L., & Stocken, P.C. (2005). Credibility of management forecasts. *The Accounting Review*, *80*(4), 1233-1260.
Williams, P.A. (1996). The relation between a prior earnings forecast by management and analyst response to a current management forecast. *The Accounting Review*, *71*(1), 103-113.

Non-Performing Loans Ratio Measurement and Determinants Assessment

Aivars Spilbergs
BA School of Business and Finance, Riga

This article examines what macroeconomic factors affect the NPL ratio and how to better measure this ratio to reduce the valuation bias caused by credit growth over the reporting period. The study is based on current publications in the context of the topic and the macroeconomic indicators available in the Eurostat database for regression analysis during the period 2001-2018. As a result of the study, it is proposed to adjust the traditionally used NPL ratio with credit growth during the reporting period. Statistical tests that were carried out with 623 regression models provide strong evidence to conclude that the adjusted NPL ratio leads to higher explanatory power of macroeconomic indicators and thus increase the level of confidence of NPL ratio predictions.

Estimates show that NPL ratio is mostly determined by some key macroeconomic variables, such as loans to GDP ratio, loan growth, unemployment, foreign investment growth, household income growth, inflation rates, and others. Macroeconomic development tendencies must be carefully considered when formulating policies in order to reduce credit risk.

Keywords: non-performing loans ratio measures, influencing factors, econometric models

INTRODUCTION

Since the last financial crisis, which has had a major impact on the banking sector, much attention has been paid to non-performing loans (NPLs), whose volumes and dynamics have an impact on both the macroeconomic environment, the availability of credit in the financial markets, and all key aspects of bank management such as provisioning in accordance with International Financial Reporting Standards (IFRS 9), developing business strategies and plans, defining risk strategies and risk appetite framework, internal capital adequacy assessment process (ICAAP), recovery plan, etc. Large and/or fast-growing NPLs can affect the bank's short-term and long-term operation through two major channels. First, NPLs cause losses to banks, thereby reducing bank profitability and can affect bank capital when non-performing loans are written off. In severe cases, high NPLs can lead to bankruptcy. Second, high NPLs attract a significant amount of resources, both human and financial, making it difficult for banks to grant new loans to companies and individuals (European Commission, 2018).

According EU Regulation No 680/2014, non-performing loans and advances are those that satisfy any of the following criteria: (a) material exposures which are more than 90 days past due; (b) the debtor is assessed as unlikely to pay its credit obligations in full without realisation of collateral, regardless of the existence of any past due amount or of the number of days past due. (EU, 2014). European Banking Authority (EBA) defines NPL ratio in the following way: “The gross NPL ratio is the ratio of the gross

carrying amount of NPLs and advances to the total gross carrying amount of loans and advances subject to the NPE definition." (EBA, 2018)

Analyzing the causal relationships of NPLs, most studies focus on macroeconomic and bank-specific determinants of NPLs, but little attention is paid to the NPLs measure as such. What does it say? First, it must be understood that the NPL ratio is significantly influenced by credit growth - higher credit growth over the reporting period, at the same NPL amount, leads to a lower NPL ratio, and vice versa, because the denominator of the NPL ratio has changed. In addition, according to studies (Peric, 2017; Shahzad, 2019), credit quality deterioration and NPL status are not 'achieved' within six months and even a year since the issuance of the new loan. A shift in time further weakens the NPL ratio's ability to present a true picture of the real situation. Consequently, especially in periods with different growth rates of loans, the NPL ratio does not fulfill one of the essential criteria to be an unbiased measure of the actual situation. The same applies to comparisons between countries and regions with different growth rates of loans - the NPL ratio does not, by its nature, reflect these different situations. The above problems, at least in part, explain why correlation coefficients with theoretically correct NPL determinants are often lower and more volatile than expected. Therefore, this study seeks to identify ways to improve the NPL ratio information coverage.

LITERATURE REVIEW

Researchers have accumulated evidence that excessive loans growth and leverage, as evidenced by the rapid rise in the loan-to-deposit ratio, leads to a deterioration in credit quality and NPLs increase (Davis, 2008; Espinoza, 2010; Foos, 2010; Klein, 2013; Peric, 2017; Ozili, 2019; Shahzad, 2019). Some research models include lagged variables of loans growth (Jakubik, 2013; Klein, 2013; Skarica, 2014; Peric, 2017).

Significant growth in non-performing loans is a bad signal, as the increase in NPLs also increases the cost of financing for banks (Bholat, 2018). Often these costs are passed on to bank customers, potentially slowing down economic growth. Louzis et al. (2012), Shahzad et al. (2019) and other researchers confirms close relationship between the NPLs rapid growth and bank solvency.

Macroeconomic conditions form a link between the business cycle and lending, as environmental changes directly affect the borrower's ability to service debt. For example, research has often found that GDP growth shows a negative correlation with NPLs, indicating a countercyclical nature of NPLs (Davis, 2008; Espinoza, 2010; Drehmann, 2011; Nkusu, 2011; Louzis, 2012; Jakubik, 2013; Klein, 2013; Makri, 2014; Skarica, 2014; Beck, 2015; Cifter, 2015, Cucinelli, 2015; Filip, 2015; Tanaskovic, 2015; Anastasiou, 2016; Beaton, 2016; Gila-Gourgoura, 2017; Kupcinskas, 2017; Peric, 2017; Koju, 2018; Petkovski, 2018; Ozili, 2019; Radivojevic, 2019; Ari, 2020; Liu, 2020; Staehr, 2020).

With rising unemployment and falling wages, which are typically seen in times of economic downturns, borrowers face greater difficulties in repaying their debt and, as a result, NPLs increase. Many researchers explicitly include unemployment in their models and find strong positive relationships between unemployment and NPLs (Nkusu, 2011; Louzis, 2012; Klein, 2013; Makri, 2014; Skarica, 2014; Cifter, 2015; Cucinelli, 2015; Filip, 2015; Anastasiou, 2016; Koju, 2018; Kupcinskas, 2017; Petkovski, 2018; Spilbergs, 2020; Staehr, 2020).

In addition to the above, the following are also considered to be important determinants of NPLs:

- inflation, since its growth reduces real wages and hence ability to meet liabilities. This is particularly important in circumstances where inflation exceeds wage growth (Davis, 2008; Nkusu, 2011; Klein, 2013; Skarica, 2014; Filip, 2015; Koju, 2018; Petkovski, 2018; Liu, 2020; Staehr, 2020);
- variable interest rates, which directly affect the ability of borrowers to pay interest, especially when the proportion of variable rate loans is significant (Davis, 2008; Espinoza, 2010; Nkusu, 2011; Louzis, 2012; Beck, 2015; Peric, 2017);
- exchange rate depreciation, which may have a negative impact on NPLs, especially in a country with flexible exchange rate regimes and high amounts of foreign currency loans, may contribute to an increase in NPLs (Jakubik, 2013; Klein, 2013; Beck, 2015; Cifter, 2015 Tanaskovic, 2015; Koju, 2018);

- the house price index - falling house prices are tightly linked to higher default rates (Nkusu, 2011; Beck, 2015; Tajik, 2015; Staehr, 2020);
- foreign direct investment, the growth of which is usually conducive to economic growth and thus has a positive impact on NPLs (Cifter, 2015; Koju, 2018).

NPL RATIO ADJUSTMENT

NPL Ratio Measure

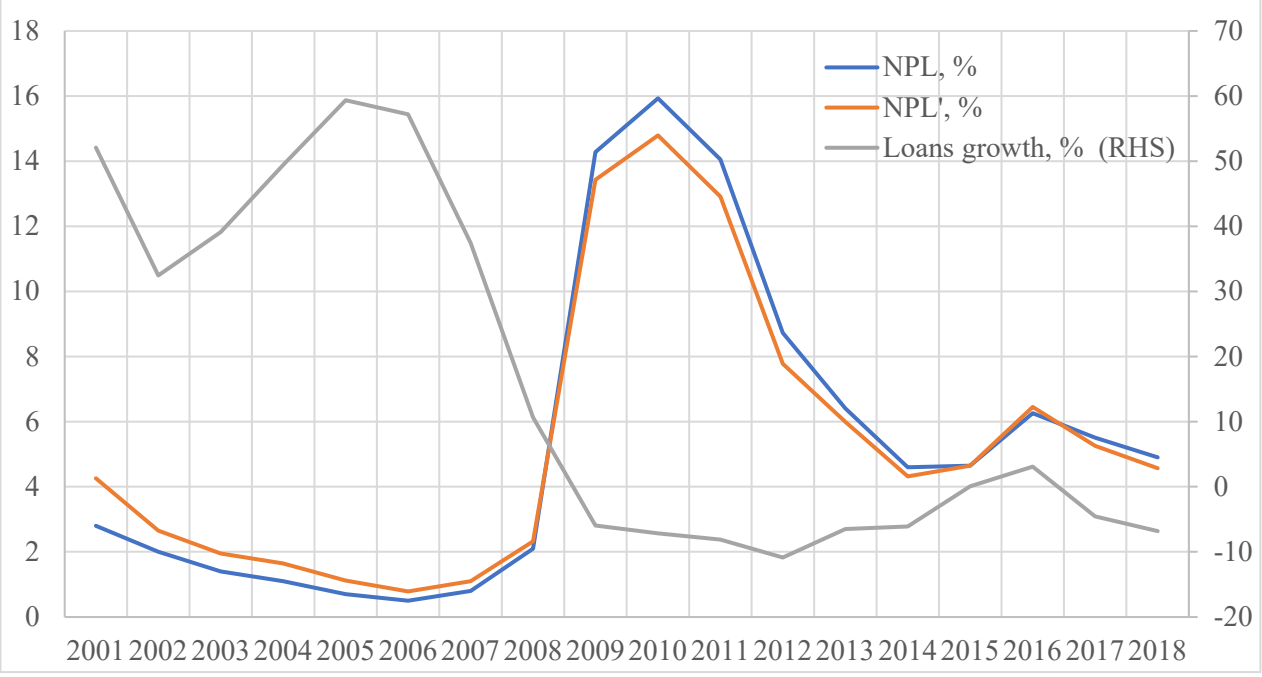
To eliminate the impact of current year loan growth and to reduce the NPL ratio valuation bias, thee proposal is to adjust the NPL ratio with the increase in loan balances ($\frac{Loan_t}{Loan_{t-1}}$) and estimate the adjusted NPL (NPL ') as follows:

$$NPL'_t = NPL_t * \frac{Loan_t}{Loan_{t-1}} = \frac{NPLA_t}{Loan_{t-1}} \quad (1)$$

where NPL_t - the unadjusted NPL ratio at the end of period *t*,
NPL_t' - the adjusted NPL ratio at the end of period *t,*
$NPLA_t$ - the NPL amount at the end of period *t,*
$Loan_t$ - loans outstanding amount at the end of period *t*,
$Loan_{t-1}$ - loans outstanding amount at the end of period *t*-1.

Using adjustment (1), the NPL increases relative to the unadjusted NPL during periods of loan growth and decreases when the loan outstanding amounts decreases. And the faster the increase or decrease, the greater the impact, see. the following Figure 1.

FIGURE 1
NPL, ADJUSTED NPL AND LOANS GROWTH IN LATVIA



As can be seen from Figure 1, the NPL adjustment provides a more conservative estimate of the NPL ratio in periods of strong credit growth and, conversely, in periods of decline. But does NPLs adjustment provide a better assessment of the influencing factors and a better quality of conclusions about the determinants of NPLs? This issue is relevant in the context of credit policy and credit strategy. To answer this question, the research hypothesized: NPL adjustment provides an opportunity to improve the stability of the regression coefficients and thus to more reasonably explain the effects of the macroeconomic environment.

NPL Ratio Associations With Risk Drivers

Based on the literature analysis, factors that could significantly influence the NPL ratio were selected for the hypothesis test and correlations were evaluated based on Eurostat data for Latvia for the period 2001-2018, see the following Table 1.

TABLE 1
MACRO INDICATORS CORRELATION WITH NPL

Indicator	Unit of measure	Denotation	Correlation with	
			NPL'	NPL
Household disposable income growth	%	*HDI*	-0,8971	-0,8859
Unemployment rate	%	*UNPL*	0,8226	0,7982
Net wages growth	%	*NWG*	-0,7812	-0,7598
Loans to GDP	%	*LtGDP*	0,6921	0,7163
Loans growth	%	*Loan*	-0,6894	-0,7243
GDP growth	%	*GDP*	-0,6785	-0,6727
Investments to GDP	%	*INV*	-0,6649	-0,6582
Private sector debt of GDP	%	*PSD*	0,6038	0,6350
House price index	%	*HPI*	-0,6007	-0,5892
Loan growth to GDP growth	%	*LgtGDPg*	-0,5950	-0,6487
Foreign direct investments growth	%	*FDI*	-0,5931	-0,5858

As can be seen from Table 1, in the 9 : 4 cases, the associations of determinants with the adjusted NPL are stronger than those of the unadjusted NPLs. The correlation coefficients in absolute values between 0,5 and 0,9 indicate moderate to strong correlations with the NPL, while their signs are consistent with those found in previous studies.

Figure 2 shows the associations of the major macro indicators with the adjusted *NPL*. Consequently, we see that household disposal income, net wages, loans and GDP growth as well investments to GDP increase, NPLs in turn decrease. And opposite, as unemployment, loans to GDP and private sector debt of GDP increase, so do *NPL*s. These observed trends, which are based on the historical data of Latvia in 2001-2018, are in line with theoretical researches.

The Model and Results

Let NPL_t' be dependent variable 'adjusted nonperforming ratio' in year *t*. Further, let $x_{1t},\ldots,x_{kt}$ denote independent variables and $b_{it},\ldots,b_{kt}$ denote regression coefficients of independent variables, than the model can be expressed as in equation:

$$NPL_t{}' = f(x_{1t}, \ldots, x_{kt}) + \varepsilon_t \quad (2)$$

where ε_t – the error term.

During the research, combining the selected factors, 623 regression models were calibrated which passed the F-test at the confidence level of 0,95 and the Durbin Watson test with α = 0,05. In the next step,

230 models were screened that passed the regression coefficient t-test at a confidence level of 0,95 (α = 0,05).

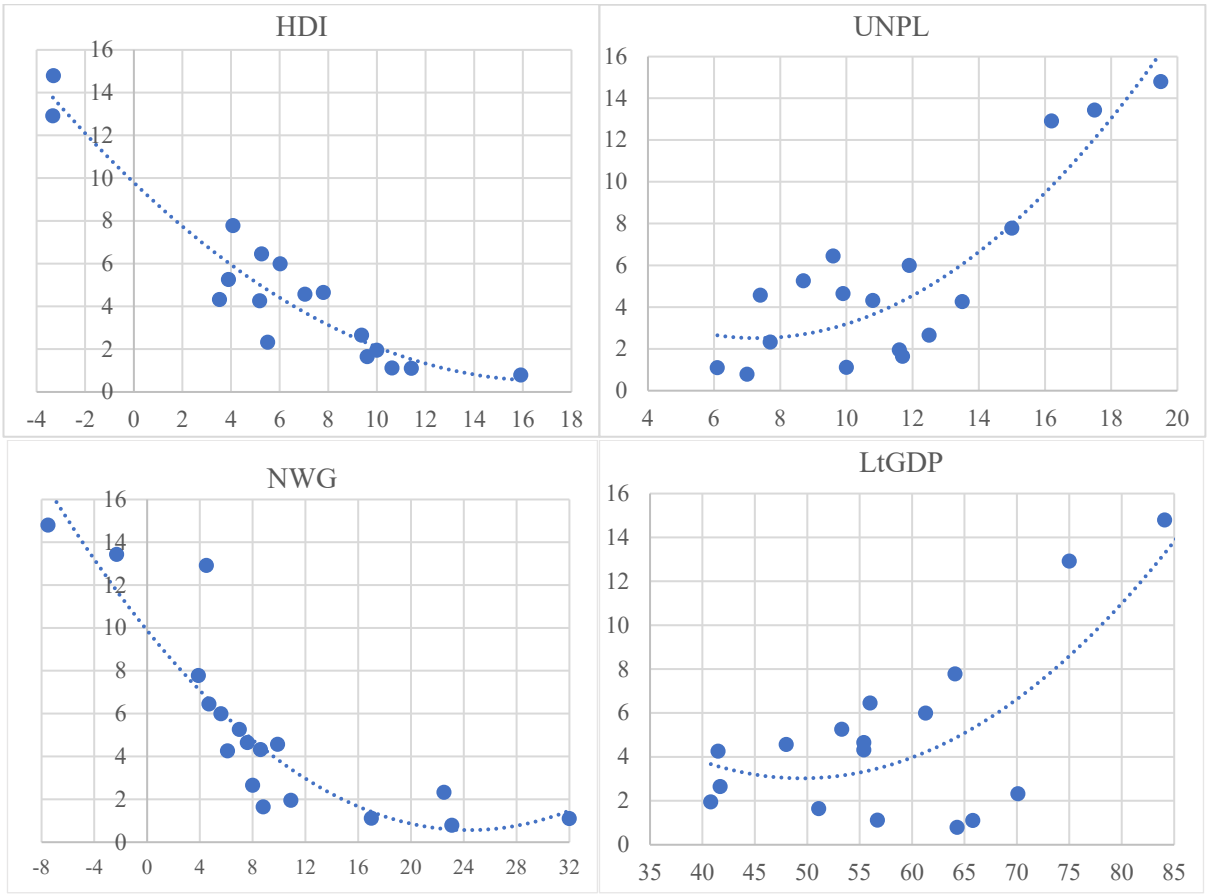
For research hypothesis testing following null and alternative hypothesis where stated:

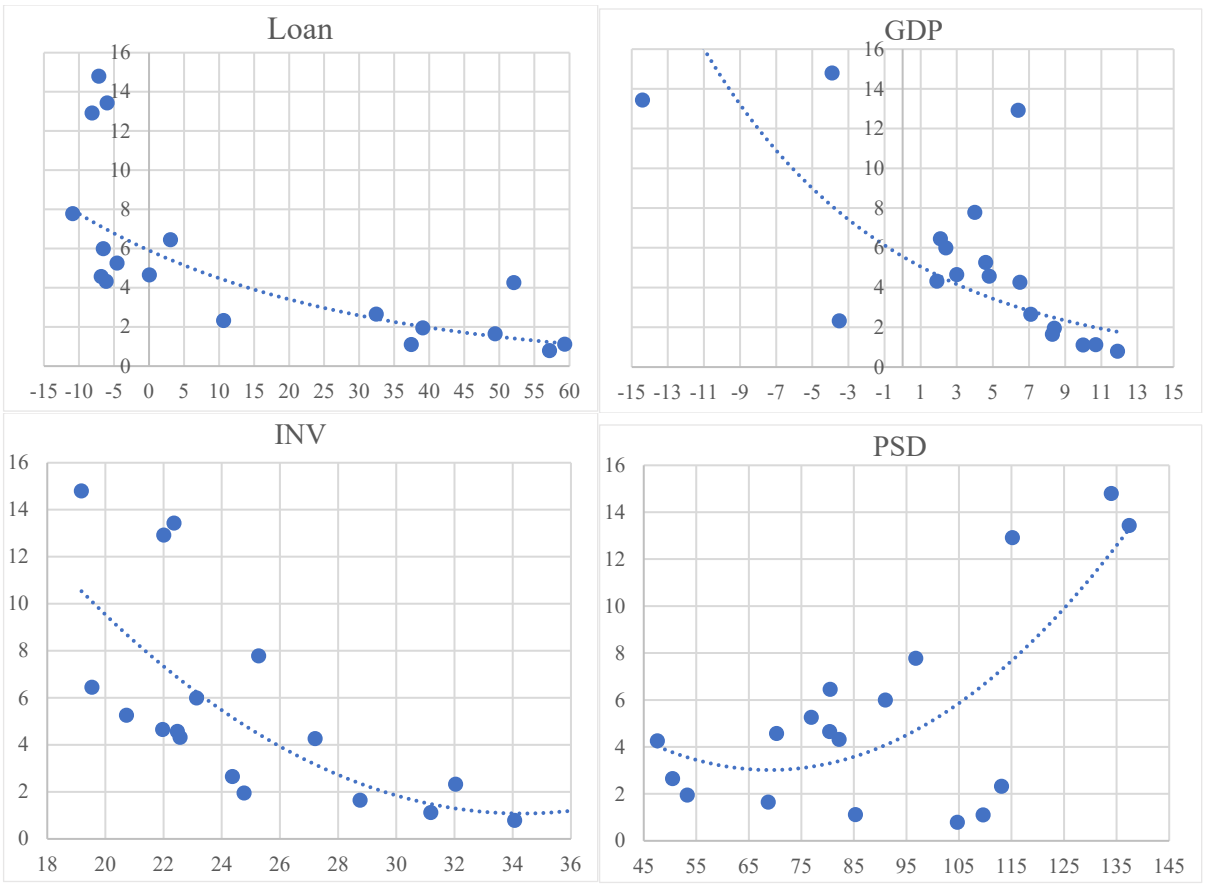
H_0: $\bar{p}_i - \bar{p}_i' \leq 0$ (3)

H_A: $\bar{p}_i - \bar{p}_i' > 0$ (4)

where $\bar{p}_i$ – the regression coefficients t-test average p-value for models with non-adjusted NPLs;
$\bar{p}_i'$ - the regression coefficients t-test average p-value for models with adjusted NPLs.

FIGURE 2
MACROECOMOMIC INDICATORS AND ADJUSTED NPL TRENDS





The following table summarizes the regression coefficients *t*-test *p*-values means and *t*-statistics for all 623 calibrated statistically significant models, as well as for the 230 screened. As shown in Table 2, for both regression model groups, *t*-stat > t_{crit} at a confidence level of 0,95 and thus null hypothesis can be rejected and alternative hypothesis accepted. From this analysis we see that there is strong evidence to conclude that adjusted *NPL* provide higher explanatory power of determinants included in regression models. This conclusion is supported also by low (<0,01) *t*-test *p*-values.

TABLE 2
REGRESSION COEFFICIENTS STATISTICS AND T-TEST RESULTS

Models	With NPL		With NPL'		*t*-statistic	df	t_{crit}	*p*-value
	Mean	Variance	Mean	Variance				
All 623	0,077223	0,018263	0,073685	0,014548	2,5092	1797	1,6457	0,0061
Selected 230	0,005318	0,000146	0,004364	0,000084	2,7874	470	1,6481	0,0028

Top 10 Best Fit Model's Statistics and Discussion

The following Table 3 summarizes the coefficients of determination, *F*-statistics and *p*-values for top 10 statistically significant models.

TABLE 3
TOP 10 REGRESSION MODEL'S STATISTICS

#	Variables	R^2	F	p - value
43	*NWG, LtGDP*	0,9125	78,238	<0,0001%
44	*NWG, PSD*	0,9106	76,387	<0,0001%
550	*UNPL, HDI, LgtGDPg*	0,9074	45,756	<0,0001%
551	*NWG, Loan, PSD*	0,8990	41,555	<0,0001%
435	*HDI, LtGDP, Loan*	0,8723	31,874	<0,001%
562	*NWG, Loan, LTR*	0,8683	30,765	<0,001%
387	*INV, Loan, LTR*	0,8605	28,775	<0,001%
490	*LtGDP, HCPI, FDI*	0,8575	28,093	<0,001%
546	*LtGDP, FDI, HPI*	0,8544	27,380	<0,001%
76	*LtGDP, HCPI*	0,8012	30,232	<0,001%

As one can see, the top 3 regression models explain more than 90%, while the top nine regression models explain at least 85% of the total NPL 's variability. This result is relatively high compared to the results reported in the studies (Espinoza, 2010; Foos, 2010; Nkusu, 2011; Klein, 2013; Skarica, 2014; Beck, 2015; Cucinelli, 2015; Filip, 2015; Tanaskovic, 2015; Beaton, 2016; Gila-Gourgoura, 2017; Kupcinskas, 2017). The *F*-test results show that the statistical stability of all top 10 models is high (<0,001%), but the probability of statistical error of the first four models is even lower than 0,0001%.

The following Table 4 summarizes the regression coefficients and *t*-test *p*-values for top 10 statistically significant models.

As one can see from Table 4, the regression coefficients *t*-test *p*-values for all top 10 models do not exceed 2,02%, but 23 out of 27 are less than 1%, indicating strong relationship between adjusted NPL's and relevant macro indicators. Meanwhile, the regression coefficient signs are as expected and consistent with those reported in most published studies. The most common macro indicators included in the top 10 models are "Loans to GDP", "Net wages growth" and "Loans growth", whose change per unit determines the change in adjusted NPL's by an average of 0,182 – 0,302; (-0,347) to (-0,176) and (-0,084) to (-0,043), depending on the other macro indicators included in a particular model, respectively.

TABLE 4
TOP 10 MODEL'S REGRESSION COEFFICIENTS AND T-TEST P-VALUES

Model 43 variable	Regression coefficient	p - value	Model 44 variable	Regression coefficient	p - value
NWG	-0,3142	0,00002%	NWG	-0,3467	0,000004%
LtGDP	0,1819	0,00015%	PSD	0,0901	0,000182%
Model 550 variable	Regression coefficient	p - value	Model 551 variable	Regression coefficient	p - value
UNL	0,4669	0,30017%	NWG	-0,3111	0,00080%
HDI	-0,3050	0,19438%	Loan	-0,0430	1,21938%
LgtGDPg	-0,3203	0,77603%	PSD	0,0736	0,00685%
Model 435 variable	Regression coefficient	p - value	Model 562 variable	Regression coefficient	p - value
HDI	-0,3778	0,08546%	NWG	-0,1764	0,48534%
LtGDP	0,3020	0,93635%	Loan	-0,0838	0,04360%
Loan	-0,1096	1,97106%	LTR	0,5979	0,04730%
Model 387 variable	Regression coefficient	p - value	Model 490 variable	Regression coefficient	p - value

INV	-0,3433	0,75501%	LtGDP	0,2319	0,00043%
Loan	-0,0722	0,43429%	HCPI	-0,4519	0,17191%
LTR	0,7844	0,00260%	FDI	-0,1098	1,68956%
Model 546 variable	Regression coefficient	*p* - value	Model 76 variable	Regression coefficient	*p* - value
LtGDP	0,2518	0,00014%	LtGDP	0,2516	0,00038%
FDI	-0,5157	2,01247%	HCPI	-0,6128	0,00905%
HPI	-0,4607	0,15385%			

CONCLUSIONS

Numerous studies on non-performing loans have been conducted since the global financial crisis, with particular attention to their macroeconomic determinants. Studies reveal that one of the important factors is the rapid loan growth in pre-crisis years, which in Latvia exceeded 48% on average in years 2001-2006. Such rapid credit growth does not allow for an objective assessment of the true quality of credit with the widely used NPL ratio. As research shows, only 1-2 years later the true ability of borrowers to make loan payments without delay appears.

This study proposes adjusting the traditionally used NPL ratio with credit growth during 12 month period. Therefore, the NPL amount at the reporting date is divided by the total outstanding loan amount one year in advance (1). The research demonstrates that the adjusted NPL ratio reduce the NPL ratio valuation bias and is more conservative in the assessment of credit quality in years of strong credit growth and thus provides more critical information for making credit decisions.

To evaluate the impact of the NPL ratio adjustment on credit risk drivers associations with macroeconomic indicators, combining 13 macro indicators selected based on literature research and correlation analysis results, 623 statistically significant regression models were generated in this study and compared with analogous regression models, where the non-adjusted NPL was included as the dependent variable. *p*-values difference in between regression coefficients *t*-test average *p*-value for models with non-adjusted and adjusted NPLs *t*-test at significance level of 0,95 provides sufficient evidence that the adjusted NPL ratio shows stronger associations with key macroeconomic indicators. Therefore, it can be concluded that the adjusted NPL ratio provides better possibilities to explain risk drivers impact on credit quality.

The quality of NPLs analysis can be improved by using multiple models in parallel to estimate expected levels of credit risk as a weighted average of multiple outcomes, for example, taking in account the availability and reliability of macro indicators forecasts, and the preference of each model to soft criteria and other considerations.

Similarly, before using developed and calibrated models, it is important not only to check their statistical stability, but also to carefully evaluate model residuals, as research (Spilbergs, 2020) shows that each model has different behaviors during different economic cycles and should not be ignored.

According to the study, the most influential macroeconomic indicators of credit quality are "Loans to GDP", "Loans growth", "Net wages growth" and others. In this respect, the study largely coincides with the previous and described in the literature.

It is important also to remember that once created econometric models cannot be considered to be of good quality over a longer period of time, as the environment is changing and therefore needs to be regularly checked and, if necessary, recalibrated.

REFERENCES

Anastasiou, D., Louri, H., & Tsionas, M. (2016). Determinants of non-performing loans: Evidence from Euro-area countries. *Finance Research Letters*, *18*, 116–119.

Ari, A., Chen, S., & Ratnovski, L. (2020). *The Dynamics of Non-Performing Loans During Banking Crises: A New Database*. ECB WP 2395.

Beaton, K., Myrvoda, A., & Thompson, S. (2016). *Non-Performing Loans in the ECCU: Determinants and Macroeconomic Impact*. IMF WP 16/229.

Beck, R., Jakubik, P., & Piloiu, A. (2015). Key Determinants of Non-performing Loans: New Evidence from a Global Sample. *Open Economies Review*, *26*(3), 525-550.

Bholat, D., Lastra, R., Markose, S., Miglionico, A., & Sen, K. (2018). Non-performing loans at the dawn of IFRS 9: regulatory and accounting treatment of asset quality. *Journal of Banking Regulation*, *19*(1), 33-54.

Cifter, A. (2015). Bank concentration and non-performing loans in Central and Eastern European countries. *Journal of Business Economics and Management*, *16*(1), 117-137.

Cucinelli, D. (2015). The Impact of Non-performing Loans on Bank Lending Behavior: Evidence from the Italian Banking Sector. *Eurasian Journal of Business and Economics*, *8*(16), 59-71.

Davis, E.P., & Karim, D. (2008). Comparing early warning systems for banking crises. *Journal of Financial Stability*, *4*(2), 89-120.

Drehmann, M., & Gambacorta, L. (2011). The effects of countercyclical capital buffers on bank lending. *Applied Economics Letters*, pp. 1–6.

EBA. (2018). *Guidelines on management of non-performing and forborne exposures* (pp. 1-130). Retrieved from https://eba.europa.eu/sites/default/documents/files/documents/10180/2425705/371ff4ba-d7db-4fa9-a3c7-231cb9c2a26a/Final%20Guidelines%20on%20management%20of%20non-performing%20and%20forborne%20exposures.pdf

Espinoza, R., & Prasad, A. (2010). *Nonperforming loans in the GCC banking system and their macroeconomic effects* (pp. 1-24). IMF WP, 10/224.

EU. (2014). *Annex V to EU Regulation No 680/2014*, *35*. Retrieved from https://eba.europa.eu/sites/default/documents/files/documents/10180/1017687/01ae88e7-b609-407a-8f70-8e41e6e4f65f/Annex%205%20%28Annex%205%20of%20EC%20Implementing%20Act%20on%20Reporting%29.doc.pdf?retry=1

European Commission. (2018). *Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms* (CRR).

Eurostat. (2020). Retrieved from https://ec.europa.eu/eurostat/data/

Filip, B. (2015). The quality of bank loans within the framework of globalization. *Procedia Economics and Finance*, *20*, 208-217.

Foos, D., Norden, L., & Weber, M. (2010). Loan growth and riskiness of banks. *Journal of Banking & Finance*, *34*(2), 2929-2940.

Gila-Gourgoura, E., & Nikolaidou, E. (2017). Credit risk determinants in the vulnerable economies of Europe: Evidence from the Spanish banking system. *International Journal of Business and Economic Sciences Applied Research*, *1*, 60-71.

Jakubik, P., & Reininger, T. (2013). Determinants of nonperforming loans in Central, Eastern and Southeastern Europe. *Focus on European Economic Integration*, *3*, 48-66.

Klein, N. (2013). *Non-performing loans in CESEE. Determinants and impact on macroeconomic performance* (pp. 1-26). IMF WP 13/72.

Koju, L., Abbas, G., & Wang, S. (2018). Do macroeconomic determinants of non-performing loans vary with the income levels of countries? *Journal of Systems Science and Information*, *6*(6), 512–531.

Kupcinskas, K., & Paskevicius, A. (2017). Key factors of non-performing loans in Baltic and Scandinavian countries: Lessons learned in the last decade. *Ekonomika*, *96*(2), 43-55.
Liu, L., Liu, Y-M., Kim, J-M., Zhong, R., & Ren, G-Q. (2020). Analysis of Tail Dependence between Sovereign Debt Distress and Bank Non-Performing Loans. *Sustainability*, *12*(747), 1-20.
Louzis, D., Vouldis, A., & Metaxas, V. (2012). Macroeconomic and bank-specific determinants of non-performing loans in Greece: A comparative study of mortgage, business and consumer loan portfolios. *Journal of Banking & Finance*, *36*, 1012-1027.
Makri, V., Tsagkanos, A., & Bellas, A. (2014). Determinants of non-performing loans- the case of Eurozone. *Panoeconomicus*, *2*, 193-206.
Nkusu, M. (2011). *Nonperforming loans and macro-financial vulnerabilities in advanced economies* (pp. 1-27). IMF WP 11/161.
Ozili, P.K. (2019). Non-performing loans in European systemic and non-systemic banks. *Journal of Financial Economic Policy*, pp. 1-19.
Peric B., & Konjusak N. (2017). How did rapid credit growth cause non-performing loans in the CEE countries? *Journal of Economics and Business*, *12*(2), 73-84.
Petkovski, M., Kjosevski, J., & Jovanovski, K. (2018). Empirical Panel Analysis of Non-Performing Loans in the Czech Republic. What are their Determinants and How Strong is Their Impact on the Real Economy? *Journal of Economics and Finance*, *68*(5), 460-490.
Radivojevic, N., Cvijanovic, D., Sekulic, D., Pavlovic, D., Jovic, S., & Maksimovic, G. (2019). Econometric model of non-performing loans determinants. *Physica A*, *520*, 481–488.
Shahzad, F., Fareed, Z., Zulfiqar, B., Habiba, U., & Ikram, M. (2019). Does abnormal lending behavior increase bank riskiness? Evidence from Turkey. *Financial Innovation*, *5*, 37. https://doi.org/10.1186/s40854-019-0152-2
Skarica, B. (2014). Determinants of non-performing loans in Central and Eastern European countries. *Financial Theory and Practice*, *38*(1), 37-59.
Spilbergs, A. (2020). Residential Mortgage Loans Delinquencies Analysis and Risk Drivers Assessment. *Emerging Science Journal*, *4*(2), 104-112.
Staehr, K., & Uuskula, L. (2020). Macroeconomic and macro-financial factors as leading indicators of non-performing loans. *Journal of Economic Studies*. https://doi.org/10.1108/JES-03-2019-0107
Tajik, M., Aliakbari, S., Ghalia, T., & Kaffash, S. (2015). House prices and credit risk: evidence from the United States. *Economic Modelling*, *51*, 123-135.
Tanaskovic, S., & Jandric, M. (2015). Macroeconomic and institutional determinants of non-performing loans. *Journal of Central Banking Theory and Practice*, *1*, 47-62.
World Bank, World Development Indicators. (2019). *Bank nonperforming loans to total gross loans (%).* Retrieved from https://data.worldbank.org/

US Banks Participation in Credit Derivatives and the Financial Crisis

Paul Abbondante
University of La Verne

Adham Chehab
University of La Verne

Yibo Xiao
University of La Verne

This paper was a study of US bank participation in credit derivatives around the financial crisis of 2008 by using data between 1997 and 2017. The results from a more detailed analysis showed that the financial crisis represented a transitional period for US banks' holdings of credit derivatives. The results indicate that US banks increased their hedging and their appetite for risk. The financial crisis also represented a structural change in type of loans held by participating US banks. The results show that US banks converged towards hedging around the financial crisis in credit derivatives.

Keywords: credit derivatives, hedge, banks, financial crisis

INTRODUCTION

As hedging instruments, credit derivatives were viewed as stabilizing instruments on the financial markets prior to the financial crisis because they provided banks with the flexibility to manage the credit risk of their assets independently of their holdings. Greenspan (2004) argued that banks were able to increase their interest income from assets while transferring the default risk to other less leveraged institutions.

To state that credit derivatives were used for hedging the banking system's credit risk may be an overstatement. As Figure 1 shows, the volume of outstanding credit derivatives exceeded total banking system assets beginning in 2005 and continued to rapidly increase through 2007. The use of credit derivatives continued to increase until the financial crisis. The literature provides varying motivations for utilizing credit derivatives by American banks. Minton, Stulz, and Williamson (2009) found that banks mainly used credit derivatives for speculation purposes rather than hedging. However, newer literature, post the financial crisis of 2007 and 2008, argued that banks used credit derivatives for hedging (Li and Marinč, 2014; Bliss, Clark, and DeLisle, 2018).

In this paper, we attempt to address the following questions using observations that encompass the financial crisis of 2007 and 2008:

(1) What are the characteristics of banks that participate in the credit derivatives market?

(2) What are the changes in the determinants of a bank's gross holdings of credit derivatives around the financial crisis?
(3) What are the changes in the determinants of a bank's net position as a guarantor or beneficiary in the credit derivatives market around the financial crisis?

The remainder of this paper is organized as follows. Section II presents a review of the disseminated literature and a discussion on the banks' holdings of credit derivatives. Section III presents a description of the sample and data sources. Section IV presents the empirical tests? and the regression results. Finally, section V presents the summary and conclusion of the paper.

LITERATURE REVIEW AND DISCUSSION

Origins of Credit Derivatives

Credit derivatives in various forms existed since the early 1990s (Smithson & Mengle 2006). Guill (2016) reported that Bankers Trust was active in developing risk management tools in the 1970s and 1980s. The Economist (2013, February 02) and Philips (2008) credits JP Morgan with creating the modern version of credit derivatives when JP Morgan sold a credit default swap to the European Bank of Reconstruction and Development to transfer the risk JP Morgan took on when it extended a $4.8 billion line of credit to Exxon to cover its potential $5 billion in punitive damages after the Exxon Valdez oil spill. According to Lanchester (2009), JP Morgan sold the credit default swap to improve its balance sheet.

Literature Survey

Since the financial crisis was a critical economic and political event in the US and the world, we review the credit derivatives literature in two sections. In the first section, we review the pre-financial crisis literature while the post-financial crisis is reviewed in the second section.

Pre-Financial Crisis Literature

Although the publication dates for some of the papers reviewed were after 2008, the sample data in these papers were observed before the financial analysis. The early part of the literature concentrated on the benefits and costs of credit derivatives and their use for transferring credit risk to third parties. Smith and Stulz (1985) showed that firms are more likely to hedge higher financial distress costs they face than other risks. The literature shows that banks use credit derivatives to hedge credit risk (Batten and Hogan, 2002; Cebenoyan and Strahan, 2004; Duffie, 2008). Brewer, Jackson, and Moser (1996) argued that hedging allows banks to increase diversification of their sources of income. Rule (2001) argued that credit derivatives are beneficial to banks because the separation of credit risk from the asset origination would provide stability to lending institutions and facilitate increased resource allocation efficiency. Duffie and Zhou (2001) showed that the source of the asymmetric information in the credit market determines the value of credit derivatives. In cases of adverse selection, banks would be better off while in the case of moral hazard, they would be worse off. The use of credit derivatives may also impact the economy in other ways.

The ability of the banks to separate the risk of the asset from holding it causes shifts in banks' behavior. In the past, investors viewed a bank's extension of credit to a corporation as a type of certification regarding the financial health of the borrower. Morrison (2005) argued that bank debt has a certification value to the financial markets. The existence of credit derivatives may cause banks to reduce the quality of their assets to sub-investment quality. Credit derivatives could reduce welfare by disintermediation in the credit market due to decreased quality and lack of certification of bank loans. In addition, the reduction of the signaling effect due to credit derivatives may lead to changes in financing decisions. Therefore, the lack of transparency would prevent banks from committing to holding the newly originated assets which leads to suboptimal investment and reduced welfare. The certification argument may no longer hold if banks are able to transfer the credit risk of corporate borrowers to third parties.

Instefjord (2005) showed that the use of credit derivatives to transfer risk depends on credit prices and the price elasticity of the underlying credit market. If the price elasticity is high, the stability of banks would

be threatened. If the price is inelastic, banks would be stabilized by credit derivatives. Instefjord (2005) also used cost of financial distress analysis and found that credit derivatives increase the benefit of risk sharing and transfer. Therefore, if banks use credit derivatives only to hedge their credit risk exposure in an elastic market, their risk increases, and they would destabilize.

The literature also studied the impact of credit derivatives on the banking sector and the economy by evaluating how innovation affects the credit market. Wagner (2007) showed that the value of innovation such as credit derivatives would depend on the state of the economy. In normal states of the economy, innovation in risk transfer would increase stability in the financial sector because banks are incentivized to reduce their risk. However, in recessions, such innovations would destabilize the financial markets because banks are incentivized to increase their risk. Rajan (2006) suggested that the world is better off due to the innovation of risk transfer and sharing because it expanded the credit market. Expanding the economy's credit capacity and, therefore, the credit market made companies and households better off.

The motivation behind the use of credit derivatives by banks was also investigated in the literature. Ashraf, Altunbas, and Goddard (2007) found that bank size was the main factor that determined the institution's participation in the credit derivatives market. They also found evidence that banks use credit derivatives to manage their credit risk exposure. Nicolò and Pelizzon (2008) studied the optimal credit derivative contract design under asymmetry of information. They proved that binary credit derivative contracts are optimal when banks are under strict capital loss requirements. Hirtle (2009) found only a limited relationship between the use of credit derivatives and increases in corporate loans and suggested that the benefits may be narrow. Minton, et al (2009) evaluated the motivation for banks to participate in the credit derivatives market. Their analysis questioned whether banks use the credit derivatives market for hedging. Using a detailed analysis of participating banks, they found evidence of speculation.

In summary, pre-financial crisis literature while arguing that credit derivatives were not always used for hedging or risk transfer, they generally were beneficial to banks and contributed to the stability of the credit markets. In addition, some argued that credit derivatives increased intermediation and helped expand the economy.

Post Financial Crisis Literature

We now turn our attention to the post financial crisis literature on credit derivatives. The newer literature mainly concentrated on the impact of the credit derivatives on the credit markets. Norden, Buston, and Wagner (2014) found that the banks' gross holdings of credit derivatives led to lower corporate credit spreads while net positions were not related to corporate loan pricing. They argued that the pricing reaction to holdings of credit derivatives was consistent with passing on the benefits of credit derivatives to their borrowers. They also did not observe a change in risk management due to the financial crisis. The literature also showed that credit derivatives were used for hedging financial risk. For example, Li and Marinč (2014) analyzed bank holding companies using observations between 1997 and 2012 and found that the use of credit derivatives was positively and significantly related to the institution's exposure to systematic risk. Their findings were consistent with Bliss, et al, (2018) who found that banks use credit derivatives in addition to other instruments to hedge financial stress risk exposure.

Luis, Rodriguez Gil, Sara, and Santomil (2015) argued that the view of credit derivatives has changed around the financial crisis from making banks sounder to taking the blame for the crisis. However, their empirical evidence showed that credit derivatives did not cause the financial crisis.

SAMPLE SELECTION AND VARIABLE DEFINITION

Sample

In this section we present the variables used in the regression analysis and their definition. We also describe the sample selection process and the data source.

To answer the three questions posed in the Introduction section of this paper, we evaluate the determinants of three dependent variables. The first dependent variable is *PARTICIPATE*. For the first regression set, we used logistic regression to determine the characteristics of banks that participated in the

credit derivatives market. Hence, *PARTICIPATE* is a dummy variable equal to 1 if the bank participated in the credit derivatives market and 0 otherwise. The second variable is *GROSS_HOLD*. We used the second regression set to analyze the determinants of a bank's gross holdings of credit derivatives. Since the analysis was performed on only a subset of banks that participated in credit derivatives, we used the 2-stage Heckman regression to account for any selection bias in the observations (Heckman, 1979). The third variable is *NETGRNTR*. We analyzed the determinants of a bank acting as a net guarantor or beneficiary in the credit derivatives market. Hence *NETGRNTR* is a dummy variable equal to 1 if the bank was a net guarantor and 0 otherwise. Since this analysis was also performed on a subset of banks, we used simultaneous logit regressions to account for any selection bias. We did not use 2-stage Heckman because the second stage would have to be OLS and we used logistic regression.

We obtained our data from the FDIC website statistics on depository institutions. The data can be accessed at https://www5.fdic.gov/sdi/download_large_list_outside.asp. The sample included all banks that reported $5 billion in total assets for any quarter between 1997 and 2017.

Table 1 shows the summary statistics. Although the number of US banks steadily declined by more than 50% between 1997 and 2017, the number of banks that reported $5 billion in total assets in any quarter for the same period decreased by about 44%. However, the number of participating banks increased by more than fivefold during the same period. Table 1 also shows that the size of the credit derivatives market exhibited rapid growth before 2008 and declined steadily afterwards. The average holdings exhibited a similar trend. Average holdings increased steadily through 2008 and started to decline afterwards. The summary statistics also show that the holding averages are much greater than the medians which indicates a concentration on credit derivatives in larger banks. The information presented in Table 1 indicates that the financial crisis years represented a structural change in the CD market.

Variables

We look to published literature to gauge the influence of the bank's portfolio on its participation in credit derivatives. Minton, Stulz, and Williamson (2009), Broccardo, Mazzuca, Yaldiz (2014), and Mattana, Petroni, & Rossi (2015) argued that the composition of the loan portfolio would impact its desire to hedge. Li, et al. (2013) found that agricultural loan (AGRILOANS) default rates are not higher or are lower than those of other bank assets. Therefore, increases in agricultural loans would require less hedging than other assets. Minton, Stulz, and Williamson (2009) argued that banks are more likely to hedge with increases in their holdings of commercial and industrial loans (COMM_IND). Ghosh, A. (2016) found that Commercial real estate loans (COMM_RE) and Construction loans (CONSTRCT) are cyclical in nature with the local economy and the overall GDP. The cyclical nature of these loans would present increased risk for the lending bank and, hence, increase the need for hedging. Rajaratnam, Beling, and Overstreet (2017) showed that banks would increase consumer loans beyond the ideal which would increase their risk. Therefore, banks with higher consumer loans would increase their participation in the credit derivatives market. Doukas and Melhem (1987) showed that the default rate in foreign loans is lower than that for domestic assets. Therefore, an increase in originating foreign loans would reduce the need for hedging. However, Minton, Stulz, and Williamson (2009) found that banks who originate foreign loans are more likely to hedge.

TABLE 1
SUMMARY STATISTICS OF BANKS PARTICIPATING IN THE CREDIT DERIVATIVE (CD) MARKET

Year	*Number of Participating Banks*	*Number of banks $5 billion total assets*	*Number of Banks on FDIC Website*	*Minimum CD holdings ,000s*	*Maximum CD holdings ,000s*	*Average CD holdings ,000s*
1997	11	373	10,946	10,000	7,526,000	1,778,741
1998	15	364	10,484	10,000	25,878,000	5,263,880
1999	20	358	10,240	10,000	41,911,000	5,705,154
2000	22	351	9,920	9,000	68,247,000	7,007,346
2001	26	352	9,630	6,793	271,673,000	16,171,481
2002	21	344	9,369	3,805	366,050,000	30,531,217
2003	25	335	9,194	1,345	577,693,000	40,023,429
2004	26	327	8,988	1,075	1,066,160,000	90,255,075
2005	29	322	8,845	65	2,301,064,000	200,748,428
2006	35	307	8,691	38	4,654,282,000	257,704,262
2007	36	294	8,544	175	7,900,570,000	440,617,651
2008	37	282	8,314	175	8,391,629,000	433,219,318
2009	35	265	8,021	301	6,079,453,000	403,203,371
2010	34	257	7,667	301	5,474,978,000	416,192,931
2011	35	249	7,366	3,097	5,775,740,000	421,683,305
2012	40	241	7,092	2,022	5,982,888,000	329,751,820
2013	43	236	6,821	47	5,334,563,000	260,233,625
2014	48	233	6,518	125	4,247,239,000	196,843,807
2015	52	223	6,191	240	2,893,039,000	134,344,765
2016	52	217	5,922	153	2,007,083,000	99,858,615
2017	58	209	5,679	85	1,664,568,000	72,151,615

The sample banks reported $5 billion in total assets for any quarter between 1997 and 2017. The data was obtained from the FDIC website, www.FDID.gov

To capture a bank's attitude towards risk, we use the bank's risk weighted assets (LN_RISK). A positive coefficient of risk assets indicates that the bank is using credit derivatives to hedge its risk. However, a negative sign indicates that the bank is using credit derivatives to speculate (Rajan, 2006; Ashraf, Altunbas, and Goddard, 2007). We also use gross loans and leases (LOAN_AST) and nonperforming loans (BAD_LOANS) to proxy the quality of the bank's portfolio. The sign of the coefficients would be similar to LN_RISK. Banks are highly regulated. Therefore, they have to maintain certain capital levels. We capture the impact of regulation on banks by using the bank's capital as a variable, Tier 1 capital (TIER_1). A positive coefficient for TIER_1 suggests that the bank is using credit derivatives when they have sufficient capital, while a negative sign indicates that the bank is attempting to hedge the possibility of default or financial distress (Ashraf, Altunbas, and Goddard, 2007).

Minton, Stulz, Williamson (2009) argued that banks would use all available instruments to manage risk. We capture the bank's use of hedging instruments using three variables, derivatives holdings including CD (DRVTVS), whether the bank holds interest rate contracts (HAS_RT), and whether the bank sells its loans (LOANSALE).

Other control variables. A more profitable bank is less likely to experience financial distress than a less profitable one. Therefore, we expect a negative coefficient for profitability (Minton, Stulz, Williamson, 2009; Broccardo, Mazzuca, Yaldiz, 2014). Following Minton, Stulz, Williamson (2009), we proxy profitability as net interest margin as a proxy for profitability (NIM). In addition, a bank's holding of liquid

assets enables the bank to take on more risk or increase its appetite for risk. We also consider the bank's investments. We use the bank's securities holdings (SC_AST) as proxy for investments. Securities are more liquid than other assets and, therefore, holding them reduces the bank's exposure to financial distress. We hypothesize that increased security holdings increase the probability the bank would act as a guarantor in the credit derivatives market.

We also control for the bank size using the natural log of the bank's total assets (LN_ASSET). According to the literature, the larger the bank the more likely it will participate in the credit derivatives market (Ashraf, Altunbas, and Goddard, 2007; Minton, Stulz, Williamson, 2009; Broccardo, Mazzuca, Yaldiz, 2014; Mattana, Petroni, & Rossi, 2015).

The variable definitions and calculations used in this paper are presented in Appendix A. To analyze the three questions posed in the introduction section of this paper, we base our hypotheses on the disseminated literature. We hypothesize that larger banks are more likely to participate in credit derivatives. We also expect hedging instruments to complement rather than substitute for each other. The type of loans banks hold is also a factor in determining a bank's holding of credit derivatives. Therefore, the paper will evaluate the impact of different types of loans on banks' participation in credit derivatives. We expect banks' holdings of credit derivatives to converge towards hedging because of the financial crisis. Finally, we also hypothesize that banks which act as net guarantors have an increased appetite for risk.

The change in the credit derivatives market around the financial crisis is also evident in Figure 1. As the figure shows, the sample banks' holdings of credit derivatives exceeded the entire banking system's loans before the financial crisis. The volume of credit derivatives holdings suggests that banks were speculating rather than hedging. This is consistent with the findings of Minton, Stulz, and Williamson (2009). Figure 1 also shows that the total assets and loans for banks with more than $5 billion in assets, while both greater than those of the participating banks, were smaller than their holdings of credit derivatives. The volumes of assets and loans appeared to be rather steady for the banks in this study as well as all the banks in the US during the 1997 through 2017 period. During the same period, the US banks holdings of credit derivatives exhibited an entirely different pattern. Figure 1 shows that there are three distinct holding patterns of credit derivatives. Bank holdings of credit derivatives steadily increased between 1997 through 2007, stayed roughly steady between 2007 and 2011, and declined after 2011. According to the Business Cycle Dating Committee, the National Bureau of Economic Research dating cycle, the Great Recession in the US started in June 2007 and ended in June, 2009. However, for the purposes of our paper, based on the banking industry's holdings of credit derivatives, we define the period 1997 through 2006 as pre-financial crisis, the period 2007 through 2011 as a transitional period, and the period 2012 through 2017 as post financial crisis.

EMPIRICAL RESULTS

In this section, we present the results of our analysis of the banks involvement in credit derivatives to answer the three questions posed by this paper in the introduction section.

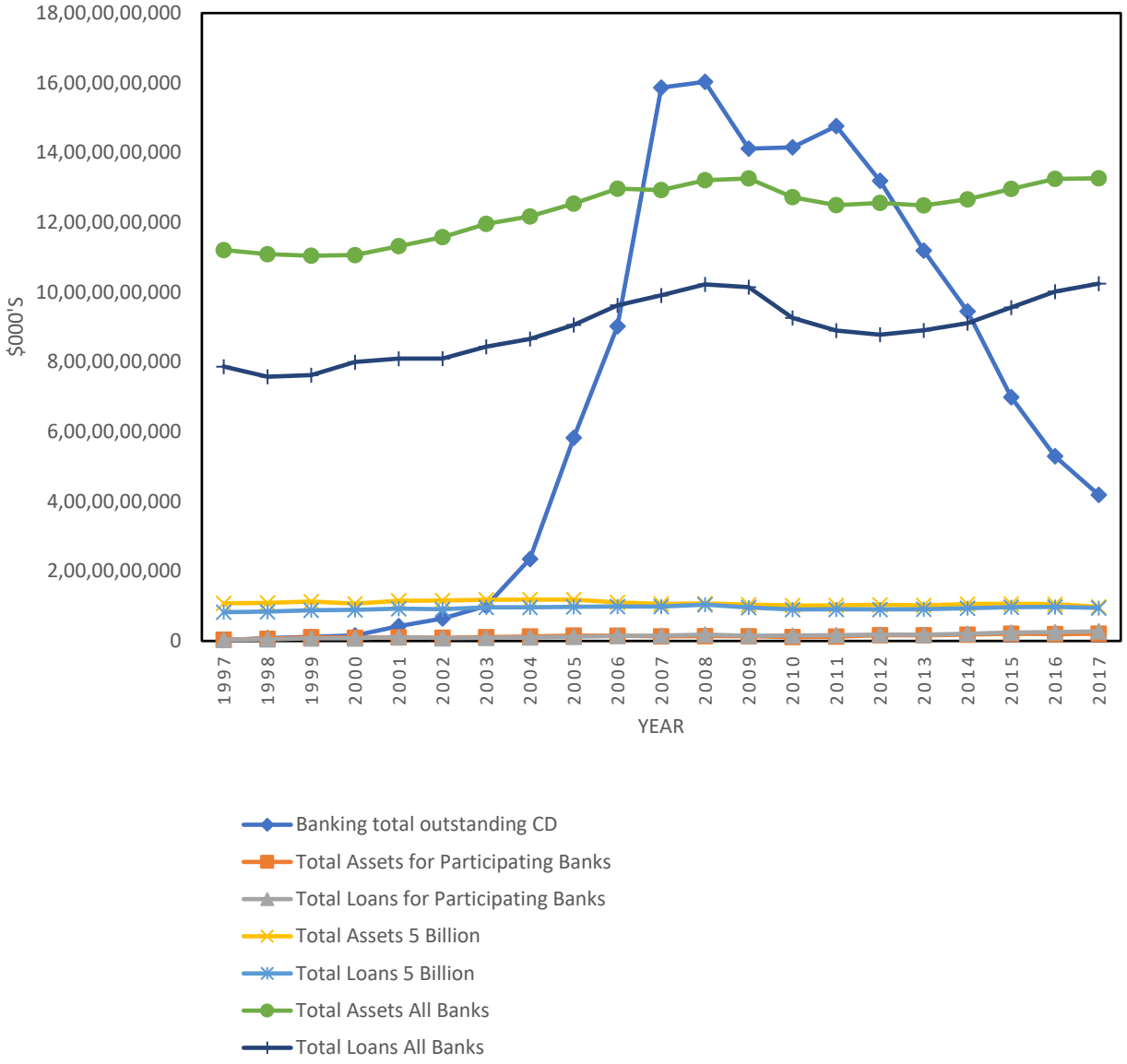
Determinants of Bank Participation in Credit Derivatives

To determine the characteristics of banks that participate in the credit derivatives market, we ran a logistic regression where the dependent variable, *PARTICIPATE*, is equal to one if the bank holds credit derivatives and zero otherwise. We used a synthesized matched pair technique. For each bank in the sample, we synthesized a matched bank by calculating the average value of the variables for nonparticipating banks with total assets equal to +/- 10% of the total assets of the sample bank. The number of banks used to create the synthesized matching bank varied for each bank.

We ran the regression as four models defined below to avoid correlation values of 0.3 or higher between the independent variables. Tier 1 capital, commercial and industrial loans, and agricultural loans were common to all four models. The additional independent variables are as follows. In model one, we used the natural logarithm of assets, liquidity, net interest margin, and derivatives holdings divided by assets. In model two, we used the natural logarithm of risk assets, securities holdings divided by assets, loans to

individuals, derivatives holdings, and construction loans. In model three, we used total loans divided by assets, securities holdings divided by assets, loans to individuals, liquidity, originates foreign loans, and construction loans. In model four, we used has-interest-rate contracts, nonperforming loans, total loans divided by assets, net interest margin, originates foreign loans, and construction loans.

FIGURE 1
SHOWS CREDIT DERIVATIVE HOLDINGS, ASSETS, AND LOANS FOR US BANKS FOR THE ANALYSIS PERIOD 1997 THROUGH 2017



The data was downloaded from the FDIC statistical data on depository institutions website https://www5.fdic.gov/sdi/download_large_list_outside.asp

Table 2 shows the results of the logistic regression. The results generally agree with the predictions of the theoretical part of the paper presented earlier. We found that banks are more likely to participate in the credit derivatives market if they utilize other hedging instruments. The coefficients for holding interest rate contracts and other derivatives were positive and significant at the 0.01 level. We also found that the size of the bank as measured in this paper, risk assets, and total assets were a positive and a significant contributor to the probability of participating in the credit derivatives market. Larger banks were more likely to participate in the credit derivatives market than smaller ones. The coefficient of Tier 1 capital was negative and significant indicating that banks were less likely to participate in the credit derivatives market with increases in their core capital. Increasing Tier 1 capital reduces the bank's risk. The results for the bank size and core capital variables indicate that larger banks are either more skilled in managing their risk which reduced their dependence on core capital or that they relied on credit derivatives to manage the risk of their assets. The results also indicate the sample banks were more likely to participate in the credit derivatives market with increases in their nonperforming loans and the size of their loan portfolios.

As hypothesized, we also found that the types of loans the bank held in its loan portfolio contributed to the likelihood of participating in credit derivatives. The results show that banks were more likely to participate in credit derivatives if they increased their holdings of wholesale loans. The coefficients for foreign and commercial and industrial loans were positive and significant. However, banks were less likely to participate in credit derivatives with increases in retail lending. The coefficients for loans to individuals, agricultural, and construction loans were negative and significant. Wholesale loans are less costly for the bank to originate but are usually larger than retail loans. We also report that banks were more likely to participate in this market if their liquidity increased and their profitability decreased. The coefficient for liquidity is positive and significant while that for profitability is negative and significant. Increases in liquidity enable a bank to fulfill credit derivatives claims from others in cases where it has acted as a guarantor. We test for which class of bank acted as a guarantor later in this paper.

Determinants of US Banks Gross Holdings

For the second step in our analysis of credit derivatives in US banks, we analyze the determinants of the gross holdings of credit derivatives. For the first stage of the Heckman regression, we used model 2 from the previous regression since it had the highest goodness-of-fit measure. The results are presented in Tables 3 through 6. Table 3 represents the regression results for the entire analysis period. Tables 4, 5, and 6 show the results for the three subperiods.

As in the case with other regressions in this paper, we ran the regression as three models to avoid correlation values of 0.3 or higher between the independent variables. Nonperforming loans, commercial and industrial loans, originated foreign loans, agricultural loans, sells loans, and liquidity were common to all models. The additional independent variables are as follows. In model one, we used commercial real estate loans and net interest margin. In model two, we used construction loans and net interest margin. In model three, we used loans to individuals.

Results for the Entire Period

As Table 3 shows, the sample banks increased their gross holdings of credit derivatives with increases in loans to consumers and to foreign entities. The coefficients of loans to consumers and foreign loans were positive and significant at the 1% level. In addition, selling loans, liquidity, and bad loans positively and significantly contributed to bank holdings of credit derivatives. The coefficient for agricultural loans was negative. However, the coefficient for agricultural loans was significant at the 10% level in only one model. Commercial real estate loans, construction loans, commercial and industrial loans, and profitability were not significant. We did not report the first stage Hickman results because they are a reproduction of model two in the first set of regressions in this paper.

Results for the Three Subperiods

We then drilled deeper into the data by analyzing banks' gross holdings of credit derivatives during the three subperiods described above. The results for the subperiods show that banks changed their credit

derivatives holdings around the financial crisis. This paper posits that the financial crisis represented a transition period for banks.

TABLE 2
LOGISTIC REGRESSION RESULTS FOR THE DETERMINANTS OF PARTICIPATION IN THE CREDIT DERIVATIVES MARKET

	Model 1	**Model 2**	**Model 3**	**Model 4**
INTERCEPT	-4.9696*** <.0001	-13.5053*** <.0001	-1.6063*** <.0001	-2.1594*** <.0001
HAS_RT				1.7006*** <.0001
LN_RISK	--	0.7524*** <.0001	--	--
LN_ASSET	0.2746*** <.0001	--	--	--
BAD_LOANS	--	--	--	1.9722*** 0.0076
LOAN_AST	--	--	0.1109*** <.0001	0.0897*** <.0001
SC_AST	--	0.3768 0.1193	-0.0085 0.9657	--
INDVIDUAL	--	-1.6077*** <.0001	-0.6057*** <.0001	--
LIQUID	0.3369** 0.0431	--	0.5520*** 0.0005	--
NIM	-0.2562*** <.0001	--	--	-0.2204*** <.0001
DRVTVS	3.1041*** <.0001	1.8478*** <.0001	--	--
HAS_FORN	--	--	1.1034*** <.0001	0.8803*** <.0001
CONSTRCT	--	-2.7621*** <.0001	-3.0277*** <.0001	-3.0122*** <.0001
TIER1	-0.0266*** <.0001	-0.0012 0.8471	-0.0150*** 0.0020	-0.0086 0.1842
COMM_IND	2.1192*** <.0001	1.7897*** <.0001	1.6443*** <.0001	1.2895*** <.0001
AGRILOANS	-6.4116*** 0.0024	-5.6904** 0.0193	-5.8956*** 0.0025	-5.6930*** 0.0077
McFadden's Pseudo R^2	0.2295	0.4176	0.2206	0.3268
Observations	6139	6139	6139	6139
***, **, * Represent significance at the 0.01, 0.05, and 0.10 level respectively				

TABLE 3
SHOWS THE SECOND STAGE OF THE 2-STAGE HECKMAN REGRESSION USED TO ESTIMATE THE DETERMINANTS OF A BANK'S GROSS HOLDINGS OF CREDIT DERIVATIVES FOR THE ENTIRE ANALYSIS PERIOD, 1997 THROUGH 2017

Heckman second stage	Model 1	Model 2	Model 3
INTERCEPT	-51.8102	-41.5741	-160.4866**
	0.6594	0.7146	0.0393
COMM_RE	56.6797	--	--
	0.8041		
CONSTRCT	--	-308.4908	--
		0.4688	
INDVIDUAL	--	--	446.0161**
			0.0246
NIM	-17.7367	-14.2482	--
	0.4249	0.5257	
BAD_LOANS	4042.0177***	4038.3914***	4091.7636***
	<.0001	<.0001	<.0001
COMM_IND	-208.6217	-205.564	-150.12
	0.1192	0.1247	0.2689
HAS_FORN	119.7501***	112.8377***	107.5206***
	0.0071	0.0069	0.0096
AGRILOANS	-3368.8605	-3315.5406	-3778.0887*
	0.1403	0.1433	0.0901
LOANSALE	4.2492E-5***	4.2281E-5***	4.1382E-5***
	<.0001	<.0001	<.0001
LIQUID	373.3637***	356.0227**	432.7728***
	0.0098	0.0134	0.0002
lambda	-34.9398	-27.8554	-42.4658
	0.3263	0.4288	0.2152
Adjusted R^2	0.3236	0.3241	0.3288
Observations, Stage 1	6139	6139	6139
Observations, Stage 2	701	701	701
***, **, * Represent significance at the 0.01, 0.05, and 0.10 level respectively			

As Tables 4, 5, and 6 show, the impact on profitability changes from negative and significant to positive and significant from before to after the crisis. The coefficient was not significant during the transition period. In addition, the coefficient on bad loans changed from not significant to positive and significant.

TABLE 4
REPORTS THE RESULTS FOR THE ANALYSIS REPORTED IN TABLE 3 BUT LIMITED TO PRE-FINANCIAL CRISIS

Heckman second stage	Model 1	Model 2	Model 3
INTERCEPT	125.566	126.028	-63.820367
	0.3051	0.3048	0.4638
COMM_RE	375.3705	--	--
	0.2741		
CONSTRCT	--	319.7199	--
		0.4594	
INDVIDUAL	--	--	441.0564**
			0.0365
NIM	-41.4552*	-35.2327	--
	0.0748	0.1135	
BAD_LOANS	-693.272	-713.753	-857.8228
	0.5676	0.5571	0.4778
COMM_IND	-40.3524	-7.3519	88.6765
	0.7791	0.9589	0.5514
HAS_FORN	23.4179	20.1405	4.2046
	0.5788	0.6321	0.9199
AGRILOANS	-8074.3793**	-8179.6909**	-9234.4700***
	0.0255	0.0242	0.008
LOANSALE	2.9336E-5***	2.9322E-5***	2.9650E-5***
	<.0001	<.0001	<.0001
LIQUID	111.4581	78.4373	212.5887
	0.4925	0.6201	0.1129
lambda	-69.3792**	-68.7923*	-84.2059**
	0.0455	0.0502	0.0155
Adjusted R2	0.3476	0.3458	0.353
Observations, Stage 1	3433	3433	3433
Observations, Stage 2	230	230	230
***, **, * Represent significance at the 0.01, 0.05, and 0.10 level respectively			

TABLE 5
REPORTS THE RESULTS FOR THE ANALYSIS REPORTED IN TABLE 3 BUT LIMITED TO THE FINANCIAL CRISIS

Heckman second stage	Model 1	Model 2	Model 3
INTERCEPT	-107.1455	-186.7337	-340.3328
	0.7401	0.5611	0.14
COMM_RE	-992.0266	--	--
	0.1479		
CONSTRCT	--	-778.1723	--
		0.4619	
INDVIDUAL	--	--	556.0186
			0.3767
NIM	1.4163	-0.8962	--
	0.9785	0.9865	
BAD_LOANS	6809.6355**	6496.4563**	6514.2878**
	0.0143	0.0197	0.0179
COMM_IND	-188.7571	-143.4996	-97.8816
	0.6265	0.7123	0.7983
HAS_FORN	192.4037	255.3888*	256.4912*
	0.1827	0.0622	0.0581
AGRILOANS	-11768	-13040	-13832
	0.1947	0.159	0.139
LOANSALE	5.4256E-5***	5.5038E-5***	5.4746E-5***
	<.0001	<.0001	<.0001
LIQUID	370.687	340.0817	497.2757
	0.3157	0.3979	0.1185
lambda	31.9992	13.1104	-5.5871
	0.7473	0.8943	0.9535
Adjusted R2	0.4009	0.3956	0.4
Observations, Stage 1	1347	1347	1347
Observations, Stage 2	177	177	177
***, **, * Represent significance at the 0.01, 0.05, and 0.10 level respectively			

TABLE 6
REPORTS THE RESULTS FOR THE ANALYSIS REPORTED IN TABLE 3 BUT LIMITED TO POST-FINANCIAL CRISIS

Heckman second stage	Model 1	Model 2	Model 3
INTERCEPT	-443.1540***	-476.0594***	-119.1197
	0.0094	0.0049	0.2209
COMM_RE	-285.1254	--	--
	0.2816		
CONSTRCT	--	-1288.1382	--
		0.1682	
INDVIDUAL	--	--	707.5014***
			0.0026
NIM	129.0527***	135.2479***	--
	0.0014	0.001	
BAD_LOANS	1696.0039	1688.1112	2409.3521**
	0.1056	0.1046	0.0212
COMM_IND	-47.2663	-14.6575	64.7188
	0.7992	0.9377	0.731
HAS_FORN	353.4062***	369.7283***	340.9970***
	<.0001	<.0001	<.0001
AGRILOANS	-902.908	-1004.53	-27.2719
	0.6534	0.6146	0.9889
LOANSALE	2.7974E-5***	2.7699E-5***	2.7263E-5***
	0.0002	0.0002	0.0003
LIQUID	551.8049***	550.0227***	156.1398
	0.0059	0.006	0.3279
lambda	-59.460805	-63.325086	-43.0509
	0.2387	0.1913	0.3531
Adjusted R2	0.3214	0.3231	0.3204
Observations, Stage 1	1359	1359	1359
Observations, Stage 2	294	294	294
***, **, * Represent significance at the 0.01, 0.05, and 0.10 level respectively			

While originating foreign loans and liquidity increased in significance, the significance of agricultural loans decreased around the financial crisis. However, the signs of the coefficients did not change. The results show that banks' use of credit derivatives converged towards hedging and increased their appetite for risk assets. We did not report the results for the first stage Heckman for the three subperiods since they are qualitatively similar to the ones for the entire period.

Determinants of US Banks Acting as Guarantor in Credit Derivatives

Finally, we analyze which banks act as net guarantors in the credit derivatives market. Again, we used model 2 from the first part of the analysis as one of the two simultaneous logistic regressions since it had the best-goodness-of-fit.

We ran the regression as four models to avoid correlation values of 0.3 or higher between the independent variables. Commercial and industrial loans, agricultural loans, and tier 1 capital were common to all models. The additional independent variables are as follows. In model one, we used total loans divided by assets, securities holdings divided by assets, loans to individuals, construction loans, originates foreign loans, and liquidity. In model two, we use the natural logarithm of assets, liquidity, derivatives, and net interest margin. In model three, we used total loans divided by assets, has-interest-rate contracts,

construction loans, originates foreign loans, nonperforming loans, and net interest margin. In model four, we used commercial real estate loans, construction loans, originates foreign loans, liquidity, derivatives, nonperforming loans, and net interest margin.

The results for the entire period are presented in Table 7 while the results for the three subperiods are presented in Tables 8 through 10.

Results for the Entire Period

As Table 7 shows, the coefficients for all types of domestic of loans investigated in this section of the paper were positive and significant. However, the coefficient for originating foreign loans was negative and significant. The coefficients for derivatives holdings and profitability were positive and significant, while Tier 1 capital and liquidity were negative and significant. The coefficients obtained in the regression indicate that the participating banks exhibited increases in their appetite for risk.

Results for the Three Subperiods

As Tables 8, 9, and 10 show, construction loans underwent a structural change. Construction loans changed from negative and significant during the pre-crisis period to positive and significant during the post crisis period. Liquidity increased in significance gradually from pre-crisis to post crisis while commercial and industrial loans lost significance over the three subperiods. The other variables did not exhibit significant changes during the three subperiods. The results confirm the findings of the previous section. US banks converged towards hedging and away from speculation in credit derivatives around the financial crisis. The increase in hedging may be a factor in US banks increasing their appetite for risk assets.

TABLE 7
REPORTS THE RESULTS OF THE LOGISTIC REGRESSION WHERE *NETGRNTR*, SET EQUAL TO 1 IF THE BANK WAS A NET GUARANTOR IN CREDIT DERIVATIVES AND 0 OTHERWISE

	Model 1	**Model 2**	**Model 3**	**Model 4**
Intercept	0.1836 0.5255	0.303 0.8443	-1.4341** 0.0234	-0.5227 0.1764
Loan_Ast	0.0643*** 0.0095	--	0.0858*** 0.0002	--
LN_Asset	--	-0.0622 0.519	--	--
HAS_RT	--	--	0.539 0.2791	--
COMM_RE	--	--	--	1.3818** 0.0344
Sc_Ast	0.7285 0.1006	--	--	--
INDVIDUAL	1.0635* 0.0667	--	--	--
LIQUID	-2.0579*** <.0001	-1.8279*** <.0001	--	-1.2141*** 0.0069

Drvtvs	--	0.8674** 0.0345	--	1.1099*** 0.0006
CONSTRCT	0.2097 0.8699	--	-0.4861 0.6941	-0.9170*** 0.454
HAS_FORN	-0.6202*** <.0001	--	-0.7083*** <.0001	-0.4893*** 0.0002
BAD_LOANS	--	--	-4.1505 0.1494	-3.7257 0.1794
NIM	--	0.1563** 0.0115	0.2971*** <.0001	0.1834*** 0.0053
COMM_IND	2.5644*** <.0001	2.2109*** <.0001	2.1353*** <.0001	2.2458*** <.0001
AGRILOANS	-17.3518** 0.0166	-12.7500* 0.063	-22.3679*** 0.0034	-21.2125*** 0.005
Tier1	-0.0832*** <.0001	-0.0567*** 0.0011	-0.0765*** 0.0002	-0.0770*** 0.0002
Rho	0.1760* 0.0787	0.2681*** 0.0048	0.1329 0.1895	0.1455 0.1839
Observations, Stage 1	6139	6139	6139	6139
Observations, Stage 2	701	701	701	701
***, **, * Represent significance at the 0.01, 0.05, and 0.10 level respectively				

TABLE 8
REPORTS THE RESULTS FOR THE ANALYSIS REPORTED IN TABLE 7 BUT LIMITED TO PRE-FINANCIAL CRISIS

	Model 1	**Model 2**	**Model 3**	**Model 4**
Intercept	-0.4432 0.3716	0.3888 0.9016	-0.752 0.5488	-1.4871** 0.0369
Loan_Ast	0.0888* 0.0873	--	0.0889* 0.0732	--
LN_Asset	--	-0.1339 0.4894	--	--
HAS_RT	--	--	-0.8018 0.4142	--
COMM_RE	--	--	--	0.6921 0.7387
Sc_Ast	-1.1384 0.2294	--	--	--
INDVIDUAL	1.2719 0.234	--	--	--
LIQUID	-1.3287* 0.0704	-0.3468 0.6882	--	-0.0653 0.9444
Drvtvs	--	-0.4563 0.6151	--	0.2224 0.7495
CONSTRCT	-3.8562* 0.0963	--	-5.3920** 0.0156	-5.0801** 0.0452

HAS_FORN	-0.4245** 0.04	--	-0.3833* 0.0686	-0.3326 0.1105
BAD_LOANS	--	--	-36.2614** 0.0147	-33.3563** 0.0241
NIM	--	0.2090* 0.0752	0.3158*** 0.002	0.2734** 0.0338
COMM_IND	3.3179*** <.0001	2.8028*** <.0001	3.1423*** <.0001	2.9670*** <.0001
AGRILOANS	22.4363 0.1988	15.7406 0.3403	5.633 0.737	11.9703 0.4741
Tier1	-0.0775** 0.011	-0.0502* 0.0916	-0.0501* 0.0937	-0.048 0.1044
Rho	0.2127 0.201	0.1725 0.3121	0.1652 0.4078	0.2634* 0.0959
Observations, Stage 1	3433	3433	3433	3433
Observations, Stage 2	230	230	230	230
***, **, * Represent significance at the 0.01, 0.05, and 0.10 level respectively				

TABLE 9
REPORTS THE RESULTS FOR THE ANALYSIS REPORTED IN TABLE 7 BUT LIMITED TO THE FINANCIAL CRISIS

	Model 1	**Model 2**	**Model 3**	**Model 4**
Intercept	-0.1029 0.8707	6.4781** 0.0392	-0.7507 0.5294	-1.0411 0.169
Loan_Ast	0.1904*** 0.0003	--	0.2040*** <.0001	--
LN_Asset	--	-0.4722** 0.0209	--	--
HAS_RT	--	--	-0.2663 0.7858	--
COMM_RE	--	--	--	3.7901** 0.0187
Sc_Ast	0.1605 0.8545	--	--	--
INDVIDUAL	1.2994 0.3618	--	--	--
LIQUID	-1.4994* 0.0914	-1.9597** 0.0113	--	-1.5135* 0.0945
Drvtvs	--	0.5316 0.5543	--	2.1773*** 0.0063
CONSTRCT	1.4157 0.583	--	1.2152 0.5801	-0.0738 0.9735
HAS_FORN	-0.8109*** 0.0025	--	-0.8945*** 0.0021	-0.3045 0.3241
BAD_LOANS	--	--	4.9271 0.3915	1.1602 0.8402
NIM	--	0.1920* 0.0521	0.1785* 0.0766	0.139 0.186

COMM_IND	2.5897*** 0.0039	3.5304*** 0.0002	2.8766*** 0.002	3.1904*** 0.0011
AGRILOANS	5.567 0.7599	18.1457 0.2942	10.3913 0.5654	8.547 0.6397
Tier1	-0.0880** 0.0323	-0.0809** 0.0419	-0.1005** 0.0253	-0.0807* 0.0677
Rho	0.0697 0.7422	0.037 0.8596	0.0065 0.9753	-0.0653 0.7775
Observations, Stage 1	1347	1347	1347	1347
Observations, Stage 2	177	177	177	177
***, **, * Represent significance at the 0.01, 0.05, and 0.10 level respectively				

TABLE 10
REPORTS THE RESULTS FOR THE ANALYSIS REPORTED IN TABLE 7 BUT LIMITED TO POST FINANCIAL CRISIS

	Model 1	**Model 2**	**Model 3**	**Model 4**
Intercept	0.3583	-3.443	-13.3555***	0.3332
	0.5524	0.1736	<.0001	0.6868
Loan_Ast	-0.0255	--	0.0072	--
	0.5316		0.847	
LN_Asset	--	0.2148	--	--
		0.166		
HAS_RT	--	--	12.3773***	--
			<.0001	
COMM_RE	--	--	--	-0.8559
				0.4274
Sc_Ast	1.6981**	--	--	--
	0.0282			
INDVIDUAL	2.5425**	--	--	--
	0.0101			
LIQUID	-2.7648***	-1.9091*	--	-1.6882**
	0.0003	0.0136		0.0463
Drvtvs	--	1.3310**	--	0.7132
		0.0264		0.1531
CONSTRCT	13.0948***	--	6.7659*	7.3962*
	0.0006		0.0665	0.0553
HAS_FORN	-0.3722	--	-0.3652	-0.324
	0.1322		0.1396	0.2197
BAD_LOANS	--	--	-13.9133**	-13.9285**
			0.0288	0.016
NIM	--	0.3751**	0.5533***	0.3199*
		0.0219	0.0001	0.0642
COMM_IND	2.0546**	2.6613***	1.5281*	1.5458*
	0.0181	0.0004	0.0856	0.0647
AGRILOANS	-33.2743***	-30.1298**	-45.6530***	-36.5506***
	0.0059	0.0125	0.0015	0.006
Tier1	-0.1077*	-0.1716***	-0.0980*	-0.1408***
	0.0515	0.0005	0.0604	0.0071

Rho	0.1024	0.3694*	-0.026	0.2045
	0.6376	0.0691	0.9034	0.4011
Observations, Stage 1	1359	1359	1359	1359
Observations, Stage 2	294	294	294	294
***, **, * Represent significance at the 0.01, 0.05, and 0.10 level respectively				

SUMMARY AND CONCLUSION

This paper was a study of US bank participation in credit derivatives. After their development, credit derivatives' popularity with US banks increased rapidly between 1997 and 2007, stayed stable between 2007 and 2011, and declined steadily after 2011.

The results for the first regression show that banks are more likely to participate in the credit derivatives market if they have interest rate contracts, are larger, have more risk assets, hold more nonperforming loans, have larger loan portfolios, hold other derivatives, originate foreign loans, have commercial and industrial loans, and have higher liquidity. We also find that banks are less likely to participate in the credit derivative market if they have higher Tier 1 capital.

The paper used 2-stage Heckman regression to investigate the determinants of gross holdings of credit derivatives by US banks. The regression was run as three models to avoid having variables with correlation coefficients of 0.3 or more in the regression. The results show that loans to consumers, foreign loans, loan sale, liquidity, and bad loans were significant contributors to gross holdings of credit derivatives. The coefficients of these variables were positive and significant at the 1% level. However, construction and commercial loans were not significant. The results for more detailed analysis of the three subperiods showed that the financial crisis represented a transitional period for US banks' holdings of credit derivatives. The results indicate that US banks increased their hedging and increased their appetite for risk assets.

Finally, we used simultaneous logistic regressions to determine which banks acted as guarantors in credit derivatives. The regression was run as four models to avoid having variables with correlation coefficients of 0.3 or more in the regression. The results show that banks with increases in all types of domestic loans, other derivatives holdings, and profitability were more likely to act as net guarantor. However, banks with increases in foreign loans, tier 1 capital, and liquidity were less likely to act as net guarantors. The results generally indicate that US banks converged towards hedging during and after the great financial crisis. Hedging may cause US banks to increase their appetite for risk assets and therefore returns.

REFERENCES

Ashraf, D., Altunbas, Y., & Goddard, J. (2007). Who transfers credit risk? determinants of the use of credit derivatives by large US banks. *The European Journal of Finance*, *13*(5), 483.

America's improved giant; AIG. (2013, February 2). *The Economist*, *406*, 59-60.

Batten, J., & Hogan, W., (2002). A perspective on credit derivatives. *International Review of Financial Analysis*, *11*(3), 251-278.

Bliss, B.A., Clark, J.A., & DeLisle, R.J. (2018). Bank risk, financial stress, and bank derivative use. *The Journal of Futures Markets*, *38*(7), 804-821.

Brewer, E.I., II, Jackson, W.E., Jr., & Moser, J.T. (1996). Alligators in the swamp: The impact of derivatives on the financial performance of depository institutions. *Journal of Money, Credit, and Banking*, *28*(3), 482-497.

Broccardo, E., Mazzuca, M., & Yaldiz, E. (2014). The use and determinants of credit derivatives in Italian banks. *The Journal of Risk Finance*, *15*(4), 417-436.

Cebenoyan, A.S., & Strahan, P.E. (2004). Risk management, capital structure and lending at banks. *Journal of Banking and Finance*, *28*(1), 19-43.

Doukas, J., & Melhem, M. (1987). Canadian banks: Risk reduction by international diversification. *Applied Economics*, *19*(12), 1561.
Duffie, D. (2008). *Innovations in credit risk transfer: Implications for financial stability*. BIS Working Paper no 255. Bank for International Settlements.
Duffee, G.R., & Zhou, C. (2001). Credit derivatives in banking: Useful tools for managing risk? *Journal of Monetary Economics*, *48*, 25-54.
Instefjord, N. (2005). Risk and hedging: Do credit derivatives increase bank risk? *Journal of Banking and Finance*, *29*(2), 333-345.
FDIC statistics on depository institutions. (n.d.). Retrieved from https://www5.fdic.gov/sdi/download_large_list_outside.asp
Ghosh, A. (2016). Do real estate loans reflect regional banking and economic conditions? *Journal of Financial Economic Policy*, *8*(1), 37-63.
Guill, G.D. (2016). Bankers Trust and the Birth of Modern Risk Management. *Journal of Applied Corporate Finance*, *28*(1), 19–29.
Greenspan, A. (2004). Economic flexibility. *Speech before the HM Treasury Enterprise Conference*. London, England. Retrieved from https://www.federalreserve.gov/boarddocs/speeches/2004/20040126/default.htm
Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica* (Pre-1986), *47*(1), 153.
Hirtle, B. (2009). Credit derivatives and bank credit supply. *Journal of Financial Intermediation*, *18*(2), 241-252.
Lanchester, J. (2009) Outsmarted. *New Yorker*, *85*(16), 83–87.
Li, S., & Marinč, M. (2014). The use of financial derivatives and risks of U.S. bank holding companies. *International Review of Financial Analysis*, *35*, 46.
Li, X., Escalante, C.L., Epperson, J.E., & Gunter, L.F. (2013). Agricultural lending and early warning models of bank failures for the late 2000s great recession. *Agricultural Finance Review*, *73*(1), 119-135.
Luis, O.G., Rodríguez Gil, L.I., Sara, C.A., & Santomil, P.D. (2015). The effect of credit derivatives usage on the risk of European banks. *Revista De Economía Mundial*, *40*.
Mattana, P., Petroni, F., & Rossi, S.P.S. (2015). A test for the too-big-to-fail hypothesis for european banks during the financial crisis. *Applied Economics*, *47*(4), 319.
Minton, B.A., Stulz, R., & Williamson, R. (2009). How much do banks use credit derivatives to hedge loans? *Journal of Financial Services Research*, *35*(1), 1-31.
Morrison, A.D. (2005). Credit derivatives, disintermediation, and investment decisions. *The Journal of Business*, *78*(2), 621-647.
National Bureau of Economic Research, Business Cycle Dating Committee. (n.d.). Retrieved from https://www.nber.org/cycles/sept2010.html
Nicolò, A., & Pelizzon, L. (2008). Credit derivatives, capital requirements and opaque OTC markets. *Journal of Financial Intermediation*, *17*(4), 444-463.
Norden, L., Buston, C.S., & Wagner, W. (2014). Financial innovation and bank behavior: Evidence from credit markets. *Journal of Economic Dynamics and Control*, *43*, 130.
Philips, M. (2008). The Monster That Ate Wall Street. *Newsweek*, *152*(14), 46.
Rajan, R.G. (2006). Has finance made the world riskier? *European Financial Management*, *12*(4), 499.
Rajaratnam, K., Beling, P., & Overstreet, G. (2017). Regulatory capital decisions in the context of consumer loan portfolios. *The Journal of the Operational Research Society*, *68*(7), 847-858.
Rule, D. (2001, June). The credit derivatives market: its development and possible implications for financial stability. *Financial Stability Review*, pp. 117-140.
Smith, C.W., & Stulz, R.M. (1985). The determinants of firms' hedging policies. *Journal of Financial and Quantitative Analysis*, *20*(4), 391.
Smithson, C., & Mengle, D. (2006). The Promise of Credit Derivatives in Nonfinancial Corporations (and Why It's Failed to Materialize). *Journal of Applied Corporate Finance*, *18*(4), 54–60.

Wagner, W. (2007). The liquidity of bank assets and banking stability. *Journal of Banking and Finance*, *31*(1), 121.

APPENDIX A

Variable definitions. The first column shows the name of the variable as used in the paper, the second column shows the definition of the variable and how it is calculated, and the third column shows the name and calculations of the variable as listed on the FDIC statistics on depository institutions website.

Variable	Definition	Calculation
PARTICIPATE	Participated in the credit derivatives (CD) market. Dummy variable = 1 if the bank participated, 0 otherwise	idctder > 0
NETGRNTR	Net Guarantor indicates that the bank's holdings of CD as a guarantor exceed its holdings as a beneficiary. Dummy variable =1 if the bank acted as net guarantor, 0 otherwise.	(ctdergty – ctderben) > 0
GROSS_HOLD	Gross CD holdings divided by total assets	ctdergty + ctderben
AGRILOANS	Farm loans divided by gross loans and leases.	lnag/lnlsgr
BAD_LOANS	Nonperforming loans divided by gross loans and leases.	nclnls/lnlsgr
COMM_IND	Commercial and industrial loans divided by gross loans and leases.	lnci/lnlsgr
COMM_RE	Commercial real estate loans divided by gross loans and leases.	lnrenres/lnlsgr
CONSTRCT	Construction loans divided by gross loans and leases.	lnrecons/lnlsgr
DRVTVS	Total derivatives including CD divided by assets	obsdir/asset
HAS_FORN	Originates foreign loans. Dummy variable = 1 if the bank originated foreign loans, 0 otherwise	lnfg > 0
HAS_RT	Has interest rate contracts. Dummy variable = 1 if the bank held interest rate contracts, 0 otherwise	rt > 0
INDVIDUAL	Loans to individuals divided by gross loans and leases.	lncon/lnlsgr
LIQUID	Liquid assets = (Cash and Balances due from depository institutions + Trading account assets + Available-for-sale securities (fair market value)) divided by total assets	(chbal+trade+scaf)/asset
LN_ASSET	Bank assets. Natural log of the bank's total assets	LN(asset)
LN_RISK	Bank risk assets. Natural log of the Bank's risk weighted assets	LN(rwajt)
LOAN_AST	Gross loans and leases divided by total assets	lnlsnet/asset
LOANSALE	Selling of bank's loans. Dummy variable = 1 if the bank sold part of its loan portfolio, 0 otherwise	Lnlssale > 0
NIM	Net interest margin as a proxy for profitability	nimy
SC_AST	Securities holdings divided by total assets	sc/asset
TIER1	Tier 1 capital divided by risk assets	rbc1aaj

Predicting Significant Operating Deficits in Municipalities Using Economic Indicators

John M. Trussel
University of Tennessee at Chattanooga

This study investigates economic activity associated with the operating results of municipalities surrounding and including the great recession of 2007-2009. The model hypothesizes that poor operating performance, as measured by a significant operating deficit, is related to four primary drivers of a local economy--employment, investment, industrial output and wealth. The findings indicate that municipalities with significant operating deficits have (statistically speaking) significantly lower job growth (employment) and GDP growth (industrial output) and marginally fewer housing permits issued (investment). The ensuing model was able to correctly predict up to 84% of the municipalities as either having significant budget deficits or not.

Keywords: operating performance, municipalities, economic indicators, adversity index

INTRODUCTION

Municipal governments provide invaluable services to the citizenry, which include fire, police, water, sewer, parks, and recreation. These municipalities also provide a wide array of public and social services, and contribute to the quality of community life. The citizenry count on municipalities to respond in times of trouble and to help maintain the quality of life; however, municipalities can do this only if they avoid poor operating performance, such as significant operating deficits (also called "budget deficits"). Thus, a municipality's ability to avoid poor operating performance directly affects its ability to sustain its current level of services. Also, the operating performance of a municipality is closely tied to the local economy. This study investigates economic activity associated with the operating performance of municipalities surrounding and including the great recession of 2007-2009. The model hypothesizes that the operating performance of a municipality is related to four primary drivers of a local economy--employment, investment, industrial output and wealth.

A municipality can continue to provide important public services only if it can avoid, poor operating performance, defined in this study as a significant operating deficit. The purpose of this study is to develop a model based on four primary factors of the local economy that can be used to predict significant operating deficits in municipalities. The model will develop and compare the economic profiles of US municipalities that have operating deficits versus those that do not. This study is important because fiscal distress is on the rise since the great recession. For example, the city administrators of Harrisburg, the capital of Pennsylvania, declared bankruptcy following the great recession.

Poor operating performance, such as significant operating deficits, occurs when there is an imbalance between the needs and resources of the people and the resources of the municipality. Such problems are an

intergovernmental problem. It can make state governments unstable, threaten the bond-ratings of state and local governments, and put pressure on state governments to intervene in delivering services when municipalities can no longer do so (Honadle 2003). These problems can also impair the willingness of businesses to move into local areas, since business decisions are often based on local taxes, services, infrastructure, and fees (Honadle 2003).

I examine the operating performance of a municipality as it is related to the local economy. This issue is of broad interest, as both scholars and professionals are interested in the topic of municipal operating performance. This study conveys important implications for future research and practice related to municipalities. The resultant model is used to develop early warning systems of significant operating deficits.

The model uses proxy measures for four primary drivers of a local economy, as utilized by Moody's Analytics—employment, investment, industrial output and wealth. Using logistical regression, these drivers are used to predict significant operating deficits. I find municipalities with significant operating deficits have (statistically speaking) significantly lower job growth (the proxy for employment) and GDP growth (industrial output), with marginally fewer housing permits (investment). However, changes in housing prices (wealth) were not statistically significant in the model. The model correctly predicts up to 84% of the municipalities as either having significant operating deficits or not. Municipalities looking to improve their operating results should first look to increase job growth, as the results indicate that such growth has the biggest impact. This model may be used by state and local government administrators in mitigating significant operating deficits, bond investors in evaluating default risk, auditors in analyzing going-concern, and others.

The remainder of the paper is organized as follows: The next section gives a background by describing the extant literature on operating results in municipalities. The section after next discusses the economic factors associated with operating results. The empirical testing the operating results model are analyzed in the next to last section, and the last section concludes the paper.

BACKGROUND ON MUNICIPAL OPERATING PERFORMANCE

Shortly after the spate of municipal emergencies in the late 1970s and early 1980s, academic and other researchers began to study the causes and measures of poor financial performance, in particular, and fiscal assessment, in general. In November 1985, the U.S. Advisory Commission on Intergovernmental Relations (ACIR) issued a widely-recognized report on fiscal health (ACIR 1985b). In this report, the ACIR purports that poor fiscal health is caused by cyclical or structural forces. Cyclical forces are short-term conditions that usually correspond to business cycles. Structural forces are long-term changes in the economy that are beyond the control of the state or local government. The ACIR (1985b) added that poor fiscal health is often the result of a complex array of economic conditions (such as unemployment rates), socioeconomic conditions (such as poverty levels), physical conditions (such as the condition of infrastructure assets), and financial factors (such as dependence on intergovernmental revenue). The primary focus of the present study is on the economic conditions associated with operating results.

Previous research addresses the issue of measuring and assessing fiscal health, which is a broader concept than just operating performance. Groves et al. (1981) use a set of financial indicators to assess the fiscal health of 24 cities. Their indicators include environmental factors such as changes in population growth, personal income levels, property values, unemployment rates, and business activity, as well as regional inflation rates. They also consider intergovernmental constraints (e.g., the extent of grants-in-aid, tax restrictions, and federal/state mandates), legislative policies and managerial practices. They develop a theoretical model, which captures all these external factors with several basic financial indicators, and they develop financial ratios for each factor. In the end, the model includes six broad categories--revenues, expenditures, earnings, debt structure, unfunded liabilities, and condition of capital plant –and approximately 30 related financial ratios. This model, developed by Groves and Valente (1994) and published by the International City/County Management Association, is a widely accepted financial monitoring tool for municipalities. The model was not tested on a large sample of municipalities.

There are other financial assessment tools, such as Brown's 10-point test to conduct a trend analysis (1993), Ammons' (1995) attempt to establish municipal benchmarks, and Kleine et al.'s (2003) 10-point scale for Michigan. There have been several other studies of fiscal health in municipalities (e.g., Brown 1996; Campbell 1990; Honadle 1998; Stevens and LaPlante 1987; Weinberg 1984).

The great recession of 2007-2009 caused a major shift in the analysis of fiscal health of municipalities. Wallace et al. (2018), for example, discuss how the recession impacted the structure, function and regulation of the municipal bond market. Following the crisis of those years, many changes were made to primary market issuance and secondary market trading practices and to bond ownership composition, bond structures, products, processes, and market participants. Murphy et al. (2018) summarize how such changes impact the future default risk of municipal bonds. These studies imply the need for predictive models of fiscal health, such as the presence of significant operating deficits.

Most of the research prior to 2009 was descriptive and not predictive in nature. The main focus was to describe the nature of fiscal issues in municipalities, not to predict the fiscal problems. Following the great recession of 2007-2009, the need arose for the models to be more predictive in nature, allowing for an early warning system for analysts. Trussel and Patrick (2009) changed the tenor of the studies to focus on the predictive ability models of fiscal issues. They use financial factors, including revenue concentration, debt usage, administrative expenditures and entity resources, to develop a model to predict fiscal health in Pennsylvania. These authors subsequently made modifications to their initial study, including expanding to municipalities in all states (Trussel and Patrick, 2012), testing special district governments (Trussel and Patrick, 2013), using only county governments (Patrick and Trussel, 2013), applying alternative statistical methods (Trussel and Patrick, 2013) and addressing socio-demographic data (Trussel and Patrick, 2014). This paper expands upon these studies by focusing on economic factors and their relationship to the significant operating deficits.

MODEL DEVELOPMENT

Moody's Analytics, a subsidiary of the Moody's Corporation that offers research for credit, economic and financial risk management, gauges the economic health of metropolitan areas in the US on a monthly basis (Dedman, 2009). The company's analysis is based on four measures of economic activity--changes in employment, housing starts, industrial production and housing prices. Moody's uses a complex method of weighting each component to come up with an overall financial rating, called the "Adversity Index." The ratings are in one of four discrete categories, either in recession, at risk of recession, recovering from recession, or expanding. The company does not have a specific method of testing the appropriateness of the classifications. That is, they have no external validation of their defined status. Dedman (2009) notes, "Components that commonly show sporadic movement are weighted less, while smoother measures are given a higher weight. The weighting is different for each metro area." The index is developed using weights developed from the economic data itself with no independent method of defining the actual fiscal status of the underlying governmental entity.

I adopt Moody's method of using these four economic factors to develop a model predicting significant operating deficits by using a statistically valid method of weighing the four factors, while testing the reliability of the resultant classifications. I use an external (independent of the four factors) definition of municipal operating performance, namely, significant operating deficits to develop the weights on the four variables and to validate the model. As with the Moody's index, I focus on metropolitan statistical areas (MSA).

Significant Operating Deficits

The purpose of this paper is to develop a model to predict significant operating deficits of municipalities using lagged economic factors as inputs. In order to achieve this goal, a definition of the operating performance of a municipality is needed. Measuring operating performance, including analyzing fiscal health in general, has been addressed in the literature using several different methods.

The Government Accountability Office (GAO) defines a poorly performing municipality as one "in which residents bear substantially higher tax burdens in order to obtain levels of public services comparable to better-off communities" (GAO 1990). DeSanto et al. (1991) define poor fiscal performance as "a persistent shortfall in cash flows…resulting from an imbalance between revenues and expenditures for given service levels" (p. 7). Kloha et al. (2005a) define poor operating performance as "a failure to meet standards in the areas of operating position, debt, and community needs and resources over successive years" (p. 314). A variety of methods have been utilized in attempt to operationalize these constructs. Raman (1982) identifies poor performing entities using bond ratings from Moody's. Trussel and Patrick (2012, 2013a, 2013b, 2014) using either a decline in spending on public services or an operating deficit to operationalize the operating performance of a municipality. Their conceptual definition is that a municipality is performing poorly if there is an imbalance between revenues and expenditures. This imbalance is best captured by the operating margin, revenues less expenditures scaled by revenues. If revenues exceed expenditures, then the municipality has an "operating surplus." When expenditures exceed revenues, an "operating deficit" occurs.

Following Trussel and Patrick (2009, 2012, 2013a) and Patrick and Trussel (2011), I define a poorly performing municipality as one that experiences a significant imbalance between revenues and expenditures. I operationalize this imbalance as a significant operating deficit. I use data from the US Census Bureau to measure a significant operating deficit for the principal municipality in the MSA. An MSA typically consists of a principal city (the largest in the area) and several smaller ones. Since the fiscal health of the smaller cities is tied to the largest city, I use the principal city in the MSA to measure the operating surplus or deficit. The annual operating margin, revenues less expenditures scaled by revenues, is used by many researchers to reflect the overall operating performance of a municipality (e.g., Trussel and Patrick, 2009). Researchers argue that a deficit must be significant for a municipality to be considered as fiscally imbalanced. Following the recommendations of Pennsylvania's Department of Economic Development (DCED, 2001), I classify a municipality as poorly performing if the annual operating deficit is significant, greater than five percent of revenues. In other words, if the municipality has an operating margin of less than negative five percent, then it is considered to be performing poorly. I provide robustness tests of other definitions of fiscal performance later.

Economic Factors

Following Moody's Analytics, I identify proxy measures of the four primary drivers of a local economy—employment, investment, industrial output and wealth. The data sources are explained below and summarized in Table 1.

Employment

I use employment data from US Bureau of Labor Statistics to measure the annual percentage change in the total number of jobs (JOBS) in a Metropolitan Statistical Area (MSA). I hypothesize that the percentage change in the number of jobs is directly related to the operating performance of the MSA, meaning that if employment increases, then the likelihood of significant operating deficits should decrease.

Investment

As a proxy for investment, I use housing construction permit data from the US Census Bureau to measure the percentage change in the annual number of new privately owned housing units (PERMITS) authorized in an MSA. I hypothesize that the percentage change in the number of permits is directly related to the operating performance of the MSA, meaning that if housing permits increase, then the likelihood of significant operating deficits should decrease.

Industrial Output

I use data from the US Bureau of Economic Analysis to measure the percentage change in gross domestic product (GDP) in an MSA. I hypothesize that the percentage change in GDP is directly related to

the operating performance of the MSA, meaning that if GDP increases, then the likelihood of significant operating deficits should decrease.

Wealth

As a proxy for wealth, I use data from the Federal Housing Finance Agency to measure the year end housing price index (PRICES) for an MSA. This index is a measure of housing prices indexed to a base year. I hypothesis that the housing price index is directly related to the operating performance of the MSA, meaning that if housing prices increase, then the likelihood of significant operating deficits should decrease.

TABLE 1
ECONOMIC FACTORS ASSOCIATED WITH OPERATING PERFORMANCE

Economic Factor	Measurement	Source
Employment (JOBS)	Annual percentage change in number of jobs	US Bureau of Labor Statistics
Investment (PERMITS)	Annual percentage change in number of permits for new privately owned housing units	US Census Bureau
Industrial Output (GDP)	Annual percentage change in GDP	US Bureau of Economic Analysis
Wealth (PRICES)	Annual index for housing prices	Federal Housing Finance Agency

EMPIRICAL TESTING

Sample

Financial data are used for the years 2005-2012, which are the years surrounding and including the great recession of 2007-2009. This allows me to test the model during years of expansion, recession and recovery. Due to the nature of the economic data, I focus my attention on the 384 to 397 Metropolitan Statistical Areas (MSA), which account for nearly 85% of the US population (Dedman, 2009). An MSA contains a core urban area with a population of at least 50,000. The data comes from different data sources as discussed in the previous section and summarized in Table 1. Combining the eight years, 2005-2012, to develop panel data gives a possible 3,176 MSA-years. Some data are missing to arrive at a final sample of 2,718 MSA-years, or 85.6% of all the possible available. Panel A of Table 2 summarizes the sampling procedures.

The breakdown of the final sample by operating performance and year is included in Panel B of Table 2. Nearly 34% of the cities in the final sample have significant operating deficits. Except for a slight decline in 2007, there is a gradual increase in the percentage of municipalities with significant operating deficits that climaxes in 2010 and begins to drop in 2011. Given that the economic recession began in 2007 and ended in 2009, there seems to be about a one year lag in the financial impact. Thus, the economic indicators were measured one year in advance. For example, if operating performance was measured in 2006 on Panel A of Table 2, then the economic indicators were measured in 2005. This will also allow for me to use the economic data for predictive purposes.

TABLE 2
SAMPLE SELECTION PROCEDURES

Panel A: Overall Sample

	Municipal Statistical Areas (MSA)	
	Number	Percent
Total MSA-years	3,176	100.0
Missing Data	(458)	(14.4)
Final Sample	2,718	85.6

Panel B: Sample Partitioned by Operating Performance and Year

Year	No Significant Deficits*	Significant Deficits*	Total	Percent with Significant Deficits
2005	224	96	320	30.0%
2006	200	114	314	36.3%
2007	257	107	364	29.4%
2008	225	105	330	31.8%
2009	213	129	342	37.7%
2010	186	156	342	45.6%
2011	231	111	342	32.5%
2012	262	102	364	28.0%
Total	1,798	920	2,718	33.8%

*A significant operating deficit is defined as an operating margin less than negative five-percent. Operating margins are revenues less expenditures scaled by revenues.

Descriptive Statistics

The descriptive statistics are included in Panel A of Table 3, including information about the population of the principal cities in the MSAs. MSAs are by definition large metropolitan areas with a core municipality of at least 50,000 people. As noted, the economic data are from MSAs, while the financial data are from the principal city within the MSA. The sample results indicate that cities with significant operating deficits (mean population of 187,172) have larger populations than do cities without significant budget deficits (146,520). This finding is consistent with other studies (e.g., Trussel and Patrick, 2013) that find that the size of a city is positively correlated with fiscal problems. As predicted, cities without significant operating deficits have higher job growth, issue more housing permits and report higher GDP growth. However, there is no significant difference in the housing price indices between the two groups.

Panel B of Table 3 includes the Pearson correlation coefficients among the economic factors. The largest coefficient was 0.666 between JOBS and GDP. This does not appear to cause any problems with the regression results.

TABLE 3
DESCRIPTIVE STATISTICS

Panel A: t-Tests

Variable*	Significant Deficit?**	Mean	Std. Dev.	*t*-statistic	*p*-value
JOBS	No	0.008	0.024	4.652	<0.001
	Yes	0.004	0.025		
PERMITS	No	(0.058)	0.358	3.201	0.001
	Yes	(0.106)	0.388		
GDP	No	0.041	0.050	4.558	<0.001
	Yes	0.032	0.048		
PRICES	No	177.237	34.464	0.563	0.574
	Yes	176.253	35.166		
POPULATION	No	146,520	207,719	(4.261)	<0.001
	Yes	187,172	259,955		

Panel B: Pearson Correlation Coefficients

	JOBS	PERMITS	GDP
PERMITS	0.222***		
GDP	0.666***	0.240***	
PRICES	0.200***	-0.091***	0.097***

*Variables are defined in Table 1.
** A significant operating deficit is defined as an operating margin less than negative five-percent. Operating margins are revenues less expenditures scaled by revenues.
***Pearson correlation coefficient is significant at the 0.01 level (two-tailed).

Logistic Regression Results

I use cross-sectional time-series (panel data) analysis to test the model of operating performance. Since the dependent variable is categorical, the significance of the multivariate model is addressed using logistic regression analysis and adjusted for autocorrelation. Using this method, the underlying latent dependent variable is the probability of significant operating deficits for municipality *i*, which is related to the observed variable, $Status_i$, through the relation:

$Status_i = 0$ if the organization does not have significant operating deficits, or
$Status_i = 1$ if the organization has significant operating deficits.

The model includes all of the independent variables from Table 1, and each is lagged by one year for predictive purposes. Recall that municipalities are classified as poorly performing if the operating deficit is greater than five-percent of annual revenues. The predicted probability of the k^{th} status for municipality *i*, $P(Status_{ik})$ is calculated as:

$$P(Status_{ik}) = \frac{1}{1+e^{-Z}} \quad (1)$$

where

$$Z_i = \alpha + \beta_1 \bar{JOBS}_{t-1} + \beta_2 \bar{PERMITS}_{t-1} + \beta_3 \bar{GDP}_{t-1} + \beta_4 \bar{PRICES}_{t-1}$$

I use data from 2005-2011 to develop the model (the estimation sample) and data from 2012 to test the model (the holdout sample). The results of the logistic regression model (adjusted for autocorrelation) are

included in Table 4. JOBS and GDP are significantly related to the probability of significant operating deficits (at the 0.05 level) with the predicted negative signs. PERMITS is marginally significant at the 0.10 level with the predicted negative sign. The other economic factor, PRICES, is not statistically significant in the multivariate model, although it does have the predicted negative sign.

TABLE 4
REGRESSION RESULTS

Variable	Coefficient	Std. Error	Wald	*p*-value	*Exp(B)*	Impact
Intercept	-0.481	0.3209	2.249	0.134		
JOBS	-5.563	2.3174	5.763	0.016	0.004	-0.054
PERMITS	-0.250	0.1517	2.709	0.100	0.779	-0.002
GDP	-2.446	1.0817	5.111	0.024	0.087	-0.024
PRICES	-0.001	0.0018	0.020	0.888	1.000	<0.001

Note: See Table 2 for a description of the independent variables. The latent dependent variable equals 0 if the municipality did not have a significant operating deficit, and 1 if it did. The last column represents the impact on the predicted likelihood of a significant operating deficit due to a 0.01 increase in the value of the covariate, except for PRICES, which represents the impact due to a one-unit increase in the value.

The results of the regression analysis also allow one to address the impact of a change in an economic factor on the likelihood of significant operating deficits. In Table 4, *Exp(B)* is the odds ratio, which is the change in the odds of the event (significant operating deficits) occurring for a one-unit change in the economic factor. A one-unit change in the price index is reasonable, but a one unit change in the other factors is not possible, since they are percentage changes. Thus, the last column in Table 4 represents the impact on the predicted likelihood of significant operating deficits due to a 0.01 increase in the value of JOBS, PERMITS and GDP, and a one-unit change in PRICES. The impact is computed as $Exp(b)^{0.01} - 1$, except for PRICES, which is just $Exp(b)-1$. JOBS and GDP have the biggest influences on the likelihood of significant operating deficits. A one-percent increase in the number of jobs will decrease the predicted likelihood of significant operating deficits by 0.05, while a one-percent increase in the GDP will decrease the likelihood by 0.024. Based on the economic factors in this model, cities attempting to reduce the likelihood of significant operating deficits will have the biggest impact by increasing the number of jobs or by increasing industrial output (GDP). Changes in the housing permits has a relatively small impact, and changes in housing prices has a negligible impact.

Predicting Significant Operating Deficits

I use the results of the logistic regression analysis to test the predictive ability of the model. The observed logistic regression equation (from Table 4) for entity *i* at time *t* is:

$$P(i,t) = 1/(1+e^{-Zi})$$

where:

$$Z_i = -0.481 - 5.563\ JOBS_{t-1} - 0.250\ PERMITS_{t-1} - 2.446\ GDP_{t-1} - 0.001\ PRICES_{t-1}$$

The predicted dependent variable, *P(i,t)* the probability of significant operating deficits for municipality *i*, is computed using the actual economic indicators for each municipality in the estimation sample. The resulting probabilities are used to classify municipalities as experiencing a significant operating deficit or not. Jones (1987) suggests adjusting the cutoff probability for classifying in two ways. Following the suggestion of Jones, I first incorporate the prior probability of a significant operating deficit and then include the expected cost of misclassification.

Using logit, the proportion of distressed municipalities in the sample must be the same as the proportion in the population to account for the prior probability of significant operating deficits. If the proportion is not the same, then the constant must be adjusted (Maddala 1991). This is more of a problem when a paired sample method is used, which is not the case here. Since I do not know the proportion of municipalities that had a significant operating deficit in the population of all municipalities, I assume that the proportion of municipalities in the sample is an unbiased estimator of the proportion in the population of all municipalities. Since 33.8% of the municipalities in the sample had significant operating deficit, I assume that the prior probability of significant operating deficits is 0.338.

The ratios of the cost of type I errors (incorrectly classifying municipalities with significant operating deficits as not having them–a false negative) to type II errors (incorrectly classifying municipalities that do not have significant operating deficits as have them–a false positive) also must be determined. The particular cost function is difficult to ascertain and will depend on the user of the information. For example, municipal bond investors may want to minimize losses (and thus type I errors); however, they will suffer an opportunity cost (type II error) if another bond is purchased offering a lower rate. In most cases, the cost of a type II error is likely to be much smaller than a type I error. Thus, I incorporate multiple relative cost ratios (and cutoff probabilities) into my analysis. Specifically, I include the relative costs of type I to type II errors of 1:1, 10:1, and 20:1 (Beneish 1999; Trussel 2002). Ratios beyond 20:1 were also considered, but there is no change in the classification accuracy of the model at cost ratios greater than 10:1.

The results of using the logit model to classify municipalities as either having significant operating deficits or not are included in Table 5 for both the estimation and the holdout samples. The cutoff probabilities presented are those that minimize the expected costs of misclassification. Following Beneish (1999), the expected costs of misclassification (ECM) are computed as:

$$ECM = P(FD)P_I C_I + [1 - P(FD)]P_{II} C_{II},$$

where P(FD) is the prior probability of having a significant operating deficit, P_I and P_{II} are the conditional probabilities of type I and type II errors, respectively, and C_I and C_{II} are the costs of type I and type II errors, respectively.

TABLE 5
THE PREDICTIVE ABILITY OF THE OPERATING PERFORMANCE MODEL INCLUDING THE EXPECTED COSTS OF MISCLASSIFICATION AND THE RELATIVE COSTS OF TYPE I ERROR TO TYPE II ERROR

	Ratio of the Cost of Type I to Type II Errors					
	Estimation Sample			Holdout Sample		
	1:1	10:1	20:1	1:1	10:1	20:1
Cutoff	0.360	0.100	0.100	0.360	0.100	0.100
Type I Error	0.463	0.000	0.000	0.969	0.000	0.000
Type II Error	0.000	0.524	0.524	0.041	1.000	1.000
Overall Error	0.161	0.342	0.342	0.305	0.716	0.716
ECM Model	0.157	0.347	0.347	0.355	0.662	0.662
ECM Naïve	0.338	0.662	0.662	0.338	0.662	0.662
Relative Costs	0.463	0.524	0.524	1.050	1.000	1.000
Overall Correct	0.839	0.658	0.658	0.695	0.284	0.284

Note: The cutoff is the probability of a significant operating deficit that minimizes the expected cost of misclassification, ECM. ECM is computed as $ECM = P(FD)P_I C_I + [1 - P(FD)]P_{II} C_{II}$, where P(FD) is the prior probability of a significant operating deficit (0.338), P_I and P_{II} are the conditional probabilities of Type I and Type II errors, respectively. C_I and C_{II} are the costs of type I and type II errors, respectively. The relative costs are the ECM Model divided by the ECM Naïve.

The validity of the model is tested on the holdout sample (2012 holdout data) using the same cutoff probabilities from the estimation sample. The results, displayed in Table 5, indicate that the model can identify cities with significant operating deficits with 66% (at a cost ratio greater than 1:1) to 84% (at a cost ratio of 1:1) of the entities in the estimation sample correctly classified. In the holdout sample, 29% to 70% of the entities are correctly classified.

To test the usefulness of the model, I compare these results to a naïve strategy. This strategy classifies all municipalities as having a significant operating deficit (or not) when the ratio of relative costs is greater than (or less than or equal to) the prior probability of having a significant operating deficit (0.338). This switch in strategy between classifying all organizations as not having a significant operating deficit to classifying all of them as having one occurs at relative cost ratios of just below 3:1 (i.e., 1 / 0.338).

If all municipalities are classified as having a significant operating deficit (not having one), then the naïve strategy makes no type I (type II) errors. In this case, P_I (P_{II}) is zero, and P_{II} (P_I) is one. The expected cost of misclassification for the naïve strategy of classifying all municipalities as not having one (having one) reduces to $0.662C_{II}$ ($0.338C_I$).

In both panels of Table 5, I also report the relative costs, which is the ratio of the ECM for my model to the ECM for the naïve strategy. Relative costs below 1.0 indicate a cost-effective model. For the estimation sample, my model consistently has a much lower ECM than the naïve strategy. In fact, the relative costs are below 53% for all levels of type I to type II errors. These results provide evidence to suggest that the classification model is extremely cost-effective in relation to a naïve strategy for all ranges of the costs of type I and type II errors. However, the results are not as good for the holdout sample with the model about the same as the naïve model.

Applying the Prediction Model

I use one of the municipalities from the sample to illustrate the model. The model allows one to predict the status of the municipality as having a significant operating deficit or not. From the results of the logistic regression, the probability of a significant operating deficit for municipality *i* at time *t*, *P(i,t)* is:

$$P(i,t) = \frac{1}{1+e^{-zi}} \quad (2)$$

where

$$Z_i = -0.481 - 5.563\ JOBS_{t-1} - 0.250\ PERMITS_{t-1} - 2.446\ GDP_{t-1} - 0.001\ PRICES_{t-1}$$

Substituting the actual variables from the example entity (in parentheses), I obtain:

$$Z_i = -0.481 - 5.563\ (-0.031) - 0.250\ (0.493) - 2.446\ (-0.018) - 0.001\ (137.16)$$
$$Z_i = -0.422$$
$$P = 1 / (1+e^{0.422})$$
$$P = 0.396.$$

Table 5, Panel A, shows that the selected municipality is predicted to have a significant operating deficit, since the actual probability (0.396) is greater than the cutoff at all levels of the ratio of type I to type II errors. The entity actually experienced a significant operating deficit. Thus, the model correctly predicted the operating performance of this municipality.

Robustness Tests

I made several assumptions when developing and testing my model and test these assumptions for robustness. For example, I defined a municipality as having a significant operating deficit if the organization has operating deficits greater than five-percent of annual revenues. I also test my model by using any decline in the operating margin (i.e., more than zero percent) and declines of more than 10% and 20% in the operating deficit. For these versions of the model, the tenor of my results does not change.

I also assumed the prior probability of a significant operating deficit in developing my prediction model. I assumed that the prior probability was 0.338 because 33.8% of the municipalities in the initial sample had a significant operating deficit. I evaluated the sensitivity of the model to other assumptions of the prior probability by using prior probabilities of 0.10, 0.20 and 0.50. These assumptions do not alter the results significantly.

CONCLUSION

Municipalities provide important public services. They provide the first level of response when it comes to public services such as safety, water, sewer, streets, parks, and recreation. Municipalities also play a large role in the quality of community life by providing a wide variety of public and social services; however, they can sustain these services only if they avoid significant operating deficits. Significant operating deficits threaten the ability of municipalities to continue to serve the citizenry and maintain essential public functions. Economic factors, such as job growth and GDP impact a municipality's operating performance. This study explicitly links four primary economic factors related to a municipality's job growth, investment, industrial output and wealth to its operating performance.

Using logistical regression, I find that municipalities with operating deficits exceeding five-percent of annual revenues, have (statistically speaking) significantly lower job growth and GDP and marginally fewer new housing permits. However, housing prices were not statistically significant in the model. I was able to correctly predict up to 84% of the municipalities as either having a significant operating deficit or not. Municipalities looking to improve their operating performance should first look to improve its job growth.

These results should interest state and municipal administrators in mitigating poor operating performance, bond investors in evaluating default risk, auditors in analyzing going-concern, and others. Further research is needed to apply this model to smaller municipalities across the US and other countries and to different time periods.

REFERENCES

Advisory Commission on Intergovernmental Relations (ACIR). (1985a). *Bankruptcies, defaults, and other local government financial emergencies*. Advisory Commission on Intergovernmental Relations, Washington, DC.

Advisory Commission on Intergovernmental Relations (ACIR). (1985b). *The states and performing poorly communities: The final report*. Advisory Commission on Intergovernmental Relations, Washington, DC.

Ammons, D. (1995). *Municipal benchmarks: assessing local health and establishing community standards*. Sage Publications, Thousand Oaks, CA.

Beneish, M. (1999, September/October). The detection of earnings manipulation. *Financial Analysts Journal*, *55*, 24-41.

Brown, K. (1993). The 10-point test of financial condition: Toward an easy-to-use assessment tool for small cities. *Government Finance Review*, *9*(6), 21-27.

Brown, K. (1996). Trends in key ratios using the GFOA financial indicators databases 1989-1993. *Government Finance Review*, *12*(6), 30-35.

Campbell, H. (1990, Summer). Comparative fiscal analysis for counties. *Public Budgeting and Finance*, 88.

Dedman, B. (2009). *How the adversity index detects trends in local economies*. Retrieved December 24, 2019, from http://www.nbcnews.com/id/29866676/ns/us_news-the_elkhart_project/#.Vnwvg1LG_qm

Department of Community and Economic Development. (2001). *Debt management handbook*. Governor's Center for Local Government Services, Harrisburg, PA.

DeSanto, F., Emmett, R., Hall T., Horton, B., Seliga, R., Stoichess, . . . Zurhellen, D. (1991). *Coping with fiscal distress in Pennsylvania's local governments: A program evaluation of act 47: A report to*

the local government commission of the general assembly of Pennsylvania. The Pennsylvania State University, State College, PA.
General Accounting Office (GAO). (1990). *Performing poorly communities: Public services declined in California as budget pressures mounted*. GAO HRD-90-95, Washington, DC.
Groves, S., & Valente, M. (1994). *Evaluating financial condition: A handbook for local government*. ICMA, Washington, DC.
Groves, S., Godsey, W., & Shulman, M. (1981). Financial indicators for local government. *Public Budgeting and Finance*, *18*(2), 5-19.
Honadle, B. (1998). Analyzing rural local governments' financial condition: An exploratory application of three tools. *Public Budgeting and Finance*, *18*(2), 69-88.
Honadle, B. (2003). The states' role in US local government financial crises: A theoretical model and results of a national survey. *International Journal of Public Administration*, *26*(13), 1431-72.
Jones, F. (1987). Current techniques in bankruptcy prediction. *Journal of Accounting Literature*, *6*, 131-164.
Kleine, R., Kloha, P., & Weissert, C. (2003). Monitoring local government fiscal health: Michigan's new 10-point scale of fiscal distress. *Government Finance Review*, *19*(3), 18-23.
Kloha, P., Weissert, C., & Kleine, R. (2005a). Developing and testing a composite model to predict local fiscal distress. *Public Administration Review*, *65*(3), 313-323.
Maddala, G. (1991). Perspective on the use of limited-dependent and qualitative variables models in accounting research. *Accounting Review*, *66*(4), 788-807.
Murphy, C., Fields, D., Lemke, D., & Chabotar, K. (2018). The future is now: Default study 2025. *Municipal Finance Journal*, *39*(2), 111-139.
Patrick, P., & Trussel, J. (2011, August). The financial indicators associated with reductions of public services by Pennsylvania municipalities. *International Journal of Business and Social Science*, *2*, 53-62.
Patrick, P., & Trussel, J. (2013). Predicting County Expenditures by Policy Area. *International Research Journal of Applied Finance*, *IV*(12), 1646-1662.
Raman, K. (1982). Alternative accounting measures as predictors of municipal poor fiscal health. *Journal of Accounting, Auditing, and Finance*, *6*(1), 44-50.
Stevens, J., & LaPlante, J. (1987). Shaping state-based financial policy in an era of change: a policy and research perspective on city size. *Policies Studies Review*, *7*(1), 61-76.
Trussel, J. (2002). Revisiting the prediction of financial vulnerability. *Nonprofit Management and Leadership*, *13*(1), 17-31.
Trussel, J., & Patrick, P. (2009, Winter). A Predictive Model of Fiscal Distress in Local Governments. *Journal of Public Budgeting, Accounting, and Financial Management*, *21*(4), 578-616.
Trussel, J., & Patrick, P. (2012). A Survival Analysis of US Municipalities in Fiscal Distress. *International Journal of Public Administration*, *35*, 620-633.
Trussel, J., & Patrick, P. (2013a, December). The Symptoms and Consequences of Fiscal Distress in Municipalities: An Investigation of Reductions in Public Services. *Accounting and the Public Interest*, *13*, 151-171.
Trussel, J., & Patrick, P. (2013b, Winter). Predicting Fiscal Distress in Special District Governments. *Journal of Public Budgeting, Accounting and Financial Management*, *24*(4), 589-616.
Trussel, J., & Patrick, P. (2014). The Socio-Demographic, Economic and Financial Profiles of Municipalities at Risk of Poor fiscal health in Pennsylvania. *Journal of Accounting, Audit, and Finance Management*, pp. 1-9.
Wallace, A., Colby, M., Doe, T., & Cosgrove, R. (2018). A decade later: The municipal market post-financial crisis through today. *Municipal Finance Journal*, *39*(2), 141-159.
Weinberg, M. (1984). Budget retrenchment in small cities: A comparative analysis of Wooster and Athens, Ohio. *Public Budgeting and Finance*, *4*, 46-57.

Assessment of Financial Risk in Uncertainty Environments

Juan Carlos Martínez Barrio
CEEI-Burgos

An operational composite indicator to evaluate the estimated recovery probability of an entrepreneur's loan, or any other business-oriented loans, given by

$$ERP= \alpha(RR) + \beta(SOA) + \gamma(COA) – [NCR] – [CI]$$

is presented. The resulted indicator evaluates the recovery probability of an entrepreneur/SME/company loan in uncertainty environments, without guarantees or collaterals. This indicator has been named "Estimated Recovery Probability", ERP. Being a totally new approach, ERP methodology has been tested in actual scenarios showing its efficiency in comparison with current risk assessment methods used by the traditional financial system, pointing out the fact that guarantees or collaterals have not been requested.

Keywords: financial risk analysis, micro finances, risk estimation

INTRODUCTION

Guarantees and collaterals are determining factors in the financial risk analysis. The lack of these forces assessment methodologies to be radically different from the methods used by the traditional banking institutions, which are mainly based in the three following principles:

- Security: Feasibility of the operation itself.
- Solvency: Profitability of the operation itself.
- Exit: Guarantee in case of default.

The absence of this last principle is an essential characteristic of the ERP methodology, focusing, therefore, the whole analysis on first two ones, mainly in the first one, the feasibility of the project itself. For the second one, solvency, it is a more a decision element rather than a component of the risk analysis. However, it should be taking into account in order to properly assess whether the given risk should be assumed or accepted. For this purpose, a comparison between the current market interest rate and the financial profitability of the operation itself given by a standard yield/profitability ratio should be done.
The "Estimated Recovery Probability" compound indicator with terms:

$$ERP= \alpha(RR) + \beta(SOA) + \gamma(COA) – [NCR] – [CI]$$

calculates an estimated probability being RR, SOA and COA indexed variables which values range from 0 to 1; α, β y γ weighting coefficients ranging from 0 to 1 as well.

The concept "estimated probability" does not refer to a pure mathematical probability calculation, for it is an estimation which value is the result of the enunciated equation. It is an estimation of the measure of the likelihood that an event will occur, but it is not calculated as the quotient:

$$P = \frac{The\ number\ of\ ways\ an\ event\ can\ occur}{The\ total\ number\ of\ possible\ outcomes}$$

Therefore, ERP ranks from 0 to 1, as a mathematical probability would be, being regularly presented as a percentage for common use.

Another way of this concept to be presented would be:

$$ELP = 1 - ERP$$

being ELP the "estimated loss probability", which values ranks from 0 to 1, regularly presented as a percentage as well.

This formula, based on a totally innovative approach, is successfully being implemented in an actual economic environment as the methodology used for the financial risk assessment of the credit operations of a micro finance fund since 01/01/2013.

ERP method or approach is not based upon any other previous financial risk analysis work, with the exception of the part related to the Structural Operational Analysis (SOA) component which is inspired in empirical tools designed for start-up evaluation.

The following sections will include a detailed description of the components of ERP method, an explanation of the weighting coefficients and the empirical testing description. This work will be concluded by some final remarks and the acknowledgment and bibliography sections.

COMPONENTS

The ERP composite indicator is formed by four main components and a corrector index:

- Risk Ratio (RR).
- Structural Operational Analysis (SOA).
- Credit Officer Assessment (COA).
- Negative Credit Record (NCR).
- Experience based corrector index (CI)

The component Risk Ratio (RR) is a statistical and financial evaluation of the risk shown for a specific company for the present and the next 3 years. Different variables are analysed to calculate this ratio, such as, economic feasibility of the project, indebtedness, equity, liquidity ratios and loan regular payments.

The Structural Operational Analysis (SOA) is a general operational analysis of the project's feasibility, but from the entrepreneur/SME/company point of view. Various characteristics and variables such as the owner's ability and skills, product's concept, knowledge of customer's need, potential customers, cash amount ratio, business plan, break-even point, information network of partners, motivation and recruitment of employees and level of ICT use.

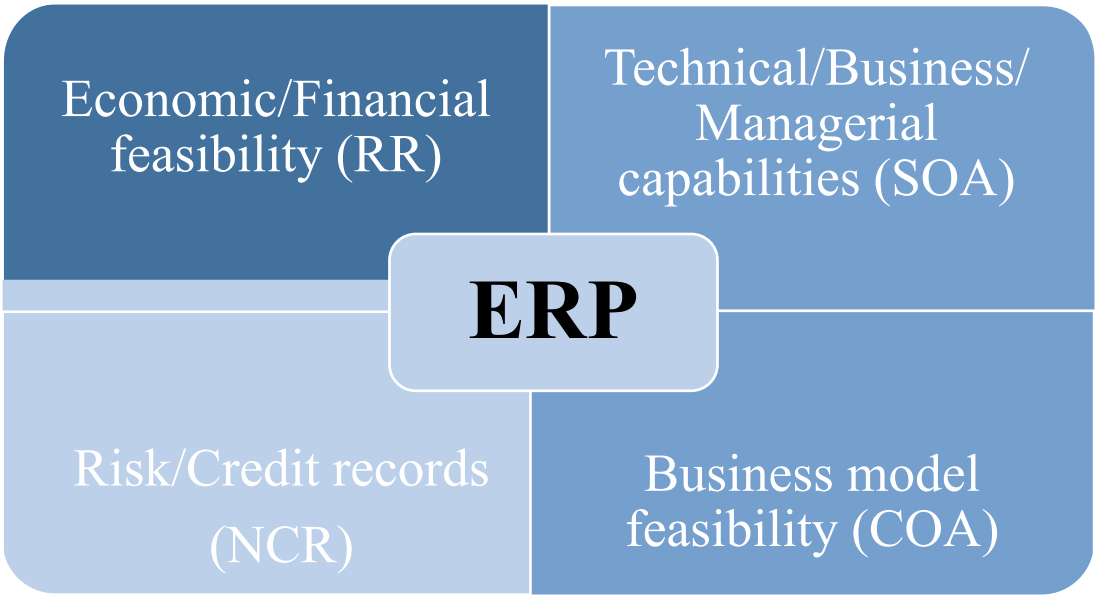
The Credit Officer Assessment (COA) is an expert evaluation of the overall business concept. The Credit Officer judge aspects such as the viability of the business, evolution of the sector, the stage of the project development, knowledge and experienced of the entrepreneur/promoter or pre-set agreements with potential customers.

The Negative Credit Record (NCR) is based on a risk and a credit records assessment that is supported, in the case of an operation based in Spain, by the CIRBE, a document issued by the Spanish Central Bank, to check the credit records of the micro credit/loan applicant.

The last component is the experience based Corrective Index (CI), which would be added to the formula after a determined number of years of implementation of the formula itself without this component, in accordance with the accuracy of the ERP value and the actual result of the operation assessed.

Components RR, SOA and COA are also adjusted by α, β and γ weighting coefficients. The first four component reflects a different axis of ERP methodology:

FIGURE 1
AXIS OF ERP METHODOLOGY



This way ERP composite indicator avoid biased approaches by integrating all aspects of the project assessed isolating the analysis itself from the subjectivity of the analyst and the Credit Officer, in order to provide the Credit Committee – the decision makers - with objective and valuable information to support their decision about the operation.

The corrector index CI will adjust the resulted value in accordance with actual data and feedback gathered as a result of a significant number of implementations, real loans assessments performed.

RISK RATIO (RR)

This first component is, in turn, also a composite sub-indicator formed by four annual values of two ad-hoc created financial ratios, which will be named, for denomination purposes, Yearly Quota Weighting (YQW) and Equity Capital Weighting (ECW):

- $YQW = \frac{Profit\ after\ taxes\ (PAT) + Loan\ yearly\ quota\ (LYQ)}{Loan\ Yearly\ Quota\ (LYQ)}$

- $ECW = \frac{Previous\ outstanding\ loans\ (POL) + Equity\ capital\ (EC) + Subsidies\ (S)}{Equity\ capital\ (EC)}$

So that,

- $RR = 0{,}5\ \{0{,}5\{\alpha YQW_1 + (1-\alpha)[0{,}5YQW_2 + 0{,}3YQW_3 + 0{,}2YQW_4]\} + 0{,}5\{\alpha ECW_1 + (1-\alpha)[0{,}5ECW_2 + 0{,}3ECW_3 + 0{,}2ECW_4]\}\}$

Being YQW_1, YQW_2, YQW_3, YQW_4, ECW_1, ECW_2, ECW_3 and ECW_4 the Yearly Quota Weighting and Equity Capital Weighting values corresponding to the first four years of forecasted activity of the project analyzed. α default value is 0,5. Maximum value reachable for RR is 0,5.

STRUCTURAL OPERATIONAL ANALYSIS (SOA)

The SOA is formed by a set of 20 multiple choice questions with five possible answers each of them, aligned with the degree of accomplishment of the question formulated. One example of these questions, and its possible answers, would be:

- Knowledge and calculation of the break-even point:
- It does not exist.
- It has not been calculated.
- It is not realistic.
- It is not updated or it is not used.
- It is realistic, accurate and updated.

So that, proposed questions tackled the following topics relating both the project and the promoter themselves:

- Accuracy and veracity of the information provided by the promoter
- Viability of the product concept
- Knowledge of customer needs
- Number of clients
- Sufficient liquidity to start-up
- Sufficient liquidity to individual needs
- Income and expenditure plan
- Cash flow plan
- Calculation of the break-even point
- Calculation of costs and prices
- Information about potential or actual partners/Networking capabilities
- Business plan
- Family support
- Training and experience of the promoters
- Sector knowledge
- Commercialization strategy
- Processes and control mechanisms
- Level of ICT use
- Planning capabilities
- Motivation for promoters

SOA value ranges from 0 to 1, being possible to express it as a percentage (0 to 100) as well.

CREDIT OFFICER ASSESSMENT (COA)

The COA is made up by a set of 14 multiple choice questions with three possible answers each of them, aligned with the degree of accomplishment of the question formulated. One example of these questions, and its possible answers, would be:

- Is the information provided by the developer/s consistent with the project presented?
- The provided information is consistent with the project.
- The provided information about the project presents inconsistencies.
- The provided information about the project, idea and market presents important inconsistencies.

Full set of questions is as follows:

- Is the information provided by the promoter/s consistent with the project presented?
- Is the loan applied relevant for the business?
- Is the project already ongoing?
- How is the evolution of the sector?
- How does competition affect the project?

- Are there any barriers for the implementation of the business?
- In a hypothetical cessation of activity, are there any barriers out?
- Are the sales seasonal?
- What is the validity of the business model in relation to the market opportunity?
- Is the investment consistent in accordance with the project submitted?
- Has the promoter/s another additional sources of income outside the business?
- What is the promoter/s personal motivation?
- Does the promoter/s personality affect the management of the business?
- Is the promoter/s actually established in the environment of the activity carried out?

COA value ranges from 0 to 1, being possible to express it as a percentage (0 to 100) as well.

NEGATIVE CREDIT RECORD (NCR)

In accordance with the official credit record presented by the loan applicant, a pre-defined value from 0 to 1 will be deducted. Therefore, this is always a negative value which will be 0 exclusively in the cases were the credit record do not show any unpaid quota from another previously taken loan and the level of risk is low, not exceeding a predetermined level.

EXPERIENCE BASED CORRECTOR INDEX (CI)

The CI component has been included into the ERP formula in order to adjust it in accordance with data and information gathered from actual implementations in relation with critical issues that could noticeably affect the result. These would be experienced tested factors and circumstances that eventually could negatively affect the correct process of the assessed project itself, such an incoherent location of the premises in accordance with the activity to be carried out or a total disregard of the promoter/entrepreneur. Therefore, the CI can be considered as a safety component in order to avoid previously detected causes or factors of potential failure not included or foreseen in the default formula. This will imply a permanent adaptation of the algorithm aligned with relevant experienced results gathered during actual implementation periods.

As for the NCR, the CI is always 0 or a negative value ranging from 0 to 1.

WEIGHTING COEFFICIENTS

α, β y γ are coefficients aligned with the relative strength assigned to each one of the three components RR, SOA and COA. α, β y γ coefficients range from 0 to 1. Default values are 0.5, 0.25 and 0.25 respectively.

EMPIRICAL TESTING

As result of different factors, most of them direct consequences of the global financial crisis initiated in 2007, SMEs and entrepreneurs were forced to face the crudest credit crunch of last decades. Besides, guarantees and collaterals required by regular financial and banking institutions in Spain were higher than ever, often impossible to be gathered. Traditional financial institutions were avoiding small credit operations which was particularly dramatic for medium - low sized investments, directly affecting entrepreneurs, start-ups and SMEs. In fact, credit facilities under 25.000.- € were very restricted and limited in Spanish regular financial markets. In commercial terms, this supposed a market gap with real needs to be covered. In accordance with this, it was decided to offer microcredit services through a fund provided by the Regional Government of Burgos to these collectives (Entrepreneurs, SMEs and self-employees), mainly aiming at start-ups, and taking into account social issues such unemployment and rural development. It was, therefore, understood that financial support, in combination with mentoring, is essential in order to foster the creation of new enterprises. The geographical reach was, initially, Burgos province. Further

extensions to the whole Castilla y Leon region would be considered. The financial instrument conceived for that purpose was a micro credit with the following characteristics:

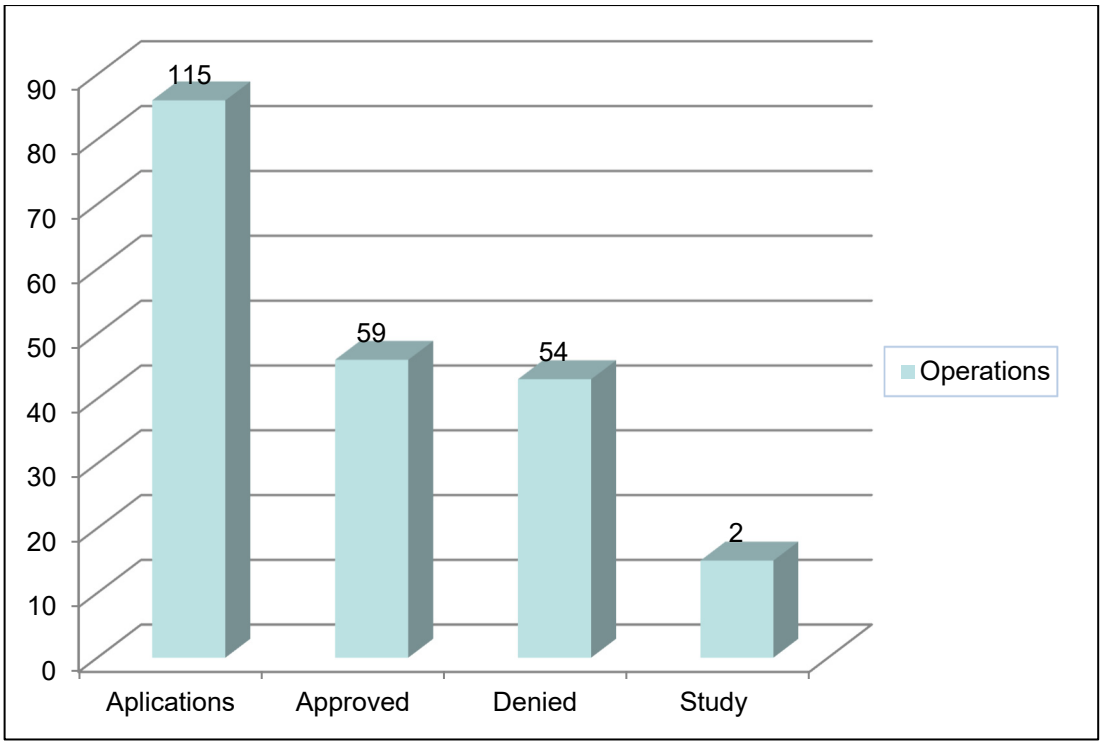
- 4/5 years term.
- 6 months grace period (Optional).
- 1,5 – 3% fix interest rate.
- Maximum amount 25000 €, including 5000 € for working capital.
- No fees charged.
- No guarantees/No collaterals.

The methodology itself is determined by this lack of guarantees or collaterals, being, therefore, based on the four subsequent pillars:

- Financial figures and ratios of the operation.
- Business project features (Business model, structure, market trends…).
- Entrepreneur management skills.
- Credit record.

The referred micro credit fund, managed by CEEI-Burgos, was raised and set up on 01/01/2013. From this moment onwards the fund was fully operative summing and initial capital of 150000€. This amount has been yearly increased till 825000€ by June 2017. The number of operations for the period of 01/01/2013 – 03/07/2017 are the following ones:

FIGURE 1
ERP OPERATIONS



For this period the following indicators of performance have been extracted:

TABLE 1
ERP PERFORMANCE INDICATORS

INDICATORS OF PERFORMANCE	
Total budget (Microfinance fund)	825.000,00 €
Fully repaid	412.645.99 €
Total leveraged investment (approved applications)	3.652.563,99 €
Total loans granted	1.097.354,88 €
Jobs created	69
PAR (NPL 30)	1,21 %
LAR 30	2,12 %
Write-off ratio	0 %
Survival projects/companies ratio	98,31%

Being NPL 30 the standard ratio to assess the quality of the portfolio indicating the portfolio at risk (PAR) past due more than 30 days. NPL stands for "Non-performing loans". NPL 30 is given by:

$$\text{NPL 30} = \frac{Total\ ammount\ at\ risk}{Outstanding\ principal\ balance\ of\ all\ loans}$$

Being: *Total amount at risk = Outstanding principal balance of all loans with at least one payment past due more than 30 days + Outstanding balance of loans that are not more than 30 days late but have been renegotiated.*

LAR, which does not differ much from PAR, is a simple indicator that counts the number of loans at risk instead of their amounts. According to this:

$$\text{LAR 30} = \frac{Number\ of\ loans\ more\ than\ 30\ days\ late}{Total\ number\ of\ outstanding\ loans}$$

Finally:

$$\text{Write-off ratio} = \frac{Value\ of\ loans\ written\ off\ during\ period}{Average\ gross\ loan\ portfolio\ during\ period}$$

These empirical results show the efficacy and efficiency of this methodology, which have been tested in actual scenarios and environments, in comparison with current financial risk assessment methods used by the traditional banking and financial system, pointing out the fact that guarantees or collaterals have not been requested in ERP field implementations which, theoretically, would increase the risk factor itself. However, the NPL 30 ratio shows figures far below the banking system ones, despite the common and general used of guarantees and collaterals in its operations.

CONCLUDING REMARKS

In this work, a new method of financial risk assessment, presented as an operational composite indicator to evaluate the estimated recovery probability of an entrepreneur's loan, or any other business-oriented loans, has been introduced.

The core aspect of this methodology is the ad hoc coined concept of "Estimated Recovery Probability" as an estimation of the measure of the likelihood that an event will occur. Not being a pure mathematical probability calculation, the value is, however, equally expressed as a percentage.

This approach integrates not only pure financial aspects of the operation but also management skills of the promoter and the business model of the project assessed itself. This analysis is determined by the lack of guarantees or collaterals, therefore focusing in the security principle understood as the feasibility of the operation itself.

ERP concept has been tested in actual environments with outstanding results in terms of accuracy in accordance with the actual developments of the assessed projects themselves. The calculated NPL 30 ratio which value is 1,21% positively stands out among average traditional banking default rates.

In order to avoid the formula to get out-dated, it includes a corrector index (CI) able to allow the adaptation of the algorithm to actual present and future circumstances that could affect the correct assessment process.

ERP method is a useful tool to be used by MFIs (Micro Finance Institutions) or any other finance institution providing credits or loans in highly uncertain environments or circumstances.

Future research will be focused in further adjustments of the "Experience based corrector index (CI)" and potential inclusion of other standard financial ratios.

CONCEPTUAL BACKGROUND

Since ERP method is a totally new approach, not based on previous works, the below listed references should be considered as general concepts not specifically mentioned in any particular parts of the text:

EFQM Excellence Model

Start – up Lifecycle Strategy Analysis by Harvard Business School

BLUES Brain Logistics Start-Up Evaluator, JIC (South Moravian Innovation Centre)

An Examination of the Relationships Between Audit Automation and Performance Measurement Implementation in Government Internal Audits

Stephen Kwamena Aikins
University of South Florida

This study examines the effects of audit automation such as the uses of Computer Assisted Audit Techniques on successful implementation and integration of performance measures into management of government internal audits. The researcher analyzed a survey of local government auditors in the USA. Results reveal that successful integration of performance measures into ongoing audit management is influenced by audit use of Computer Assisted Audit Techniques. Additionally, successful implementation and integration of performance measures into audit management are influenced by the use of audit time reporting system, presence of dedicated information technology audit staff, and posting audit reports on audit websites.

Keywords: government internal audit, performance measurement, information technology, computer assisted audit techniques, audit automation

INTRODUCTION

The objective of this study is to investigate the uses and effect of information technology (IT) such as audit automation on successful implementation and integration of performance measures into the management of government internal audit departments. Performance measurement system supports the management control system as a formal information-based routines and procedures that managers use to maintain or alter patterns in organizational activities (Jamil & Mohamed, 2013; Simons, 1994). Ideally, a well-developed, broad-based performance measurement system could be used to improve organizations' strategic focus (Kaplan & Norton, 1996). Performance measurement has an important role to play in the efficient and effective management of an organization (Kennerly & Neely, 2002), and performance measures allow organizations to allocate economic responsibilities and decision rights, set performance targets, and reward target achievements (Merchant & Van der Stede, 2007). Through their monitoring and risk mitigation role, internal auditors can add value through audit reviews (Erasmus & Coetzee, 2018; Vadasi et al., 2019) to ensure IT is effectively utilized in their organizations to enable adequate and accurate measurement and linking of performance to organizational outcomes.

Successful use of IT to implement and integrate performance measurements into operations could result in audit productivity and overall quality performance. In a global IT audit benchmarking study, GAIN (2009) found 63 percent of survey respondents use Computer Aided Audit Technique (CAAT) tools such as data extraction software, 76 percent use data analysis software and 52 percent use automated working paper software. When asked to indicate how the software has improved their capabilities and also provide best practice linked to the use of the software identified, many respondents stated the use of the above-mentioned software has led to improved productivity and efficiency of work. The results also showed some

of the common software used include ACL, Excel and MS Access (GAIN, 2009). Other studies (e.g. Stacoikas & Rupsys, 2005; Coderee, 1993) have also concluded the use of IT has impact on audit effectiveness, and others (e.g. Devaraj & Kohli, 2003; Doll & Torkzadeh, 1998) found a link between IT usage and organizational performance.

Although the above-mentioned studies make significant contribution toward the literature on audit automation and IT impact on organizational performance, the extent to which the use of IT influence successful integration of performance measures into the management of government audit department has been largely ignored. Investigating the impact of IT on successful implementation and integration of performance measures into audit management is important because facilitation of such integration by IT will enable audit to standardize the needed benchmarks to strengthen audit performance management.

THEORETICAL FRAMEWORK

The literature on IT and audit automation abounds with several studies on uses and impacts of IT on internal audit value creation regarding performance, quality, efficiency and effectiveness (e.g. Stacoikas & Rupsys, 2005; Saikh, 2004; Deloitte & Touche, 1996; Debrecency et al., 2005; Jackson, 2004; Moorthy et al., 2011; Kim et al., 2009). Many scholars (e.g. Devaraj & Kohli, 2003; Doll & Torkzadeh, 1998) have discussed the significance of the link between IT usage and organizational performance. In a study that examined the performance impact of information technology, Devaraj and Kohli (2003) concluded there is a general support for the proposition that the greater the actual usage of technology, the financial and quality performance of hospitals. Thus, for IT impact to occur, it is imperative that usage is tied to organizational performance metrics (Devaraj & Kohli, 2003). Indeed, tying IT to organizational performance metrics implies IT serving as an enabler for performance measurement data to be integrated into management processes.

The relationship between IT use and successful implementation, as well as successful integration, of performance measures into audit management activities can be explained using the theory of resource based view (RBV). This is because IT is an organizational resource that can be strategically managed and deployed to enhance organizational performance. RBV suggests that the resources possessed by an organization are the primary determinants of its performance, and may contribute to sustainable competitive advantage of the organization (e.g. Hoffer & Schendel, 1978; Wenerfelt, 1984). According to Barney (1991), the concept of resources includes all assets, capabilities, organizational processes, firm attributes, information, knowledge, etc. controlled by an organization that enable it to conceive of and implement strategies that improve its efficiency and effectiveness (Barney, 1991; Daft, 1983). Grant (1991) argues resources are inputs into the production process. They include skills of individual employees, finance, capital equipment, etc. On their own, few resources are productive, which implies productive activity requires the cooperation and coordination of teams of resources. In this respect, Powell and Dent-Micallef (1997) maintain that as a resource, information and communication technology (ICT) alone does not provide sustainable competitive advantage: its use along with complementary human and organizational resources such as flexible culture, integration of ICT and the organization's strategy, is what allows organizations to obtain competitive advantage.

Breznik (2012) suggests IT can be seen as a bundle of resources, not merely a single resource. Ross et al. (1996) recognized the composition of three IT resources: (1) human resources; (2) technology base; and (3) relationships between IT and business management that can lead to expected business value. From the standpoint of government internal audit, technology base resource include having various information systems with different functionalities, including audit sampling, workpaper file, audit workpaper formats, audit time reporting, report development and a computer assisted audit (CAAT) for data extraction and analysis. Additionally, IT human resources include having a dedicated IT audit staff with the capability of effectively utilizing the above-mentioned systems to enhance audit performance. Gustavson and Sundström (2018) argue enhanced human resources and educational skills are among the essential requirements for good auditing in public organizations. Carr (2003) asserts that the economic and strategic IT impact comes from continual innovation of IT and that many organizations have gained important advantages through the

innovative use and exploration of IT. For example, in the case of government internal audit, the use of CAAT software for data extraction and analysis can help to integrate performance measures into audit management and enhance overall audit performance by focusing audit attention on areas of significance and high risks, and by helping to set a benchmark for assessing the extent to which audit is able to detect frequency of auditee errors and fraudulent transactions.

Similarly, the use of audit time reporting system can help audit management to determine the percentage of auditors' time spent on specific tasks and auditors' contribution toward achievement of audit goals. Considering the fact that cooperation and coordination of teams of IT resources can enhance productivity (Barney 1991), and that human and other resources can complement IT to provide meaningful advantage (Powell and Dent-Micallef 1997), the case can be made that successful implementation and integration of performance measures into the management of the audit function can be achieved if IT resources are utilized by audit personnel. Therefore, we can expect positive relationships between audit staff uses of IT resources in the form of automated tools like CAAT, audit time reporting system, audit sampling system, audit workpaper format, audit report development system as well as audit workpaper file and successful implementation, as well as integration of performance measures into management of government internal audit.

RBV theory also discusses IT capability to help explain IT impact on organizational performance. Capability is the capacity for a team of resources to perform some task or activity (Grant, 1991). Bharadwaj (2000) distinguishes capability from resources and explain "IT capabilities are abilities that mobilize and deploy IT resources, such as automated tools, in combination with other resources and capabilities (p.171). Amit & Schoemaker (1993) also distinguish resources from capabilities by defining resources as stocks of available factors that are owned or controlled by an organization, which are converted into final products or services. Capabilities, in contrast, refer to an organization's capacity to deploy resources, usually in combination, using organizational processes to produce a desired outcome (Amit & Schoemaker, 1993). Explaining how CAAT-based programs can help automate certain audit functions, Coderee (1993) argued an automated tool such as CAAT has several benefits for audit planning and reporting because it can be deployed to increase audit coverage, improve integration of audit skills, strengthen independence of auditing from information system functions, foster greater credibility and increase cost-effectiveness through the development of reusable computerized techniques.

Independence is the cornerstone of internal audit (Al-Akra et al. 2016) and helps to demonstrate capability of the auditor. Consequently, the degree of perceived independence affects the role of internal audit mechanisms (Drogalas et al (2020) including capability of the audit function. This implies the presence of IT auditor capability enables resources to begin to be utilized effectively, and the potential for the creation of output and outcome arises. Auditors' cognitive and technical skills constitute an important capability affecting internal control implementation and effectiveness (Petridis et al., 2019; Mahadeen et al. 2016). Based on the arguments of RBV, if government internal audits have the capability manifested by the presence of a dedicated IT audit staff utilizing audit IT resources, such as CAAT and audit time reporting system, such efforts could produce desired outcomes regarding successful implementation, as well as successful integration of performance measures into the management of the audit function. Therefore, we can expect positive relationships between the presence of dedicated IT audit staff, uses of automated tools and successful implementation, as well as integration of performance measures into the management of the audit function.

Staciokas & Rupsys (2005) explored the implications of IT for internal audit functions and analyzed the advantages of internal audit in organizational governance. The authors concluded although IT resources like automation of workpaper files has drawbacks such as network breakdown and data loss, the use of Computer Assisted Audit Tools (CAAT) increased internal audit effectiveness. In another study, Debrecency et al. (2005) concluded one of the most widely deployed CAATs is generalized audit software (GAS), examples of which are the audit command language (ACL), interactive data extraction and analysis (IDEA) and Panaudit Plus. In its annual benchmarking study that determined performance measurements for internal audit activities worldwide, GAIN (2009) reported that 76 percent of the survey respondents measure the quality, effectiveness and efficiency of internal audit processes. Therefore, if resources like

CAAT increases audit effectiveness as stated above, then we can expect a positive relationship between the use of CAAT and successful implementation and integration of the measures of effectiveness such as the use of audit issue tracking system to determine the number of audit recommendations implemented.

Shaikh (2004) explored the impact of e-commerce on audit processes and methodologies, the application of technologies that may assist auditors in improving audit quality, and how to use CAATs effectively with emerging technologies. The author disclosed that emerging technologies such as object-oriented distributed middleware, Internet security technology and intelligent agents constitute software frameworks that facilitate electronic auditing – a process where some audit tasks are conducted electronically over the Internet with the support of ITs (Moorthy et al., 2011). From this finding, it is obvious that IT resources that facilitate electronic auditing such as work paper formats, work paper filing and report development systems improve audit performance from the standpoint of efficiency and effectiveness because they speed up supervisory review and reduce the number of audit travels and audit costs. Audit efficiency through enhanced scientific approach to internal control reviews and risk management can lead to improved organizational performance and management processes (Danescu et al., 2015; Khalid et al. 2017). Therefore, to the extent that audit time reporting management (i.e., planned versus actual reporting time), for example, is audit performance metric, then we can expect internal auditors' use of the above-mentioned audit IT resources to generate audit reports for electronic auditing to have direct impact on successful implementation and integration of that performance measure through speedy supervisory review.

Government internal auditors' specific uses of a resource such as Internet technology could also have impact on successful implementation and integration of performance measures into audit management activities. The implementation of internal audit in the public sector is of paramount importance because it helps to improve governance processes of local government organizations (Noraini et al. 2018). The exponential growth of the Internet has contributed immensely to electronic, web-based Internet reporting in both government and the private sector. Currently, many government internal audit departments have web presence and post audit reports online for public access. In a study that explored the widespread organizational reporting on the Internet and its implication for the auditing profession, Khadaroo (2005) found a significant increase of use of the Internet to supply information to the public. Brown (2008) argues transparency and public access to performance management results and data will provide external pressure to ensure its sustainability and will also have the potential to create a positive force to reward and support improved results. From government internal audit perspective, having web presence and posting audit and performance measurement reports online ensures transparency and aids in the accountability of auditors to citizens and elected officials. This could facilitate the monitoring of auditor performance and motivate the internal audit departments to develop and successfully implement performance measurements. Based on the discussion above, the following hypothesis are developed for testing:

H1. *There is a positive relationship between the use of audit sampling system and successful implementation, as well as successful integration, of performance measures into audit management activities.*

H2. *There is a positive relationship between the use of audit workpaper filing system and successful implementation, as well as successful integration, of performance measures into audit management activities.*

H3. *There is a positive relationship between the use of audit workpaper formats and successful implementation, as well as successful integration, of audit performance measures into audit management activities.*

H4. *There is a positive relationship between the use of audit time reporting system and successful implementation, as well as successful integration, of performance measures into audit management activities.*

H5. *There is a positive relationship between the use of audit issue tracking system and successful implementation, as well as successful integration, of performance measures into audit management activities.*

H6. *There is a positive relationship between the use of audit report development system and successful implementation, as well as successful integration, of performance measures into audit management activities*

H7. *There is a positive relationship between the use of CAAT and successful implementation, as well as successful integration, of performance measures into audit management activities.*

H8. *There is a positive relationship between the presence of dedicated IT audit staff and successful implementation, as well as successful integration, of performance measures into audit management activities.*

H9. *There is a positive relationship between having an audit web site and successful implementation, as well as successful integration, of performance measures into audit management activities.*

H10. *There is a positive relationship between posting audit report on audit web site and successful implementation, as well as successful integration, of performance measures into audit management activities.*

METHODOLOGY

This research utilizes data from a 2012 survey conducted by the Association of Local Government Auditors (ALGA). The survey was sent to audit heads of 263 audit shop members and 114 returned the survey, yielding a 43 percent response rate. Fifty-two percent of the respondents work for city government, 27 percent work for county government, and 21 percent work for other types of governmental entities, including schools/Universities, utilities and transportation. Twelve percent of the respondents are in the Midwest of the United States, 18 percent in the pacific west, 12 percent in the mountain west, 5 percent in the northeast, 46 percent in the south east and south west and 5% are international. Twenty-five percent of respondents have audit staff of between 1-2, 34 percent have 3-5 staff, 21 percent have 6-10 staff, 11 percent have 11-15 staff, and 9 percent have 16 or more staff. Based on the profile presented above, the case can be made that the survey respondents constitute adequate and fairly balanced representation of the local governments and regions of the United States, as well as departmental size in terms of the number of audit staff.

The areas surveyed include performance measurements and uses of IT, i.e. audit automation such as CAAT to help measure the effect of audit automation on successful implementation and integration of performance measurements into management of the government internal audit function. Respondents were asked whether their government organization uses performance measurements and whether their audit departments use performance measurements. Additionally, they were asked to state on a seven point scale, their agreements regarding the following statements: 1) Your department has developed and implemented performance measures; and 2) Your department has integrated performance measures into management of the department. The scale was as follows: 1= Completely Disagree; 2 = Strongly Disagree; 3 = Somewhat Disagree; 4 = Neutral; 5 = Somewhat Agree; 6 = Strongly Agree; and 7 = Completely Agree. Respondents were also asked to state on a seven point scale, their agreement regarding the following question: Overall, to what extent do you agree the following have been successful in your department? 1) development and implementation of performance measures; and 2) integration of performance measures into the management of the audit function.

Respondents were also provided a list of automated audit tools and asked the following question: What automated tools does your audit department use? Select all that apply, or specify other. The following are

the list of automated audit tools provided to respondents: Audit Sampling, Work Paper Formats, Work Paper File System, Audit Time Reporting, Issue Tracking, Report Development, Computer Aided Audit Techniques (ACL, IDEA, etc.). Additionally, respondents were asked to indicate whether they have dedicated IT audit staff, whether they have a website and whether their reports are posted on the websites. Descriptive statistics was used to analyze the uses of performance measures and Chi-Square tests were used to determine the relationships between uses of automated audit tools and successful implementation and integration of performance measurements into the management of the audit functions.

RESEARCH FINDINGS

Use of Performance Measures

Ninety one respondents answered the questions on whether their audit departments use performance measures and 67 (74%) responded in the affirmative. A total of 86 respondents provided responses regarding the success of developing and implementing performance measures, and of integrating performance measures into the management of the audit function. Table 1 provides details of responses provided. Twenty six respondents said they somewhat agree that development and implementation of performance measures in their audit department has been a success, 22 indicated they strongly agree and 6 noted they completely agree. Together, these also represent 63% of the 86 respondents. Twenty five respondents said they somewhat agree that integration of performance measures into their audit functions has been a success, 20 indicated they strongly agree and 7 noted they completely agree. Together, these represent 60% of the 86 respondents. Clearly, a strong majority of the respondents believe implementation and integration of performance measures in their audit departments have been a success.

TABLE 1
SUCESSFUL IMPLEMENTATION & INTEGRATION OF PERFORMANCE MEASURES

	Success of Developing and Implementing Performance Measures in Audit Department	Success of Integrating Performance Measures into Audit Functions	N
Completely Disagree	4	5	86
Strongly Disagree	4	6	86
Somewhat Disagree	5	5	86
Neutral	19	18	86
Somewhat Agree	26	25	86
Strongly Agree	22	20	86
Completely Agree	6	7	86

Use of Audit Automation

Table 2 shows the details of respondents' use of various types of software applications for audit automation. As can be seen from Table 2, 46% of respondents use audit sampling system, 54% use audit work paper formats,

TABLE 2
USES OF AUTOMATED TOOLS, DEDICATED IT AUDIT STAFF & WEB PRESSENCE

Automated Tool, Dedicated IT Audit Staff and Web Presence	Frequency	Percentage of Total Survey Respondents	N
Use Audit Sampling System	52	46%	114
Use Work Paper Formats	62	54%	114
Use Work Paper File System	50	44%	114
Use Audit Time Reporting	51	45%	114
Use Issue Tracking	36	32%	114
Use Report Development	28	25%	114
Use Computer Assisted Audit Techniques-CAAT (ACL, IDEA etc.)	57	50%	114
Use Other Automated Tools	107	94%	114
Has Dedicated IT Audit Staff	26	23%	114
Has Audit Website	70	61%	114
Post Audit Report on Website	51	45%	114

44% use work paper file system and 45% use audit time reporting system. The results in Table 2 also show only 32% of respondents use issue tracking and even a smaller 25% use automation for report development.

An issue tracker is an important tool that enables auditors to monitor resolution of outstanding recommendations issued, and to make informed decisions regarding the audit client's internal controls as well as follow-up audit activities. Additionally, the results reveal that 50% of respondents use CAAT data extraction and analysis tools such as ACL and IDEA. Among other capabilities, these tools help to extract and analyze data to enable auditors to form an opinion on audit evidential matter. Ninety four percent of respondents also use other automated tools for audit. These tools typically include but are not limited to Microsoft Excel and Access. As noted in Table 2, only 23% of the survey respondents said they have dedicated IT audit staff. This could be due to the fact that many of audit shops outsource their IT audits. The results also show while a majority of 61% of respondents have audit websites, only 45% post their reports on the websites.

IT Use, Performance Measures Implementation and Integration

Chi-Square tests of independence were used to test the relationships between IT and successful implementation, as well as successful integration, of performance measures. Table 3 shows the details of the test of the relationships between the use of automated tools, having dedicated IT audit staff as well as web presence, and the successful implementation of performance measurements. As noted in Table 3, Successful implementation of Performance Measurements is influenced by the use of: Audit Time Reporting System ($p = 0.016$), Dedicated IT Audit Staff ($p = 0.017$), as well as having Audit Reports on Website ($p = 0.039$). These confirm research hypotheses 4, 8 and 10. Good performance measurements measure efficiency of audit activity such as actual audit hours versus planned hours. Using an audit automated tool such as audit time reporting system helps to track audit hours and determine whether actual hours worked measure favorably against planned hours. The result in Table 3 therefore reveals that audit use of a time reporting system provides that monitoring mechanism which influences successful implementation of the efficiency measure of actual versus planned audit hours.

TABLE 3
CHI-SQUARE TEST – IMPACT OF IT USE ON PERFORMANCE MEASURES (PM) IMPLEMENTATION

Uses of Audit Automated Tools, Having IT Audit Staff, And Audit Web Presence	**Successful Implementation of PM**		
	Pearson Chi-Square		
	Value	df	Asymp. Sig. (2-sided)
Use of Audit Sampling System	11.660	6	.070
Use of Audit Work Paper Formats	1.893	6	.929
Use of Work Paper File System	.646	6	.996
Use of Audit Time Reporting	15.612	6	**.016**
Use of Issue Tracking	10.194	6	.117
Use of Report Development	11.774	6	.067
Use Computer Assisted Audit Techniques-CAAT (ACL, IDEA etc.)	10.101	6	.120
Use of Other Automated Software	8.622	6	.196
Has Dedicated IT Audit Staff	15.468	6	**.017**
Has Audit Website	11.035	6	.087
Posts Audit Reports on Website	13.277	6	**.039**

Having dedicated IT audit staff usually helps to maintain an in-house expertise in the use of CAAT tools like ACL to effectively perform data extraction for analysis and efficient audits. This helps audit to sort, view and analyze large amounts of data in order to identify internal control weaknesses and risks inherent in the audit clients' operations. The result in Table 3 implies that by having dedicated IT audit staff, a government internal audit department most likely has the capability to extract and analyze high volume of data that results in successful implementation of efficiency measures such as the breadth of audit coverage, and degree of identifying high risk transactions for targeted audit. Posting reports of audit activities online helps to ensure transparency, accountability and public awareness of audit effectiveness. Based on the results in Table 3, public access to performance management results over the Internet influences government internal auditors' successful implementation of performance measures in order to meet the expectation of accountability from a well-informed public about audits' activities.

Table 4 shows the details of the test of the relationships between the use of automated tools, having dedicated IT audit staff as well as web presence and the successful integration of performance measures. The results shown reveal that Successful Integration of Performance Measures is a function of the use of: Audit Time Reporting ($p = 0.005$), Computer Assisted Audit Techniques like ACL and IDEA ($p = 0.032$), as well as having Dedicated IT Audit Staff ($p = 0.017$), and Audit Reports on Website ($p = 0.010$). These results confirm research hypotheses 4, 7, 8 and 10.

These results indicate that the same factors (audit time reporting, dedicated audit staff and audit report on website) that influence successful implementation of performance measures of government internal audit departments also impact successful integration of performance measures into the management of the audit function. However, unlike successful implementation of performance measures, successful integration into the management of the audit function is also significantly influenced by CAAT data extraction and analysis software such as ACL.

TABLE 4
CHI-SQUARE TEST – IMPACT OF IT USE ON PERFORMANCE MEASURES (PM) INTEGRATION

Uses of Audit Automated Tools, Having IT Audit Staff, And Audit Web Presence	Successful Integration of PM		
	Pearson Chi-Square		
	Value	df	Asymp. Sig. (2-sided)
Use of Audit Sampling System	10.180	6	.117
Use of Audit Work Paper Formats	4.609	6	.595
Use of Work Paper File System	.595	6	.892
Use of Audit Time Reporting	18.588	6	**.005**
Use of Issue Tracking	10.653	6	.100
Use of Report Development	11.856	6	.065
Use Computer Assisted Audit Techniques-CAAT (ACL, IDEA etc.)	13.833	6	**.032**
Use of Other Automated Software	7.527	6	.275
Has Dedicated IT Audit Staff	15.468	6	**.017**
Has Audit Website	11.496	6	.074
Posts Audit Report on Website	15.170	6	**.010**

DISCUSSION

The result of this study reveals that many local government auditors use performance measures and majority of those believe their implementation and integration of performance measures into audit management functions have been successful. The findings also reveal that 94% of local government internal auditors use some form of automated tools for their audit work, although less than half of them (46%) use audit sampling, 54% of use audit work paper formats, and 44% use work paper filing system. The use of audit work paper format and work paper filing system are a bit less than the findings of the GAIN (2009) worldwide benchmarking study which showed a little more than half of study respondents (52%) use software to automate their working papers. Automating audit work papers can provide standardization and consistency of audit processes, which could lead to high productivity and efficiency of work, thereby facilitating the development of benchmarks for measuring audit performance. Additionally, the use of automated work paper software can reduce the time required for audit file reviews, help with better organization of audit information and enable review of audit work papers from remote locations. Despite its findings, GAIN (2009) concluded there is the need for improvement in the use of audit software because most of its study respondents do not use software to detect or investigate fraud, perform control self-assessment, monitor compliance activities and assess risks for the annual audit plan. The relatively low use software to automate working papers may be due to budgetary constraints of some audit departments stemming from inadequate allocation of resources. Therefore, the result of this study appears consistent with that of Jackson (2004) who found several limitations in implementing audit software, including cost implications and auditor resistance to training.

As revealed in Tables 3 and 4, the use of an audit time reporting system is significantly related to successful implementation, as well as successful integration of performance measures into the management of the audit function ($p = 0.016$ and $p = 0.005$ respectively). Effective time management goes to the core of audit efficiency and workload management. Having a time reporting system imply possession of an IT resource that accurately tracks audit hours can help monitor audit performance by comparing planned hours

to actual hours in order to determine audit efficiency and put in place needed corrective measures to improve performance. Therefore, the finding that the use of audit time reporting system significantly impact both the implementation and integration of performance measures into audit management is consistent with the RBV theory (e.g. Hoffer & Schendel, 1978; Wenerfelt, 1984) that resources possessed by an organization are the primary determinants of its performance. Efficient time management can also enhance audit cost-effectiveness and workload management by ensuring that time savings on work performed are utilized on more engagements. In those government agencies where internal audit operates as a cost center, an effective time reporting system will help to determine billable hours to auditees.

Tables 3 and 4 also reveal that there is a statistically significant relationship between having a dedicated IT audit staff and successful implementation and integration of performance measures into the management of the audit function ($p = 0.017$). IT audit staff usually are individuals who possess the capability and expertise in the use of data extraction and analytical software. As indicated above, their use of data analytics helps to increase efficiency by auditing more areas and reviewing higher volumes of data while using fewer resources. This finding is consistent with RBV arguments (e.g. Grant, 1991) that productive activity requires the cooperation and coordination of teams of resources, and that capability such as the ability to mobilize IT resources can lead to desired outcome (Bharadwaj, 2000; Amit & Schoemaker, 1993). The implication here is that automated tools like CAAT are, on their own, not productive unless they are coordinated and utilized by human resources like dedicated IT auditors with the capability to effectively utilize them for efficient risk-based audit, and enhanced audit performance. Additionally, the finding confirms the argument of Ross et al. (1996) that the composition of IT resources include not only the technology base but also the human resources.

The research findings also reveal that a majority of 61% of respondents have audit websites while 39% post the reports of their activities on audit websites. As revealed in Tables 3 and 4, there are statistically significant relationships between posting report of audit activities on audit website and successful implementation as well as successful integration of performance measures into the management of the audit function. By posting reports online, the results of audit performance measurement activities become available to relevant stakeholders such as citizens as well as elected and appointed government officials. Granted that these stakeholders view the audit function as effective and efficient based on its outputs that are placed in the public domain, the credibility of the audit function will not only be boosted but will also place it in a better position to compete effectively for much needed funding to help automate many of its activities. Therefore, the statistically significant relationship between posting audit reports online and successful implementation and integration of performance measures into audit management is consistent with the arguments of Brown (2008) and the National Performance Measurement Advisory Commission that public reporting will provide external pressure to have the potential to create a positive force to reward and support improved results.

According to the research findings, while only 23% of respondents have dedicated IT audit staff, 50% of them use CAAT data extraction and analysis software like ACL and IDEA. As revealed in Table 4, the research results show local government auditors' use of CAAT is significantly related to successful integration of performance measures into the management of the audit function ($p = 0.032$). The significance of this finding lies in the fact that these CAAT software help to analyze entire population rather than a sample in order to focus audit attention on areas of significance and high risks and to detect frequency of errors and fraudulent transactions. In so doing, they enable audit management to get a broader view of audit efficiency in terms of audit coverage, and provide the needed data and knowledge to standardize and manage measures of audit performance such as the degree to which audit identifies clients' internal control weaknesses. Thus, the use of data analytics help to automate and facilitate more efficient internal audit processes, thereby helping to identify and manage risks more promptly, effectively and efficiently. Therefore, the finding of significant relationship between the use of CAAT and successful integration of performance measures is also consistent with the argument of RBV theory that resources are primary determinants of performance. Additionally, it is consistent with the argument of Coderee (1993) that CAAT has several benefits for audit planning and reporting because it can increase audit coverage, improve

integration of audit skills, and increase cost-effectiveness. Furthermore, it is consistent with that of Staciokas & Rupsys (2005) that the use of CAAT increases audit effectiveness.

An important finding regarding local government audit use of automation is the low percentage of respondents that use issue tracking system. This may be due to the fact that 94% of respondents use other automated tools, as Microsoft Excel and Access can also be used to track issues and outstanding recommendations. The finding from this study regarding the use of issue tracking appears consistent with some prior studies. In a study that examined the determinants of auditee adoption of audit recommendations, Aikins (2012) found that local government auditors are not keen on maintaining an issue tracker to enable them document control weaknesses identified in their audits. That study also found that while they perform follow-up audits to ensure audit recommendations are being implemented, they seemed indifferent regarding the timeliness of follow-up audits. Most importantly, that study found that the use of issue tracker and follow-up audits are significantly related to auditee adoption of audit recommendations. The implication here is that automation of audit issue tracking could enhance the effectiveness of government internal auditors because managers are more inclined to implement audit recommendations to address weaknesses identified in the course of the audit, and enhance public accountability, as long as government auditors document, track and follow-up to verify management's actions on those recommendations.

This study is limited in that it focuses on the uses of IT and their effects on successful implementation and integration of performance measures into audit functions. The study did not focus on the systems for measuring government internal audit performance. Additionally, the best practices associated with the creation and use of performance metrics was outside the scope of this study. Furthermore, the data analyzed in this study was sourced from a survey conducted several years before the study, and it is possible that an analysis of new data could yield different results. These limitations notwithstanding, the findings are useful in that there is very little empirical research on the effect of information technology on successful implementation and integration of performance measures into government internal audit management. Further research is required to focus exclusively on the systems for measuring government internal audit performance. Future research direction may include examination of the extent to which audit performance measures align with other elements of broader management control systems such as planning and administrative controls, and their effects on successful use of IT to integrate performance measures into audit management. For example, audit planning sets the goal of the audit function and the standards to be achieved, and directs auditor efforts and expected behaviour. Audit administration provides governance regarding line of accountability, rewards and monitoring of auditor behaviour to ensure conformance to standards. Future research into the alignment between the audit function's goals, standards and performance measures, and their effects on successful integration of performance measures into audit management will be a useful study.

CONCLUSION

Overall, the findings in this study suggest that the use of information technology such as CAAT, time management system and online reporting of audit activities can enhance government internal auditors' ability to measure the efficiency and effectiveness of their operations. In an era where government internal audit resources are limited due to fiscal stress of state and local governments, the efficient management of all aspects of government to enhance accountability becomes prevalent. This study contributes to the literature on government performance measurement because it adds technological perspective to the literature on the determinants of successful implementation of performance measurements from the standpoint of government internal audits. The study also contributes to the theory and practice of public administration. From theoretical perspective, it confirms the link between technology uses and organizational performance as noted in prior studies. From practical perspective, the findings suggest that government internal auditors who want to improve their performance will benefit from appropriate use of information technology to integrate their activities and improve implementation of selected measures in order to enhance efficiency. For this to happen, the limitation regarding the use of technology by government internal auditors has to be addressed.

REFERENCES

Al-Akra, M., Abdel-Qader, W., & Billah, M. (2016). Internal auditing in the Middle East and North Africa: A literature review. *Journal of International Accounting, Auditing and Taxation*, *26*, 13-27

Bourne, M. (2008). Performance measurement: Learning from the past and projecting the future. *Measuring Business Excellence*, *12*(4), 67-72

Brown, M.F. (2008). *Challenges to implement performance management*. National Performance Measurement Commission. Retrieved November 6, 2013, from http://www.gfoa.org/downloads/PMCommissionIssuePaper4.pdf

Coderee, D.G. (1993). *Automating the audit function*. Retrieved November 21, 2013, from http://www.findarticles.com/p/articles/aut0_audit

Debreceny, R., Lee, S.L., Neo, W., & Toh, J.S. (2005). *Employing generalized audit software in the financial services sector*. Retrieved January 14, 2014, from http://raw.rutgers.edu/docs/wcars/8wcars/contassurronaldjdaigle_files/DebrecenyCAATSBanks_paper.pdf

Danescu, T., Prozan, M., & Prozan, R.D. (2015). The valances of the internal audit in relationship with the internal control–corporate governance. *Procedia Economics and Finance*, *26*, 960–966

Devaraj, S., & Kohli, R. (2003). Performance impacts of information technology: Is actual usage the missing link? *Management Science*, *49*(3), 273-289.

Deloitte, T. (1996). *The Current Impact of Information Technology on Internal Auditing Departments*. Retrieved November 21, 2013, from http://www.theiia.org/ecm/tech.cfm?doc_id=859

Doll, W.J., & Torkzadeh, G. (1998). Developing a multidimensional measure of system-use in an organizational context. *Inform. Management*, *33*(3), 171-185.

Drogalas, G., Petridis, K., Petridis, N.E., & Zografidous, E. (2020). Valuation of internal audit mechanism in the decision support department of the local government organizations using mathematical programming. *Annals of Operations Research*. Retrieved, August 27, 2020, from https://link.springer.com/article/10.1007/s10479-020-03537-4

Erasmus, L., & Coetzee, P. (2018). Drivers of stakeholders' view of internal audit effectiveness: Management versus audit committee. *Managerial Auditing Journal*, *33*(1), 90–114.

Global Audit Information Network [GAIN]. (2008). Performance monitoring/quality assurance programs in the internal audit activity. *The Institute of Internal Auditors*, pp. 1-9.

GAIN. (2009a). IT audit benchmarking study: Executive summary and report. *The Institute of Internal Auditors Research Foundation*, pp. 1-76.

GAIN. (2009b). Knowledge Report: Measuring internal audit performance. *The Institute of Internal Auditors*, pp. 1-20.

Grembergen, W.V. (2004). *Strategies for Information Technology Governance*. Hershey, PA: Idea Group Inc.

Gustavson, M., & Sundström, A. (2018). Organizing the audit society: Does good auditing generate less public sector corruption? *Administration & Society*, *50*(10), 1508–1532.

Flanagan, C. (2005). *An investigation into the performance measurement practices of Irish Hotel Groups*. Unpublished M.Sc. Thesis. Dublin Institute of Technology, Dublin.

Frigo, M.L. (2002). *A balanced scorecard framework for internal auditing departments*. Altamonte Spring, FL: The Institute of Internal Auditors.

Harrison, M.J., & Datta, P. (2007). An empirical assessment of user perceptions of feature versus application level usage. *Communications of the Association for Information Systems*, *20*, 300-321.

Haas, L.D. (2001, June). The bottom line – value of internal audit department to firms. *Internal Auditor*, pp. 40-41.

Hatkamir, M., & Harris, P. (2005). Performance measurement practice in an independent hotel context: A case study approach. *International Journal of Contemporary Hospitality Management*, *17*(1), 39-50.

Hill, K.L., Highful, T., Driver, S., Garner, M., Goodson, S., MacCabe, D.J., & Young, H.S. (2009). *Performance measures for internal audit functions: A research project*. The Institute of Internal Auditors Research Foundation.

Jackson, R.A. (2004). Get the most out of audit tools. *Internal Auditor*, *61*(4), 36-44.

Jamil, C.Z.K., & Mohamed, R. (2013). The effect of management control system on performance management system at small medium hotel in Malaysia. *International Journal of Trade, Economics and Finance*, *4*(4), 202-208.

Kaplan, R.S. (1996). *The balance scorecard- Translating strategy into action*. Boston, MA: Harvard Business School Press.

Kellen, V. (2003). *Business performance measurement: At the crossroads of strategy, decision-making, learning and information visualization*. Faculty, DePaul University, Chicago: IL.

Kennerly, M.P., & Neely, A.D. (2002). A framework for the factors affecting the evolution of performance measurement systems. *International Journal of Operations and Production Management*, *22*(11), 1222-1245.

Khalid, A.A., Haron, H.H., &Masron, T.A. (2017). Relationship between internal Shariah audit characteristics and its effectiveness. *Humanomics*, *33*(2), 221–238.

Khadaroo, I. (2005). Corporate reporting on the Internet: some implications for the auditing profession. *Managerial Auditing Journal*, *20*(6), 578-591.

Khan, M.N.A., Baharun, R., Rahin, K.A., & Zakauan, N. (2011). An empricial evidence of performance measurement of audit firms in Malaysia. *International Business Research*, *4*(4), 191-198.

Kim, H.J., Mannino, M., & Nieschwietz, R.J. (2009). Information technology acceptance in the internal audit profession: Impact of technology features and complexity. *International Journal of Accounting Information Systems*, *10*(4), 214-228.

KPMG. (2004). *Building a success for internal audit: The balanced scorecard*. Singapore.

Kurtén, R. (2009). *IT-styrning: Vad ledning och styrelse behöver känna till om informationsteknologi*. Bulls Graphics, Halmstad.

Mahadeen, B., Al-Dmour, R.H., Obeidat, B.Y., & Tarhini, A. (2016). Examining the effect of the organization's internal control system on organizational effectiveness: A Jordanian empirical study. *International Journal of Business Administration*, *7*(6), 22–41.

Merchant, K.A., & Van der Stede, W. (2007). *Management control systems: Performance measurement, evaluation and incentives*. London, UK: Pearson/ Prentice Hall.

Moorthy, M.K., Jain, S.P. & Selangor, S.D.E. (2011). The impact of information technology on internal auditing. *African Journal of Business Management*, *5*(9), 3523-3539.

Munir, R., Perera, S., & Baird, K. (2011). An analytical framework to examine changes in performance measurement systems within the banking sector. *Australasian Accounting Business and Finance Journal*, *5*(1), 93-115.

Neely, A.D., Adams, C., & Kennerly, M. (2002*). The performance prism: The scorecard for measuring and managing business success*. Financial Times, Prentice-Hal: London.

Neely, A., Gregory, M., & Platts, K. (2005). Performance measurement system design: A literature review and research agenda. *International Journal of Operations, Production Management*, *25*(12), 1228-1263.

Noraini, S., Zaini, J.A., Mustaffha, N., & Norhanizah, J. (2018). Internal audit effectiveness in Zakat Institutions from the perspective of the auditee. *Management & Accounting Review*, *17(*3), 17-35.

Petridis, K., Drogalas, G., & Zografidou, E. (2019). Internal auditor selection using a TOPSIS/non-linear programming model. *Annals of Operations Research*. Retrieved August 27, 2020, from https://link.springer.com/article/10.1007%2Fs10479-019-03307-x

Shaikh, J.M. (2004). Ecommerce Impact: Emerging technology – Electronic auditing. *Managerial Auditing Journal*, *20*(4), 408-421.

Simons, R. (194). How new top manager use control systems as levers of strategic renewal. *Strategic Management Journal*, *15*(3), 169-189.
Staciokas, R., & Rupsys, R. (2005). *The internal audit role in organizational government*. Retrieved January 20, 2014, from http://www.academicjournals.org/journal/AJBM
Rupsys, R., & Boguslauskas, V. (2007). Measuring performance of internal auditing: Empirical evidence. *Engineering Economics*, *5*, 9-14.
Van Wijk, E.R. (2006, October). Collecting performance data: A CAE seeks a new benchmarking strategy to gauge his audit shop's effectiveness. *Internal Auditor*, pp. 30-32.
Yang, D.C., & Guan, L. (2004). The evolution of IT auditing and internal control standards in financial statement audits: The case of the United States. *Managerial Auditing Journal*, *19*(4), 544-555.
Vadasi, C., Bekiaris, M., & Andrikopoulos, A. (2019). Corporate governance and internal audit: An institutional theory perspective. *Corporate Governance*, *20*(1), 175–190.

The Role of Government Ownership in the Capital Structure of State-Owned Enterprises

Candra Chahyadi
Eastern Illinois University

Using a sample of 340 privatized firms from 51 countries, I examine their capital structure after privatization and analyze the capital structure determinants of privatized firms. The results indicate that privatized firms have a target leverage ratio and it is independent of the remaining government ownership. The leverage choice of privatized firms is affected by a high level of information asymmetry and government's future privatization agenda. In general, my study confirms the argument that privatized firms are unique and need to be separated from the sample of large firms when conducting international capital structure studies.

Keywords: privatization, capital structure, government ownership, initial leverage, corruption

INTRODUCTION

Although privatization has become a core economic policy in more than 100 countries and privatized firms account for a very significant portion of the privatizing country's GDP, it is very surprising that we still do not know much about the capital structure of privatized firms. Considering the size and the growing significance of privatized firms in countries in which privatization has taken place, the lack of activity in this area is unexpected. To the best of my knowledge, this is the first study examining the capital structure and the capital structure determinants of privatized firms.

Figure 1 shows that privatized firms are the largest in many countries. In 17 of the 27 Organization for Economic Co-operation and Development (OECD) and developing countries listed, privatized firms are the largest.

FIGURE 1
SIZE OF PRIVATIZED FIRMS IN 27 COUNTRIES

Country	Largest	Second largest	Third largest
Australia		X	x
Austria		X	x
Brazil	x	X	
Britain	x		
China	x	X	x
Czech Republic	x	X	x

Finland		X	x
France	x	X	x
Germany	x		
Greece	x		x
Hungary	x	X	x
India	x		
Israel		X	x
Italy	x	X	x
Japan		X	x
Malaysia		X	x
Mexico	x		x
Netherlands		X	
Norway	x	X	x
Poland	x	X	x
Portugal	x	X	x
Russia	x	X	x
Singapore	x		
South Africa		X	
South Korea			x
Spain	x		x
Taiwan		X	

This figure shows the size of privatized firms in 27 OECD and developing countries (Source: William L. Megginson presentation, constructed from Business Week, Top 200 Emerging-Market companies)

Privatized firms share some similarities with other large non-privatized firms (henceforth large firms), such as size and the economic importance to the privatizing country. However, it would be logically flawed for us to posit that just because privatized firms are like large firms in size and that size is an important capital structure factor, privatized firms will automatically have the same capital structure as those of large firms. There are at least three reasons why privatized firms differ from large firms and we should be very cautious not to treat privatized firms as simply a subset of large firms.

First, privatized firms and large firms take very different paths to become as large as they are. Privatized firms' size is due primarily to government intervention and protection. Most privatized firms are in strategic industries with heavy government involvement, leading to rapid and sustained growth; it usually does not take long for these state-owned enterprises (SOEs) to become very large in size. Conversely, most large firms, which have usually been in the business for a long time and have built a strong reputation, increase in size gradually over years of sustained growth. Kayhan and Titman (2007) show that firms' histories can affect their capital structures. Thus, the difference in firm maturity (i.e., how long a firm has been in the business) between privatized and large firms can lead to differences in capital structure.

Second, although privatized firms become more profitable after privatization, their level of profitability is still significantly lower than that of large firms. Many capital structure studies have confirmed that profitability is inversely related to the leverage ratio. Thus, differences in the profitability levels between privatized and large firms may result in differences in capital structure.

Third, unlike large firms, partially privatized firms (privatized firms in which government still retains a portion of ownership) may not always seek to maximize shareholder wealth. They may operate according to a hidden government political agenda that does not align with the goal of profit maximization. The trade-off theory of capital structure argues that a firm can maximize its value by choosing the optimal leverage ratio. Although the debate about the theory of capital structure is still ongoing, the trade-off theory of capital structure has gained some momentum recently especially after considering costly adjustment costs. If the trade-off theory holds, then partially-privatized firms, which represents a significant portion of the sample, would have a different capital structure from large firms.

One fundamental question about privatized firms' capital structure is how the capital structure of privatized firms empirically evolves after these SOEs are privatized. Does the evolution of privatized firms' capital structure indicate that privatized firms have a persistent capital structure or does the capital structure move randomly after privatization (i.e., because they might not seek an optimal capital structure)? Lemmon, Roberts, and Zender (2008) find that a firm's initial leverage is a very important capital structure determinant. Will the initial leverage of privatized firms also be important in explaining the long-term capital structure? Privatized firms invariably experience a significant drop in their leverage ratios because prior to privatization, they do not have any external equity, causing their leverage ratios to be superficially high and after privatization these ratios will drop as the firms start getting external equity. Will this post-privatization leverage ratio be able to explain the capital structure of privatized firms?

Many other important empirical questions can be raised: Will the capital structure determinants observed in the sample of U.S. firms also be the determinants in the sample of privatized firms? How important is the remaining government ownership in determining a privatized firm's capital structure? How do country-specific factors (e.g., corruption, economic development, legal system) affect the leverage choice of a firm?

Fan, Titman, and Twite (2012) suggest that institutional factors are critical determinants of firms' financial structures. Therefore, I examine whether country-specific factors such as corruption index, economic development, and legal system of a country affect the capital structure of privatized firms.

The remainder of this dissertation is organized as follows. I present the hypothesis development in Section 2. Section 3 contains data and sample selection. Section 4 presents the methodology used in this study and Section 5 presents my results. Finally, I conclude in Section 6.

HYPOTHESIS DEVELOPMENT

Before privatization, SOEs might not always seek to optimize the firm value due to the soft budget constraint. However, newly privatized SOEs are forced by the market to optimize their firm values and the wealth of their shareholders. The newly encountered threat of bankruptcy and market discipline make privatized firms more prudent in managing their debt level and force them to optimize their capital structure (i.e., to maximize firm value) and the wealth of their shareholders. Because of the privatized firms' new objective to maximize their firm values and the wealth of their shareholders (through optimizing their capital structure), I argue that the capital structure of privatized firms does not evolve randomly but, rather, has a persistent capital structure like other large firms do, leading to my first hypothesis.

H1: *Like large firms, privatized firms do have a persistent capital structure.*

The relation between remaining government ownership and leverage within partially privatized firms has not received much attention. Remaining government ownership may be an important capital structure determinant for privatized firms because the level of government ownership could potentially affect the perceived probability of bankruptcy. That is, privatized firms with a higher level of government ownership may be perceived as less likely to go bankrupt than firms with the lower level of government ownership. Borisova (2011), using a sample of European privatized firms, finds that a higher level of remaining government ownership leads to a lower cost of debt in partially privatized firms. This result is the basis of my second hypothesis:

H2: *The government ownership variable is an important capital structure determinant for privatized firms.*

Lemmon, Roberts, and Zender (2008) find that initial leverage captures about 90 percent of the variation in leverage and subsumes the significance of the traditional capital structure determinants. They further contend that this result holds for both public and private firms. Because privatized firms are either fully or partially privatized, Lemmon, Roberts, and Zender's result indirectly suggests that initial leverage

is a very important capital structure determinant for privatized firms. Therefore, my third hypothesis is as follows:

H3: *The initial leverage variable is an important capital structure determinant for privatized firms, and adding this variable reduces the significance of the traditional determinants.*

Fan, Titman, and Twite (2012) find that firms in countries with a higher level of corruption are more leveraged. However, whether this result holds true for privatized firms is unclear. On the one hand, privatized firms in more corrupt countries could use more debt (obtained from state-owned banks) because they can deliberately choose not to repay the debt knowing that there are a few or no adverse consequences from defaulting. On the other hand, privatized firms could use less debt because the government wants to send a message to the financial market that privatized firms are prudent in managing debt and will not expropriate the debt holders. This is particularly important especially if government still has firms on its privatization agenda. Therefore, the relation between corruption and leverage in privatized firms remains an empirical issue. Regardless, I assume capital structure is affected by corruption level, leading to my fourth hypothesis:

H4: *Corruption level is an important capital structure determinant for privatized firms.*

Considering the different institutional structures in developed and developing countries, it is important that we determine whether the leverage of privatized firms in developed countries differs from that in developing countries. Demirgüç-Kunt and Maksimovic (1999) compare the capital structure of firms in developed and developing countries and find that firms in developed countries use more long-term debt than firms in developed countries. Although Booth, Aivazian, Demirgüç-Kunt, and Maksimovic (2001), using a sample of the largest companies in each of the ten developing countries, provide evidence that leverage choices of firms in developing countries are affected by the same variables as in developed countries, they argue that differences persist across countries. Thus, I hypothesize:

H5: *Privatized firms in developed countries use more leverage than privatized firms in developing countries.*

La-Porta, Lopez-de-Silanes, Shleifer, and Vishny (1997) find that a country's investor protections are positively correlated with the size of debt and equity markets. Giannetti (2003) finds that institutions that favor creditor rights and have stricter enforcements are associated with higher level of leverage. Similar to the previously unknown relation between the economic development of a country and leverage decisions for privatized firms, the relation between a country's legal system and leverage decisions for privatized firms has never been investigated. I examine this relation and, based on the notion that a stronger legal system encourages more use of debt, I hypothesize:

H6: *Privatized firms in common law countries use more leverage than privatized firms in civil law countries.*

DATA AND SAMPLE SELECTION

I construct the sample of privatized firms using data from William Megginson's privatization appendix, the Privatization Barometer, and the World Bank Privatization databases. Accounting data are obtained from the Compustat Global database, and the 12 industry classification data are from the Kenneth French's Web site. Table 1 provides the variable definitions and predicted signs.

The sample includes only SOEs that are privatized through the share issue privatization (SIP) method because SIP firms provide accounting data after privatization and they are the most significant privatized firms in terms of size and relative importance to the privatizing nations' economies. In addition, SIP

offerings are secondary offerings (except in China and Russia) so the proceeds go directly to the government, not to the SOE. Thus, any improvements in performance reported after divestiture must be related to the changes in incentives, regulation, macroeconomic policy, or ownership structure rather than to cash injections into the firm from a new offering.

I require all firms to have more than one year of consecutive data and that all firm-years have no missing data for the book value of total assets. Data for Canada go back as far as 1987; data for all other countries in the sample start in 1991 or later. Therefore, I use data only from 1991 to 2005 for my analyses. My final sample consists of 340 privatized firms from 51 countries and has 3,013 firm-year observations spanning 1991 to 2005.

Consistent with many international capital structure studies, I consider two measures of leverage: the long-term-debt-to-total-asset (LDA) and the total-debt-to-total-asset (TDA) ratios. However, because the TDA ratio contains trade credit (from the short-term debt) whose determinants could be under the influence of completely different determinants, I mainly focus on the LDA ratio.

The capital structure determinants used in this study are size, defined as the natural log of the book value of total assets; collateral, defined as the ratio of fixed assets to the book value of total assets; profitability, defined as the ratio of operating income before depreciation to the book value of total assets; and median industry leverage, defined as the industry's median long-term debt to the book value of total assets ratio in a specific year. Frank and Goyal (2009) consider these variables to be among the most important. The data for government ownership variable are collected from the World Bank's Privatization database and the Privatization Barometer database along with William Megginson's privatization appendix. Missing data are obtained manually. Table 2 presents the summary statistics of the sample.

To determine whether a country is a developed or a developing country, I rely primarily on the gross domestic product (GDP) per capita criteria and obtain the data from the World Bank database. I define a developed country as a country that has a GDP per capita over USD 12,000. To determine a country's legal system, I use the Central Intelligence Agency (CIA)'s *World Factbook*. In some cases, it is not easy to determine a country's legal system. For example, many countries adopt mixed systems, such as Morocco, which employs both Islamic law and French and Spanish civil law systems. In such cases, I defer to the classification developed by La-Porta, Lopez-de-Silanes, Shleifer, and Vishny (1998). The corruption data used in this study are obtained from the Web site of Transparency International (TI). After collecting both firm-specific and country-specific data, I construct a correlation matrix, as presented in Table 3, which shows the correlation among factors.

METHODOLOGY

Evolution of the Capital Structure of Privatized Firms

To study the evolution of the capital structure of privatized firms, I first form leverage quartiles beginning in 1991, the first year that the sample contains sufficient firms (n=125) through 2005. After I sort and rank the firms by their leverage ratio, I divide them into quartiles and denote them as very high (VH), high (H), medium (M), and low (L). Using the same set of leverage quartiles, I continue by averaging every leverage quartile in 1992 and so forth. I do not re-sort the leverage quartiles for each year (i.e., to keep the components of each quartile constant) so that I can examine whether firms that initially have a high level of leverage continue to use a relatively high level of leverage in the long run. I then plot the average of each leverage quartile over time to show the evolution of the capital structure of privatized firms.

Main Specification

To test the first hypothesis, I regress leverage on the lagged variable of leverage and on the lagged traditional determinants of capital structure such as size, profitability, median industry leverage, and collateral.

$$y_{it} = \alpha + \beta\, y_{it-1} + \gamma\, x_{it-1} + e_{it} \quad (1)$$

where y_{it} is the leverage ratio and y_{it-1} is the one-year lagged variable of leverage ratio, for firm *i* at time *t* and x_{it-1} is the vector of lagged independent variables that includes size, profitability, median industry leverage, and collateral. If privatized firms have a persistent capital structure, after controlling for firm characteristics, β should be significantly positive and to test other hypotheses, I add a variable and test for its significance by looking at the t-statistic.

Although other capital structure studies often exclude financial services and utilities firms from their sample, I include these firms because they represent a significant portion of the sample (approximately thirty percent) and many are very important privatized firms. Including these firms in my sample provides a clearer picture of how the capital structure of privatized firms evolves. However, to address any concerns that my results could be skewed by the inclusion of financial services and utilities firms, I separate these firms from the main sample and rerun all the analyses with the new sample.

RESULTS

The median for LDA and TDA ratios are plotted on Figures 2 and 3, respectively. Two observations are worth noting. First, the figures show that, on average, privatized firms use more long-term debt and less total debt than large firms. Privatized firms may use more long-term debt than large firms because, in the long run, financially troubled privatized firms are considered less likely to go bankrupt (the probability of bankruptcy is not as important in the short run as it is in the long run). Another possible explanation is that privatized firms do not need to use short-term debt as much as large firms. Rajan (1992) argues that shorter maturity debt limits the time period during which a borrower can exploit creditors without being in default. In this case, partially privatized firms in particular do not need to take shorter maturity debt to convince creditors that the firm will not exploit them because the government is perceived as less likely to defraud creditors. This could explain Borisova's (2011) finding that the cost of debt is negatively related to the remaining level of government ownership. Second, I find similar, albeit weaker, evidence to support Demirgüç-Kunt and Maksimovic's (1999) finding that large firms in developed countries use more long-term debt than firms in developing countries. My weaker evidence, however, might be a result of a smaller sample (especially in developed countries, where the number of privatized firms is usually smaller than large firms).

Figure 3 also shows that large firms usually use more total debt (as a percentage of total assets) and less equity than privatized firms. This finding supports the pecking order theory that when firms need external financing, they prefer debt to equity. Conversely, privatized firms use less total debt and more equity (as percentages of total assets) than large firms, which is consistent with Megginson, Nash, Netter, and Poulsen (2004), who find that SIPs are more likely when the capital market in the country is less developed. Therefore, privatized firms' use of more equity than debt is not surprising, and their decision to issue more equity than debt – which may be driven by the government's desire to improve the condition of their less-developed capital markets – is understandable.

Figure 5 shows a much stronger pattern of the leverage convergence when TDA is used as the leverage measure. During the first nine years, from 1991 to 1999, the average of leverage quartiles strongly converges, and the leverage convergence continues although not as strongly. Based on this evidence, privatized firms appear to care about their target leverage ratio and seek to optimize their capital structure.

Table 4 presents results for all regression models. I find size, interestingly, has a negative sign in all models, and in more than half of the models the negative coefficients are statistically significant at the five percent level. This finding contradicts the usual sign for size found in capital structure studies of large firms. However, as previously mentioned in chapter 1, despite similarities between privatized firms and large firms, they have some distinct differences. Unlike large firms, privatized firms have more of the agency problem. Before SOEs are privatized, managers own no ownership, which give them more incentive to consume perquisites at the government's expense. In addition, in many countries, state employees have more job security than their counterparts in large firms, which further exacerbates the problem because they do not risk losing their jobs even when their productivity level is unacceptably low. Therefore, after SOEs are privatized, smaller (larger) privatized firms use more (less) leverage.

The lagged variable of leverage is very significant even at one percent level (t-statistic=78.35). This result suggests that, like large firms, the capital structure of privatized firms is persistent and does not evolve randomly across time.

Table 4 also shows that the government ownership variable has an insignificantly positive sign, suggesting that higher government ownership leads to higher level of leverage use. This finding, although insignificant, confirms the assumption that financially troubled privatized firms with higher remaining government ownership are perceived to be less likely to go bankrupt thereby allowing firms to incur a lower cost of debt.

The initial leverage variable is significantly positive, supporting Lemmon, Roberts, and Zender's (2008) finding that initial leverage is a very significant factor. However, my result does not support their suggestion that initial leverage substantially subsumes the significance levels of traditional capital structure determinants.

I find a significantly positive relation between the corruption index of a country and leverage, which means that as a country becomes less corrupt (indicated by higher TI index), privatized firms in that country use more leverage. This positive relation between corruption index and leverage found in privatized firms is not illogical. When a country has a higher level of corruption, the government might be forced to pressure managers of privatized firms to use less leverage if the government still has a full agenda of privatizations pending (government needs to convince the potential creditors that they will not be expropriated).

After controlling for the corruption level, I reexamine the impact of the level of remaining government ownership on leverage choice. In the presence of corruption, the result shows a more significant coefficient of government ownership than the previous findings. My result shows that government ownership is now positively significant, suggesting that higher government ownership leads to a higher use of leverage because creditors face less risk of default.

The economic development dummy has a significantly positive coefficient. This finding, which is consistent with Demirgüç-Kunt and Maksimovic (1999) and Booth, Aivazian, Demirgüç-Kunt, and Maksimovic (2001), supports the notion that privatized firms in developed countries use more leverage than privatized firms in developing countries. When I regress leverage on per capita GDP (as an alternative measure to the economic development dummy), the result is even stronger. The significantly positive coefficient means that privatized firms in countries with higher (lower) per capita GDP use more (less) leverage.

The legal system dummy has a significantly positive coefficient, suggesting that privatized firms in common law countries use more leverage than in civil law countries. That is, privatized firms are more likely to use leverage in countries with a stronger legal environment. When I combine all firm-specific and country-specific factors in model 10, the result does not change significantly.

To address any potential concerns that my results are driven by the sample that includes financial services and utilities firms, I exclude them from the sample for robustness check and rerun all regression models.

For robustness check, I now exclude financial services and utilities firms. The results are relatively similar to those with financial services and utilities firms. One striking result is that privatized firms no longer have a negative sign but they are not statistically significantly positive. This result is very surprising as size has been known as a very important capital structure determinant.

CONCLUSIONS

I investigate the evolution and determinants of privatized firms' capital structure. My study contributes to the relatively scant literature on international corporate finance on privatized firms at a time when privatized firms are becoming increasingly important in almost every country in the world. This study also is the first to examine the long-term capital structure of state-owned enterprises (SOEs) following privatization.

I find that privatized firms, in general, use more long-term debt but less total debt than large firms. This result might be because, in the long run, privatized firms are considered to be less likely to go bankrupt – a

factor that is more important in the long run than in the short run. Another possible explanation is that privatized firms do not need to use short-term debt to signal investors that they will not defraud the creditors (see Rajan, 1992). I also find that privatized firms use less total debt and more equity (as percentages of total assets) than large firms, which might be due to the government's desire to improve their less-developed capital markets (see Megginson, Nash, Netter, and Poulsen, 2004).

I find that size is not an important capital structure determinant of privatized firms. Furthermore, I find that privatized firms have a persistent capital structure and that the capital structure of privatized firms does not change randomly across time. These results are important as they shed some light on the issue regarding whether privatized firms carefully manage their leverage level. My study also indicates that privatized firms have a target leverage ratio; this finding is independent of the level of remaining government ownership, indicating that as soon as governments start selling off their ownership, the former SOEs will start seeking their optimum capital structures.

Initial leverage is also an important capital structure determinant for privatized firms, but the inclusion of the initial leverage variable does not reduce the significance levels of the traditional capital structure determinants. Therefore, these initial-leverage-related findings do not fully support Lemmon, Roberts, and Zender (2008). Corruption is significantly positively related to leverage, which indicates that privatized firms in a less corrupt country use more leverage. This might be because in more corrupt countries, governments might be forced to pressure managers of privatized firms to use less leverage if governments still have a full agenda of privatizations pending (i.e., governments of more corrupt countries prefer not to use as much leverage because privatized firms can expropriate the creditors, and governments do not have full control to avoid creditor expropriation). The finding that privatized firms in developed countries use more long-term debt than in developing countries is consistent with Demirgüç-Kunt and Maksimovic (1999) and Booth, Aivazian, Demirgüç-Kunt, and Maksimovic (2001), who find that large firms in developed (developing) countries use more (less) long-term debt. When per capita GDP is used as an alternative measure of economic development, I find a significantly positively relation.

In conclusion, this study presents many new findings about the evolution and determinants of privatized firms' capital structure. Considering the less developed literature of privatized firms and international capital structure, there exists a very important need for further exploration into how privatized firms make their financial policies.

REFERENCES

Booth, L., Aivazian, V., Demirgüç-Kunt, A., & Maksimovic, V. (2001). Capital structures in developing countries. *Journal of Finance*, *61*(1), 87-130.

Borisova, G. (2011). Does government ownership affect the cost of debt? Evidence from privatization. *Review of Financial Studies*, *24*(8), 2693-2737.

Demirgüç-Kunt, A., & Maksimovic, V. (1999). Institutions, financial markets, and firm debt maturity. *Journal of Financial* Economics, *54*, 295-336.

Diamond, D. (1991). Monitoring and reputation: The choice between bank loans and directly placed debt. *Journal of Political Economy*, *99*(4), 689-721.

Fan, J., Titman, S., & Twite, G. (2012). An international comparison of capital structure and debt maturity choices. *Journal of Financial and Quantitative Analysis*, *47*(1), 23-56.

Flannery, M., & Rangan, K. (2006). Partial adjustment toward target capital structures. *Journal of Financial Economics*, *79*, 469-506.

Frank, M., & Goyal, V. (2009). Capital structure decisions: Which factors are reliably important? *Financial Management*, *38*(1), 1-37.

Giannetti, M. (2003). Do better institutions mitigate agency problems? Evidence from corporate finance choices. *Journal of Financial and Quantitative Analysis*, *38*, 185-212.

Graham, J., & Harvey, C. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics*, *60*, 187-243.

Kayhan, A., & Titman, S. (2007). Firms' histories and their capital structures. *Journal of Financial Economics*, *83*(1), 1-32.
LaPorta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (1997). Legal determinants of external finance. *Journal of Finance*, *52*(3), 1131-1150.
LaPorta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (1998). Law and finance. *Journal of Political Economy*, *106*(6), 1113-1155.
Leary, M., & Roberts, M. (2005). Do firms rebalance their capital structures? *Journal of Finance*, *60*(6), 2575-2619.
Lemmon, M., Roberts, M., & Zender, J. (2008). Back to the beginning: persistence and the cross-section of corporate capital structure. *Journal of Finance*, *63*(4), 1575-1608.
Megginson, W., Nash, R., Netter, J., & Poulsen, A. (2004). The choice of private versus public capital markets: Evidence from privatizations. *Journal of Finance*, *59*(6), 2835-2870.
Rajan, R. (1992). Insiders and outsiders: the choice between informed and arm's length debt. *Journal of Finance*, *47*, 1367-1400.

APPENDIX

FIGURE 2
THE MEDIAN LDA RATIOS OF PRIVATIZED AND LARGE FIRMS

This figure plots the median LDA ratio of privatized and large firms in each of 51 countries. The TDA ratio is defined as the total debt over the book value of total assets of a firm.

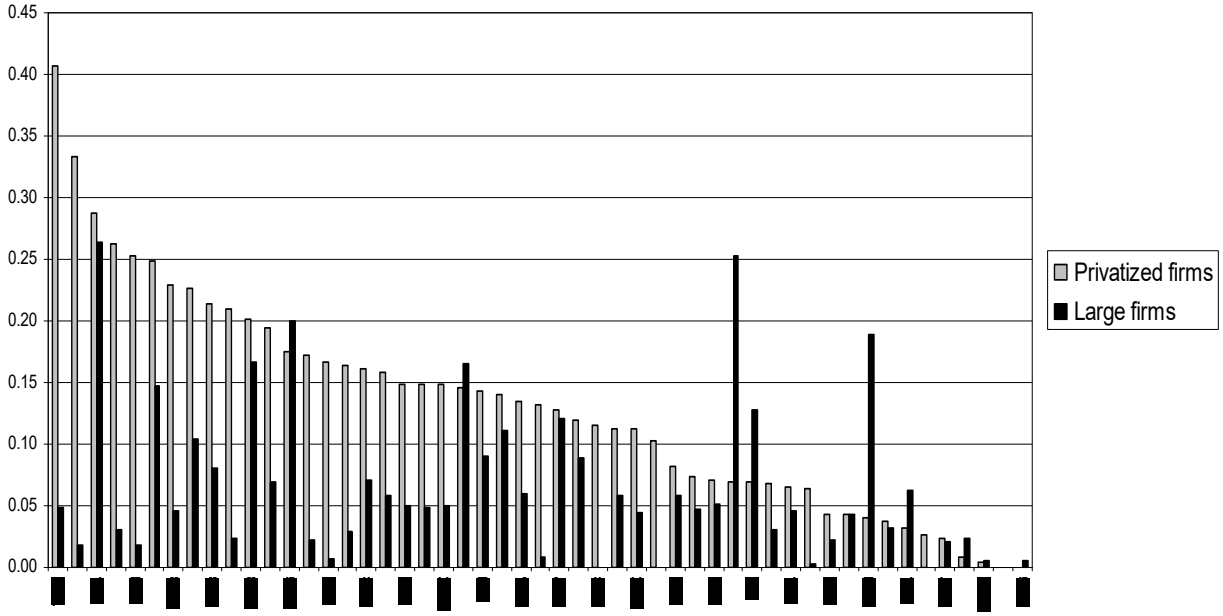


FIGURE 3
THE MEDIAN TDA RATIOS OF PRIVATIZED AND LARGE FIRMS

This figure plots the median TDA ratio of privatized and large firms across 51 countries. The TDA ratio is defined as the total debt over the book value of total assets of a firm.

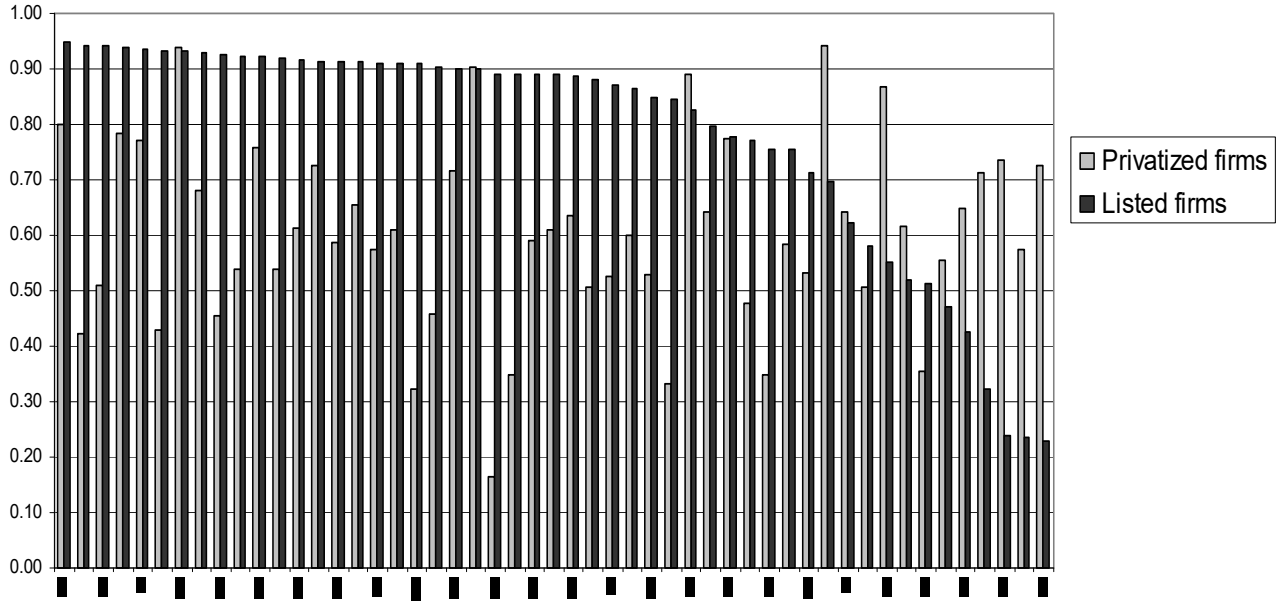


FIGURE 4
THE EVOLUTION OF LDA QUARTILES ACROSS TIME

This figure plots the average LDA ratio of every leverage quartile for privatized firms during 1991-2005. The leverage quartiles are formed in 1991 and each quartile contains a constant number of firms.

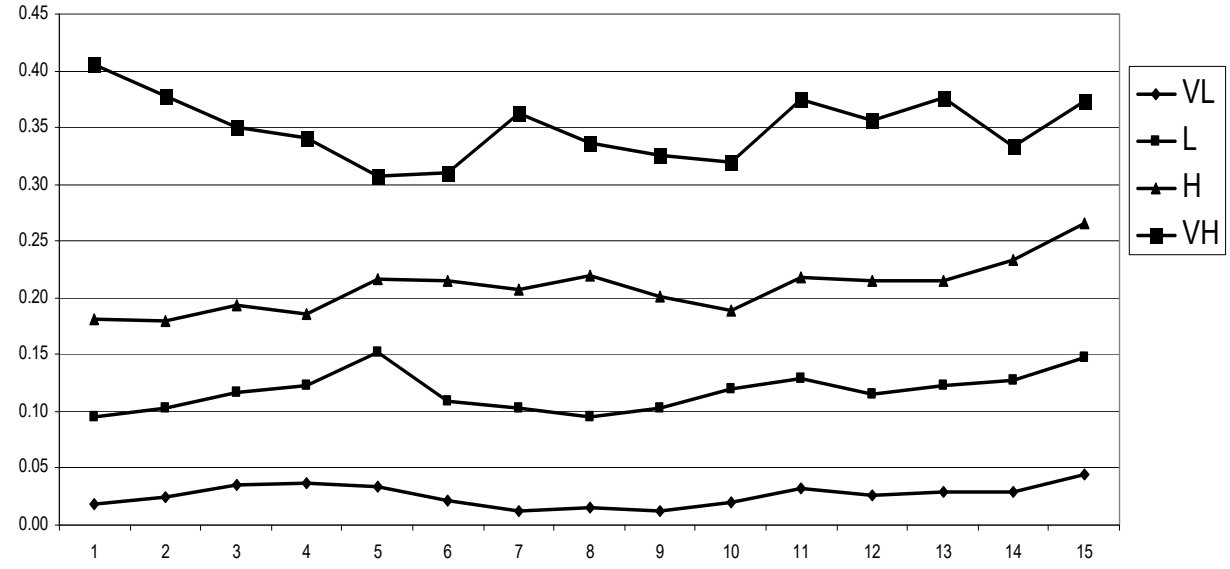


FIGURE 5
THE EVOLUTION OF TDA QUARTILES ACROSS TIME

This figure plots the average TDA ratio of every leverage quartile for privatized firms during 1991-2005. The leverage quartiles are formed in 1991 and each quartile contains a constant number of firms.

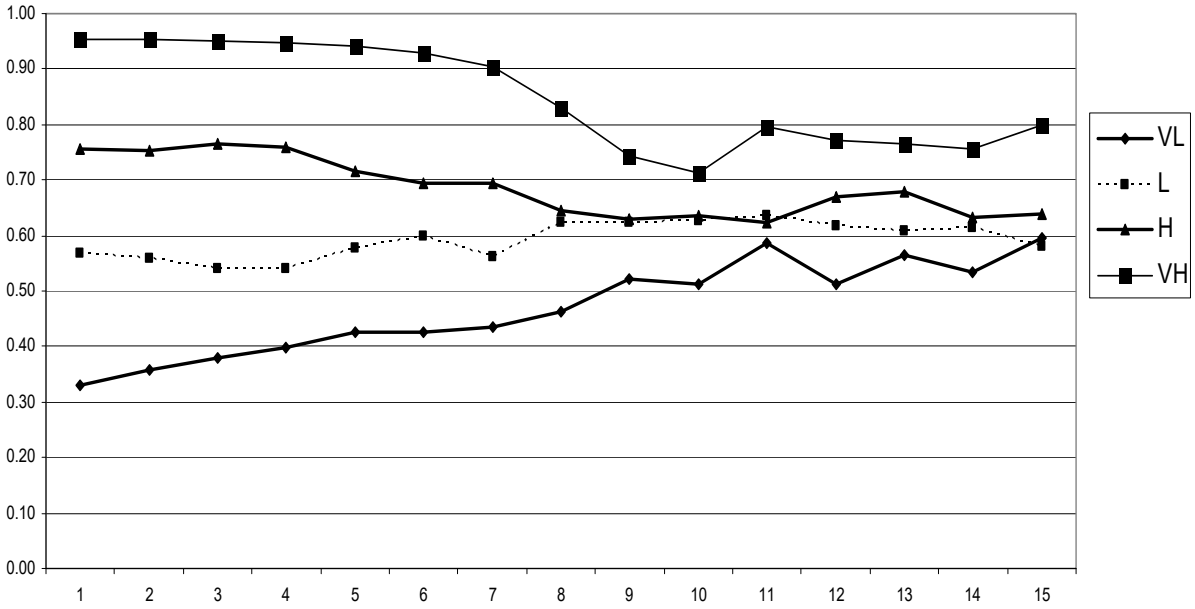


TABLE 1
VARIABLE DEFINITIONS AND PREDICTED SIGNS

Variable	Description	Predicted Sign
lda	Long-term debt/book value of total assets	
size	Natural log of the book value of total assets where assets are deflated by the GDP deflator	(+)
collat	Net fixed assets / book value of total assets	(+)
profit	Operating income before depreciation / book value of total assets	(-)
med	median of firms' ldas by SIC code and by year	(+)
laglda	the lagged variable of lda ratio	(+)
govtown	% of remaining government ownership	(+)
initlev	initial leverage	(+)
corrupt	the corruption index where a higher value represents a more corrupt country	(+)
d_dev	a dummy that takes a value of 1 if a country is a developed country and 0 otherwise	(+)
gdp	deflated gdp number	(+)
d_leg	a dummy that takes a value of 1 if a country is a common law country and 0 otherwise	(+)

Firm characteristic data are obtained from Compustat Global and Thomson One Banker, privatization data are from William L. Megginson's appendix and other privatization databases, corruption data are from Transparency International, economic development data are from World Bank and IMF databases, and legal system data are from CIA World Factbook and LLSV (1998) paper.

TABLE 2
SUMMARY STATISTICS

Variables	N	Mean	Standard Deviation	Minimum	Maximum
lda	3431	0.1766	0.1509	0.0000	0.9694
size	3438	9.2959	2.1783	2.4912	16.0691
collat	3113	0.7905	0.5348	0.0000	5.1576
profit	3436	0.0382	0.0740	-1.8103	0.5962
med	3893	0.6305	0.2102	0.0000	1.8857
govtown	2745	0.4349	0.3164	0.0000	0.9920
initlev	2821	0.1621	0.1538	0.0000	0.7482
corrupt	3215	0.3813	0.2157	0.0000	0.8300
d_dev	4154	0.7198	0.4492	0.0000	1.0000
gdp	4026	26.6102	1.2830	21.8960	29.2385
d_leg	4138	0.2343	0.4230	0.0000	1.0000

This table provides number of observations, mean, standard deviation, minimum, and maximum values of each variable. **LDA** is the ratio of long-term debt to book value of total assets. **Size** is the natural log of book value of total assets. **Collat** is the collateral variable defined as the ratio of net fixed assets to book value of total assets. **Profit** is the profitability variable defined as the ratio of operating income before depreciation to book value of total assets. **Med** is the median industry leverage variable, measured as the median LDA ratios in the firm's industry. **Govtown** is the government ownership variable defined as the percentage of remaining government ownership. **Initlev** is the initial leverage variable defined as the firm's LDA when the firm shows up in the sample. **Corrupt** is the corruption index variable obtained from Transparency International. **D_dev** is the economic development dummy variable that takes a value of 1 if a country is a developed country and 0 otherwise. **Gdp** is the GDP variable defined as the natural log of deflated GDP number. **D_leg** is the legal system dummy that takes a value of 1 if a country is a common law country and 0 otherwise.

TABLE 3
CORRELATION MATRIX

Variables	lda	laglda	size	collat	profit	med	govtown	initlev	corrupt	d_dev	gdp	d_leg	indust
lda	1.000												
laglda	**0.834**	1.000											
size	-0.001	0.008	1.000										
collat	**0.275**	**0.258**	**-0.281**	1.000									
profit	**-0.160**	**-0.086**	**-0.090**	**0.053**	1.000								
med	**0.211**	**0.163**	**0.380**	**-0.439**	**-0.339**	1.000							
govtown	0.010	0.019	**0.074**	-0.026	-0.037	0.012	1.000						
initlev	**0.680**	**0.643**	**0.048**	**0.244**	**-0.079**	**0.129**	**0.167**	1.000					
corrupt	**-0.176**	**-0.177**	**0.077**	**-0.056**	**0.043**	**-0.081**	**0.127**	**-0.260**	1.000				
d_dev	**0.153**	**0.150**	-0.004	**-0.036**	**-0.103**	**0.303**	**-0.074**	**0.207**	**-0.716**	1.000			
gdp	**0.143**	**0.137**	**-0.057**	**0.057**	**-0.092**	**0.158**	**-0.163**	**0.133**	**-0.060**	**0.373**	1.000		
d_leg	**0.132**	**0.124**	**-0.145**	**0.097**	**0.066**	**-0.152**	**-0.286**	**0.105**	**-0.265**	0.026	**-0.076**	1.000	
indust	**0.104**	**0.086**	**0.131**	**-0.184**	**-0.078**	**0.234**	-0.032	**0.110**	**-0.053**	**0.109**	0.027	**-0.038**	1.000

This table provides correlation matrix for the sample. Variables are defined in Table 5. Bolded number indicates statistical significance at the 5 percent level.

TABLE 4
YEAR FIXED-EFFECTS REGRESSION RESULTS (INCLUDING FINANCIAL SERVICES AND UTILITIES FIRMS)

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
size	**-0.003**	**-0.002**	**-0.005**	**-0.006**	-0.002	**-0.003**	-0.002	-0.002	-0.001	**-0.002**
	-(2.40)	*-(2.44)*	*-(3.19)*	*-(5.86)*	*-(1.30)*	*-(2.02)*	*-(1.92)*	*-(1.77)*	*-(1.21)*	*-(2.28)*
collat	**0.115**	**0.030**	**0.103**	**0.046**	**0.114**	**0.103**	**0.114**	**0.115**	**0.115**	**0.013**
	(22.10)	*(9.31)*	*(16.75)*	*(9.87)*	*(20.03)*	*(15.79)*	*(21.99)*	*(21.93)*	*(22.45)*	*(3.18)*
profit	**-0.126**	**-0.121**	**-0.115**	**-0.099**	**-0.095**	**-0.082**	**-0.123**	**-0.112**	**-0.126**	**-0.131**
	-(3.51)	*-(5.85)*	*-(2.80)*	*-(3.32)*	*-(2.50)*	***-(1.97)***	*-(3.43)*	*-(3.12)*	*-(3.55)*	*-(5.49)*
med	**0.235**	**0.066**	**0.239**	**0.127**	**0.224**	**0.234**	**0.225**	**0.222**	**0.245**	**0.042**
	(16.06)	*(7.55)*	*(13.18)*	*(10.02)*	*(13.84)*	*(12.21)*	*(14.88)*	*(14.98)*	*(16.90)*	*(3.41)*
laglda		**0.792**								**0.777**
		(78.35)								*(47.54)*
govtown			0.010			**0.022**				-0.010
			(1.12)			*(2.20)*				*-(1.61)*
initlev				**0.580**						**0.106**
				(40.79)						*(6.78)*
corrupt					**0.080**	**0.103**				0.024
					(6.25)	*(7.05)*				*(1.67)*
d_dev							**0.017**			0.007
							(2.77)			*(1.08)*
gdp								**0.008**		0.000
								(4.11)		*(0.35)*
d_leg									**0.051**	**0.010**
									(8.88)	*(2.21)*
Intercept	-0.027	-0.007	-0.000	**0.029**	**-0.078**	**-0.082**	**-0.038**	**-0.240**	**-0.061**	0.024
	-(1.74)	*-(0.74)*	*-(0.01)*	*(2.13)*	***-(4.07)***	***-(3.62)***	*-(2.38)*	*-(4.38)*	*-(3.91)*	*(1.67)*
N	3110	3013	2285	2552	2542	1963	3110	3089	3104	1569

This table presents results of leverage on firm-specific and country-specific factors using annual data (including financial services and utilities firms) from 1991 to 2005. Model 1 is the main model using panel regression with year fixed-effects where leverage is regressed on traditional capital structure determinants. Models 2–9 add a unique variable or variables to Model 1: Model 2 adds the lagged variable of leverage; Model 3, government ownership; Model 4, initial leverage; Model 5, corruption; Model 6, government ownership and corruption; Model 7, an economic development dummy; Model 8, gross domestic product; and Model 9, a legal system dummy. Model 10 includes all independent variables. Variables are defined in Table 5. t-statistics are in parentheses and bolded numbers represent statistical significance at 5% level.

Managerial Ability and Analyst Forecast Behavior: Large Sample Evidence

Shiyou Li
Texas A&M University Commerce

Prior literature documents that the managerial ability, derived from frontier analysis, is positively associated with accounting quality (Demerjian, Lev, Lewis, and McVay 2013; Baik, Farber and Lee 2011). In addition, prior literature indicates that number of analysts following a firm is positively associated with accounting quality, and analyst forecast dispersion is negatively associated with accounting quality (Lang and Lundholm 1996; Irani and Karamanou 2003). I examine the relation between managerial ability and the number of analysts following a firm as well as analyst forecast dispersion. I find that managerial ability is positively associated with analyst following a firm and negatively associated with analyst forecast dispersion. In addition, the effects are more pronounced after Sarbanes–Oxley Act of 2002, and are more pronounced for firms with complicated financial reports. Collectively, our findings support the notion that the managerial ability is relevant to analysts' decision making.

Keywords: managerial ability, analyst forecast

INTRODUCTION

Prior research suggests that managerial ability is informative to the financial information users and relevant to their decisions. For example, Baik, Farber and Lee (2011) test the relation between stock return and managerial ability and indicate that investors are more responsive to the news in management earning forecast associated with higher managerial ability; Krishnan and Wang (2014) suggest that managerial ability is informative to auditors and managerial ability is relevant to auditors' decisions. I investigate whether the managerial ability is informative to the analysts and relevant to analysts' decisions. Specifically, I examine the relation between the managerial ability and analyst forecast behavior. I use two important properties of analyst forecast behavior: the number of analysts following a firm and the dispersion of analyst forecast.

Prior research suggests that earnings quality is positively associated with managerial ability (Demerjian, Lev, Lewis, and McVay 2013), and that the management earnings forecasts for firms with high –ability CEOs are more frequent and accurate (Baik, Farber and Lee 2011). Prior literature also suggests that higher quality disclosure increases analysts' following and reduces analyst forecast dispersion (e.g., Lang and Lundholm, 1996; Healy et al., 1999; Botosan and Harris, 2000; Core, 2001; Francis et al., 2002; Roulstone, 2003; Liu et al. 2014). I bridge the gap by examining the association between managerial ability and analyst forecast behavior. I provide large sample evidence that managerial ability is positively associated with analyst following a firm and negatively associated with analyst forecast dispersion. In addition, the results are more pronounced after Sarbanes–Oxley Act of 2002, and are more pronounced for

firms with complicated financial reports. Collectively, our findings support the notion that the managerial ability is relevant to analysts' decisions.

The rest of the study proceeds as follows. Next section describes background and hypothesis development. This section is followed by a discussion of the research design. I then describe the data and sample, followed by the presentation of empirical results. The final section summarizes the findings.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Prior literature documents that managerial ability is positively associated with accounting quality. Demerjian, Lev, Lewis, and McVay (2013) suggest that earnings quality is positively associated with managerial ability. Specifically, more able managers are associated with fewer subsequent restatements, higher earnings and accruals persistence, lower errors in the bad debt provision, and higher quality accrual estimations. The results are consistent with the premise that managers can and do impact the quality of the judgments and estimates used to form earnings. Baik, Farber and Lee (2011) indicate that likelihood of management earnings forecast issuance increases in CEO ability. This result is consistent with Trueman's 1986 theory (Trueman 1986) that high-ability managers are more likely than low-ability managers to signal their ability to anticipate changes in their firm's prospects. In addition, forecast accuracy increases in CEO ability. These findings are consistent with the notion that forecasts associated with high ability CEOs reflect information about their ability to anticipate changes in their firms' underlying economics.

Financial analysts use accounting information (among other sources) to base their recommendations (Krishnan, Li and Wang 2013; Easton 2007). Prior literature suggests that higher quality disclosure increases analysts' following and reduces analyst forecast dispersion (e.g., Lang and Lundholm, 1996; Healy et al., 1999; Botosan and Harris, 2000; Core, 2001; Francis et al., 2002; Roulstone, 2003; Liu et al. 2014). The corresponding prediction I test is:

H1: *Managerial ability is positively associated the number of analysts following a firm and negatively associated with the analyst forecast dispersion.*

SOX imposes considerably greater potential penalties on chief executive officers (CEOs) and chief financial officers (CFOs) who engage in financial wrongdoing. Therefore, risk-averse managers are likely to report lower earnings by reducing discretionary accruals following SOX (Lobo and Zhou 2010). They document that firms subject to SOX are more conservative in financial reporting in the post-SOX period as evidenced by lower signed discretionary accruals, the Ball and Shivakumar (2005) conditional conservatism measure, and the Penman and Zhang (2002) unconditional conservatism measure. The corresponding prediction I test is:

H2: *The associated between managerial ability and the number of analysts following a firm, and between the managerial ability and analyst forecast dispersion is more pronounced after SOX.*

Next, I test whether the impact of the managerial ability on the forecasting variables is stronger for filers with complicated operation and complicated financial reports. We conjecture that managerial ability would have more pronounced effects on analysts' activities for filers with complicated operation and complicated financial reports, on the premise that the able managers would give analysts greater and higher quality information about the filers. The corresponding prediction I test is:

H3: *The associated between managerial ability and the number of analysts following a firm, and between the managerial ability and analyst forecast dispersion is more pronounced for firms with complicated financial reports.*

EMPIRICAL METHOD

Empirical Models

Following prior literature (e.g., Lang and Lundholm, 1996; Liu et al., 2014), I use the models below to test the effect of managerial ability on analysts' forecast activities:

$$FOLLOW_{it} = \alpha_0 + \alpha_1 MA_{it} + \alpha_2 SIZE_{it} + \alpha_3 EPS_{it} + \alpha_4 LOSS_{it} + \alpha_5 LEVERAGE_{it} + \alpha_6 VOLATILITY_i + \sum Year + \sum Industry + \varepsilon_{it} \quad (1)$$

$$DISPERS_{it} = \alpha_0 + \alpha_1 MA_{it} + \alpha_2 SIZE_{it} + \alpha_3 EPS_{it} + \alpha_4 LOSS_{it} + \alpha_5 LEVERAGE_{it} + \alpha_6 VOLATILITY_i + \sum Year + \sum Industry + \varepsilon_{it} \quad (2)$$

where $FOLLOW_{it}$ is natural log of the number of analyst followings for firm *i* in quarter *t* (see Table 1for variable definitions). $DISPERS_{it}$ is analyst forecast dispersion for firm *i* in quarter *t,* which is defined as standard deviation of analyst forecasts for firm *i* in quarter *t.* MA_{it} is managerial ability. See section 3.2 for detailed discussion. SIZEit is the natural log of firm *i*'s total assets at the end of quarter *t*. EPSit is firm *i*'s earnings per share (EPS) at the end of quarter *t.* while LOSSit equals one if EPSit is negative and 0 otherwise. LEVERAGEit is the leverage ratio (total liability divided by total assets) of firm *i* at the end of quarter *t*. VOLATILITYi is the volatility of stock returns, which equals the standard deviation of monthly stock returns for 12 months in the year. I further control for year and industry effects. See Appendix 1 for variable definitions.

$$FOLLOW_{it} = \alpha_0 + \alpha_1 MA_{it} + \alpha_2 SOX*MA_{it} + \alpha_3 EPS_{it} + \alpha_4 LOSS_{it} + \alpha_5 LEVERAGE_{it} + \alpha_6 VOLATILITY_i + \sum Year + \sum Industry + \varepsilon_{it} \quad (3)$$

SOX is a dummy variable equal 1 when year is greater or equal to 2002 and 0 otherwise. The variable of interest is the interaction of SOX and MA.

$$DISPERS_{it} = \alpha_0 + \alpha_1 MA_{it} + \alpha_2 SOX*MA_{it} + \alpha_3 EPS_{it} + \alpha_4 LOSS_{it} + \alpha_5 LEVERAGE_{it} + \alpha_6 VOLATILITY_i + \sum Year + \sum Industry + \varepsilon_{it} \quad (4)$$

$$FOLLOW_{it} = \alpha_0 + \alpha_1 MA_{it} + \alpha_2 HIGH\_SIZE*MA_{it} + \alpha_3 EPS_{it} + \alpha_4 LOSS_{it} + \alpha_5 LEVERAGE_{it} + \alpha_6 VOLATILITY_i + \sum Year + \sum Industry + \varepsilon_{it} \quad (5)$$

HIGH_SIZE is a dummy variable equal to 1 when rank is equal to 3 or 4 when partition the full sample into five quintiles (ranks: 0, 1, 2, 3 and 4) based on SIZE. The variable of interest is the interaction of HIGH_SIZE and MA.

$$DISPERS_{it} = \alpha_0 + \alpha_1 AER_{it} + \alpha_2 HIGH\_SIZE*MA_{it} + \alpha_3 EPS_{it} + \alpha_4 LOSS_{it} + \alpha_5 LEVERAGE_{it} + \alpha_6 VOLATILITY_i + \sum Year + \sum Industry + \varepsilon_{it} \quad (6)$$

Measure of Managerial Ability (MA)

Our main measure of managerial ability, MA, is developed by Demerjian et al. (2012). This measure is widely used by researchers (e.g., Baik et al. 2011; Demerjian et al. 2013; Krishnan and Wang 2014) and outperforms all other measures.

Demerjian et al. (2012) use data envelopment analysis (DEA) to estimate firm efficiency within industries, comparing the sales generated by each firm, conditional on the following inputs used by the firm: Cost of Goods Sold, Selling and Administrative Expenses, Net PP&E, Net Operating Leases, Net Research and Development, Purchased Goodwill, and Other Intangible Assets. Demerjian et al. (2012) use DEA to solve the following optimization problem:

$$Max_v\theta = Sales\ (v_1\ CoGS + v_2\ SG\&A + V_3\ PPE + v_4\ OpLease + v_5\ R\&D + v_6\ Goodwill + v_7\ OtherIntan \quad (7)$$

The efficiency measure that DEA produces, takes a value between 0 and 1. The efficiency measure generated by the DEA estimation is attributable to both the firm and the Manager. Demerjian et al. (2012) therefore modify the DEA generated firm efficiency measure by purging it of key firm-specific characteristics expected to aid or hinder management's efforts, including firm size, market share, positive free cash flow, and firm age, which aid management, and complex multi-segment and international operations, which challenge management. They estimate the following Tobit regression model by industry:

$$Firm\ Efficiency = \alpha_1 + \alpha_2\ Ln\ (Total\ Assets) + \alpha_3\ Market\ Share + \alpha_4\ Positive\ Free\ Cash\ Flow + \alpha_5\ Ln\ (Age) + \alpha_6\ BusinessSegmentConcentration + \alpha_7\ Foreign\ Currency\ Indicator + \alpha_8\ Year\ Indicators + \varepsilon \quad (8)$$

The residual from the estimation is the MA-Score, which is attributable to the management team and rely on as our main measure of managerial ability.

DATA AND SAMPLE

The measure of managerial ability is developed in Demerjian et al. (2012). I start with all 190,843 firm-year observations of managerial ability data for fiscal years 1980 through 2012. Financial data is from Compustat. Observations with no financial data are eliminated. This reduced the sample to 181,505 observations. Stock price data is from CRSP and analyst forecast data is from IBES. Observations with no CRSP and IBES data were eliminated. This further reduced the sample to 168,579.

EMPIRICAL RESULTS

Table 1 shows the summary statistics. The mean analyst following is 1.5416, the mean analyst forecast dispersion is 0.1292, the mean MA is 0.00103, and mean SIZE is 6.03657.

TABLE 1
SUMMARY STATISTICS

Variables	Mean	Median	Standard Deviation
FOLLOW	1.54160	1.60943	0.95509
DISPER	0.12920	0.04000	0.12404
MA	0.00103	0.01222	0.14895
SIZE	6.03657	5.80700	1.89088
MA*SOX	0.00139	0.00000	0.11133
MA*HIGH_SIZE	0.00691	0.00000	0.09236
EPS	0.02273	0.03028	0.45484
LEVERAGE	0.42507	0.42507	0.26589
VOLATILITY	0.14824	0.12570	0.09834

$FOLLOW_{it}$ = The natural log of number of analyst following for firm i in quarter t.
$DISPERS_{it}$ = Analyst forecast dispersion, which is the standard deviation (STDEV) of analyst forecasts.
MA_{it} = Managerial ability. See section 3.2 for detailed discussion.
$SIZE_{it}$ = The natural log of firm i's total assets at the end of quarter t.
$HIGH\_SIZE_{it}$ = Dummy variable equal to 1 when rank is equal to 3 or 4 when partition the full sample into five quintiles (ranks: 0, 1, 2, 3 and 4) based on SIZE.
EPS_{it} = Earnings per share for firm i in quarter t.
$LEVERAGE_{it}$ = The leverage ratio of firm i at the beginning of quarter t.
$VOLATILITY_{it}$ = The volatility of stock returns, equals the standard deviation of monthly stock returns at year t.

Table 2 shows the regression results of impact of managerial ability (MA) on the forecasting variables. The coefficient of MA on FOLLOW is 0.40417, positive and significant at 1% level, suggesting managerial ability is positively related to the number of analyst following a firm. The coefficient of MA on DISPER is -0.0574, negative and significant at 1% level, suggesting managerial ability is negatively related to analyst forecast dispersion. In sum, these results are in line with H1. The coefficients for other variables are generally consistent with existing literature.

TABLE 2
ANALYSIS OF IMPACT OF MANAGERIAL ABILITY (MA) ON ANALYST FOLLOWING (FOLLOW) AND FORECAST DISPERSION (DISPER)

Independent Variables	Dependent Variables	
	FOLLOW	DISPER
Intercept	-0.47601	-0.0592
	(0.01)***	(0.01)***
MA	0.40417	-0.0574
	(0.01)***	(0.01)***
SIZE	0.36211	0.00736
	(0.01)***	(0.01)***
EPS	2.3E-06	-0.0505
	(0.01)***	(0.01)***
LOSS	0.05748	0.10943
	(0.01)***	(0.01)***
LEVERAGE	-0.51008	0.09018
	(0.01)***	(0.01)***
VOLATILITY	0.29650	0.42726
	(0.01)***	(0.01)***
Year Effect	Yes	Yes
Industry Effect	Yes	Yes
Adj. R^2	0.16	0.07

$FOLLOW_{it}$ = The natural log of number of analyst following for firm i in quarter t.
$DISPERS_{it}$ = Analyst forecast dispersion, which is the standard deviation(STDEV) of analyst forecasts.
MA_{it} = Managerial ability. See section 3.2 for detailed discussion.
$SIZE_{it}$ = The natural log of firm i's total assets at the end of quarter t.
EPS_{it} = Earnings per share for firm i in quarter t.
$LOSS_{it}$ = Dummy variable that equals 1 if EPS at the end of year t is negative and 0 otherwise.
$LEVERAGE_{it}$ = The leverage ratio of firm i at the beginning of quarter t.
$VOLATILITY_{it}$ = The volatility of stock returns, equals the standard deviation of monthly stock returns at year t.

Table 3 shows regression results of impact of managerial ability (MA) on forecasting variables after SOX. In this case, coefficients on MA show the impact of MA on forecasting variables before SOX: MA is positively related to analyst following and negatively related to analyst forecast dispersion, both at 1% significance level. The coefficients on MA*SOX shows the incremental impact of MA on forecasting variables after SOX: the coefficient of MA on analyst following is 0.46364 (0.32875+0.13489) and the coefficient of MA on analyst forecast dispersion is -0.06085 (-0.0526-0.00825). These results provide support for our H2.

TABLE 3
ANALYSIS OF IMPACT OF MANAGERIAL ABILITY (MA) ON ANALYST FOLLOWING (FOLLOW) AND FORECAST DISPERSION (DISPER) AFTER SOX

Independent Variables	Dependent Variables FOLLOW	DISPER
Intercept	-0.47482	-0.0593
	(0.01)***	(0.01)***
MA	0.32875	-0.05260
	(0.01)***	(0.01)***
MA*SOX	0.13489	-0.00825
	(0.01)***	(0.04)**
SIZE	0.36196	0.00738
	(0.01)***	(0.01)***
EPS	2.30E-6	-0.0505
	(0.01)***	(0.01)***
LOSS	0.05734	0.10945
	(0.01)***	(0.01)***
LEVERAGE	-0.50970	0.09013
	(0.01)***	(0.01)***
VOLATILITY	0.29435	0.42735
	(0.01)***	(0.01)***
Year Effect	Yes	Yes
Industry Effect	Yes	Yes
Adj. R^2	0.15	0.07

$FOLLOW_{it}$ = The natural log of number of analyst following for firm i in quarter t.
$DISPERS_{it}$ = Analyst forecast dispersion, which is the standard deviation (STDEV) of analyst forecasts.
MA_{it} = Managerial ability. See section 3.2 for detailed discussion.
$SIZE_{it}$ = The natural log of firm i's total assets at the end of quarter t.
EPS_{it} = Earnings per share for firm i in quarter t.
$LOSS_{it}$ = Dummy variable that equals 1 if EPS at the end of year t is negative and 0 otherwise.
$LEVERAGE_{it}$ = The leverage ratio of firm i at the beginning of quarter t.
$VOLATILITY_{it}$ = The volatility of stock returns, equals the standard deviation of monthly stock returns at year t.
SOX = Dummy variable equal 1 when year is greater or equal to 2002 and 0 otherwise.

Table 4 shows regression results of impact of managerial ability (MA) on forecasting variables for firms with complicated financial reports. In this case, coefficients on MA show the impact of MA on forecasting variables for firms with simpler financial reports: MA is positively related to analyst following and negatively related to analyst forecast dispersion, both at 1% significance level. The coefficients on MA*HIGH_SIZE shows the incremental impact of MA on forecasting variables for firms with complicated financial reports: the coefficient of MA on analyst following is 0.52574 (0.29796+0.22778) and the coefficient of MA on analyst forecast dispersion is -0.07379 (-0.0476-0.02619). These results provide support for our H3.

TABLE 4
ANALYSIS OF IMPACT OF MANAGERIAL ABILITY (MA) ON ANALYST FOLLOWING (FOLLOW) AND FORECAST DISPERSION (DISPER) FOR FIRMS WITH COMPLICATED FINANCIAL REPORTS

Independent Variables	Dependent Variables	
	FOLLOW	DISPER
Intercept	0.16854	-0.0141
	(0.01)***	(0.01)***
MA	0.29796	-0.04760
	(0.01)***	(0.01)***
MA*HIGH_SIZE	0.22778	-0.02619
	(0.01)***	(0.04)**
EPS	0.00001	-0.0505
	(0.01)***	(0.01)***
LOSS	-0.26817	0.10349
	(0.01)***	(0.01)***
LEVERAGE	0.26973	0.10722
	(0.01)***	(0.01)***
VOLATILITY	-0.12661	0.39582
	(0.01)***	(0.01)***
Year Effect	Yes	Yes
Industry Effect	Yes	Yes
Adj. R^2	0.06	0.07

$FOLLOW_{it}$ = The natural log of number of analyst following for firm i in quarter t.
$DISPERS_{it}$ = Analyst forecast dispersion, which is the standard deviation (STDEV) of analyst forecasts.
MA_{it} = Managerial ability. See section 3.2 for detailed discussion.
$SIZE_{it}$ = The natural log of firm i's total assets at the end of quarter t.
$HIGH\_SIZE_{it}$ = Dummy variable equal to 1 when rank is equal to 3 or 4 when partition the full sample into five quintiles (ranks: 0, 1, 2, 3 and 4) based on SIZE.
EPS_{it} = Earnings per share for firm i in quarter t.
$LOSS_{it}$ = Dummy variable that equals 1 if EPS at the end of year t is negative and 0 otherwise.
$LEVERAGE_{it}$ = The leverage ratio of firm i at the beginning of quarter t.
$VOLATILITY_{it}$ = The volatility of stock returns, equals the standard deviation of monthly stock returns at year t.

CONCLUSION

I investigate the association between managerial ability, derived from frontier analysis, and the number of analyst following a firm, as well as the analyst forecast dispersion. I bridge the gap in the existing literature: prior literature documents that the managerial ability is positively associated with accounting quality (Demerjian, Lev, Lewis, and McVay 2013; Baik, Farber and Lee 2011). In addition, prior literature indicates that number of analysts following a firm is positively associated with accounting quality, and analyst forecast dispersion is negatively associated with accounting quality (Lang and Lundholm 1996; Irani and Karamanou 2003). I examine the relation between managerial ability and the number of analysts following a firm as well as analyst forecast dispersion. I find that managerial ability is positively associated with analyst following a firm and negatively associated with analyst forecast dispersion. Additionally, the effects are more pronounced after Sarbanes–Oxley Act of 2002, and are more pronounced for firms with complicated financial reports. Collectively, our findings support the notion that the managerial ability is relevant to analysts' decision making.

REFERENCES

Baik, B., Farber, D.B., & Lee, S.S. (2011, Winter). CEO Ability and Management Earnings Forecasts. *Contemporary Accounting Research*, *28*(5), 1645-1668.

Ball, R., & Shivakumar, L. (2005). Earnings Quality in UK Private Firms: Comparative Loss Recognition Timeliness. *Journal of Accounting and Economics*, *39*(1), 83–128.

Botosan, C., & Harris, M. (2000). Motivations for a change in disclosure frequency and its consequences: An examination of voluntary quarterly segment disclosures. *Journal of Accounting Research*, *40*, 21-40.

Core, J.E. (2001). A review of the empirical disclosure literature: discussion. *Journal of Accounting and Economics*, *31*(1-3), 441-456.

Demerjian, P., Lev, B., & McVay, S. (2012). Quantifying managerial ability: A new measure and validity tests. *Management Science*, *58*(7), 1229–1248.

Demerjian, P.R., Lev, B., Lewis, M.F., & McVay, S.E. (2013). Managerial ability and earnings quality. *The Accounting Review*, *88*(2), 463-498.

Easton, P. (2007). Estimating the cost of capital implied by market prices and accounting data. *Foundations and Trends in Accounting*, *2*(4), 241–364.

Francis, J., Schipper, K., & Vincent, L. (2002). Earnings announcements and competing information. *Journal of Accounting and Economics*, *33*(3), 313-342.

Healy, P., Hutton, A., & Palepu, K. (1999). Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Research*, *16*, 485-520.

Irani, A.J., & Karamanou, I. (2003). Regulation Fair Disclosure, analyst following and analyst forecast dispersion. *Accounting Horizons*, *17*(1), 15-29.

Krishnan, G.V., & Wang, C. (2014). The Relation between Managerial Ability and Audit Fees and Going Concern Opinions. *Auditing: A Journal of Practice & Theory*. doi: http://dx.doi.org/10.2308/ajpt-50985

Krishnan, J., Li, C., & Wang, Q. (2013). Auditor industry expertise and cost of equity. *Accounting Horizons*, *27*(4), 667-691.

Lang, M., & Lundholm, R.J. (1996). Corporate disclosure policy and analyst behavior. *The Accounting Review*, *71*(4), 467-492.

Liu, C., Wang, T., & Yao, L.J. (2014). XBRL's Impact on Analyst Forecast Behavior: An Empirical Study. *Journal of Accounting and Public Policy*, *33*(2014), 69-82.

Lobo, G.J., & Zhou, J. (2010). Changes in Discretionary Financial Reporting Behavior Following the Sarbanes-Oxley Act. *Journal of Accounting, Auditing and Finance*, *25*(1), 1-26.

Penman, S., & Zhang, X. (2002). Accounting Conservatism, the Quality of Earnings, and Stock Returns. *The Accounting Review*, *77*(2), 237–264.

Roulstone, D. (2003). Analyst following and market liquidity. *Contemporary Accounting Research*, *20*, 552-578.

Trueman, B. (1986). Why do managers voluntarily release earnings forecasts? *Journal of Accounting and Economics*, *8*(1), 53–71.

APPENDIX 1
VARIABLE DEFINITIONS

Variables	Definition
$FOLLOW_{it}$	The natural log of number of analyst following for firm i in quarter t.
$DISPERS_{it}$	Analyst forecast dispersion, which is the standard deviation (STDEV) of analyst forecasts.
MA_{it}	Managerial ability. See section 3.2 for detailed discussion.
$SIZE_{it}$	The natural log of firm i's total assets at the end of quarter t.
$HIGH\_SIZE_{it}$	Dummy variable equal to 1 when rank is equal to 3 or 4 when partition the full sample into five quintiles (ranks: 0, 1, 2, 3 and 4) based on SIZE.
EPS_{it}	Earnings per share for firm i in quarter t.
$LOSS_{it}$	Dummy variable that equals 1 if EPS at the end of year t is negative and 0 otherwise.
$LEVERAGE_{it}$	The leverage ratio of firm i at the beginning of quarter t.
$VOLATILITY_{it}$	The volatility of stock returns, equals the standard deviation of monthly stock returns at year t.
SOX	Dummy variable equal 1 when year is greater or equal to 2002 and 0 otherwise.

Piercing the Autonomy in Payment Undertakings: Fraud and Others?

Eliza Zhangminmin Xue
Hong Kong Baptist University

This article reviews and challenges the position taken in the current English law in documentary payment undertakings that autonomy stands as a cardinal rule and fraud as the only exception. While admitting that the autonomy principle which secures a smooth, speedy and dependable documentary payment remains the backbone of the international financing system, it submits that it would adversely affect the integrity of law in a broader sense if such justifiable grounds of exception as illegality, nullity and unconscionability are entirely disregarded. It is necessarily beneficial to the international trade and the instrumental payment system as a whole if a principled and incremental approach would be adopted by courts when weighing the strength of justification for each individual new ground, rather than shutting a blind eye to their potential merits.

Keywords: documentary payment undertakings, fraud, illegality, nullity, unconscionability

INTRODUCTION

It is the current English law in documentary payment undertakings, as represented in *Edward Owen Engineering Ltd v. Barclay's Bank International Ltd* [1978] QB 159, that autonomy stands as a cardinal rule and fraud as the only exception. However, it remains debatable whether fraud should be the only exception and to what extent new grounds, such as illegality, nullity and unconscionability, should be allowed to pierce the autonomy.

It is submitted in this paper that it is in the interest of commercial efficiency for the courts to adopt a principled and incremental approach in adopting new grounds of exception to autonomy by weighing their strength of justification, rather than to show a firm reluctance to allow room for recognition when the trend of their gradual expansion is inevitable, as witnessed in the judicial practice among different jurisdictions.

This paper is divided into three parts: First, the autonomous nature of documentary payment undertakings as stated in the relevant international rules and the English case law is discussed. Secondly, it examines the fraud exception and its high threshold for proving fraud in the English law. Thirdly, some judicial practice which endeavor to extend the narrow boundary of the fraud exception would be explored, with an aim to evaluate whether it is necessarily beneficial to documentary payment undertakings when new grounds are gradually recognized alongside with the fraud exception.

THE *RAISON D'ÊTRE* OF BANKS' UNDERTAKING: IRREVOCABILITY AND AUTONOMY

An Assured Right to Be Paid

The whole purpose to bring banks and their credit in is to bridge the 'distrust divide' in the international trade. By issuing a credit or a demand guarantee, the bank creates an assured right to be paid to the beneficiary, before the seller parts with control of the goods in the case of letters of credit, and before the buyer proves any breach by the seller in the case of demand guarantees. They are autonomous in nature and are separate from and independent of the underlying contract. The right and duty of banks to make payment under the documentary payment undertakings do not at any rate depend on performance of beneficiaries under the underlying contract.

The fundamental rules of irrevocability and autonomy are stipulated in both the Uniform Customs and Practice for Documentary Credits 2007 (UCP 600) and the Uniform Rules for Demand Guarantees 2010 (URDG 758). UCP 600 art 4(a) provides that 'a credit by its nature is a separate transaction from the sale' and is 'not subject to claims or defenses by the applicant'. URDG 758 art 5 contains similar words for a demand guarantee. The irrevocable undertaking to pay and its independence of the sale determine that banks deal with documents and not with goods, services or performance of the contract, as prescribed in UCP 600 art 5 for a letter of credit and URDG 758 art 6 for a demand guarantee. Moreover, UCP 600 art 34 expressly excludes banks' liability for 'the form, sufficiency, accuracy, genuineness, falsification or legal effect of any document'.

It is such assurance from banks that injects life into the commercial utility. As described by Kerr J. in *RD Harbottle (Mercantile) Ltd v. National Westminster Bank Ltd* [1978] QB 146, the irrevocable obligations assumed by banks are 'the life-blood of international commerce', and 'trust in international commerce could be irreparably damaged' (p. 155) if they are not allowed to be honored. That system of financing these operations, as Jenkins L.J. observed in *Hamzeh Malas & Sons v. British Imex Industries Ltd* [1958] 2 KB 127, 'would break down completely' (p. 129) if the dispute arising from the underlying contract has the effect of freezing the sum promised to be paid. Lord Denning chose to regard these demand guarantees as virtually promissory notes payable on demand. In *Edward Owen Engineering Ltd*, his lordship summarized the general position: 'A bank which gives a performance guarantee must honor that guarantee according to its terms. It is not concerned in the least with the relations between the supplier and the customer; nor with the question whether the supplier has performed his contracted obligation or not; nor with the question whether the supplier is in default or not. The bank must pay according to its guarantee, on demand, if so stipulated, without proof or conditions. The only exception is when there is clear fraud of which the bank has notice' (pp. 170-1). Messages from the English courts reinforce that irrevocability and autonomy are the *raison d'être* of banks' undertaking.

A 'Bargained-for Risk Redistribution Device'

It follows that the risks on the underlying contract have been redistributed: the risk of non-payment has been reversed from the seller to the buyer by a letter of credit, and the risk of non-performance from the buyer to the seller by a demand guarantee. The 'pay first, sue later' situation, as pointed out by Professor Benjamin, strengthens the beneficiary's negotiating position in the subsequent claim on the underlying contract when it possesses the funds paid by the bank, and they are regarded as a 'bargained-for risk redistribution device' (2018, paras. 23-076, 24-002). The courts have shown marked reluctance to interfere with such freely assumed commercial risk in a normal arm's length commercial transaction.

Compared with letters of credit, demand guarantees are more vulnerable to abusive calling since they are 'virtually promissory notes payable on demand' (*Edward Owen Engineering Ltd*, p.170). Though the URDG 758 purports to provide a fetter or disincentive on abusive calling of the guarantee in art 15(a) and 15(b) requiring a statement of the seller's breach and its nature to be presented with the guarantee, its effect could be eroded by the high threshold of proving subjective dishonesty of the beneficiary when making the demand, let alone the exclusion allowed by URDG 758 art 15(c). As confessed by Kerr J. in the *Harbottle* case, the vulnerability to abusive calling in demand guarantees results 'as though they represented a

discount in favor of the buyers' (p. 150). The fraud exception, therefore, is devised to strengthen the immunity from such abusive calling.

A TIGHTLY DEFINED FRAUD EXCEPTION BASED ON THE *EXTURPI CAUSA* PRINCIPLE

It is a matter of general principle and public policy that the courts shall not allow fraud-afflicted claims. It is the *exturpi causa* principle, i.e., fraud unravels all, that provides the rock foundation to the fraud exception. However, it is in the vital interest of preserving the value of certainty in commercial transactions to define the boundary of the fraud exception. Both UCP and URDG are silent about the fraud exception, and proper recourse could only be made to the national law. It is widely recognized that the English approach is extremely strict.

A Delineation Between Innocent and Non-Innocent Beneficiaries

An enshrined formulation of fraud exception is provided by Lord Diplock in *United City Merchants (Investments) Ltd v. Royal Bank of Canada*, *The American Accord* [1983] 1 AC 168, which declared that 'there is one established exception: that is, where the seller, for the purpose of drawing on the credit, fraudulently presents to the confirming bank documents that contain, expressly or by implication, material representations of fact that to his knowledge are untrue' (p. 183). Two elements define the success of a fraud claim: clear evidence of subjective dishonesty of the beneficiary and the bank's knowledge of it. An honest but mistaken belief would not suffice. Nor will a fraud which belongs to a third party, such as carriers.

Such a delineation between innocent and non-innocent beneficiaries has been criticized to be 'problematic' in that 'fraud is fraud, regardless of who the perpetrator is' (Donnelly, 2008, p. 323). It might undermine the documentary integrity if a beneficiary who is ignorant of a third party's fraud is still entitled to payment even though such fraud renders the relevant document worthless. Professor Goode comments that such a formulation was 'seriously flawed' and 'untenable' (Goode, 2016, p. 1062-3).

A High Threshold of Proving Fraud

The fraud must be very clearly established. In *Edward Owen Engineering*, Lord Denning MR confessed that 'the banks will rarely, if ever, be in a position to know whether the demand is honest or not', and 'at any rate they will not be able to prove it to be dishonest' (p. 170). Only cases of truly compelling evidence of fraud would suffice. Banks by no means stand in an investigation position to detect suspicions or to make inquiries.

The threshold for the applicant in proceedings for an interim injunction based on fraud is a high one. As Ackner L.J. formulated in *United Trading Corp SA v. Allied Arab Bank Ltd* [1985] 2 Lloyd's Rep. 554, it must be established that 'it is seriously arguable that, on the material available, the only realistic inference is that the defendant could not honestly have believed in the validity of its demands on the performance bonds' (p. 561). A test involving a three-stage enquiry has been laid down by the House of Lords in *American Cyanamid Co v. Ethicon Ltd* [1975] AC 396, namely: (1) the evidence of the merits, (2) the adequacy of damages as an alternative remedy and, (3) the overall balance of convenience. It has been noted by the Privy Council in *Alternative Power Solution Ltd v. Central Electricity Board* [2015] 1 WLR 697 that the result of the balance of convenience 'will almost always militate against the grant of an injunction' (para. [79]).

The combined effect of difficulties brought by the substantial and procedural requirements as to establish fraud and to justify an interim injunction is that 'a successful plea of fraud appears to be illusory' (Chong, 1990, p. 416) Fraud, as observed by Professor Goode, is 'the one least likely to succeed' (2016, p. 1057) among all the defenses to payment. The irresistible necessity arising from the general *exturpi causa* principle and the almost unconquerable castle of proving fraud generate a collective appeal to the adoption of an expanded view of the orthodox fraud exception.

A TREND MOVING FROM STRICT APPROACH TOWARDS ENFORCING BROAD STANDARDS?

The tightly bound fraud exception in the English law has felt the pressure of the force of extension by judicial practice at home and abroad. No matter how arguable the development would be, it has been observed a trend moving 'towards enforcing broad standards of conduct which appeal to public perceptions of fairness and justice' (Mugasha, 2004, 516). Across the spectrum of strength of justification, there lies illegality or violation of public policy on the strong end, unconscionability on the other end, and nullity in the center involving heavy debates.

Illegality or Violation of Public Policy: A Strongly Justified Breakthrough?

It seems a natural and irresistible inference from the *exturpi causa* principle that if fraud unravels all, illegality or violation of public policy does the same. The first sign of recognition of illegality in the underlying transaction is noticed in *Group Jose Re v. Walbrook Insurance Co Ltd* [1996] 1 WLR 1152, where Staughton L.J. admitted that there must be cases when illegality can affect a letter of credit, and illegality would be a defence if it is clearly proved. His position obtained positive echo from Colman J. in *Mahonia Ltd v. JP Morgan Chase Bank* [2003] EWHC 1927 where he considered it wrong to follow 'a rigid inflexibility in the face of strong countervailing public policy considerations' (para. [68]). In his view, if a beneficiary should not be allowed to profit from his own fraud, 'it is hard to see why he should be permitted to use the courts to enforce part of an underlying transaction which would have been unenforceable on grounds of its illegality' (para. [68]).

Even not to be regarded as an independent exception, illegality could at least be seen as a broadened fraud exception in the sense that a fraud could be perpetrated against the public as a whole rather than merely against the bank as a party to the credit (Enron, 2004). It seems a preferable development to enable illegality to be a valid ground within or without the collective label of fraud exception, which could hardly be seen as a threat to the lifeblood of international commerce. Otherwise, banks' payment undertaking might be abused to provide financial aid to illegal transactions.

Nullity: A Service or Disservice to Documentary Integrity?

Under the tightly laid boundary of the fraud exception in the English law, nullity of an apparently complying document will not defeat the beneficiary's claim for payment if its dishonesty is not clearly proved. In *Montrod Ltd v. Grundkötter Fleischvertriebs GmbH* [2002] 1 All ER 257, where a forged signature rendered an inspection certificate null and void, Porter L.J. refused to extend the law by creating a general nullity exception based on 'sound policy reasons' (para. [58]). A nullity exception, in Porter L.J.'s view, is 'susceptible of precision, involves making undesirable inroads into the principles of autonomy and negotiability', and 'would thus undermine the system of financing international trade by means of documentary credits' (para. [58]). Though the *Montrod* decision has undergone strong criticism (Hooley, 2002; Todd, 2008), it stands the current English law representing an unleaking boundary to the fraud exception.

In contrast with the English position, the Singapore Court of Appeal in *Beam Technology (MFG) Pte Ltd v. Standard Chartered Bank* [2003] 1 SLR 597 held that the bank was entitled to refuse to pay if 'a material document required under the credit is forged and null and void and notice of it is given within that period' (p. 610) However, the court confessed that 'it is not possible to define when is a document a nullity' (pp. 610-1) and choose to resort to the standard of reasonableness.

The academic responses to these two approaches are heavily divided. Those who advocate the Singapore approach believe that it is a 'distortion of the autonomy principle' if the beneficiary is allowed to collect payment by tendering 'worthless pieces of paper' (Goode, 2016, p. 1063) as long as he acts *bona fide*, and recognition of a nullity exception is necessary 'for maintaining the integrity of the law as a whole' (Donnelly, 2008, p. 342) when the trust and sanctity of financing international trade via payment instruments could be maintained. Those who choose to safeguard the English approach endeavor to remind that the law governing documentary undertakings is rule-oriented and commercial policy and efficiency

considerations that demand prompt honor of instruments are of paramount importance (Mugasha, 2004, p. 537). Further, it is observed that even the Singapore cases have not offered any definition of nullity that is reasonably sensible and workable, the lack of which is considered 'fatal to the proposition that there should be a nullity exception' (Ren, 2015, p. 19).

On the balance of the above stated arguments, it seems better for the courts to adopt a principled and incremental approach which, on the one hand, does not open the wide door to an undefining nullity exception; on the other hand, does not shut the window to ad hoc situations which deserve strong policy considerations.

Unconscionability: 'Too Interventionist for English Tastes'?

Unconscionable demand as a defense to autonomy has not been embraced by the English courts. An obvious reason is that the notion of unconscionability is too wide and discretional to be a workable or measurable concept in the context of documentary payment which requires a high degree of certainty, efficiency, and predictability. Though some recent English cases, such as *TTI Team Telecom International Ltd v. Hutchison 3G UK Ltd* [2003] 1 All ER 914 and *Simon Carves Ltd v. Ensus UK Ltd* [2011] EWHC 657, hint (in *dicta*) that the previous reluctance to apply a concept of unconscionability may not last forever, it is commented to be 'too interventionist for English tastes' (Goode, 2016, p. 1065).

Nonetheless, the unconscionability jurisdiction has gained judicial support in both Singapore and Australia. In *Dauphin Offshore Engineering v. Sultan* [2000] 1 SLR 657, the Singapore Court of Appeal described the notion of unconscionability as unfairness, as distinct from dishonesty or fraud, or conduct of a kind so reprehensible or lacking in good faith that a court of conscience would either restrain the party or refuse to assist the party. The court sees no reason why it should be so sacrosanct and inviolate as not to be subject to the court's intervention except on the ground of fraud. The approach has been confirmed by the same court in *BS Mount Sophia Pte Ltd v. Join-Aim Pte Ltd* [2012] 3 SLR 352 where the test in a nuanced fashion has been emphasized. Similarly, an Australian court in *Olex Focas Pty Ltd v. Skodaexport Co Ltd* (1996) 134 FLR 331 considered unconscionability in a statutory context and injected such requirements as reasonable behavior in accordance with ordinary human standards and that reasonable expectations of others should not be defeated when one exercises his strict legal rights.

The elaborations provided by the Singaporean and Australian courts could not have alleviated the worry of its overuse and the courts' accompanying discretionary power. It has been criticized that when the courts is prepared to examine the issue of unconscionability, they have to dive into the underlying transaction details which shall be kept outside the realm of autonomy of letters of credit or demand guarantees (Mugasha, 2004, p. 520). The resultant eroding effect is believed to be more a disservice than a service to the business efficiency and certainty.

CONCLUSION

The autonomy principle which secures a smooth, speedy and dependable documentary payment remains the backbone of the international financing system. To uphold its efficiency as far as can be, the fraud exception has been tightly defined in the English law and a high threshold is adopted. However, it would adversely affect the integrity of law in a broader sense if such justifiable grounds of exception as illegality, nullity and unconscionability are entirely disregarded. It is necessarily beneficial to the international trade and the instrumental payment system as a whole if a principled and incremental approach would be adopted by courts when weighing the strength of justification for each individual new ground, rather than shutting a blind eye to their potential merits.

REFERENCES

Bridge, M. (2018). *Benjamin's Sale of Goods* (10th ed.). Sweet & Maxwell.

Chong, W.S. (1990). The abusive calling of performance bonds. *Journal of Business Law*, *5*(6), 414-427.

Clarke, M.A., Hooley, R., Munday, R.J.C., Sealy, L.S., Tettenhorn, A.M., & Turner, P.G. (2017). *Commercial Law: Text, Cases, and Materials* (5th ed.). Oxford University Press.

Donnelly, K. (2008). Nothing for nothing: A nullity exception in letters of credit? *Journal of Business Law*, *23*(4), 316-343.

Enonchong, N. (2006). The autonomy principle of letters of credit: An illegality exception? *Lloyd's Maritime and Commercial Law Quarterly*, *33*(3), 404-425.

Nathanson, C.P. (2004). Enron, letters of credit and the autonomy principle. *Journal of International Banking and Financial Law*, *19*(6), 204-209.

Goode, R., Miles, R., & McKendrick, E. (2016). *Goode on Commercial Law* (5th ed.). Penguin Books.

Hooley, R. (2002). Fraud and letters of credit: Is there a nullity exception? *Cambridge Law Journal*, *61*(2), 279-281.

Mugasha, A. (2004). Enjoining the beneficiary's claim on a letter of credit or bank guarantee. *Journal of Business Law*, *19*(5), 515-538.

Ren, J. (2015). A nullity exception in letter of credit law? *Journal of Business Law*, *30*(1), 1-19.

Todd, P. (2008). Non-genuine shipping documents and nullities. *Lloyd's Maritime and Commercial Law Quarterly*, *35*(4), 547-573.

How Do Household Financial Obligations Impact the Equity Premium?

Pedram Jahangiry
Utah State University

In this paper, two specific channels are proposed to investigate how household financial obligations impact the equity premium. Preference channel and borrowing constraints channel. Preferences are defined over households' consumption relative to their financial obligations. The model also introduces dynamic borrowing constraints, using financial obligation ratio as a proxy. A novel feature of the model is that in states of high marginal utility, the borrowing constraint binds and making it more difficult for households to smooth consumption. In addition, in these states, households become more risk averse. This dual mechanism both amplifies the risk premia and makes it time varying.

Keywords: asset pricing, equity premium puzzle, incomplete markets, household financial obligation ratio

INTRODUCTION

Mehra and Prescott (1985), modified a Lucas (1978) type exchange economy to reconcile the standard neoclassical macroeconomic theory with US data on the equity premium. In a consumption-based asset pricing setup, they specify an explicit two-state Markov process for consumption growth and calculate the price of the consumption claim and the risk-free rate. They find that under reasonable parameterization, the model is able, at most, to generate an equity premium of about 0.35 % as opposed to the 6 % premium observed in the data. They call it the "equity premium puzzle" and argue that the mean stock excess return calculated in their calibrated economy is too low, unless the coefficient of relative risk aversion is raised to implausibly high values.

"The ink spilled on the equity premium would sink the Titanic" (John Cochrane, 2008). There is no easy way to summarize the huge literature on the equity premium puzzle. Nevertheless, there is consensus among researchers that, if any, only an absolute asset pricing model would be able to explain the equity premium rather than a portfolio-based model. Absolute pricing model refers to the asset pricing models that use macroeconomic variables such as consumption, labor income, GDP growth and interest rate. After all, portfolio models are relative asset pricing models and are not able to answer questions like why the average returns are what they are or why the expected market return varies over time? To answer these questions, we need to construct a macroeconomic-based asset pricing model. Note that the most basic absolute pricing model - the standard consumption-based model - performs poorly in explaining the historical equity premium puzzle and it fails to explain the cross-sectional variations of expected returns. Hence, proposing a macroeconomic-based asset pricing model with the ability to explain the equity risk premium and the cross-sectional variations of excess returns has been always in the center of attentions among macro-finance researchers.

In the past few decades, there have been many efforts to resolve the puzzle. Various generalizations have been proposed to address the shortcomings of the standard consumption-based model. One approach

to solve the puzzle is to investigate the preference channel and start with the dependency of marginal utility of consumption on some other variables besides today's consumption. Many current frameworks including habit formation, consumption commitments, heterogenous preferences and employing non-separable utility functions allow for this dependency and indeed this is what has been extensively done in the literature.

Another approach to solve the equity risk premium puzzle is by focusing on the borrowing constraint channel and considering incomplete markets. There is also a large literature on asset pricing with borrowing constraints and margin requirements.

By introducing a single counter-cyclical macroeconomic variable, this paper is the first one connecting the two large literature on asset pricing with borrowing constraints and preference specifications. This paper also contributes to the literature on the relation between households' financial obligations and the equity risk premium.

In my version of the consumption-based asset pricing model, household financial obligations impact the equity risk premium via two separate channels. The first one is the preference channel where individuals' preferences are defined over consumption relative to financial obligations. The framework is analogs to habit formation models where the utility function depends on consumption relative to some habit level (Abel, 1990), (Constantinides, 1990) and (Campbell and Cochrane, 1999). However, in my model the persistence level is observable which is a big advantage over other standard habit models. With this setup, the marginal utility derived from the model is directly related to two components. One is of course the household's consumption and the other one is household financial obligations.

The preference channel is important because one possibility for consumption variation is the potential impact of household's debt level on its preferences. If households are averse to holding large amounts of debt relative to their income, a decline in income will prompt larger declines in consumption among highly indebted households in order to restore the desired debt-to-income ratio for a wide range of loss functions (Scott Baker, 2015). From the other hand, there is evidence that individuals who are more likely to face income uncertainty or to become liquidity constrained exhibit a higher degree of risk aversion in the presence of uninsurable risk (Guiso and Paiella, 2008). By defining households' preferences over consumption relative to financial obligations, the preference channel captures the time varying risk aversion behavior of individuals by featuring fluctuations of consumption net of financial obligations over the business cycles.

The second channel that I investigate in this paper is the borrowing constraint channel which most asset pricing models abstract from. In an infinite horizon aggregate household economy, financial obligations ratio - defined as total debt payments, housing payments and auto lease payments divided by total disposable income - will act as the borrowing constraint in this model. I show that the equity risk premium implied by the model is quiet sensitive to financial obligations ratio as a proxy for borrowing constraints.

But why should the model work? What are the ingredients of the model enabling it to explain and generate the observed equity risk premium in U.S data? The intuition is straight forward. As consumption rises in good times, households take on more debt and the debt payments gradually increase. In bad times consumption falls and households will delever slowly. Thus, debt payments move slowly, following consumption. Now consider a household who has taken a specific level of debt which it must repay. In recessions, as income declines towards this specific level of debt payments, the household reduces its consumption because it is scared, because of risk and risk aversion. Indeed, to make sure that the household can repay its debt payments, it exhibits more risk aversion and takes on less risk. This will decrease the demand for risky assets and increases the demand for precautionary savings in recessions. During booms however, consumption gets further away from the slow-moving financial obligations and hence the investor will become less risk averse and take on more risk. Thus, lower ratio of consumption relative to financial obligations in bad times and higher ratio in good times will directly impact the marginal utility of the household and make the pricing kernel more volatile. This is the households' time varying risk aversion that lead to desires to shift the composition of a portfolio from risky to risk free assets, this is what generates the equity premium.

Also note that in bad times, because of lower income, households face a higher financial obligations ratio. This means that lenders are less willing to lend them in these states of the economy. Thus, households'

borrowing constraints become binding exactly when they want to smooth consumption. Now since they cannot borrow to consume, they have to either invest less in or liquidate more of their assets. In either case, they will start by the risky assets rather than risk free assets (recall that the model proposed in this paper features a time varying risk aversion and people are more risk averse in recession). The decrease in demand for risky asset is much faster than the decrease in demand for risk free asset. This generates a higher risk premium during recessions. This intuition is consistent with our findings when we relax the borrowing constraint. As we let the household borrow more in bad times, the equity premium shrinks. Time varying borrowing constraints are basically other factors that make households to shift from risky assets to risk free assets.

In summary, a novel feature of the model is that in states of high marginal utility (i.e. in recessions, when income falls) lenders are reluctant to lend, the borrowing constraint binds and making it more difficult for households to smooth consumption exactly when they need to do so. In addition, in these states, households become more risk averse. This dual mechanism both amplifies the risk premia and makes it time varying.

The economic variable explored in this paper - financial obligations ratio - provides a fresh opportunity to investigate the determinants of asset risk. As a start, in a separate research, Jahangiry documents that the risk associated with the aggregate household financial obligations is an economy-wide risk and it is significant for explaining the variations in cross-section of stock returns. Conditioning down on financial obligations ratio, the financial obligations capital asset pricing model (FCAPM) proposed by Jahangiry, survives a wide range of classical econometric and diagnostic tests on explaining the variations in average returns across 25 portfolios formed based on size and book to market ratio. In another study, Jahangiry take one step further and test the predictability of stock returns/excess returns with household's obligations ratio. Using U.S stock market data, he shows that household's debt service ratio is able to predict market returns at short horizon and over business cycle frequencies. Jahangiry argues that mean deviations from debt service ratio is a better forecaster of future returns both in-sample and out-of-sample than several other popular forecasting variables.

The rest of the paper is as follows. In Section 2, I will discuss in details what is the household financial obligations and why does it matter. In section 3, I set up the model and derive the fundamental equations of asset pricing. Section 4 discusses the data and estimations. The results of the model are proposed in Section 5 and section 6 concludes.

HOUSEHOLD FINANCIAL OBLIGATIONS

Household financial obligations include total debt payments (mortgage debt payments plus consumer debt payments) and total financial commitments (rent, lease, insurance and property tax payments). Financial obligations affect households' optimization behavior. There is an extensive documentation in economic psychology literature studying the psychological impacts of being in debt suggesting that financial obligations are associated with high levels of anxiety and stress, (Brown, Taylor and Price 2005), (Richardson, Elliott and Roberts 2013). And more importantly this impact is independent of the poverty with which it is often associated, (Jenkins, Bhugra, Bebbington and Farrell 2008), (Meltzer, Bebbington, Brugha, and Dennis 2011). Financial obligations also impact households' budget constraints. In particular, if these obligations are high relative to income, and it is not possible to roll over the debt, then borrowers have to cut back on expenditure to avoid default. There is evidence that high financial obligations reduce expenditure at the micro level. The negative effect of a high debt service burden on consumption of households has been shown by (Olney, 1999), (Johnson and Li, 2010), (Dynan, 2012) and (Juseliuse and Drehmann, 2015).

Further the household financial obligations position is really important in determining whether the household is constrained from optimal consumption smoothing. The fact that a household may have been able to borrow in the past does not imply that it can borrow as much in the future. However, household financial obligations alone do not tell us much about this borrowing capability. Imagine two households with the same amount of financial obligations but different levels of income. Most probably the lenders are

willing to lend to the household with higher income. Hence, we need to define a more indicating variable as a direct proxy of borrowing constraints and that's the financial obligations ratio which we are going to present in next section.

What Is Financial Obligations Ratio (FOR) and Why Does It Matter?

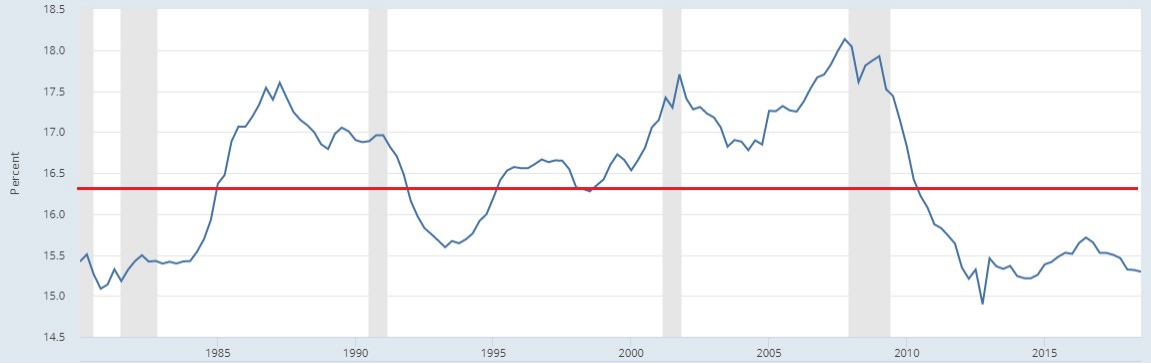
Households' Financial obligations ratio is defined as the households' total financial obligations divided by their total disposable income. Financial obligations ratio consists of two parts,

1. **Total debt service ratio**, which is equal to total debt payments divided by total disposable income. Debt payments include all the mortgage debt payments and consumer debt payments including auto loans, student loans and consumer credit cards.
2. **Total financial commitment ratio**, which is equal to total financial commitments divided by total disposable income. Financial commitments include all the rent payments, lease payments, insurance and property tax payments of the homeowners.

Properties of Financial Obligations Ratio

Since 1980 onward, Federal Reserve Board has reported the financial obligations ratio for U.S households. FIGURE 1 shows that household financial obligations ratio is a time varying macroeconomic variable with the average of 16.40%. As FIGURE 1 suggests, financial obligations ratio tends to move counter-cyclically over the business cycles.

FIGURE 1
FINANCIAL OBLIGATIONS RATIO AS A PERCENT OF DISPOSABLE INCOME



While the economy stays in good times, consumers keep spending more and hence increasing their financial obligations, now when the economy is hit by a negative shock (recession) this is the time that financial obligations are already high and people cannot smooth their consumption exactly when they need to do so. Hence what we observe in the data is that financial obligations ratio is high almost in the early stage of every recession because households are carrying a huge amount of obligations from previous good old days and then the ratio decreases as the economy recovers and households delever.

Higher financial obligations ratio also implies less investment in risky assets. This is due to the fact that when households are overextended, even a small income shortfall prevent them from smoothing consumption and making new investments (Drehmann and Juselius, 2012). FIGURE 1 also shows that almost after every recession, the financial obligations ratio pulls back to lower levels which is true because of the households' higher income during booms. There is also another reason in the following manner, when household obligations ratios are high and unemployment is rising, lenders may respond to the expected increase in defaults by limiting the availability of credit and this leads to lower aggregate payments and finally lower obligation ratio. Thus, financial obligations ratio has a **counter-cyclical** property. Shaded areas in FIGURE 1 indicate US recession periods. Other important properties of financial obligations ratio which make it the variable of interest in this paper are as follows:

- All the components of the financial obligation's ratio are **observable** so when working with data, there is no need to come up with questionable proxies.
- Financial obligations ratio is **directly** related to the interest rate. By construction, the higher the interest rate, the higher the payments and the higher the financial obligations. This explicit dependency establishes a direct link between obligation ratios and predictability of stock market returns. (Juselius and Drehmann, 2015) argue that "the average lending rate reflects not only current interest rate conditions, but also past money market rates, past inflation and interest rate expectations as well as past risk and term premia. This implies that the lending rate, and hence the debt service ratio, is chiefly influenced by current and past monetary policy decisions".
- Financial obligations ratio captures the burden of obligations on the households more accurately than the established debt-to-GDP ratio. More specifically financial obligations ratio accounts for changes in interest rates and also maturities that affect households' repayment capacity.
- (Drehmann and Juselius, 2012) found that the debt service ratio, which is the main part of financial obligations ratio, produces a very reliable early warning signal ahead of systemic banking crises. In the context of absolute asset pricing this is really important because we are looking for a conditioning down variable which is correlated with business cycles specially with bad times.
- Financial obligations ratio can be used as a direct indicator of borrowing constraints. (Johnson and Li, 2010) tested the proposition that a higher debt service ratio increases the likelihood of credit denial. So, household's obligation ratio is a critical input for lending institutions in order to provide the households with more leverage.

Given the aforementioned properties for financial obligations ratio and the fact that expected returns vary with business cycles, we are going to explore how these obligations impact the equity risk premium.

THE MODEL

Environment

I will consider an infinite horizon endowment economy in which the agents are endowed with an uninsurable stochastic income at each period. The agents in this economy are:

1. Large number of homogeneous households
2. A lending institution

In this paper, I use a modified version of Greenwook-Hercowitz-Huffman utility structure which enable us to represent the aggregate households with a representative agent. Thus, we have a representative agent environment with an external supply of debt provided by lending institutions. This can be thought of as small open economy. It is assumed that the lending institution is aware of the income distribution of the representative agent. There are three assets and two markets in this economy. The assets are one perishable consumption good and two durable assets, an inside security and an outside security (debt instrument). The inside security provides dividends according to an exogenous stochastic process and the outside security is exogenously supplied by the lending institution. The environment is summarized in TABLE 1.

TABLE 1
THE MODEL ENVIRONMENT

Capital Market	Debt Market
Households	Households + A lending institution
One inside security	One outside security
Zero net supply	Positive net supply
No trade equilibrium	Exogenous prices

The two markets in the model are:

1. The **capital market** which is a market for allocating idiosyncratic risk among households. In this capital market everything is in zero net supply. Households can trade contingent claims (inside security) among themselves but since all the households are identical, the prices are going to be shadow prices for no-trade equilibrium.
2. The **debt market** in which there is an outside supply of debt (outside security) provided by a lending institution. I am not going to model the supply side of the debt market. The lending rate is an exogenously specified rate. The households can borrow from this lending institution up to a certain amount defined by their financial obligations ratio.

Constraints

In each period, households are endowed with a stochastic exogenous income which they can either consume or invest in inside security. They are allowed to borrow against their stochastic income and use it only for consumption purposes. The model will have a non-stationary environment due to non-stationarity of the stochastic aggregate income and the exogenous dividends of inside security. However, the exogenous stochastic borrowing rates are stationary. Consumption, financial obligation, dividends, aggregate disposable income, prices of the equity and the risk-free bond are all denominated in units of the consumption good.

In the model, for each period t, C_t is the aggregate consumption. Y_t is the aggregate disposable income and X_t represents the dividends generated by the inside security. By making the wage income process exogenous, I abstract from the labor-leisure trade-off. This means that the labor is supplied inelastically and the labor leisure choice is not modeled. D_t is the debt service level borrowed by the representative agent at the gross rate R_t^d. I assume that the lending institutional arrangements can issue and redeem debt instruments. This lending institutions exogenously set a lending cap θ_t on each household. Hence θ_t provides a state dependent upper bound for the households' borrowing capacity. Z_t is a non-negative amount of investment in inside security with an ex-dividend price of p_t at time t. The agent faces the following constraints:

- Budget constraint

$$C_t + p_t Z_t + D_{t-1} R_{t-1}^d \leq Y_t + (p_t + X_t) Z_{t-1} + D_t \tag{1}$$

- Financial obligations constraint

$$D_t R_t^d \leq \theta_t Y_t \quad , \quad D_t \geq 0 \tag{2}$$

Inequality (1) is the budget constraint that the agents face in each period. The agent comes into the period with a stochastic wage income Y_t. There is also income from securities purchased in last period. Agent can liquidate Z_t amount of inside security at price p_t with the dividend X_t. Furthermore the agent can borrow against his stochastic income at the amount of D_t. These are resources of funds. Now the left-

hand side of (1) will show us how does the agent spend the available funds. First of all, he consumes C_t, then he can purchase inside security to take over to the next period and finally he has to pay interest on debt that he had borrowed. R_{t-1}^d is the gross return on debt instrument and it means that the agent has to payback whatever he has borrowed in last period plus the interest. At $t = 0$ the representative agent is born debt free $D_{-1} = 0$, and he is endowed with nothing but a stochastic income Y_0 i.e, $Z_0 = 0$.

Inequality (2) is the borrowing constraint. It basically indicates how much the agent can borrow against its labor income. This financial obligation ratio constraint (2) is a type of constraint that we observe in the economy. Interest payments on debt over income is a number that lenders would rather to see it below some certain levels like 1/3 or 1/4 and this number varies from time to time. (2) is a constraint that I impose in this model and that's one of the innovations of the model, the borrowing constraint channel. The innovation is introducing a State dependent, time-varying borrowing constraint. I am making the borrowing capacity to be state dependent. In bad times financial obligations ratio gets higher (due to negative income shock) and agent's borrowing capability thereafter shrinks. This force the agent to reduce its consumption in bad times even further. Note that the income Y_t, dividends X_t and obligation ratio cap θ_t are all exogenous stochastic variables. For calculation purposes, we work with detrended income y_t and dividend growth x_t which are determined by the following Markov processes:

$$y_{t+1} = (1 - \rho_y)\overline{y} + \rho_y y_t + \epsilon_{t+1}^y \tag{3}$$

$$x_{t+1} = (1 - \rho_x)\overline{x} + \rho_x x_t + \epsilon_{t+1}^x \tag{4}$$

In equations (3) and (4), $\overline{y}$ and $\overline{x}$ are the averages of detrended aggregate income and dividend growth respectively. ρ_y and ρ_x are the auto correlations and epsilons are the relevant shocks associated with y and x. Finally, the exogenous process for financial obligations ratio cap θ_t is defined as in (5).

$$E_t(\theta_{t+1}|y_t) = f(y_t) \tag{5}$$

Equation (5) implies that the process of θ is totally pinned down by the process of y. This assumption is very intuitive as the lending institution is the one that sets θ exogenously and it is aware of household's income distribution. Timing of the constraints are as follows. At time t, the agent knows at what rate he will be able to borrow, so R_t^d is measurable with respect to time t. According to the financial obligations' constraint in (2), D_t is also measurable at time t. Thus, the only random variables here are Y and θ. Note that there will be no default in this model. Under my parameterization, the agent can always reduce his consumption such that he has a positive net worth. In other word he will never have a realization of Y so low at which he cannot pay off his debt by reducing consumption. And besides, lenders are going to choose θ conditional on some expectation of future income of the agent. So, if lenders' conditional expectation of future income is low, they will lower the θ to make sure that agent is going to be able to pay off his debts. Hence the lender is building the expectations of θ_{t+1} based on y_t. It means that when expected income is low, the lender will react by decreasing θ_{t+1} conditional on y_t such that the conditional expectation of θ is lower than it's expected value i.e. $E_t(\theta_{t+1}|Y_t) \leq \overline{\theta}$. In short, in this model, the lender is the one monopolist and everybody else is a price taker. This monopolist has some expectations of agents' income and is going to reduce the amount that agents can borrow, precisely when they would like to borrow next period. This is the building block of my model and it basically shows that how the model generates a more volatile marginal utility. In bad times the effective consumption (consumption net of financial obligations) is smaller and in good times it is bigger than the standard consumption in Mehra-Prescott world. Thus, with this set up, everything will be conditional on Y_t. The variations in marginal rate of substitution is going to determine the returns on inside security and hence this is the extra variations in dividends that is going to generate a higher equity premium.

The Preferences

The utility function presented here will show us how the financial obligations ratio impacts the equity risk premium via the preference channel. In my model, the agent's preference is defined over consumption relative to financial obligations G. This is a behavioral set up which is analogous to habit formation models; In the sense that, while in the latter model "*the distance from consumption habit*" gives the agent utility, in the former case "*the distance from financial obligation*" does the same job. More specifically, I will use the simple power utility function defined over the representative agent's *effective consumption* C^*, where C^* is the consumption net of financial obligations incorporating the distance from the financial obligations. Effective consumption is defined as $C_t^* = C_t - G_t$ where $G_t = D_{t-1}R_{t-1}^d$ is the financial obligations that the agent carry over to period t from the last period. Hence the utility function of the agent is,

$$U(C_t^*) = \frac{C_t^{*(1-\gamma)}}{1-\gamma} = \frac{1}{1-\gamma}(C_t - G_t)^{(1-\gamma)} \tag{6}$$

Equation (6) suggests that a household with lower financial obligations will have a higher effective consumption and hence receives a higher utility. This behavioral set up is chosen because there is an extensive documentation in economic psychology literature studying the psychological impacts of being in debt. Indeed, financial obligations are associated with high levels of anxiety and stress (Brown, Taylor and Price, 2005), (Richardson, Elliott and Roberts, 2013). And more importantly This impact is independent of the poverty with which it is often associated (Jenkins, Bhugra, Bebbington and Farrell, 2008), (Meltzer, Bebbington, Brugha, and Dennis, 2011). This is the behavioral reason of why we should include financial obligations in the utility function. There is also a structural reasoning behind the choice of effective consumption as in difference form $(C - G)$, rather than ratio form $(\frac{C}{G})$. I pick the difference form because it generates a time varying relative risk aversion. However, there is no consensus on pro-cyclicality or counter-cyclicality of the relative risk aversion. Risk aversion is counter-cyclical in habit formation models (in recession, consumption surplus ratio is lower so risk aversion is higher) and pro-cyclical in happiness maintenance models (in good times you become more risk averse as you wish good days never end). Time varying risk aversion plays an important role in determining the equity premium, especially during recessions. Because "recessions are phenomena of risk premiums, risk aversion, risk bearing capacity and desires to shift the composition of a portfolio from risky to risk free assets, a flight to quality, not a phenomenon of intertemporal substitution, a desire to consume more tomorrow vs. today" (John Cochrane, 2016). Also note that in this model, $C_t - G_t$ is always positive. People slowly develop financial obligations, so consumption is always greater than debt obligations (no- default assumption), indeed financial obligations form the trend in consumption.

With this specification in (6), the coefficient of relative risk aversion is going to be (7)

$$RRA = -C\frac{u''(C)}{u'(C)} = \gamma\left(\frac{1}{\frac{C-G}{C}}\right) = \frac{\gamma}{S} \tag{7}$$

where $S = \frac{C-G}{C}$ is the consumption surplus. This is analogous to (Campbell and Cochrane, 1999) habit model with financial obligations replacing the consumption habit. It is very important to emphasize that the persistence level in my model is observable which is a big advantage to Campbell and Cochrane external habit model in which the habit level is not observable. However, the idea is the same, in bad times, as consumption or the surplus consumption ratio S decreases, agent's relative risk aversion rises i.e. the same proportional risk to consumption is a more fearful event when consumption starts closer to financial obligations, G.

The link between consumption surplus S and the financial obligations ratio is straight forward. Financial obligations G is like a slow-moving habit in this model. In recessions, when a negative shock to the aggregate income is realized, it will increase the current financial obligations ratio of the representative

agent. According to the budget constraint (1) and the borrowing constraint (2), the agent has no option but to decrease its consumption C_t. This is true because of three reasons. First, there is a no-trade equilibrium, the representative agent doesn't have the option of liquidating its assets. Second, I am implicitly assuming that the agent cannot default on its debt payments so it has to pay back $D_{t-1}\ R_t^d$ in full and third, in bad times the financial obligations ratio is already capped so the agent cannot leverage any more. This will make the consumption closer to slow-moving obligation G and hence reduce the consumption surplus S. Hence in recessions (negative income shocks), financial obligations ratio is high, consumption surplus is low and relative risk aversion is high. This enables the model to deliver a time-varying, recession-driven equity risk premium.

Stochastic Sequential Problem

The representative agent maximizes the following sequential problem:

$$Max \quad E_0\{\sum_{t=0}^{t=\infty} \beta^t\, U(C_t^*)\} \tag{8}$$

where $U(C_t^*) = \frac{C_t^{*1-\gamma}}{1-\gamma}$ and $C_t^* = C_t - G_t$. Subject to the budget constraint (1), borrowing constraint (2), exogenous stochastic processes (3)-(5) and the non-negativity constraints.

$$G_t = D_{t-1}R_{t-1}^d \tag{9}$$

$$C_t + p_t Z_t + D_{t-1}R_{t-1}^d \leq Y_t + (p_t + X_t)Z_{t-1} + D_t \tag{10}$$

$$D_t R_t^d \leq \theta_t Y_t \tag{11}$$

$$y_{t+1} = (1-\rho_y)\overline{y} + \rho_y y_t + \epsilon_{t+1}^y \tag{12}$$

$$x_{t+1} = (1-\rho_x)\overline{x} + \rho_x x_t + \epsilon_{t+1}^x \tag{13}$$

$$E_t(\theta_{t+1}|y_t) = f(y_t) \tag{14}$$

$$C_t \geq 0 \quad , \quad D_t \geq 0 \quad , \quad Z_t \geq 0 \tag{15}$$

$$\text{Given } X_0, Y_0, R_0^d \quad , \quad Z_0 = D_{-1} = 0 \tag{16}$$

For simplification purposes, I assume the exogenous lending rate is constant and equal to $R_t^d = \overline{R}^d$. The transversality condition holds for financial obligations which implies that the shadow value of debt service must be equal to zero in the limit. We also assume Inada conditions on effective consumption. The economy is completely specified by the preference parameters β, γ and realization of the stochastic processes followed by $\Lambda_t = (y_t, x_t, \theta_t)$. The equilibrium is defined as the sequences of consumption $\{\widetilde{C_t}\}$, investment $\{\widetilde{Z_t}\}$, borrowing decisions $\{\widetilde{D_t}\}$ of the representative agent and the prices p_t such that:

1. Taking the prices and exogenous vector Λ as given, sequences of consumption, investment and borrowing decisions, optimize the households' lifetime expected utility.
2. Consumption, capital and debt markets clear in all periods. Note that there is no need to clear the debt market; as mentioned earlier, debt service rates are exogenously determined.

Pricing Kernel

Given the utility function in (6), the intertemporal marginal rate of substitution (The pricing Kernel) is:

$$M_{t+1} = \beta \frac{U_c(C^*_{t+1})}{U_c(C^*_t)} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} \tag{17}$$

where $S_t = \frac{C_t - G_t}{C_t}$. The pricing Kernel is related to consumption growth and consumption surplus which depends on financial obligations and is an implicit state variable in my model. We can now calculate moments of the marginal rate of substitution (7) and find asset prices. Taking the first order conditions (FOC) with respect to C_t, D_t and Z_t and combining the results will give us the "fundamental equations of asset pricing". From FOC ($[C_t], [Z_t]$) we get the price of inside security:

$$p_t = E_t \left(\beta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} (p_{t+1} + X_{t+1})\right) \tag{18}$$

Now with equation (18) in hand we can derive the price of any inside security given its dividend. To calculated the equity risk premium, we need the price of risk-free bond q_t. Using (18) and the fact that no-coupon treasury bonds are traded in discounted values, then the price of risk-free bond is equal to:

$$q_t = E_t \left(\beta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} 1\right) \tag{19}$$

Market clearing condition imply that $Z_t = 0 \quad \forall t \geq 0$; this is because every household is the same and the equilibrium outcome must be the no-trade outcome. We are interested in finding the prices that support this no trade outcome. From market clearings and budget constraint, the equilibrium consumption sequence is:

$$C_t = Y_t + D_t - D_{t-1} R^d_t \tag{20}$$

Now we are only one step away from finding an explicit form solution for the equity price and risk-free bond price. Fortunately, it is easy to show that the debt service ratio constraint in (2) is binding. Note that at each period t, the utility function U_t is strictly increasing in consumption C_t, so the budget constraint and the borrowing constraint binds to guarantee that the agent is maximizing. However, one may argue that as financial obligations increase, the effective consumption will also decrease. This is not true here, since current obligations G_t is a function of last period debt service D_{t-1} and not the current borrowings D_t. This means that the representative agent will cap the amount of borrowings. This will pave the way for calculating the equity premium and the risk-free rate implied by the model. Now using the Equilibrium consumption path in (20) accompanying with equations (18) and (19), it gives us the fundamental equations of asset pricing at the equilibrium. Note that it is convenient to define $w_t = \frac{p_t}{X_t}$ as the price-dividend ratio because it will allow us to write down the equilibrium equity returns in terms of dividend growth, which is stationary and not the dividend itself, which is non-stationary. Thus, by dividing both sides of (18) to X_t, we can rewrite (18) as:

$$w_t = \beta E_t \left\{\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} (1 + w_{t+1})\left(\frac{X_{t+1}}{X_t}\right)\right\} \tag{21}$$

The right-hand side of (8) is a conditional expectation, by applying the Low of Iterated Expectations we are allowed to take X_t into the conditional expectation. Where, $\frac{X_{t+1}}{X_t}$ is the dividend growth. Next we can use (19) and (21) to solve for the risk-free rate R^f_t, Equity returns R^e_t and consequently the Equity Premium EP_t. Starting with the risk-free rate and using (19) we have:

$$R_t^f = \frac{1}{q_t} = \frac{1}{\beta E_t\left(\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{S_{t+1}}{S_t}\right)^{-\gamma}\right)} \quad (22)$$

The equity return can be derived as:

$$R_{t+1}^e = E_t\left(\frac{p_{t+1}+X_{t+1}}{X_t}\right) = E_t\left(\frac{w_{t+1}X_{t+1}+X_{t+1}}{w_t X_t}\right) = E_t\left\{\left(\frac{X_{t+1}}{X_t}\right)\left(\frac{1+w_{t+1}}{w_t}\right)\right\} \quad (23)$$

where w_t is defined as in (21). And finally, the equity premium is simply defined as (23) minus (22):

$$EP_t = E_t(R_{t+1}^e) - R_t^f \quad (24)$$

At the end, if the exogenous variables follow Markov processes, then solving functional equations in (22) and (23) is just solving a finite system of linear equations. Thus, defining the exogenous stochastic processes for the aggregate real income and dividend growth rate, we can now test the model by comparing the observed equity premium and risk-free rates in U.S data to the ones implied by the model.

It is worth emphasizing that the model deals with non-stationary environment because of non-stationarity of aggregate dividends and aggregate income. This enables the model to generate a non-stationary equilibrium consumption path, a non-stationary equity prices and a stationary equity premium which is consistent with the data. Note that the risk-free bond prices generated by (19) are stationary because q_t only depends on consumption growth and financial obligations growth which are both stationary

DATA AND ESTIMATIONS

The numerical goal is to compare the equity risk premium observed in U.S data with the one generated by the model presented in this paper. I use annual data for the period between 1980 and 2018 to report the equity premium observed in the data. TABLE 2 summarizes the data source.

TABLE 2
DATA SOURCE

Variables	Data source (1980-2018)
S&P composite prices and dividends	Robert J. Shiller Data
The 1-month T-bill returns	CRSP: Center for Research in Security Prices
Consumer Price Index (CPI)	U.S. Bureau of Labor Statistics
Aggregate income per capita	National Income and Product Accounts (NIPA)
Household Financial Obligations Ratio	Federal Reserve Bank of St. Louis

We also need to estimate the equity premium implied by the model in a calibrated endowment economy. Generally, the solutions to asset pricing equations (18) and (19) are not available in an analytically simple closed form. However, there are instances where calculation of the exact solution or a good approximation to that, is possible. In my model the binding borrowing constraints does the job and make it possible to solve for the closed form solutions. Indeed, solutions to functional equations (22) and (23) depends on values of three exogenous stochastic processes for y_t, x_t and θ_t. Note we have already assumed that R_t^d is constant and equal to $\overline{R^d}$. More specifically, I will use weighted average annual rate of commercial bank interest rate on credit cards and 1-Year adjustable rate mortgage, the average gross borrowing rate would be equal to 1.085. A typical problem is to characterize the price of an asset, where the law of motions for exogenous stochastic state variables are AR (1) processes. Now we need to check whether the stochastic processes for y and x are stationary over time and then estimate them with an AR(1) process. TABLE 3

summarized the results for stationarity tests for detrended aggregate income and divided growth using two different methods, namely Augmented Dicky-Fuller unit root test (ADF), and Kwiatkowski-Phillips-Schmidt Shin (KPSS) stationarity test.

As TABLE 3 suggests, the ADF test statistics for detrended income y and aggregate dividend growth x are equal to -5.343 and -5.042 respectively, meaning that we can reject the null hypothesis (Null: y_t and x_t exhibit unit root property). Also using KPSS test, we cannot reject the null hypothesis that the variables are stationary over time as the test statistics are smaller than the critical values. Thus, it is reasonable to estimate the stochastic processes for income and dividend growth with an autoregressive process AR(1). The last step to derive a numerical closed form solution and generating the equity risk premium implied by the model is to discretize AR(1) processes for income and divided growth. I will use (Rouwenhorst, 1995) technique to discretize the AR processes. In this paper, Rouwenhorst method is preferred to (Tauchen, 1986) approach because of the following reasons. First, the residuals of both AR(1) processes pass the ARCH heteroskedasticity test i.e. we can reject heteroskedasticity of residuals. Second, I am going to discretize the economy with a 2-state Markov process for each y and x. When the number of states is small (equal to 2 here), the Rouwenhorst technique outperforms the Tauchen approach. Note that the AR(1) stochastic processes for income and dividend growth are as follow:

$$EP_t \quad = E_t(R^e_{t+1}) - R^f_t \tag{25}$$

$$x_{t+1} = (1-\rho_x)\overline{x} + \rho_x x_t + \epsilon^x_{t+1} \tag{26}$$

TABLE 3
STATIONARITY TEST OR EXOGENOUS VARIABLES

	ADF test statistic		KPSS test statistic	
	y_t	x_t	y_t	x_t
	-5.34	-5.04	0.243	0.273
1% level	-3.64		0.74	
5% level	-2.96		0.46	
10% level	-2.62		0.35	

In my model, the stochastic process for financial obligations cap is fully determined by the process for aggregate income. TABLE 4 summarized the estimated parameters of AR (1) processes.

TABLE 4
AR(1) ESTIMATIONS

Parameter/Variable	**Description**	
y	Normalized labor income per capita (detrended)	1 , $\rho_y = 0.64$, $\epsilon_y = 0.08$
x	Dividend growth	$.032$, $\rho_x = 0.61$, $\epsilon_x = 0.05$
θ	Financial obligations ratio cap	$\bar{\theta} = 0.165, \quad \sigma_\theta = 0.009$

The economy is completely specified by realization of the joint stochastic process followed by the aggregate real income and dividend growth. I model the joint process of the aggregate income and the dividend growth as a time-stationary Markov chain with a nondegenerate, unique, stationary probability distribution. Starting with the real aggregate income, I assume that y_t follows a 2-states Markov chain (y, Q, π^Y) where y is the state vector, Q is a 2-by-2 transition matrix and π^Y is the probability distribution.

The two states are **High** and **Low** standing for high and low aggregate income during boom and recession respectively. So, I define the states vector y as:

$$y = (y^H, y^L) = (\mu^Y + \delta^Y, \mu^Y - \delta^Y) \tag{27}$$

where μ^Y is the long run aggregate income and δ^Y is its standard deviation. The probability distribution π^Y is defined as:

$$\pi_{ij}^R = Pr(y_{t+1} = y_j | y_t = y_i) \tag{28}$$

Then the transition matrix Q is symmetric with:

$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} = \begin{bmatrix} \pi_{11}^Y & 1 - \pi_{11}^Y \\ 1 - \pi_{22}^Y & \pi_{22}^Y \end{bmatrix} = \begin{bmatrix} \phi^Y & 1 - \phi^Y \\ 1 - \phi^Y & \phi^Y \end{bmatrix} \tag{29}$$

where ϕ^Y is the long run auto correlation coefficient of y. Next, I assume that dividend growth x_t also follows a 2-states Markov process (x, P^k, π^X) where x is the state vector, P^k is a 2-by-2 transition matrix where $k: H, L$ meaning that transition matrix for x depends on whether y is in high or low state. Also π^X is the probability distribution. Again the 2 states are **High** and **Low** standing for high and low dividend growth. I define the states vector x as:

$$x = (x^H, x^L) = (\mu^X + \delta^X, \mu^X - \delta^X) \tag{30}$$

where μ^X is the long run aggregate dividend growth and δ^X is its standard deviation. The probability distribution π^X is defined as:

$$\pi_{ij}^X = Pr(x_{t+1} = x_j | x_t = x_i) \tag{31}$$

Then the transition matrix P^k is symmetric with:

$$P^k = \begin{bmatrix} \pi_{11}^X & 1 - \pi_{11}^X \\ 1 - \pi_{22}^X & \pi_{22}^X \end{bmatrix} = \begin{bmatrix} \phi_k^X & 1 - \phi_k^X \\ 1 - \phi_k^X & \phi_k^X \end{bmatrix} \tag{32}$$

where ϕ_k^X is the long run auto correlation coefficient of x conditional on whether y is in it's high $(k: y_H)$ or low $(k: y_L)$ states. So, the exogenous joint stochastic processes for y and x follows a 4-state coupled Markov chain with a 1-by-4 probability distribution π and a 4-by-2 state matrix yx:

$$yx = \begin{bmatrix} yx_{HH} \\ yx_{HH} \\ yx_{HH} \\ yx_{HH} \end{bmatrix} = \begin{bmatrix} y_H & x_H \\ y_H & x_L \\ y_L & x_H \\ y_L & x_L \end{bmatrix} \tag{33}$$

and a 4-by-4 transition matrix S_{YX} :

$$S_{YX} = \begin{array}{c} \\ HH \\ HL \\ LH \\ LL \end{array} \begin{array}{c} \begin{array}{cccc} HH & HL & LH & LL \end{array} \\ \begin{pmatrix} s_{11} & s_{12} & s_{13} & s_{14} \\ s_{21} & s_{22} & s_{23} & s_{24} \\ s_{31} & s_{32} & s_{33} & s_{34} \\ s_{41} & s_{42} & s_{43} & s_{44} \end{pmatrix} \end{array} \tag{34}$$

where s_{ij} can be calculated directly from transition matrices Q and P. For example, this is how we derive $s_{12} = S_{YX_{HL,HH}}$,

$$s_{12} = S_{yx_{HL,HH}} = \mathbf{Pr}(yx_{t+1} = yx_{HL} | yx_t = yx_{HH}) = \mathbf{Pr}\left(\begin{matrix} y_{t+1} = y_H \\ x_{t+1} = x_L \end{matrix} \middle| \begin{matrix} y_t = y_H \\ x_t = x_H \end{matrix}\right) \tag{35}$$

$$s_{12} = \Pr(x_{t+1} = x_L, | x_t = x_H, y_t = y_H) * \Pr(y_{t+1} = y_H, | y_t = y_H) = P_{12}^H * q_{11} \tag{36}$$

where P_{12}^H is the 1, 2 elements of the matrix P^H. Similar reasoning leads to the following transition matrix for S_{YX}.

$$S_{YX} = \left(\begin{array}{cc|cc} P_{11}^H q_{11} & P_{12}^H q_{11} & P_{11}^L q_{12} & P_{12}^L q_{12} \\ P_{21}^H q_{11} & P_{22}^H q_{11} & P_{21}^L q_{12} & P_{22}^L q_{12} \\ \hline P_{11}^H q_{21} & P_{12}^H q_{21} & P_{11}^L q_{22} & P_{12}^L q_{22} \\ P_{21}^H q_{21} & P_{22}^H q_{21} & P_{21}^L q_{22} & P_{22}^L q_{22} \end{array}\right) = \begin{pmatrix} P^H q_{11} & P^L q_{12} \\ P^H q_{21} & P^L q_{22} \end{pmatrix} \tag{37}$$

It is easy to confirm that S_{YX} is a bona-fide transition matrix. In summary, to solve the model we need to estimate the parameters in TABLE 5.

TABLE 5
ESTIMATED PARAMETERS AND VARIABLES

Item	Description
β	Constant discount factor - time preference
γ	Coefficient of relative risk aversion
θ	Exogenous predetermined upper bound for FOR
μ^Y	Long run average annual aggregate income (detrended) *y*
δ^Y	Standard deviation of annual *y*
ϕ^Y	First order autocorrelation coefficients of annual *y*
μ^X	Long run average aggregate dividend growth *x*
δ^X	Standard deviation of annual *x*
ϕ_k^X	First order autocorrelation coefficients of annual *x* for *k: H, L.*

The acceptable value for β based on different macroeconomic models ranges from 0.95 to 0.99. I set $\beta = 0.98$ and let the coefficient of relative risk aversion γ vary from 1 to 3. In the result section of the paper I show that -unlike in the standard consumption-based model- our model is able to get close to the

equity risk premium observed in U.S data even for these small values of risk aversion. The rest of the parameters and variables are estimated as illustrated in TABLE 6.

TABLE 6
ESTIMATED VALUES FOR THE MARKOV PROCESSES

Parameter	Estimated value
μ^Y	1.00
δ^Y	0.06
ϕ^Y	$\frac{1+\rho_y}{2} = \frac{1+0.64}{2} = 0.82$
μ^X	1.04
δ^X	0.07
ϕ_k^X	$\frac{1+\rho_x}{2} = \frac{1+0.61}{2} = 0.805$

RESULTS

In this section, I show that the model is able to get close to equity risk premium observed in U.S data while keeping the risk-free rate low. TABLE 7 summarizes the results for different specifications of the model. As TABLE 7 suggests, the model presented in this paper completely outperform the standard consumption model in explaining the equity risk premium observed in the data. The equity premium puzzle states that the mean excess return calculated in the standard consumption-based model is too low unless the coefficient of relative risk aversion is implausibly high. For the standard model this number is 20 which makes no economic sense. However, in this paper, the model generates an equity premium of 4.62 % while the utility curvature is only 3. This is consistent with the intuition of our model. In bad times, as consumption gets closer to household's financial obligation, people become more risk averse (as they have to pay back their obligations) and take on less risk. This leads to less investment in the risky asset and higher equity risk premium eventually. Another advantage is that unlike the standard model, this model keeps the risk-free rate relatively stable and low. For the range of relative risk aversion between 1 and 3, the risk-free rate induced by the model varies between 1.87%. and 0.56% This is true because in our setup, as the coefficient of relative risk aversion increases, the precautionary savings dominate the intertemporal substitution effect faster than the standard model and generate a lower interest rate.

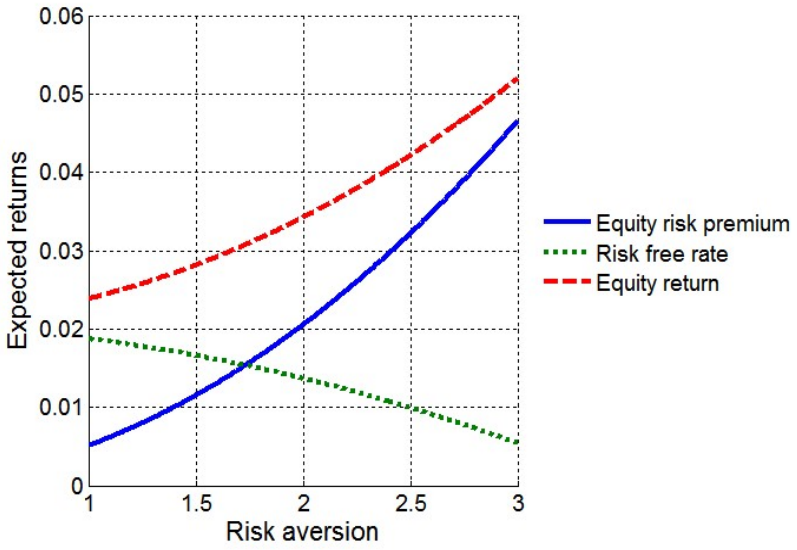
Expected Returns and the Utility Curvature γ

FIGURE 2 shows how the equity premium, risk free rate and equity returns vary with the utility curvature γ. As the γ increases, the agents become more risk averse to any bet. In this model, the precautionary savings plays an important role because households are afraid of bad times in which the financial obligations ratio is high. Since the households are restricted by the borrowing constraint and they cannot leverage more due to the already-capped financial obligations ratio, they will demand more of precautionary savings (treasury bond investments) to smooth their consumption for bad states of the economy.

This higher demand for risk free bonds will pull up the bond price q_t and thus decrease the risk-free rate. This is the dotted line showing the risk-free rates for different values of risk aversion between 1 and 3. Having invested more in treasury bonds, the demand for equity investment will decrease and the expected equity returns will go up for the higher values of γ. The dashed line shows the equity returns for different values of risk aversion. Hence, by generating higher equity returns and lower risk-free rates, the model is

able to get close to the equity risk premium observed in the data. The solid line in FIGURE 2 depicts the equity risk premium generated by the model versus the different values for utility curvature.

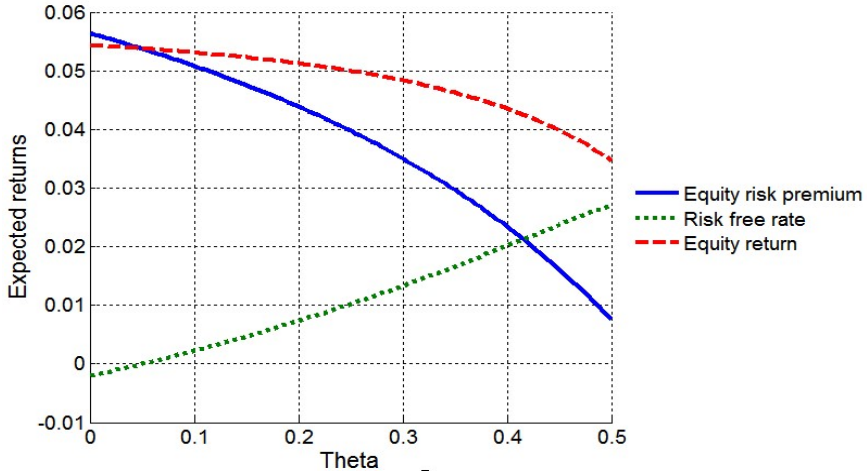
FIGURE 2
EXPECTED RETURNS AND THE COEFFICIENT OF RISK AVERSION γ



Expected Returns and the Borrowing Constraint

FIGURE 3 shows how the equity premium, the risk-free rate and equity returns vary with the household's obligation ratio θ. Recall that financial obligations ratio is a direct indicator of the borrowing constraint in our model, so relaxing the borrowing constraint is equivalent to increasing the household financial obligations ratio cap θ. FIGURE 3 indicates that as we relax the borrowing constraint, the equity risk premium shrinks which is numerically consistent with results in (Constantinides, Donaldson and Mehra, 2002).

FIGURE 3
EXPECTED RETURNS AND THE HOUSEHOLD'S OBLIGATION RATIO θ



The intuition is straight forward. According to the financial obligation's ratio constraint (the borrowing constraint) in (2), households choose a debt service level that cap their financial obligations ratio at any time. Hence as this θ increases, agents can smooth their consumption much easier via larger borrowings. This leads to a decrease in the demand for precautionary savings (treasury bond investments), driving down bond prices q_t and increasing the risk-free rate. This is the dotted line in FIGURE 3. From the other hand, as θ increases, the households have more funding resources for their consumption purposes and they can make more investments. The level of equity investment will increase (as more borrowings are to be consumed and the investment portfolio is more toward equity investment rather than bond investment) leading to lower equity returns. The dashed line shows the equity returns for different range of values for θ. These two effects will decrease the equity premium as the level of θ increases, The solid line in FIGURE 3.

Linearizing the Pricing Kernel

In this section, I explore how the model proposed in section 3 is able to generate a more volatile kernel while keeping the risk-free interest rate low. The linearized version of the stochastic discount factor is:

$$M_{t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} = \beta\exp\left\{(-\gamma)\left(\ln\frac{C_{t+1}}{C_t}\right)\right\}\exp\left\{(-\gamma)\left(\ln\frac{S_{t+1}}{S_t}\right)\right\}, \quad (38)$$

where C_t is aggregate consumption and S_t is the surplus consumption ratio defined as $S_t = \frac{C_t - G_t}{C_t}$. G_t is aggregate financial obligations in period t. Taking the logarithm from both sides and letting $c_t = \ln(C_t)$ and $s_t = \ln(S_t)$ result in the following:

$$\ln M_{t+1} = \ln\beta + (-\gamma)(c_{t+1} - c_t) + (-\gamma)(s_{t+1} - s_t) \quad (39)$$

Now, some simplifying assumptions need to be made. For aggregate U.S. data on per capita consumption of nondurables and services, a good approximation to the data is the following model that makes the growth in the log of per capita consumption a random walk with drift:

$$c_t = \mu_c + c_{t-1} + \sigma_c\epsilon_t, \quad \text{where} \quad \epsilon_t \quad i.i.d. \sim N(0,1) \quad (40)$$

Assuming that the growth in the surplus consumption ratio also follows a random walk,

$$s_t = \mu_d + s_{t-1} + \sigma_d\epsilon_t, \quad \text{where} \quad \epsilon_t \quad i.i.d. \sim N(0,1) \quad (41)$$

Note that μ_d and σ_d are the drift term and standard deviation term of the random walk process for the surplus consumption ratio. Now using (39) and (40) in (41) result in the following:

$$\ln M_{t+1} = \ln\beta + (-\gamma)(\mu_c + \sigma_c\epsilon_t) + (-\gamma)(\mu_d + \sigma_d\epsilon_t) \quad (42)$$

Because ϵ_t is $i.i.d. \sim N(0,1)$, then $\ln M_{t+1}$ is also normally distributed with mean μ and variance σ^2:

$$\mu = \ln\beta + (-\gamma)\mu_c + (-\gamma)\mu_d \quad (43)$$

$$\sigma^2 = (-\gamma)^2\sigma_c^2 + (-\gamma)^2\sigma_d^2 \quad (44)$$

Now consider the following property of normal distribution: If $\log X \sim N(\mu_x, \sigma_x^2)$, then $E(X) = \exp(\mu_x + \frac{\sigma_x^2}{2})$ and $std(X) = E(m)\sqrt{\exp(\sigma^2) - 1}$. From this property $E(M_{t+1})$ and $\sigma(M_{t+1})$ can be derived as follows:

Now $E(M) = \exp(\mu + \frac{\sigma^2}{2})$ (45)

$$\sigma(M) = E(M)\sqrt{\exp(\sigma^2) - 1} \quad (46)$$

Having $E(M)$ and $\sigma(M)$ in hand, I can now derive an equation for gross risk-free rate R_F and explain the intuition behind:

$$R_f = E(M)^{-1} \Rightarrow \ln(R_f) = \ln(1 + r_f) = -\ln E(M) = -\ln(\exp(\mu + \frac{\sigma^2}{2})) = -\mu - \sigma^2/2 \quad (47)$$

Using equations (45) and (46), the approximate risk-free rate is:

$$r_f \approx -\ln\beta + \gamma\mu_c + \gamma\mu_d - \gamma^2 \frac{\sigma_c^2}{2} - \gamma^2 \frac{\sigma_d^2}{2} \quad (48)$$

Equation (48) has some important implications. There are five terms in this equation that according to the set-up of my model can be interpreted as follows:

1. $-\ln\beta$: As β, the time discount factor decreases, agents become less patient and require higher interest rates to substitute consumption over time. For example, if β is calibrated to 0.99, this means that approximately 1% of the risk-free rate is due to time preferences.
2. $\gamma\ \mu_c$: For $\gamma > 0$, this implies that as consumption growth increases, individuals should be compensated with higher interest rates to sacrifice today's consumption for tomorrow's consumption.
3. $\gamma\ \mu_d$: For $\gamma > 0$, this implies that in recessions, when consumption gets close to financial obligations, the surplus consumption ratio decreases and investors require higher interest rates.
4. $-\gamma^2\ \frac{\sigma_c^2}{2}$: Analogous to standard consumption-based models, this part of equation (B.9) can be interpreted as *precautionary savings*. The coefficient of consumption growth volatility is negative, implying that as consumption growth becomes more volatile, precautionary savings push the interest rate down.
5. $-\gamma^2\ \frac{\sigma_d^2}{2}$: This term adds up to the precautionary savings part of equation $(B.9)$ due to economic uncertainties. As the volatility of the surplus consumption ratio increases, demand for safer assets increases which leads to lower interest rates. This is what enables my model — unlike the standard consumption-based model — to generate lower risk-free rates for higher coefficients of risk aversion.

CONCLUSION

In an infinitely lived representative agent endowment economy, this paper addresses how the household financial obligations impact the equity risk premium. The impact is studied under two different channels, the preference channel and the borrowing constraint channel. Financial obligations ratio is a counter-cyclical indicating variable that affects agents' marginal utility of consumption and reinforces its counter-cyclicality over business cycles. This is the driving force of the model. At the equilibrium, by specifying an explicit Markov process for consumption growth, in a non-stationary environment, I derive the expected returns on equity and the risk-free bond and calculate the equity risk premium. I show that in a reasonably calibrated economy, the model is able to generate the observed equity premium in U.S. data while keeping the risk-free rate low.

REFERENCES

Abel, A.B. (1990). Asset prices under habit formation and catching up with the joneses. *American Economic Review*, *80*, 38–42.

Barro, R. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics.*

Brown, T.K., & Price, S.W. (2005). Debt and distress: evaluating the psychological cost of credit. *Journal of Economic Psychology*, *42*, 642–663.

Campbell, J.Y., & Cochrane, J.H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, *107*, 205–251.

Chetty, R., & Szeidl, A. (2007). Consumption commitments and risk preferences. *The Quarterly Journal of Economics*, *122*(2), 831–877.

Cochrane, J.H. (2005). *Financial markets and the real economy*. Working Paper 11193, National Bureau of Economic Research.

Cochrane, J.H. (2016). *Macro-finance*.

Constantinides, G.M., & Duffie, D. (1996). Asset pricing with heterogeneous consumers. *Journal of Political Economy*, *104*, 219–240.

Constantinides G. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy*, *98*, 519–543.

Donaldson Constantinides, G.M., & Mehra, R. (2002). Junior can't borrow: A new perspective on the equity premium puzzle. *Quarterly Journal of Economics*, *118*, 269–296.

Dickey, D.A., & Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*, 427–431.

Dynan, K. (2012). Is a household debt overhang holding back consumption? *Brookings Papers on Economic Activity*.

Dynan, K., & Edelberg, W. (2013). The relationship between leverage and household spending behavior: Evidence from the 2007-2009 survey of consumer finances. *Federal Reserve Bank of St. Louis Review*, *95*, 425–448.

Epstein, L.G., & Zin, S.E. (1991). Substitution, risk aversion and the temporal behavior of asset returns. *Journal of Political Economy*, *99*, 263–286.

Jahangiry, P. (2019). *Households' financial obligations and the cross-section of stock returns: FCAPM*. Working paper.

Jahangiry, P. (2020). *Predictability of stock returns with household's obligation ratio*. Working paper.

Johnson, K.W., & Li, G. (2007). Do high debt payments hinder household consumption smoothing? *Finance and Economics Discussion Series (FEDS)*, *52*.

Johnson, K.W., & Li, G. (2010). The debt-payment-to-income ratio as an indicator of borrowing constraints: Evidence from two household surveys. *Journal of Money, Credit and Banking*, *42*, 1373–1390.

Juselius, M., & Drehmann, M. (2012). Do debt service costs affect macroeconomic and financial stability? *BIS Quarterly Review*, pp. 21–34.

Juselius, M., & Drehmann, M. (2015). Leverage dynamics and the real burden of debt. *BIS Quarterly Review*.

Lucas, R. (1978). Asset Prices in an Exchange Economy. *Econometrica*, *46*(6),1429–45.

Mehra, R., & Prescott, E.C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, *15*, 145–161.

Piazzesi, S., & Tuzel. (2007). Housing, consumption and asset pricing. *Journal of Financial Economics*, *83*, 581–612.

Richardson, E.P., & Roberts, R (2013). The relationship between personal unsecured debt and mental and physical health: A systematic review and meta-analysis. *Clinical Psychology Review*, *33*, 1148–1162.

Rytchkov, O. (2014). Asset Pricing with Dynamic Margin Constraints. *Journal of Finance*, *69*(1).

Sahalia, P., & Yogo. (2004). Luxury goods and the equity premium. *Journal of Finance*, *59*, 2959–3004.

APPENDIX

From Federal Reserve's Board: The household Debt Service Ratio (DSR) is the ratio of total required household debt payments to total disposable income. The DSR is divided into two parts: Mortgage DSR and Consumer DSR. The Mortgage DSR is total quarterly required mortgage payments divided by total quarterly disposable personal income. The Consumer DSR is total quarterly scheduled consumer debt payments divided by total quarterly disposable personal income. The Mortgage DSR and the Consumer DSR sum to the DSR. Quarterly values for the Debt Service Ratio are available from 1980 forward.

The limitations of current sources of data make the calculation of the ratio especially difficult. The ideal data set for such a calculation would have the required payments on every loan held by every household in the United States. Such a data set is not available, and thus the calculated series is only an approximation of the debt service ratio faced by households. Nonetheless, this approximation is useful to the extent that, by using the same method and data series over time, it generates a time series that captures the important changes in the household debt service burden. The series are revised as better data or improved methods of estimation become available. To create the measure, payments are calculated separately for revolving debt and for each type of closed-end debt, and the sum of these payments is divided by disposable personal income as reported in the National Income and Product Accounts. For revolving debt, the assumed required minimum payment is 2-1/2 percent of the balance per month. This estimate is based on the January 1999 Senior Loan Officer Opinion Survey, in which most banks indicated that required monthly minimum payments on credit cards ranged between 2 percent and 3 percent, a ratio that apparently had not changed substantially over the previous decade.

Payments on closed-end loans, which are calculated for each major category of closed-end loan, are derived from the loan amount outstanding, the average interest rate, and the average remaining maturity on the stock of outstanding debt. Estimates of the amount of mortgage debt are taken from the Federal Reserve Board's Z.1 Financial Accounts of the United States statistical release, and estimates of outstanding consumer debt are taken from the Federal Reserve's G.19 Consumer Credit statistical release. For consumer debt, a more detailed breakdown by type of closed-end loan is obtained using internal Federal Reserve estimates and data from the Federal Reserve's Survey of Consumer Finances (SCF). Interest rates on closed-end consumer loans are obtained from the Federal Reserve Board's G.19 Consumer Credit and G.20 Finance Companies statistical releases, the SCF, and additional proprietary data sources. An estimate of the interest rate on the stock of outstanding debt is obtained by weighting the recent history of interest rates using information on the age of outstanding loans in the SCF. The interest rate on the stock of outstanding mortgage debt is an estimate provided by the Bureau of Economic Analysis. Maturity series for consumer debt are taken from the SCF. Maturity series for mortgage debt are calculated using data from Lender Processing Services and Mortgage Bankers Association.

The **financial obligations Ratio** is a broader measure than the Debt Service Ratio. It includes rent payments on tenant-occupied property, auto lease payments, homeowners' insurance, and property tax payments. These statistics are obtained from the National Income and Product Accounts.

Multipliers of Financial Literacy in Germany and Russia?

Viktoriia Dembinskaite
Saint-Petersburg State University of Economics

Eduard Gaar
Technical University of Darmstadt

Tatjana Nikitina
Saint-Petersburg State University of Economics

Dirk Schiereck
Technical University of Darmstadt

In Germany and Russia, the majority of the population shows significant deficits in financial literacy. To overcome this situation both countries need people who are able to share and teach their knowledge. Economic students could be such multipliers. Therefore, it is important to understand the qualification of these students as ambassadors of financial literacy.

We examine the financial competence of German and Russian economic students. Based on a survey with international-comparable questions we find that most students perform indeed better than the international average. But only few students are able to answer all questions correctly with gender and age specific divergences. Women perform better at numeracy but worse at financial questions and younger students perform worse than older. Finally, we show that students with better understanding and education in finance are more likely to assign a probability to future crisis and that Russian students are more pessimistic than German students.

Keywords: financial literacy, financial competence, economic students, retirement planning, crisis expectation

INTRODUCTION

Many studies provide evidence that a large proportion of the population knows little about finance and that many individuals are unfamiliar with basic economic concepts, such as risk diversification in stock investing, inflation, and interest compounding (OECD, 2006, 2008). This lack of knowledge is one reason why individuals avoid dealing with topics like personal retirement saving or equity investing. Emerging a retirement saving system from a more or less complete government provision to more private individual investing enforces on individuals the responsibility to save, invest and consume reasonable over a lifetime

cycle (Lusardi & Mitchell, 2011b) and depends on individuals possessing the skills needed to manage their financial responsibilities well. Research from Lusardi and Mitchell (2017) and van Rooij et al. (2011a) into retirement savings behaviour shows that financial skills, retirement planning and retirement income are strongly related and that retirement planning is a strong predictor of wealth. Higher levels of financial knowledge are associated with increased stock market participation (Yoong, 2010; van Rooij et al., 2011b), higher private retirement savings (Bucher-Koenen, 2009), greater portfolio diversification (Guiso & Jappelli, 2008) and increased wealth holdings (Lusardi & Mitchell, 2007; van Rooij et al., 2012; Behrman et al., 2012). Further, Gerardi et al. (2010) show that mortgage delinquency rates are higher among borrowers with poor numerical ability (using the same measure of numeracy we use here) and Banks and Oldfield (2007) link poor numeracy with low savings.

Based on the relationship between financial literacy and retirement savings it is hardly surprising that there is overall a low level of financial knowledge in countries like Germany and Russia. Bucher-Koenen and Lusardi (2011) document this low level of financial literacy in Germany, and there is an ongoing debate how to overcome this undesired state of knowledge. Given the large number of less informed, often older people politicians are looking for multipliers who share their superior knowledge and teach financial literacy. Students of economics could be such multipliers. Consequently, we analyse whether a university education in economics leads to superior investment knowledge to enable people for an effective retirement planning. To address this question, we conduct a comparable survey (Lusardi & Mitchell, 2017; Gerardi et al., 2010) for assessing numeracy, basic and sophisticated financial literacy at top universities in Germany and Russia respectively. This survey design facilitates to compare our results directly with outcomes from earlier studies in other countries. The findings show a good general financial knowledge but still difficulties when it comes to more sophisticated questions and answering all questions correctly.

In search for a simple proxy to forecast who might be especially qualified to communicate financial topics we analyse the link between mathematical and financial knowledge. The known good mathematical education of Russian students indicates good numeracy results. Our findings confirm these math skills but do not show a link to financial knowledge. Actually a country comparison shows worse results for Russian economic students than for general populations.

The rest of the study is organized as follows. In section 2 we introduce the theoretical framework for financial literacy and present patterns often analysed in the context of financial literacy. Section 3 provides information on data and methodology, before section 4 shows the results. Section 5 summarizes the findings and concludes.

THEORETICAL FRAMEWORK FOR FINANCIAL LITERACY

Financial planning of household finances is an important type of non-market production that requires its own form of human capital, especially financial literacy. Like human capital, financial literacy accumulation is purposive based on its costs and benefits. Public and scholarly interest in financial literacy and informed financial decision-making is increasing in part because of the poor financial outcomes that are associated with low levels of financial literacy: problems with debt (Lusardi & Tufano, 2009b) and lack of retirement planning (Lusardi & Mitchell, 2007, 2017), among others. Especially the connection between financial literacy and retirement planning is of particular importance. Even after accounting for a large set of economic characteristics and circumstances the results that those who are more financially literate are more likely to plan for retirement are striking consistent throughout international studies (Lusardi & Mitchell, 2011b). Given that there is no evidence that people invest much in financial knowledge, partially because of a lack of adult education programs in several countries, a few papers have begun to examine the decision to acquire financial literacy and to study the links between financial knowledge, saving and investment behaviour, especially incentives to invest, (Delavande, et al., 2008; Jappelli & Padula, 2013; Hsu, 2016; Lusardi et al., 2013).

The notion that financial knowledge is a form of human capital was introduced in Delavande et al. (2008), which related the production of human capital to portfolio choice. Usually studies analysing financial literacy harness a two period approach of saving and portfolio allocation across different assets

and allow the acquisition of human capital in form of financial knowledge. The results suggest that individuals acquire knowledge mostly when it becomes relevant (Hsu, 2016), that wealth and financial literacy are strongly correlated (Banks & Oldfield, 2007; van Rooij et al., 2007) and that in countries with generous social security benefits the incentives to save and accumulate wealth and to invest in financial literacy are smaller (Jappelli & Padula, 2013).

Several studies show that these patterns are consistent through different countries or stages of economic development and that financial illiteracy is widespread even when financial markets are well developed as in the U.S., Germany, the Netherlands, Sweden, Italy, Japan, and New Zealand (Lusardi & Mitchell, 2011b). However, there are differences across countries. For example, where people score high on math and science tests, they also tend to score high on questions measuring numeracy (e.g. the Programme for International Student Assessment; (OECD, 2005)). Furthermore, people are more knowledgeable about inflation if their country has experienced it recently. For example, Italians are more likely to answer the question on inflation correctly. Conversely, in a country like Japan that experienced deflation, fewer people answer the inflation question correctly. If a country like Sweden experienced a pension privatization in recent times the people are more aware of risk diversification. Whereas, Russians and people born in East Germany know less about it (Lusardi & Mitchell, 2011).

More studies concerned with geographic and population related differences within countries show large racial and ethnic differences in the U.S. For example, Whites and Asians are consistently more likely to be financially knowledgeable compared to African Americans and Hispanics (Lusardi & Mitchell, 2007a, 2007b, 2011b). In Italy the Northern and Central regions score higher than the Southern regions, though not all of the Northern regions show high levels of financial knowledge (Fornero & Monticone, 2011). In Russia and Romania people living in urban areas tend to be more financially literate than people living in rural areas (Klapper et al., 2013; Beckmann, 2013). And in the Netherlands other religious beliefs seem to have an influence on financial knowledge meaning that another than the main religion (including Muslims and other smaller religious groups) are less financial literate (Lusardi & Mitchell, 2011b). At the same time, studies find that U.S. citizens tend to display low levels of financial literacy (Bernheim, 1998; Hilgert et al., 2003; Lusardi & Tufano, 2009). Financial illiteracy and financial mistakes are particularly widespread among older Americans (Lusardi & Mitchell, 2011b; Agarwal et al., 2009). Recent government policies, including the establishment of the Consumer Financial Protection Bureau, aim to increase financial literacy among the public (Hsu, 2016).

When reviewing the available literature certain factors are particularly consistent over countries. For this reason, we deal specifically with them in the following.

Gender

One striking feature of the empirical data on financial literacy is the large and persistent gender difference. Not only are older men generally more financially literate than older women, but similar patterns also show up among younger respondents as well (Lusardi et al., 2010; Lusardi & Mitchell, 2017; Lusardi & Tufano, 2009, 2015). Moreover, these gaps persist across both the basic and the more sophisticated literacy questions (Lusardi et al.,2010; Hung et al., 2009). One twist on the differences by sex, however, is that while women are less likely to answer financial literacy questions correctly than men, they are also far more likely to say they "do not know" as an answer to a question, a result that is strikingly consistent across countries. This awareness of their own lack of knowledge may make women ideal targets for financial education programs. Lusardi and Mitchell (2008) show for example that only 61.9% of all women answer the interest rate question correct, whereas 70.6% answer the inflation and only 47.6% of all woman are able to answer the risk diversification question correctly. With the exceptions of Russia and East Germany where women and men are equally illiterate (Lusardi & Mitchell, 2011b) these sex differences in financial literacy are so persistent and widespread across surveys and countries, several researchers seek to explain them. Hsu (2016) proposes that some sex differences may be rational, with specialization of labour within the household leading married women to build up financial knowledge only late in life (close to widowhood). Chen and Volpe (2002) and Mandell (2008) show further that even women in high school and college are usually less financial literate and confirm the observations that single woman in charge of their own finances

have lower finance knowledge. Fonseca et al. (2012) suggest that women may acquire financial literacy differently from men, while Bucher-Koenen et al. (2017) point to a potentially important role of self-confidence that differs by sex. Brown and Graf (2013) also show that sex differences are not due to different interests in finance and financial matters between women and men. To shed more light on women's financial literacy, Mahdavi and Horton (2014) examined alumnae from a highly selective U.S. women's liberal arts college. Even in this talented and well-educated group, women's financial literacy was found to be very low. In other words, even very well educated women are not particularly financially literate, which could confirm Fonseca et al. (2012) that women may acquire financial literacy differently from men. Nevertheless, this debate is far from closed, and additional research will be required to better understand these observed differences.

Age

A study by Agarwal et al. (2009) which focused on financial mistakes shows that these are most predominant among the young and old subsample, groups which usually display the lowest financial knowledge. Age patterns are notable, in that financial knowledge follows an inverted U-shaped pattern, being lowest for the young and the older groups, but peaks in the middle of the life cycle. Lusardi and Mitchell (2011a) show for US-Americans over 50 years a very low performance on basic financial literacy questions with 75% correct answers for the numeracy, 56% correct on the inflation and only 52% correct answers for the risk diversification question which are used to test gender differences and is also used in our survey. At the same time Lusardi and Mitchell (2017) show in another study how very young respondents score worse than middle age people, 60% correct versus 69% for older respondents on the risk diversification question. This is a finding which is robust across countries (Lusardi & Mitchell, 2014). Additionally, it is of interest that older people give themselves very high scores regarding their own financial literacy, despite scoring poorly on the basic financial literacy questions (Lusardi & Mitchell, 2011a; Lusardi & Tufano, 2015) and not just in the United States, but other countries as well (Lusardi & Mitchell, 2011c). Similarly, Finke et al. (2016) develop a multidimensional measure of financial literacy for the old and confirm that, though actual financial literacy falls with age, people's confidence in their own financial decision-making abilities actually increases with age. The mismatch between actual and perceived knowledge might explain why financial scams are often perpetrated against the elderly (Deevy et al., 2012).

Education

In all countries, higher educational attainment is strongly correlated with financial knowledge, but even at the highest level of schooling, financial literacy tends to be low (Lusardi & Mitchell, 2008). Moreover, education is not a good proxy for financial literacy. That is, when education and financial literacy are included in multivariate regression models, both tend to be statistically significant, indicating that financial literacy has an effect above and beyond education. Financial literacy is also higher among those who are working, and in some countries among the self-employed, compared to those who do not work. This difference may in part result from financial education programs offered in the workplace (as in the United States); it could also be the effect of learning from colleagues or skills acquired on the job (Lusardi & Mitchell, 2011a). Furthermore, they show how the percentage of correct answers for the compound interest (inflation, risk diversification) question rises with better education up to 81.2% (85.1%, 70.2%) correct answers for a level of education corresponding to "College and More". Christiansen et al. (2008) use a large register-based panel data set containing detailed information on Danish investors' education attainment, and financial and socioeconomic variables. The authors show that stock-holding increases if individuals have completed an economics education program and if an economist becomes part of the household. To sort out the double causality between portfolio choice and the decision to become an economist, Christiansen et al. (2008) use better access to education due to the establishment of a new university, as an instrument for economics

There are substantial differences in financial knowledge by education: specifically, those without a college education are much less likely to be literate about basic financial literacy concepts, as reported in several U.S. surveys and across countries (Lusardi & Mitchell, 2007a, 2011c). Moreover, numeracy is

especially poor for those with low educational attainment (Christelis et al., 2010; Lusardi, 2012). How to interpret the finding of a positive link between education and financial literacy has been subject to some debate in the economics literature. One possibility is that the positive correlation might be driven by cognitive ability (McArdle et al., 2009), implying that one must control on measures of ability when seeking to parse out the separate impact of financial literacy. Lusardi, Mitchell, and Curto (2010) did find a positive correlation between financial literacy and cognitive ability among young NLSY respondents, but they also showed that cognitive factors did not fully account for the variance in financial literacy. That means that substantial heterogeneity in financial literacy remains even after controlling on cognitive factor.

Understanding

Not only are there patterns in measured financial literacy, but we also can compare what people actually know with their self-assessed understanding of finance. So for example that women tend to indicate a higher self-rated understanding when approaching widowhood (Hsu, 2016). Across countries younger people know very little and acknowledge it. By contrast, older people consistently rate themselves as very well-informed although they are actually less literate than average. There are also important international differences in self-reports: in the U.S. a majority of respondents give themselves high scores, whereas in Japan people rate themselves quite low (Lusardi & Mitchell, 2011b).

Other Patterns

There are numerous other empirical regularities in the financial literacy literature that are, again, persistent across countries. Financial savvy varies by income and employment type, with lower-paid individuals doing less well and employees and the self-employed doing better than the unemployed (Lusardi & Tufano, 2015; Lusardi & Mitchell, 2011c). These findings hold across age groups and many different financial literacy measures (Lusardi & Mitchell, 2017). The literature also points to differences in financial literacy by family background. For instance, Lusardi, Mitchell, and Curto (2010) link financial literacy of 23-28-year old NLSY respondents to characteristics of the households in which they grow up controlling for a set of demographic and economic characteristics. Respondents' financial literacy is also significantly positively correlated with parental education (in particular, that of their mothers), and whether their parents hold stocks or retirement accounts when the respondents are teenagers. Mahdavi and Horton (2014) report a connection between financial literacy and parental background; in this case, fathers' education is positively associated with their female children's financial literacy. In other words, financial literacy may well get its start in the family, perhaps when children observe their parents saving and investing habits, or more directly by receiving financial education from parents (Chiteji & Stafford, 1999; Li 2014; Shim et al., 2009). Other studies note a nationality gap in financial literacy, with foreign citizens reporting lower financial literacy than the native born (Brown & Graf, 2013). Or differences in financial literacy according to religion (Alessie et al., 2011) and political opinions (Arrondel et al., 2013).

To summarize, while financial illiteracy is widespread, it is also concentrated among specific population subgroups and demographic and self-assessed characteristics in most countries studied to date.

DATA AND SURVEY DESIGN

Survey Design

We interview 63 students of an economics master course in Germany at the Technical University of Darmstadt and a mix of 59 students of an economics bachelors' and masters' degree course in St. Petersburg in Russia at the UNECON. Our survey aims to determine the link between an education at university level with demographic and self-assessed characteristics, numeracy and financial skills, the tendency to think about the future and the differences between Russia and Germany.

First the demographic and self-assessed part contains questions about the gender and age of the students. The next part asked about the self-assessment of the students regarding their previous education in finance on a four step scale from "Hardly at all" to "A lot" and their understanding in the finance field on a scale from 1 (very low) to 7 (very high).

The numeracy and financial skill questions are divided in three parts: a numeracy test, a basic and a sophisticated financial literacy test. The numeracy test contains five questions about simple mathematical tasks regarding multiplying, dividing, percentages and fractions. The basic literacy part contains five questions as well and asks for knowledge about numeracy, compound interest, inflation, time value of money and money illusion. The eight sophisticated literacy questions are on the function of stock markets, the knowledge of mutual funds, the relation between interest rates and bond prices, the safety of company stocks and mutual funds, about risky assets, about long period returns, about volatility and about risk diversification.

The last two questions are concerned with the tendency to predict the future. We ask how likely the respondents estimate another crisis in 5 and in 25 years.

Data

In total 122 students answer the survey. Divided per country the sample consists of 65 German and 59 Russian respondents with at least a partly background in economics. The sex is unequally distributed between the subsamples.

TABLE 1
SURVEY SAMPLE DEMOGRAPHIC AND SELF-ASSESSMENT CHARACTERISTICS

Survey respondent population							
	Germany	Russia	Total		Germany	Russia	Total
Gender				*Understanding of finance?*			
Female	15	31	46	1 (very low)	0	2	2
Male	47	28	75	2	7	2	9
Prefer not to answer	1	0	1	3	14	4	18
				4	21	12	33
				5	17	16	33
Age				6	5	9	14
17	0	2	2	7 (very high)	0	10	10
18	0	8	8				
19	0	8	8	*Education devoted to financial education?*			
20	0	10	10				
21	1	7	8	A lot	8	16	24
22	15	5	20	Some	38	29	67
23	14	13	27	A little	18	10	28
24	15	3	18	Hardly at all	0	4	4
25	11	0	11				
26	4	1	5				
27	1	1	2				
28	1	0	1				
30	0	1	1				
32	1	0	1				

Notes: Table shows responses to survey questions. Missing statements are listed in the table.

Whereas Russians are mostly female the German group consists mainly of male students. Because we ask students the distribution of age is quite young with the oldest student from Germany with 32 years and the youngest student from the Russian subsample with 17 years. Regarding the self-assessed understanding

of finance, the Russian group seems to rate itself better than the Germans whereas the German group avoids extreme statements like a very low and very high understanding. For the education devoted to finance the picture is similar. Although the Russian group is smaller the amount of students stating that they receive "A lot" of finance education is higher compared to the German group. Table 1 shows summarized both samples.

RESULTS

As a first step we analyse the three sets of questions, numeracy, basic and sophisticated financial literacy, separately. The numeracy questions are designed to test mathematical skills like division, multiplication, percentages and fractions. Nearly 95% of the students answer each of the five questions correctly and over 90% answer all five questions correctly. The mean is 4.9 correct questions for the German and Russian subsample respectively. The question with the most incorrect answers is Q5 which asks how many people out of 1000 will get a disease if the chance of getting one is 1 in 10. Q5 reframes Q2 which asks for a 10 percent probability of getting a disease. Taking into account that Q2 has the second most incorrect answers it seems that percentages are most difficult to answer whereby the differences to the other questions are in general small. Table 2 reports the results for every question, for all five question together and for each question split up per gender, age, understanding of finance and education devoted to finance. The questions were generally answered correctly with a rate of 100% with a few exceptions.

TABLE 2
NUMERACY QUESTIONS

a.) Percent correct by numeracy question

	Q1 (%) DE/RU	Q2 (%) DE/RU	Q3 (%) DE/RU	Q4 (%) DE/RU	Q5 (%) DE/RU
Correct	96.9/98.3	96.9/98.3	96.9/100	98.5/98.3	96.9/94.9

b.) Summary of correct responses to all numeracy questions

	Five (%) DE/RU	Four (%) DE/RU	Three (%) DE/RU	Two (%) DE/RU	One (%) DE/RU	None (%) DE/RU	Mean DE/RU
Correct	93.8/91.5	3.1/6.8	1.5/1.7	0/0	0/0	1.5/0	4.9/4.9

Notes: Table shows correct responses by demographic characteristics and in aggregate form. The numeracy questions were designed to test concepts of fractions, percentages, division, multiplication and simple probability.

Following Lusardi and Mitchell (2017) and van Rooij et al. (2011b) the financial literacy questions are summarized in two parts. The first five questions test basic concepts of numeracy, compound interest, inflation, time value of money and money illusion. The second part consists of eight questions and aims to measure more sophisticated concepts like volatility, differences between bonds and stocks, long period returns and risk diversification, which, among other things, are relevant for retirement planning.

The results for the basic literacy questions are worse compared to the numeracy part. Q1, which asks about interest rates, is the only question answered in more than 90% correct from both subsamples. The biggest discrepancy between the groups shows question Q2 about compound interest. On the other side Q4 about the time value of money seems to cause the most difficulties for both groups of students. More than one quarter of the Russian group gets it wrong. This result is in line with Bateman et al. (2012) who ask the same questions in an Australian survey and got the worst results for the time value of money question as well.

TABLE 2
NUMERACY QUESTIONS (CONTINUED)

c.) Percent correct by numeracy question and demographics					
	Q1 (%) DE/RU	Q2 (%) DE/RU	Q3 (%) DE/RU	Q4 (%) DE/RU	Q5 (%) DE/RU
Gender					
Female	100/100	100/96.8	100/100	100/100	100/93.5
Male	100/96.4	97.9/100	97.9/100	100/96.4	97.9/96.4
Prefer not to answer	0/-	100/-	100/-	100/-	100/-
Age					
17	-/100	-/100	-/100	-/100	-/100
18	-/100	-/100	-/100	-/100	-/100
19	-/90.9	-/100	-/100	-/100	-/100
20	-/100	-/100	-/100	-/100	-/90.0
21	100/100	100/100	100/100	100/100	100/85.7
22	100/80.0	100/80.0	100/100	100/80.0	100/100
23	92.9/100	92.9/100	100/100	100/100	92.9/92.3
24	100/100	100/100	100/100	100/100	100/100
25	100/-	100/-	100/-	100/-	100/-
26	100/100	100/100	100/100	100/100	100/100
27	100/100	100/100	100/100	100/100	100/100
28	100/-	100/-	100/-	100/-	100/-
30	-/100	-/100	-/100	-/100	-/100
32	100/-	100/-	0/-	100/-	100/-
Understanding of finance?					
1 (very low)	-/100	-/100	-/100	-/100	-/100
2	100/100	100/100	100/100	100/100	100/100
3	100/75.0	100/100	100/100	100/75.0	100/75.0
4	100/100	100/100	100/100	100/100	100/100
5	100/100	100/100	100/100	100/100	100/90.0
6	80/100	80/100	80/100	100/100	80/100
7 (very high)	-/100	-/90.0	-/100	-/100	-/100
Education devoted to financial education?					
A lot	100/93.8	100/100	87.5/100	100/93.8	100/100
Some	97.4/100	97.4/96.6	100/100	100/100	97.4/93.1
A little	100/100	100/100	100/100	100/100	100/90.0
Hardly at all	-/100	-/100	-/100	-/100	-/100

Notes: Table shows correct responses by demographic characteristics and in aggregate form. The numeracy questions were designed to test concepts of fractions, percentages, division, multiplication and simple probability.

From the German subsample 72.3% answer all five questions, with a mean of 4.6, correct whereas only 27.1% of the Russians, with a mean of 3.8 correct answers, get all questions right. Table 3 shows the correct answers per question and the share of students who answered all five questions correct. However, Table 3c breaks down responses by demographic characteristics. Female students answer the questions about numeracy and money illusion better than male students whereas males answer questions about compound interest, inflation and time value of money better than females. Comparing the results with other studies female respondents with an economics education answer better. For example, Lusardi and Mitchell (2008) show that woman in general answer the interest rate and inflation question 61.9% and 70.6% correctly whereas our results indicate 100% and 93.3% correct answers for the German subsample. The Russian subsample confirms it partly with 96.8% and 67.7% correct answers. Considering students who indicate a high understanding of finance usually the German group answers better than the Russian. The general pattern shown by Lusardi and Mitchell (2017) that very young people answer worse than middle age respondents is also seen in Table 3 with only correct answers from the age of 26 on. Regarding the education, respondents who indicate "A little" score better than students with "A lot" of education in finance. Comparing to other studies (Lusardi & Mitchell, 2011a) students with an economic background perform still better than general populations for the compound interest and inflation question.

Table 4 presents answers to the sophisticated financial literacy questions and shows some difficulties for the concepts of the relation between interest rates and bond prices, for knowledge of mutual funds and long period returns of different assets. All this questions are answered largely incorrect. That means that even students with an economics background don't know how bond prices behave depending on interest rates and whether stocks give a higher return than bonds or saving accounts. Similar patterns show Lusardi and Mitchell (2008) who state that even at the highest level of schooling financial literacy tends to be low. 18.5% of the German group and 1.7% of the Russian students answer all possible eight questions correct. A gender comparison shows that men are in 5 out of 8 cases better than women and older students usually better than younger. The effect that women perform worse than men is in line with Lusardi and Mitchell (2011b), Chen and Volpe (2002) and Mandell (2008) who state that women in high school and college show worse results than men. But compared to woman in a general population they still perform better with 60% correct answers for example for the risk diversification question than 47.6% correct in general (Lusardi & Mitchell, 2008). The results vary in general more over gender, age, understanding and education and are worse compared to the basic literacy questions.

TABLE 3
BASIC FINANCIAL LITERACY QUESTIONS

a.) Percent correct by basic financial literacy question

	Q1 (%) DE/RU	Q2 (%) DE/RU	Q3 (%) DE/RU	Q4 (%) DE/RU	Q5 (%) DE/RU
Correct	98.5/90.0	92.3/58.3	95.4/76,7	86.2/61.7	84.6/85.0

b.) Summary of correct responses to all basic financial literacy questions

	Five (%) DE/RU	Four (%) DE/RU	Three (%) DE/RU	Two (%) DE/RU	One (%) DE/RU	None (%) DE/RU	Mean DE/RU
Correct	72.3/27.1	20.0/37.3	4.6/25.4	0/8.5	1.5/0	1.5/1.7	4.6/3.8

c.) Percent correct by basic financial literacy question and demographics

	Numeracy Q1 (%) DE/RU	Compound interest Q2 (%) DE/RU	Inflation Q3 (%) DE/RU	Time value of money Q4 (%) DE/RU	Money illusion Q5 (%) DE/RU
Gender					
Female	100/96.8	93.3/58.1	93.3/67.7	80.0/58.1	93.3/90.3
Male	100/85.7	93.6/60.7	97.9/89.3	91.5/67.9	85.1/82.1
Prefer not to answer	100/-	100/-	100/-	0/-	0/-
Age					
17	-/100	-/0	-/0	-/100	-/100
18	-/87.5	-/50.0	-/62.5	-/25.0	-/87.5
19	-/100	-/75.0	-/100	-/37.5	-/75.0
20	-/90.0	-/70.0	-/70.0	-/70.0	-/80.0
21	100/71.4	100/71.4	100/100	100/42.9	0/85.7
22	100/100	93.3/40.0	100/100	93.3/80.0	86.7/100
23	100/100	100/46.2	100/76.9	85.7/84.6	92.9/92.3
24	100/66.7	93.3/66.7	100/33.3	86.7/100	80/66.7
25	100/-	90.9/-	90.9/-	81.8/-	81.8/-
26	100/100	100/100	100/100	100/100	100/100
27	100/100	100/100	100/100	0/100	100/100
28	100/-	100/-	100/-	100/-	100/-
30	-/100	-/100	-/100	-/0	-/100
32	100/-	0/-	0/-	100/-	100/-
Understanding of finance?					
1 (very low)	-/0	-/50.0	-/0	-/50.0	-/50.0
2	100/100	85.7/0	85.7/50.0	71.4/50.0	85.7/100
3	100/75.0	100/50.0	100/100	92.9/50.0	78.6/100
4	100/100	90.5/75.0	100/66.7	85.7/75.0	85.7/83.3
5	100/95.0	100/55.0	100/75.0	94.1/60.0	94.1/75.0
6	100/100	80.0/66.7	80.0/88.9	80.0/66.7	80/100
7 (very high)	-/90.0	-/60.0	-/100	-/60.0	-/100
Education devoted to financial education?					
A lot	100/87.5	87.5/75.0	87.5/75.0	87.5/87.5	75/93.8
Some	100/96.6	92.1/48.3	97.4/82.8	89.5/58.6	86.8/86.2
A little	100/90.0	100/70.0	100/80.0	83.3/40.0	88.9/80.0
Hardly at all	-/75.0	-/50.0	-/50.0	-/50.0	-/75.0

Notes: Table shows correct responses by demographic characteristics and in aggregate form.

TABLE 4
SOPHISTICATED FINANCIAL LITERACY QUESTIONS

a.) Percent correct by sophisticated financial literacy question

	Q1 (%) DE/RU	Q2 (%) DE/RU	Q3 (%) DE/RU	Q4 (%) DE/RU	Q5 (%) DE/RU	Q6 (%) DE/RU	Q7 (%) DE/RU	Q8 (%) DE/RU
Correct	95.4/86.7	60.0/46.7	55.4/31.7	76.9/53.3	75.4/66.7	67.7/38.3	78.5/71.7	95.4/80.0

b.) Summary of correct responses to all sophisticated financial literacy questions

	Eight (%) DE/RU	Seven (%) DE/RU	Six (%) DE/RU	Five (%) DE/RU	Four (%) DE/RU	Three (%) DE/RU	Two (%) DE/RU	One (%) DE/RU	None (%) DE/RU	Mean DE/RU
Correct	18.5/1.7	35.4/18.6	24.6/15.3	6.2/30.5	1.5/13.6	4.6/8.5	4.6/5.1	1.5/5.1	3.1/1.7	6.0/4.8

c.) Percent correct by sophisticated financial literacy question and demographics

	Function of Stock Market Q1 (%) DE/RU	Knowledge of Mutual Funds Q2 (%) DE/RU	Relation between Interest Rates and Bond Prices Q3 (%) DE/RU	Safer: Company stock or Mutual Fund Q4 (%) DE/RU	Risky Assets Q5 (%) DE/RU	Long period Returns Q6 (%) DE/RU	Volatility Q7 (%) DE/RU	Risk Diversification Q8 (%) DE/RU
Gender								
Female	86.7/83.9	33.3/35.5	60.0/35.5	60.0/61.3	73.3/74,2	53.3/38.7	53.3/71.0	93.3/80.6
Male	100/92.9	70.2/60.7	53.2/28.6	83/46.4	78.7/60.7	72.3/39.3	87.2/75.0	97.9/82.1
Prefer not to answer	100/-	100/-	100/-	100/-	0/-	100/-	100/-	100/-
Age								
17	-/50.0	-/0	-/50.0	-/0	-/0	-/50.0	-/0	-/50.0
18	-/62.5	-/25.0	-/0	-/37.5	-/62.5	-/25.0	-/62.5	-/50.0
19	-/87.5	-/37.5	-/25.0	-/37.5	-/62.5	-/62.5	-/50.0	-/62.5
20	-/90.0	-/40.0	-/40.0	-/40.0	-/60.0	-/50.0	-/60.0	-/70.0
21	100/85.7	100/57.1	0/14.3	100/85.7	100/57.1	100/28.6	100/57.1	100/100
22	100/100	60.0/60.0	60.0/40.0	86.7/40.0	80.0/60.0	80.0/40.0	80.0/100	93.3/100
23	100/100	78.6/61.5	64.3/53.8	78.6/69.2	78.6/84.6	78.6/30.8	92.9/100	100/100
24	93.3/100	46.7/33.3	53.3/33.3	73.3/66.7	80.0/100	60.0/33.3	60.0/100	100/100
25	90.9/-	63.6/-	63.6/-	72.7/-	72.7/-	81.8/-	90.9/-	90.9/-

26	100/100	50.0/100	50.0/100	75.0/100	50.0/100	0/0	75.0/100	100/100
27	100/100	100/100	0/0	0/100	100/100	0/0	100/100	100/100
28	100/-	100/-	0/-	100/-	100/-	100/-	100/-	100/-
30	-/100	-/100	-/0	-/100	-/100	-/100	-/100	-/100
32	100/-	0/-	0/-	100/-	0/-	0/-	0/-	100/-
Understanding of finance?								
1 (very low)	-/50.0	-/0	-/0	-/50.0	-/50.0	-/0	-/100	-/50.0
2	71.4/50.0	28.6/50.0	42.9/50.0	42.9/0	14.3/50.0	28.6/0	42.9/50.0	71.4/50.0
3	100/100	64.3/50.0	50.0/25.0	71.4/75.0	78.6/75.0	57.1/75.0	71.4/25.5	100/75.0
4	100/91.7	76.2/33.3	71.4/16.7	76.2/50.0	81.0/58.3	71.4/16.7	95.2/75.0	100/83.3
5	100/80.0	47.1/45.0	47.1/25.0	94.1/45.0	100/60.0	88.2/35.0	82.4/85.0	100/75.0
6	100/100	80.0/55.6	60.0/55.6	100/66.7	60.0/88.9	80.0/55.6	80.0/66.7	100/100
7 (very high)	-/100	-/70.0	-/50.0	-/70.0	-/80.0	-/60.0	-/70.0	-/90.0
Education devoted to financial education?								
A lot	100/100	62.5/68.8	62.5/31.3	87.5/68.8	75.0/75.0	75.0/43.8	87.5/81.3	100/100
Some	97.4/89.7	65.8/48.3	63.2/34.5	81.6/55.2	89.5/69.0	76.3/34.5	84.2/72.4	97.4/82.8
A little	94.4/80.0	50.0/30.0	38.9/20.0	66.7/40.0	50.0/70.0	50.0/50.0	66.7/70.0	94.4/60.0
Hardly at all	-/50.0	-/0	-/50.0	-/25.0	-/25.0	-/25.0	-/50.0	-/50.0

Notes: Table shows correct responses by demographic characteristics and in aggregate form.

Comparing these results to international studies with the same questions in Table 5 TU Darmstadt students perform better in four basic financial literacy questions and in two out of four comparable sophisticated financial literacy questions.

TABLE 5
INTERNATIONAL COMPARISON OF FINANCIAL LITERACY RESPONSES

	Basic financial literacy					Sophisticated financial literacy			
	Numeracy (%)	Compound interest (%)	Inflation (%)	Time value of money (%)	Money illusion (%)	Risky assets (%)	Long period returns (%)	Volatility (%)	Diversification (%)
Australia[a]	88.4	71.8	78.4	54.9	86.7	64.1	54.9	76.7	73.3
US-ALP[b]	91.8	69.0	87.1	73.8	78.4	80.2	62.3	88.3	74.9
US-NFCS[c]	64.9	n.a.	64.3	n.a.	n.a.	n.a	n.a	n.a	n.a
New Zealand[d]	86.0	n.a.	81.0	n.a.	n.a.	n.a	n.a	n.a	n.a
Germany[e]	82.4	n.a.	78.4	n.a.	n.a.	n.a	n.a	n.a	n.a
Netherlands[f]	84.8	n.a.	76.9	n.a.	n.a.	n.a	n.a	n.a	n.a
UNECON St. Petersburg	90.0	58.3	76.7	61.7	85.0	66.7	38.3	71.7	80.0
TU Darmstadt	**98.5**	**92.3**	**95.4**	**86.2**	84.6	75.4	**67.7**	78.5	**95.4**

Notes: Table shows correctly answered respondents in percentages of survey. n.a., not applicable. In bold font are the values that are higher than the international comparison. The table is taken from Bateman et al. (2012) and adjusted with given results. [a]CenSoc-UNSW survey of 1199 superannuation account holders, May 2010; [b]American Life Panel (ALP) (Lusardi & Mitchell, 2017); [c]National Financial Capability Study (NFCS) (Lusardi, 2011b); [d]ANZ-New Zealand Retirement Commission Financial Knowledge Survey (Retirement Commission, 2009); [e]SAVE 2009 survey (Bucher-Koenen & Lusardi, 2011); [f]Dutch Central Bank Household Survey (Alessie et al., 2011).

Comparing the Russian and German group of students directly in Table 6 the results show that the German group performs in most comparison better even though not all results are significant. One pattern that emerges is that the Russian group seems to be overall better at the numeracy questions even though not significant whereas the basic and sophisticated financial literacy questions show significant better results for the German group. The only significant result indicating better Russian students is for students with a high self-assessed understanding of finance for numeracy questions.

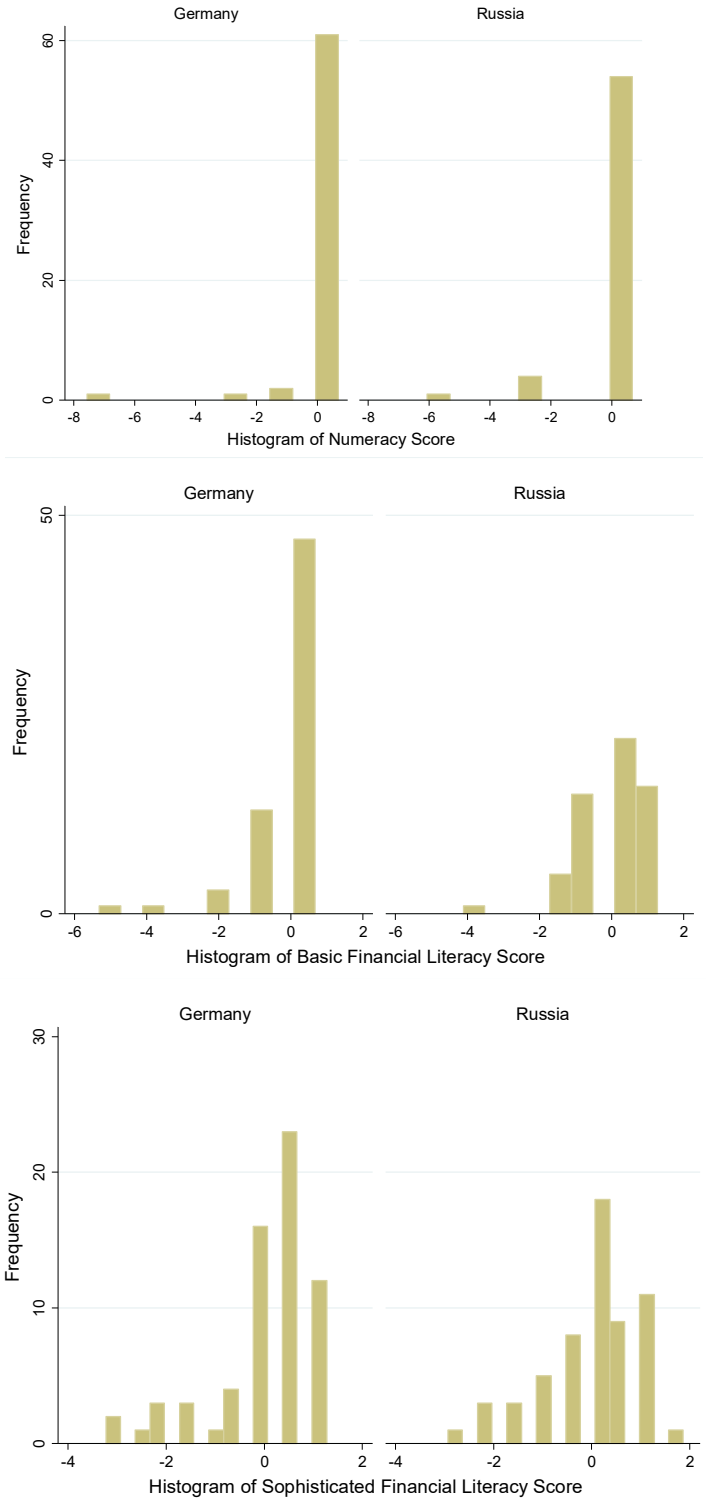
TABLE 6
DIFFERENCE BETWEEN THE GERMAN AND RUSSIAN SUBSAMPLE OVER DEMOGRAPHIC AND SELF-ASSESSMENT CHARACTERISTICS

	T-Test		
	Numeracy	Basic Financial Literacy	Sophisticated Financial Literacy
Gender			
Female	1.2399	2.8311***	0.4892
Male	0.5033	4.1842***	4.1503***
Prefer not to answer	-	-	-
Understanding of finance?			
1 (very low)	-	-	-
2	-	1.1578	0.2006
3	3.1909***	3.5821***	0.9774
4	-	2.1878**	4.6482***
5	1.3367	4.8413***	4.0124***
6	-2,9692**	-0.0393	0.7564
7 (very high)	-	-	-
Education devoted to financial education?			
A lot	0	0.6325	1.2778
Some	0.2934	4.3029***	4.3053***
A little	1.3628	3.6259***	1.0872
Hardly at all	-	-	-
Overall	-0.3706	4.4634***	3.5632***

Notes: *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. The columns report T-statistics for the test of equality of relative correct answers for the numeracy (basic, sophisticated) questions between the German and Russian subsample.

Index Construction

FIGURE 1
HISTOGRAMS OF NUMERACY, BASIC AND SOPHISTICATED FINANCIAL LITERACY



In contrast to Gerardi et al. (2010) who include only numeracy and basic financial literacy questions and Lusardi and Mitchell (2017) and van Rooij et al. (2011a) who only ask for basic and financial literacy questions this study includes all three. Banks (2010) states that numeracy questions are closer linked with cognition, whereas basic literacy questions refer to common economic decisions (inflation and time value of money) and sophisticated literacy questions relate to more financial specialised knowledge (Bateman et al., 2012). We include all three categories and investigate any significant differences between these skillsets at different demographic and self-assessed characteristics.

In order to compare the different skillsets and to analyse the connections to the demographic and self-assessed characteristics we follow the approach of Bateman et al. (2012) and create a score for every respondent and every set of questions. This score results by counting the correct answers per group of questions and standardising them using the sample means and standard deviations to make three indices per respondent.

Figure 1 shows the histograms of the numeracy, basic and sophisticated financial literacy score for the German and Russian subsample. The distributions are negatively skewed and present a slightly more balanced distribution for the sophisticated literacy score.

Financial Competence and General Demographics

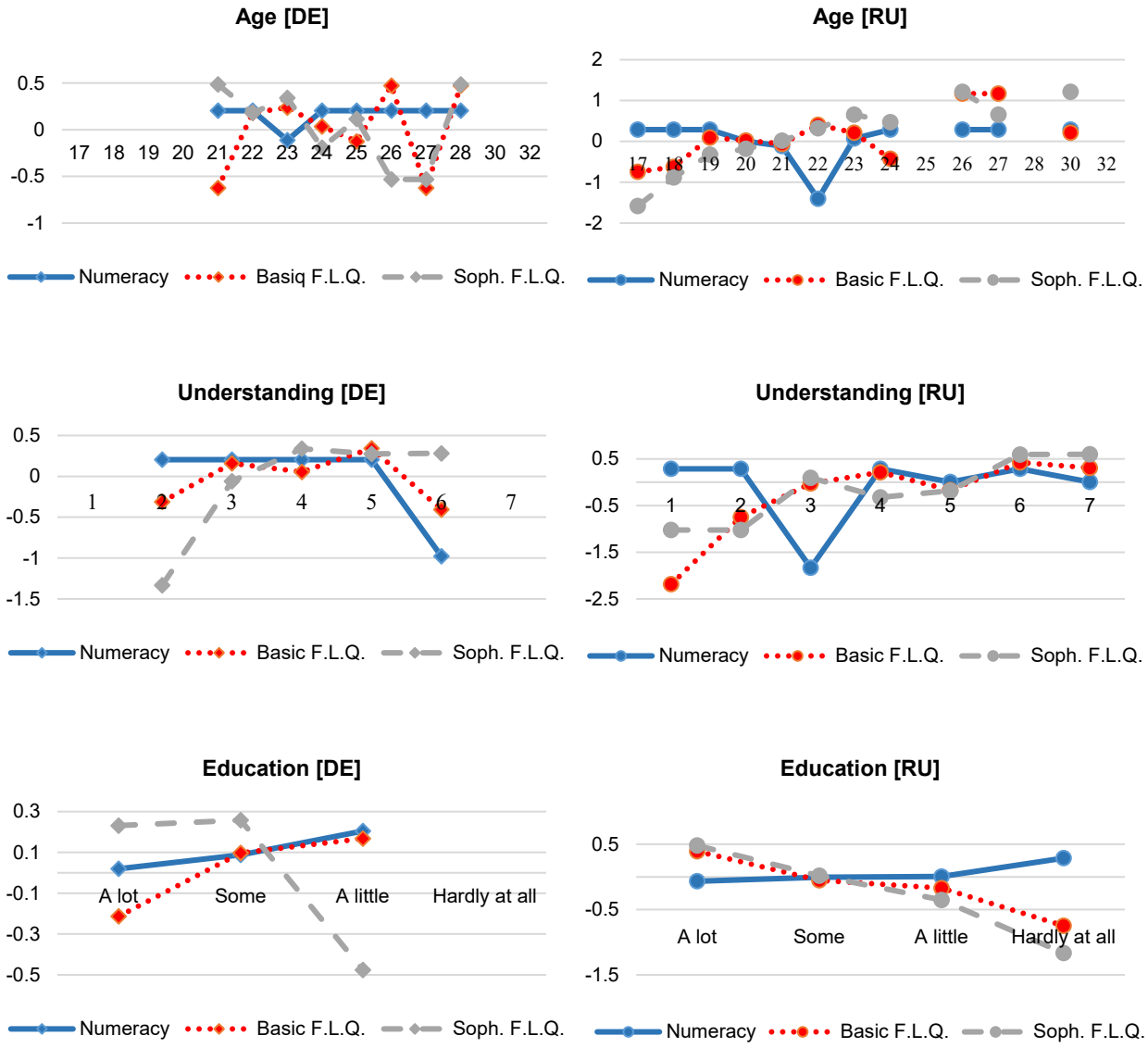
Figure 2 shows five sets of graphs covering demographics, education, self-assessed financial competence and share market expectations. On the left is the German subsample, on the right the Russian. The y-axis shows the average of the numeracy, basic and financial literacy index scores and the horizontal axis the specification of the respective variable. Where there is no natural order for the x-axis variable, we display bars, where the blue bar is average numeracy, the red bar is average basic literacy and the grey bar is average sophisticated literacy. For variables that have an order, we show lines, where the solid blue line displays numeracy scores, the dotted red line the basic financial literacy score and the dashed grey line sophisticated financial literacy scores.

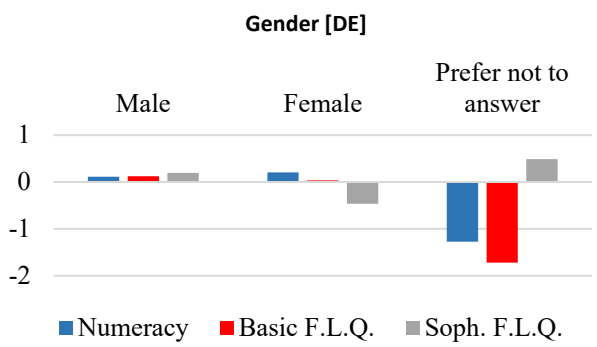
For each graph, we test the influence of the demographic and self-assessed characteristic on each score and report the results in Table 7. In a first step, we test the restriction that in a regression of the numeracy index score (and basic and sophisticated financial literacy score) on a constant and indicator variable, the coefficients on the indicators are jointly zero. These three standard F-tests indicate significant change in average numeracy or literacy score as the horizontal axis category varies. The second step is a Wald test for equality (at each horizontal axis category) of the coefficients of the three regressions for overall differences between the three indices.

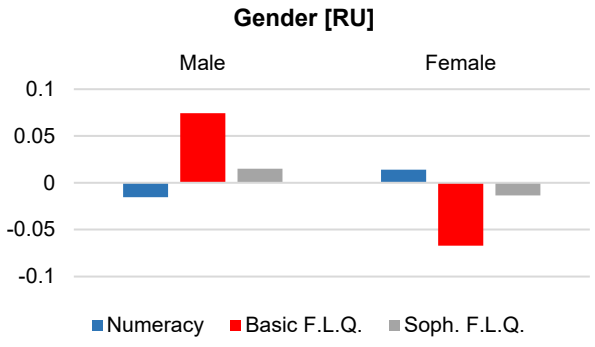
An important feature of the relationship between financial competence and gender is the difference between the sophisticated literacy score and the other two measures. For males the sophisticated literacy score is always positive and for women always negative independent from the numeracy and basic literacy scores. That means that male respondents score better in sophisticated literacy questions, on average, than female respondents. Further numeracy and basic literacy varies significantly per age for the German group whereas the sophisticated financial literacy scores rise significantly with age for the Russian group. The following figures show the scores per level of understanding and education. The understanding is inverted u-shaped and significant for all scores for the German group. That means that respondents with a high self-assessed understanding of finance score worse than respondent with an average understanding. The scores vary for the Russian group and are significant for all index scores. Again, the respondents with the highest scores do not indicate that they also have the best understanding exhibiting the overconfidence in self-assessment which is regularly seen in survey responses (see, e.g. Agnew & Szykman, 2005). (In our survey the self-assessment question comes before the financial literacy questions, so the responses should not be affected by respondents' feelings about how well or badly they answer.) The picture for the education of the German group is similar. Respondents who say that they have "A lot" education in the field of finance score worse than respondent who indicate "Some" education. The Russian group shows the best literacy scores for the students with the most education whereas students with nearly no education in finance score better in the numeracy part. The four graphs in the last row refer to financial competence scores with responses to two survey questions relating to expectations for another financial crisis. This follows the

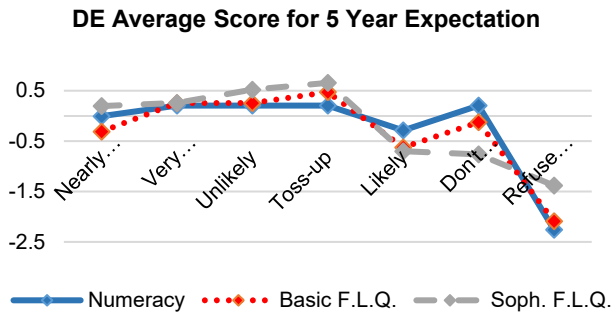
analysis of Bateman et al. (2012) and evaluates in the next section these graphs and explore the relationship between respondent characteristics and expectations using first a standard logit and second an ordered logit model.

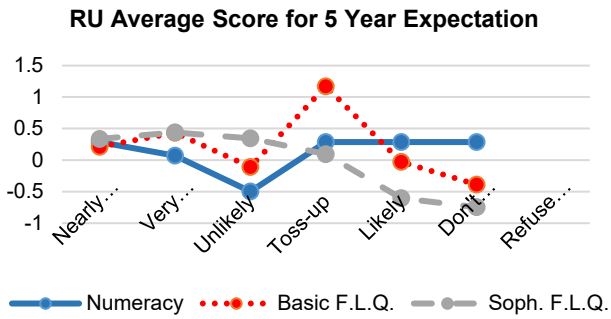
FIGURE 2
AVERAGE FINANCIAL COMPETENCE SCORE BY DEMOGRAPHIC AND SELF-ASSESSED CHARACTERISTICS

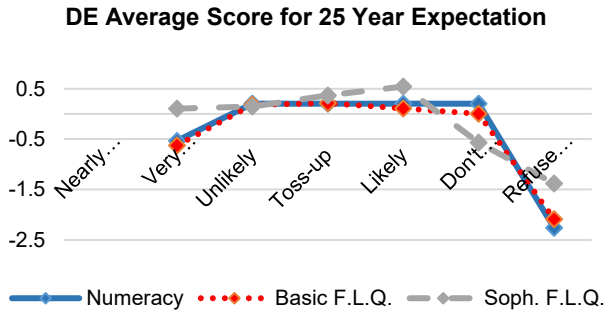


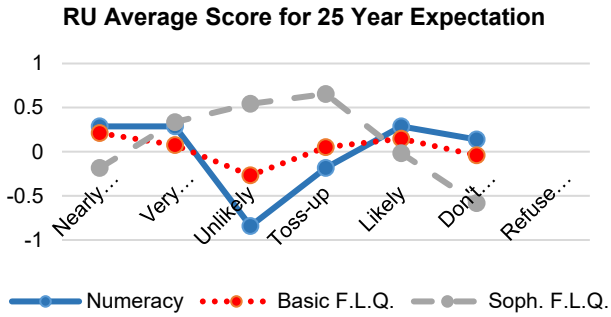












Competence and Crisis Expectations

The global financial crisis of 2007–2009 created many problems for stock holders and thus for many retirement savers, especially for those who were shortly before retirement. At the time of the survey, 13 years after the crisis and after an almost continuous growth for many investors it seemed like another crisis is just a matter of time. This raises the question of how the crisis influences the attitudes and behaviour of future investors. Because of that the last two questions of our survey address respondents' views, on the likelihood of another large stock market crash. We use these answers to evaluate the role of demographics, self-assessed characteristics and financial literacy in expectations formation. Further we are interested in the association of those variables to optimistic or pessimistic views of future markets. We assess expectations about future shocks with two questions. They describe a sharp fall in share prices and then ask how likely it is that share prices will suffer another similar sized loss in the next 5 and 25 years. The answers fall into seven categories, with each category attributing increasing probability to another sharp stock market decline within the next years, and then separate categories for "Don't know" and "Refuse to answer". The respondents could assign probabilities ranging from "Nearly impossible - less than a one in a hundred chance" to "Likely - a greater than one in two chance".

The last graphs in Figure 2 show for the German group that low financial competence is associated with two extreme responses to these questions: extreme optimism and uncertainty ("Don't know"). Respondents with poor basic financial literacy are especially likely to fall into these categories, but poor numeracy is also associated with choosing "Don't know" (uncertainty). For all three scores, we see a significant influence on share market optimism and equality between all three scores (Table 7)

The link between poor financial literacy and optimism matches research into retirement preparation and pension expectations in the Netherlands. Alessie et al. (2011) observe that Dutch respondents with lower financial literacy have difficulty forming realistic expectations of retirement replacement rates and are uncertain of at what age to retire. The Russian students don't show that clear pattern. Instead, students with a poor numeracy score choose often "Unlikely" whereas a poor sophisticated index score is again associated with the two extreme responses.

TABLE 7
VARIATION WITHIN AND BETWEEN FINANCIAL COMPETENCE INDICES

	F-Test			Wald Test
	Numeracy DE/RU	Basic Financial Literacy DE/RU	Sophisticated Financial Literacy DE/RU	Joint Equality DE/RU
Gender	5.85***/0.01	2.81*/0.29	3.07*/0.01	26.77***/0.33
Age	2.15**/1.23	1.22/0.88	0.98/2.98***	38.65**/58.16***
Understanding of finance?	113.59***/3.08**	9.56***/2.79**	7.9***/2.43**	683.03***/51.60***
Education devoted to financial education?	0.61/0.13	0.66/1.72	4.33**/3.91**	15.5**/14.67
Share market crash (5 years)	4.04***/1.39	3.56***/1.37	5.02***/3.87***	59.5***/33.61***
Share market crash (25 years)	5.42***/2.07*	3.71***/0.22	4.09***/2.94**	51.2***/35.26***

Notes: *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. First three columns report F-statistics for the test of the restriction that in a regression of the numeracy (basic, sophisticated) index the coefficients on the indicators are jointly zero for the German and Russian subsample respectively. Final column reports Chi2 statistics for the test that the constants and coefficients of the three regressions, at each horizontal axis category, are equal for the German and Russian subsample respectively.

To further investigate the connections between competence and expectations, we model respondents' subjective evaluations of the likelihood of another severe stock market decline within the next years in a standard logit model and test whether respondents with special characteristics are more likely to make a statement about the future. In a second step we deploy an ordered logit model to test whether respondents with special characteristics are more optimistic or pessimistic about the future. We include age, gender, education devoted to finance, self-assessed understanding of finance, numeracy and the two financial literacy scores, as possible indicator variables.

TABLE 8
LOGIT AND ORDERED LOGIT REGRESSION OF FINANCIAL LITERACY ON STOCK MARKET EXPECTATIONS FOR THE NEXT 5 YEARS FOR THE GERMAN SUBSAMPLE

	I	II					
	Probability	Don't know/ Refuse to answer	Nearly impossible	Very unlikely	Unlikely	Toss-up	Likely
Gender							
Female	71.4%***	28.7%***	13.8%***	31.7%***	19.2%**	3.4%	3.3%
Male	80.0%***	18.6%***	11.0%***	32.5%***	26.8%***	5.5%*	5.6%*
Prefer not to answer	-	42.4%	15.1%***	26.5%	12.2%	1.9%	1.8%
Age							
17	-	-	-	-	-	-	-
18	-	-	-	-	-	-	-
19	-	-	-	-	-	-	-
20	-	-	-	-	-	-	-
21	-	4.6%	3.6%	20.6%	42.5%***	15.5%	13.2%
22	85.7%***	12.3%**	8.3%**	33.3%***	33.9%***	7.3%	5.0%
23	84.6%***	29.4%***	14.1%***	34.1%***	17.8%***	2.7%	1.7%
24	71.4%***	25.0%**	13.2%**	35.4%***	20.9%**	3.4%	2.2%
25	81.8%***	13.3%*	8.8%**	34.0%***	32.5%***	6.7%	4.6%
26	50.0%***	28.2%	13.9%**	34.5%***	18.7%	2.9%	1.8%
27	-	99.95***	2.01e-5%	1.83e-5%	4.37e-6%	5.31e-7%	3.21e-7%
28	-	20.3%	11.7%	36.0%***	24.9%	4.3%	2.8%
30	-	-	-	-	-	-	-
32	-	1.76e-13%	1.5e-13%	1.14e-12%	7.54e-12%	1.49e-11%	100%***
Understanding of finance?							
1 (very low)	-	99.99%***	1.04e-4%	8.81e-5%	2.22e-5%	2.91e-6%	2.43e-6%
2	50.0%**	49.1%**	15.6%***	23.8%**	9.1%	1.3%	1.1%
3	64.3%***	28.9%***	14.6%***	32.9%***	18.0%**	3.0%	2.7%
4	85.0%***	15.4%**	10.2%**	33.5%***	29.1%***	6.1%	5.7%
5	88.2%***	17.3%**	11.1%**	34.1%***	27.1%***	5.4%	5.0%
6	-	5.8%	4.6%	22.3%	38.9%***	13.2%	15.3%
7 (very high)	-	-	-	-	-	-	-
Education devoted to financial education?							
A lot	-	12.9%*	9.2%*	32.0%***	31.1%***	7.4%	7.5%
Some	83.3%***	16.3%***	10.8%***	33.6%***	27.5%***	5.9%*	5.8%*
A little	58.8%***	37.5%***	16.0%***	29.2%***	13.2%**	2.2%	2.0%
Hardly at all	-	-	-	-	-	-	-

Average change of probability in percent points for an one unit increase in the score							
Numeracy	+7.2%	-6.9%	-2.0%	-1.1%	+5.5%	+1.7%	+1.9%
Basic literacy	+8.9%**	-9.0%**	-2.7%	-0.3%	+7.2%**	+2.2%	+2.5%
Sophisticated literacy	+15.9%***	-14.1%***	-4.3%**	-0.8%	+11.4%***	+3.6%	+4.2%

Notes: *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Column I reports probabilities of a logit regression whereas Column II shows ordered logit regressions. Demographic and self-assessment characteristics show an absolute probability value and the index scores an increase of probability when the index score increases in one unit.

Table 8 shows the results for the 5 year expectations of the German group. For reasons of space we show at this point only the results for the German group and only for the five year expectations. The other tables for the Russian group show similar results and are in the appendix. The table reports absolute probabilities for the standard logit (column "I") and ordered logit (column "II") estimation of respondents' expectations, where the reference category is "Don't know/refuse to answer". We put the two categories together because there were too few answers to model "Refuse to answer" separately. For example, the probability for making any statement about the future is 71.4 % when you're female and 80% when male. While no clear pattern can be identified for age and no direct comparison is possible due to the different distribution there is a pattern for the variables understanding of finance and education devoted for finance. Both variables show increasing probabilities for an increasing level of understanding and education indicating that students who are more in contact with finance are more likely to assign a probability to future crisis expectations. The same goes for the index scores even though the interpretation is a little bit different. For the index scores the tables show not absolute probabilities but an increase of probabilities when the index scores increases for one unit. That means that a student who increases his basic financial literacy score by one unit is 8.9% more likely to make a statement about the next five years. Students with higher sophisticated financial literacy are even 15.9% more likely to attribute a specific probability to another crash rather than expressing ignorance or uncertainty by choosing "Don't know". This result is in line with Arrondel (2018) who shows a positive correlation between financial literacy and the propensity to formulate a specific financial plan for a French sample.

The column "II" shows the results for an ordered logit model and thus the absolute probabilities for choosing one of the six answers for the demographic characteristics and an increase in probability for the index scores. An overall comparison between the German and Russian group indicates that Russians are more pessimistic. Overall the main category with highest probability is "Unlikely" whereas Germans mainly choose "Very unlikely". In Germany are males more pessimistic whereas in Russia females are mainly pessimistic. Regarding the understanding of finance and education it seems that students with a high level of understanding and education are more pessimistic. As the index scores increase it becomes more likely that one chooses a pessimistic option with the most improvement for the category "Unlikely". The results for the 25 years are similar except that they are more pessimistic for both groups up to the point that no one from the German students who evaluated the future says that it's "Nearly impossible" that there will be no crisis in the next 25 years. Overall, better financial competence, a better understanding of finance and more education in finance reduce uncertainty in favour of risk quantification. This result is in line with previous studies like Epstein (1999) and Halevy (2007) who show that uncertainty averse people are "not probabilistically sophisticated".

CONCLUSION

Our study documents how a formal education in economics affects financial knowledge and how mathematical knowledge is able to forecast financial literacy. The results are derived from a survey asking students from Germany and Russia an internationally established set of questions about numeracy, basic financial literacy, sophisticated financial literacy and future crisis expectations. The German group shows

superior results in four out of five comparable basic financial literacy questions and in two out of four comparable sophisticated financial literacy questions. Overall, economic students perform better than average individuals but even with that background are only few able to answer all eight sophisticated literacy questions correct (Germany: 18.4%; Russia: 1.7%) which is in line with Lusardi and Mitchell (2008) who show that the level of knowledge is even among well educated people quite low. At the same time, we show that good mathematical skills not necessarily indicate superior financial knowledge. The Russian group shows overall better numeracy results but significant worse financial literacy responses compared to the German group.

Female students perform better on the numeracy part. Regarding the financial literacy questions we confirm former studies from Lusardi and Mitchell (2008), Chen and Volpe (2002), Mandell (2008), and Hung et al. (2009) that women are less financially literate than men. Lusardi and Mitchell (2011b) show differences in self-assessed characteristics like understanding or education of finance between countries. We also find these differences in our dataset between Russians and Germans.

If financial literacy is key to a private and responsible retirement planning and considering that our results indicate that even individuals with an economic background have problems answering all questions correctly it is not surprising that the need of improving financial knowledge is gaining momentum. The success of a partly shifting from a far reaching public to a more private retirement planning depends highly on the skills people have or acquire through time going further than a university education in economics.

REFERENCES

Agarwal, S., Driscoll, J.C., Gabaix, X., & Laibson, D. (2009). The age of reason: Financial decisions over the life cycle and implications for regulation. *Brookings Papers on Economic Activity*, *2009*(2), 51-117.

Agnew, J.R., & Szykman, L.R. (2005). Asset allocation and information overload: The influence of information display, asset choice, and investor experience. *The Journal of Behavioral Finance*, *6*(2), 57-70.

Alessie, R., Van Rooij, M., & Lusardi, A. (2011). Financial literacy and retirement preparation in the Netherlands. *Journal of Pension Economics & Finance*, *10*(4), 527-545.

Arrondel, L., Debbich, M., & Savignac, F. (2013). Financial literacy and financial planning in France. *Numeracy*, *6*(2), 8.

Arrondel, L. (2018). Financial Literacy and Asset Behaviour: Poor Education and Zero for Conduct? *Comparative Economic Studies*, *60*(1), 144-160.

Banks, J. (2010). Cognitive function, financial literacy and financial outcomes at older ages: Introduction. *The Economic Journal*, *120*(548), F357-F362.

Banks, J., & Oldfield, Z. (2007). Understanding pensions: Cognitive function, numerical ability and retirement saving. *Fiscal Studies*, *28*(2), 143-170.

Bateman, H., Eckert, C., Geweke, J., Louviere, J., Thorp, S., & Satchell, S. (2012). Financial competence and expectations formation: Evidence from Australia. *Economic Record*, *88*(280), 39-63.

Beckmann, E. (2013). Financial literacy and household savings in Romania. *Numeracy*, *6*(2), 9.

Behrman, J.R., Mitchell, O.S., Soo, C.K., & Bravo, D. (2012). How financial literacy affects household wealth accumulation. *American Economic Review*, *102*(3), 300-304.

Bernheim, D.D. (1998). Financial illiteracy, education, and retirement saving (No. 96-7). Wharton School Pension Research Council, University of Pennsylvania.

Brown, M., & Graf, R. (2013). Financial literacy and retirement planning in Switzerland. *Numeracy*, *6*(2), 2-23.

Bucher-Koenen, T. (2009). *Financial literacy and private old-age provision in Germany–Evidence from SAVE 2008*. MEA Discussion Paper No. 192-2009. Retrieved from https://ssrn.com/abstract=1534664

Bucher-Koenen, T., & Lusardi, A. (2011). Financial literacy and retirement planning in Germany. *Journal of Pension Economics & Finance*, *10*(4), 565-584.

Bucher-Koenen, T., Lusardi, A., Alessie, R., & Van Rooij, M. (2017). How financially literate are women? An overview and new insights. *Journal of Consumer Affairs*, *51*(2), 255-283.
Chen, H., & Volpe, R.P. (2002). Gender differences in personal financial literacy among college students. *Financial Services Review*, *11*(3), 289-307.
Chiteji, N.S., & Stafford, F.P. (1999). Portfolio choices of parents and their children as young adults: Asset accumulation by African-American families. *American Economic Review*, *89*(2), 377-380.
Christelis, D., Jappelli, T., & Padula, M. (2010). Cognitive abilities and portfolio choice. *European Economic Review*, *54*(1), 18-38.
Christiansen, C., Joensen, J.S., & Rangvid, J. (2008). Are economists more likely to hold stocks? *Review of Finance*, *12*(3), 465-496.
Cole, S.A., Sampson, T.A., & Zia, B.H. (2009). *Financial literacy, financial decisions, and the demand for financial services: evidence from India and Indonesia* (pp. 09-117) Cambridge, MA: Harvard Business School.
Commonwealth of Australia. (2010). *Stronger Super*. Commonwealth of Australia, Canberra, ACT.
Deevy, M., Lucich, S., & Beals, M. (2012). Scams, schemes & swindles. *Financial Fraud Research Center*.
Delavande, A., Rohwedder, S., & Willis, R.J. (2008). Preparation for retirement, financial literacy and cognitive resources. *Michigan Retirement Research Center Research Paper*.
Deutsches Aktieninstitut. (2019). Aktionärszahlen des Deutschen Aktieninstituts 2018 (Stand: 2018).
Epstein, L.G. (1999). A definition of uncertainty aversion. *The Review of Economic Studies*, *66*(3), 579-608.
Finke, M.S., Howe, J.S., & Huston, S.J. (2016). Old age and the decline in financial literacy. *Management Science*, *63*(1), 213-230.
Fonseca, R., Mullen, K.J., Zamarro, G., & Zissimopoulos, J. (2012). What explains the gender gap in financial literacy? The role of household decision making. *Journal of Consumer Affairs*, *46*(1), 90-106.
Fornero, E., & Monticone, C. (2011). Financial literacy and pension plan participation in Italy. *Journal of Pension Economics & Finance*, *10*(4), 547-564.
Gerardi, K., Goette, L., & Meier, S. (2010). *Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data*. Federal Reserve Bank of Atlanta Working Paper No. 2010-10, Federal Reserve Bank of Atlanta, Atlanta, GA.
Guiso, L., & Jappelli, T. (2008). *Financial Literacy and Portfolio Diversification*. EUI Working Paper No. ECO 2008/31, European University Institute, Florence, Italy.
Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica*, *75*(2), 503-536.
Hastings, J.S., & Tejeda-Ashton, L. (2008). *Financial literacy, information, and demand elasticity: Survey and experimental evidence from Mexico* (No. w14538). National Bureau of Economic Research.
Hilgert, M.A., Hogarth, J.M., & Beverly, S.G. (2003). Household financial management: The connection between knowledge and behavior. *Fed. Res. Bull.*, *89*, 309.
Hsu, J.W. (2016). Aging and strategic learning: The impact of spousal incentives on financial literacy. *Journal of Human Resources*, *51*(4), 1036-1067.
Hung, A., Parker, A.M., & Yoong, J. (2009). *Defining and measuring financial literacy*.
Jappelli, T., & Padula, M. (2013). Investment in financial literacy and saving decisions. *Journal of Banking & Finance*, *37*(8), 2779-2792.
Klapper, L., Lusardi, A., & Panos, G.A. (2013). Financial literacy and its consequences: Evidence from Russia during the financial crisis. *Journal of Banking & Finance*, *37*(10), 3904-3923.
Li, G. (2014). Information sharing and stock market participation: Evidence from extended families. *Review of Economics and Statistics*, *96*(1), 151-160.
Lusardi, A. (2009). U.S. Household Savings Behavior: The Role of Financial Literacy, Information and Financial Education Programs. In C. Foote, L Goette, & S. Meier (Eds.), *Policymaking Insights from Behavioral Economics* (pp.109-149). Federal Reserve Bank of Boston.

Lusardi, A. (2012). Numeracy, Financial Literacy, and Financial Decision-Making. *Numeracy* 5(1).
Lusardi, A., Michaud, P.C., & Mitchell, O.S. (2013). Optimal financial knowledge and wealth inequality (No. w18669). *National Bureau of Economic Research. doi*, *10*, w18669.
Lusardi, A., & Mitchell, O.S. (2007a). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, *54*(1), 205-224.
Lusardi, A., & Mitchell, O.S. (2007b). Financial literacy and retirement planning: New evidence from the Rand American Life Panel. *Michigan Retirement Research Center Research Paper No. WP*, *157*.
Lusardi, A., & Mitchell, O.S. (2008). Planning and financial literacy: How do women fare? *American Economic Review*, *98*(2), 413-17.
Lusardi, A., & Mitchell, O.S. (2011a). Financial Literacy and Planning: Implications for Retirement Wellbeing. In *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, edited by Olivia S. Mitchell and Annamaria Lusardi (pp. 17–39). Oxford: Oxford University Press.
Lusardi, A., & Mitchell, O.S. (2011b). Financial literacy around the world: an overview. *Journal of Pension Economics & Finance*, *10*(4), 497-508.
Lusardi, A., & Mitchell, O.S. (2011c). Financial Literacy and Retirement Planning in the United States. *Journal of Pension Economics and Finance*, *10*(4), 509–525.
Lusardi, A., & Mitchell, O.S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, *52*(1), 5-44.
Lusardi, A., & Mitchell, O.S. (2017). How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness. *Quarterly Journal of Finance*, *7*(03), 1750008.
Lusardi, A., Mitchell, O.S., & Curto, V. (2010). Financial literacy among the young. *Journal of Consumer Affairs*, *44*(2), 358-380.
Lusardi, A., & Tufano, P. (2009). Teach workers about the perils of debt. *Harvard Business Review*, *87*(11), 22-24.
Lusardi, A., & Tufano, P. (2015). Debt literacy, financial experiences, and overindebtedness. *Journal of Pension Economics & Finance*, *14*(4), 332-368.
Mahdavi, M., & Horton, N.J. (2014). Financial knowledge among educated women: Room for improvement. *Journal of Consumer Affairs*, *48*(2), 403-417.
Mandell, L. (2008), Financial Education in High School. In A. Lusardi (Eds.), *Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Saving Programs* (pp. 257-279). Chicago: University of Chicago Press.
McArdle, J.J., Smith, J.P., & Willis, R.J. (2011). Cognition and Economic Outcomes in the Health and Retirement Survey. In D.A. Wise (Editor), *Explorations in the Economics of Aging* (pp. 209-236) Chicago: University of Chicago Press, 209-236.
Organisation for Economic Cooperation and Development Publishing. (2005). *Improving financial literacy: Analysis of issues and policies*. Organisation for Economic Co-operation and Development.
Organisation for Economic Cooperation and Development. (2006). *The Importance of Financial Education*. OECD Policy Brief, July 2006. OECD, Paris.
Organisation for Economic Cooperation and Development. (2008). *Improving Financial Education and Awareness on Insurance and Private Pensions*. OECD, Paris.
Retirement Commission. (2009). *2009 Financial Knowledge Survey, Report prepared for the ANZ and New Zealand Retirement Commission*. Retrieved from http://financialliteracy.org.nz/sites/default/files/webfm/Research/r-fksurvey-2009.pdf.
Shim, S., Xiao, J.J., Barber, B.L., & Lyons, A.C. (2009). Pathways to life success: A conceptual model of financial well-being for young adults. *Journal of Applied Developmental Psychology*, *30*(6), 708-723.
Van Rooij, M.C., Kool, C.J., & Prast, H.M. (2007). Risk-return preferences in the pension domain: are people able to choose? *Journal of Public Economics*, *91*(3-4), 701-722.

Van Rooij, M.C., Lusardi, A., & Alessie, R.J. (2011a). Financial literacy and retirement planning in the Netherlands. *Journal of Economic Psychology*, *32*(4), 593-608.
Van Rooij, M., Lusardi, A., & Alessie, R. (2011b). Financial literacy and stock market participation. *Journal of Financial Economics*, *101*(2), 449-472.
Van Rooij, M. C., Lusardi, A., & Alessie, R. J. (2012). Financial literacy, retirement planning and household wealth. *The Economic Journal*, *122*(560), 449-478.
Yoong, J. (2011). Financial illiteracy and stock market participation: Evidence from the RAND American Life Panel. *Financial literacy: Implications for retirement security and the financial marketplace*, *76*.

APPENDIX

TABLE 9
LOGIT AND ORDERED LOGIT REGRESSION OF FINANCIAL LITERACY ON STOCK MARKET EXPECTATIONS FOR THE NEXT 25 YEARS FOR THE GERMAN SUBSAMPLE

	I	II					
	Probability	Don't know/ Refuse to answer	Nearly impossible	Very unlikely	Unlikely	Toss-up	Likely
Gender							
Female	71.4%***	34.2%***	-	7.5%**	23.8%***	24.0%***	10.5%**
Male	73.3%***	23.2%***	-	6.1%**	23.1%***	30.8%***	16.8%***
Prefer not to answer	-	46.0%	-	7.9%**	21.8%**	17.6%	6.7%
Age							
17	-	-	-	-	-	-	-
18	-	-	-	-	-	-	-
19	-	-	-	-	-	-	-
20	-	-	-	-	-	-	-
21	-	9.5%	-	3.3%	18.4%	37.6%***	31.2%
22	78.6%***	12.3%**	-	4.0%*	21.3%***	37.0%***	25.4%**
23	76.9%***	27.2%***	-	7.0%**	27.4%***	27.1%***	11.3%*
24	71.4%***	27.2%***	-	7.0%**	27.4%***	27.1%***	11.3%*
25	72.7%***	31.0%***	-	7.5%**	27.4%***	24.5%***	9.6%*
26	50.0%**	36.5%	-	8.0%*	26.7%***	21.1%	7.7%
27	-	100%***	-	1.54e-5%	2.64e-5%	1.01e-5%	2.6e-6%
28	-	9.5%	-	3.3%***	18.4%	37.6%***	31.2%
30	-	-	-	-	-	-	-
32	-	45.9%	-	8.3%**	24.3%**	16.2%	5.3%
Understanding of finance?							
1 (very low)	-	100.0%	-	1.34e-5%	2.16e-5%	8.84e-6%	2.54e-6%
2	-	-	-	-	-	-	-
3	33.3%*	64.0%	-	7.4%	16.5%	9.1%	3.0%
4	64.3%***	22.8%	-	6.6%	25.3%	29.7%	15.7%
5	85.0%***	18.7%	-	5.8%	24.0%	32.4%	19.2%
6	76.5%***	25.5%	-	7.0%	25.8%	27.9%	13.8%
7 (very high)	-	26.5%	-	7.2%	25.9%	27.3%	13.2%

Education devoted to financial education?							
A lot	-	17.3%*	-	5.4%*	23.4%***	32.9%***	21.1%**
Some	77.8%***	22.4%***	-	6.4%**	25.3%***	29.7%***	16.2%***
A little	52.9%***	41.0%***	-	8.4%**	24.6%***	18.6%***	7.4%*
Hardly at all	-	-	-	-	-	-	-
Average change of probability in percent points for an one unit increase in the score							
Numeracy	+7.2%	-13.0%	-	-1.9%	-1.9%	+8.1%	+8.6%
Basic literacy	+7.8%	-8.5%*	-	-1.1%	-1.0%	+5.1%*	+5.6%
Sophisticated literacy	+17.4%***	-16.6%***	-	-2.4%*	-3.4%	+9.8%***	+12.6%***

Notes: *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Column I reports probabilities of a logit regression whereas Column II shows ordered logit regressions. Demographic and self-assessment characteristics show an absolute probability value and the index scores an increase of probability when the index score increases in one unit.

TABLE 10
LOGIT AND ORDERED LOGIT REGRESSION OF FINANCIAL LITERACY ON STOCK MARKET EXPECTATIONS FOR THE NEXT 5 YEARS FOR THE RUSSIAN SUBSAMPLE

	I	II					
	Probability	Don't know/ Refuse to answer	Nearly impossible	Very unlikely	Unlikely	Toss-up	Likely
Gender							
Female	74.2%***	22.9%***	11.0%***	22.2%***	33.9%***	1.9%	8.0%**
Male	71.4%***	31.4%***	12.8%***	22.2%***	26.9%***	1.4%	5.3%*
Prefer not to answer	-	-	-	-	-	-	-
Age							
17	-	99.9%***	1.49e-5%	1.04e-5%	5.15e-6%	1.66e-7%	4.57e-7%
18	75.0%***	15.8%*	10.7%*	23.7%***	40.1%***	2.4%	7.3%
19	37.5%**	58.2%***	14.6%**	15.4%**	10.4%	0.4%	1.0%
20	90.0%***	19.9%**	12.4%**	24.8%***	35.3%***	1.9%	5.6%
21	57.1%***	26.1%*	14.3%**	25.0%***	29.2%**	1.4%	4.0%
22	80.0%***	23.9%	13.8%**	25.1%***	31.1%**	1.5%	4.5%
23	92.3%***	14.1%**	9.9%**	22.9%***	42.2%***	2.7%	8.2%
24	66.7%**	27.8%	14.7%**	24.9%***	27.7%	1.3%	3.7%
25	-	-	-	-	-	-	-
26	-	23.7%	13.7%	25.1%***	31.4%	1.5%	4.5%
27	-	9.95e-12%	9.17e-12%	3.43e-11%	4.41e-10%	1.8e-10%	100%***
28	-	-	-	-	-	-	-
30	-	41.9%	16.2%***	21.4%*	17.8%	0.7%	2.0%
32	-	-	-	-	-	-	-

Understanding of finance?							
1 (very low)	-	99.9%	1.54e-5%	1.17e-5%	6.48e-6%	2.25e-7%	8.02e-7%
2	50.0%	34.5%	14.2%	22.1%	23.8.%	1.1%	4.2%
3	-	17.8%	10.3%	21.9%	37.9%	2.4%	9.7%
4	75.0%***	16.5%	9.8%	21.4%	39.2%	2.6%	10.5%
5	65.0%***	26.8%	12.9%	23.0%	29.7%	1.5%	5.9%
6	77.8%***	30.4%	13.6%	22.8%	26.8%	1.3%	5.0%
7 (very high)	90.0***	25.0%	12.5%	23.0%	31.2%	1.7%	6.5%
Education devoted to financial education?							
A lot	81.3%***	29.2%***	13.1%	23.1%***	27.8%***	1.4%	5.3%
Some	72.4%***	21.8%***	11.4%	23.0%***	34.2%***	1.9%	7.7%*
A little	80.0%***	18.9%**	10.4%	22.4%***	36.9%***	2.2%	9.1%
Hardly at all	25.0	74.4%***	9.4%	9.2%	5.9%	0.2%	0.8%
Average change of probability in percent points for an one unit increase in the score							
Numeracy	-	6.9%	1.3%	-0.2%	-5.5%	-0.5%	-2.1%
Basic literacy	+9.7%*	-4.8%	-1.0%	+0.1%	+3.9%	+0.4%	+1.5%
Sophisticated literacy	+17.4%***	-11.0%**	-2.0%*	+0.4%	+8.4%**	+0.8%	+3.5%

Notes: *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Column I reports probabilities of a logit regression whereas Column II shows ordered logit regressions. Demographic and self-assessment characteristics show an absolute probability value and the index scores an increase of probability when the index score increases in one unit.

TABLE 11
LOGIT AND ORDERED LOGIT REGRESSION OF FINANCIAL LITERACY ON STOCK MARKET EXPECTATIONS FOR THE NEXT 25 YEARS FOR THE RUSSIAN SUBSAMPLE

	I	II					
	Probability	Don't know/ Refuse to answer	Nearly impossible	Very unlikely	Unlikely	Toss-up	Likely
Gender							
Female	74.2%***	26.4%***	3.1%	11.7%***	17.7%***	11.0%**	30.0%***
Male	60.7%***	38.7%***	3.8%	12.8%***	16.4%***	8.8%**	19.6%***
Prefer not to answer	-	-	-	-	-	-	-
Age							
17	-	100%	1.74e-6%	3.74e-6%	2.60e-6%	8.47e-7%	1.09e-6%
18	75.0%***	29.9%	4.1%	14.4%	20.3%	10.8%	20.4%
19	37.5%**	56.2%	4.6%	13.1%	13.0%	5.3%	7.8%
20	80.0%***	19.8%	3.2%	12.2%	20.8%	13.3%	30.7%
21	42.9%**	44.7%	4.7%	14.6%	16.6%	7.4%	11.9%
22	60.0%***	45.8%	4.7%	14.5%	16.3%	7.2%	11.4%
23	92.3%***	15.2%	2.6%	10.5%	19.8%	14.1%	37.8%
24	66.7%**	30.8%	4.2%	14.5%	20.2%	10.6%	19.7%

25	-	-	-	-	-	-	-
26	-	6.33e-7%	1.33e-7%	6.30e-7%	1.87e-6%	2.53e-6%	100%
27	-	6.33e-7%	1.33e-7%	6.30e-7%	1.87e-6%	2.53e-6%	100%
28	-	-	-	-	-	-	-
30	-	38.0%	4.6%	14.9%	18.5%	8.9%	15.1%
32	-	-	-	-	-	-	-
Understanding of finance?							
1 (very low)	-	100%	7.71e-6%	1.73e-5%	1.28e-5%	4.77e-6%	7.20e-6%
2	50,0%	27.5%	3.5%	12.3%	18.0%	11.2%	27.5%
3	-	30.6%	3.7%	12.7%	17.8%	10.6%	24.7%
4	66.7%***	26.6%	3.4%	12.1%	18.0%	11.4%	28.5%
5	60.0%***	32.3%	3.8%	12.8%	17.6%	10.3%	23.3%
6	66.7%***	41.5%	4.1%	13.1%	15.9%	8.4%	16.9%
7 (very high)	90.0***	18.0%	2.6%	10.0%	17.2%	12.6%	39.6%
Education devoted to financial education?							
A lot	62.5%***	38.5%***	3.9%	12.8%***	16.7%***	9.1%**	19.1%**
Some	75.9%***	24.6%***	3.1%	11.3%***	18.1%***	11.8%	31.1%***
A little	70.0%***	27.7%**	3.4%	11.9%***	18.1%***	11.2%**	27.8%**
Hardly at all	25.0%	74.3%***	3.0%	7.8%	7.2%	3.0%	4.9%
Average change of probability in percent points for a one unit increase in the score							
Numeracy	-5.5%	+0.5%	+0.0%	+0.0%	-0,0%	-0.0%	-0.4%
Basic literacy	1.3%	-2.2%	-0.1%	-0,2%	+0.2%	+0.4%	+1.9%
Sophisticated literacy	+17%***	-11.9%**	-0.5%	-0.9%	+1.2%	+2.0%*	+10.2%**

Notes: *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively. Column I reports probabilities of a logit regression whereas Column II shows ordered logit regressions. Demographic and self-assessment characteristics show an absolute probability value and the index scores an increase of probability when the index score increases in one unit.