The Journal of Business Inquiry 2014, 13, Issue 2, 70-80
http:www.uvu.edu/woodbury/jbi/articles
ISSN 2155-4072

THE JOURNAL OF BUSINESS INQUIRY

Volume 13 Issue 2 2014

Utah Valley University Woodbury School of Business

ISSN 2155-4072 (online)

Report of the Editor of *The Journal of Business Inquiry* For the Year 2014, Volume 13

The Year 2014 was another good year for *The Journal of Business Inquiry (JBI)*. Volume 13 published ten articles. We received many high-quality papers with a 10.12 percent acceptance rate. The articles were written by authors, whose primary affiliations include 35 institutions from 6 countries - **Canada**, **Iran**, **Qatar**, **Saudi Arabia**, **South Korea** and the **United States**. Turnaround time took, with 29.77 percent of the editorial decisions, less than or 30 days, 46.33 percent between 31 and 90 days, 23.9 percent, between 181 and XXXX days.

The ISI Impact Factor Value of *The Journal of Business Inquiry* is 2.374 for the year 2013-14.

On behalf of *The Journal of Business Inquiry*, I would like to thank Professor Diane Tardif of the University of Ottawa for copy editing the articles in this issue and Ann Mecham, as well as Lisa Walker, for their administrative assistance and for formatting the articles. Special thanks go to Aaron Barrett for the IT services he provided.

The Journal of Business Inquiry would also like to thank the following individuals, who served as referees over the past year. Some of these individuals refereed more than one paper. Without their help, *JBI* could not fulfill its mission.

Dr. Ronnie Clayton: Jacksonville State University
Dr. Glenn Huie: Jacksonville State University
Dr. Susan Madison: Utah Valley University
Dr. Milton Mayfield: Texas A&M International University
Dr. Gary Mercado: Utah Valley University
Dr. Abdus Samad: Utah Valley University
Dr. Ramendra Thakur: University of Louisiana at Lafayette
Dr. Chris Westley: Jacksonville State University
Dr. Letty Workman: Utah Valley University

Amir Kia

Editorial Board

Meeting Planners' Perception on Convention Destination Attributes: Empirical Evidence from Six Major Asian Convention Cities

By YANG HUO*

This paper examined meeting planners' perception on Asian convention destinations and their attributes. Meeting Planners, Certified Meeting Professionals (CMP), were asked to indicate (1) how they plan to select destination by checking the attributes of a destination, (2) meeting venues and overnight accommodations. This study will contribute to the literature on the meeting and convention management since it deals with the planner's aspects in the context of his/her role in the site selection procedure as well as strategic event/meeting planning process. Descriptive statistics using frequency distribution and mean performance scores showed that Tokyo is ranked as the first preference followed by Hong Kong, Singapore, and Beijing. Furthermore, findings indicate that meeting facilities are very important attributes attracting and pulling meeting planners as well as their meetings, conventions, and exhibitions to the city. As this study is one of the first cross-national empirical tests of its kind to compare the important convention destination attributes, the findings of this study strengthen the destination management organizations (DMOs) in explaining their meeting attractiveness within the context of hosting more meetings to their cities.

Keywords: Meeting Planners, CMP, DMO, Attributes, Destination, Strategic Meeting Planning

JEL Classification: L83

I. Introduction

The importance of hosting a meeting and convention to a city has emerged as a crucial strategic decision for stakeholders of a destination since the numbers of meetings, attendees, and spending create a positive economical multiplier impact to a destination. Meeting, convention, and exhibition (MCE) industry has emerged as one of the most important segments of many Asian countries. Previous studies on destination image and attributes have concentrated on state or single country but none on multiple cities in Asia. For example, in 2010, the MCE industry contributed some HK$35.8 billion (US$4.6 billion), increased by 18.5 percent from 2008, to the local economy equivalent to 2.1 percent of Hong Kong's total GDP, while generating the equivalent of 69,150 full-time jobs, increased by 13.4 percent from 2008 (Hong Kong Exhibition Convention Industry Association Report, 2011). The MCE industry has emerged as one of the largest and fastest growing sector. Therefore, the stakeholders in Hong Kong are formulating new growth strategies for the long-term success of both the industry and Hong Kong.

*Associate Professor, Hospitality and Tourism Management, Woodbury School of Business, Utah Valley University, 800 West University Parkway, Orem, UT 84058-5999. E-mail: huoya@uvu.edu. Phone: (801) 863-8070, Fax: (801) 863-7396.

Hong Kong, Seoul, Tokyo, Shanghai, and Singapore are but a few examples of Asian convention destinations inducing a larger share of the conventions to their cities as they recognize their positive economic multiplier effects to their communities. The meeting planners play an important role in determining the site selection and their decisions are very critical to the cities hosting the meetings and conventions.

The issue of convention destination attributes and image has recently received considerable attention in the academic literature. Chacko and Fenich (2000) stated marketers and policy makers in these destinations are keenly aware that it is important to understand the key destination attributes reviewed by meeting planners in the site selection process, as the meeting planners' decisions are depended upon the destination's attributes and play pivotal roles since it is the real source of competitive advantage of destinations because they allow destinations to differentiate themselves (Beerli and Marin 2004; García *et al.*, 2012; Bregoli, 2013). Past studies on the site selection and destination attributes focus on one single destination and few destinations. Even those studies are exclusively centered in Western regions: North America and UK/Europe (Weber and Ladkin, 2003; Baloglu and Love, 2005; Lee and Back, 2005; Mair and Thompson, 2009). Therefore, the objective of this study is to explore and determine meeting planners' perceptions on destination attributes for six major convention cities in Asia: Seoul, Tokyo, Beijing, Shanghai, Hong Kong, and Singapore where the convention and meetings industry has emerged as one of the most important sectors in the city and resulting in keen competition among potential host destinations sites (Crouch and Louviere, 2004; Chen, 2006). By using taxonomy of destination attributes developed by Crouch and Ritchie (1997) and extended by Chacko and Fenich, (2000) and Chen (2006), this study aims to identify the most influential determinant factor toward the site selection as determined by meeting planners and its subsequent attributes. To the best knowledge of the author such a determination has not been done in the existing literature.

The next section describes the literature review on the meeting and convention planners' perception on destination attributes and is followed by research methodology, results, and discussions. This paper concludes with a discussion of the meeting planner's perception on the destination attributes and the significance of its relationship with the destination's stakeholders such as destination management organizations (DMOs) in explaining their meeting attractiveness within the context of hosting more meetings to their cities.

II. Literature Review

For meeting and convention planners, choosing a convention site and venue involves many complex factors since it determines whether the meeting objective and return on investment (ROI) are successfully achieved or not, as a meeting and convention planner plays a very important role on behalf of boards of directors or committees and provides very valuable inputs and information into what city will best provide the highest value of the stakeholder of the affiliation. Clark, Evans, and Knutson (1998) indicated that as members of the buying center, association meeting planners play a key role in including convention cities in the final consideration set in associations' decision making process because of their expert and information power. Therefore, meeting and convention planners' perception on the possible convention sites becomes crucial and should be of great interest not only of convention venues but also of associations (Baloglu and Love, 2005).

Content analysis of previous research on the destination attributes showed that promotional appeal of the city and destination service have significant effect in the overall ratings of destination (Chacko and Fenich, 2000; Kim and Kim, 2003; Crouch and Louviere, 2004; Wu and Weber, 2005;

Chen, 2006; Chiang *et al*., 2012). Crouch and Ritchie (1997) developed a taxonomy for classifying the multitude of site attributes which are classified into eight categories and Oppermann (1996) and Chacko and Fenich (2000) using importance-performance analysis, looked into convention destination attributes of meeting planners and illustrated how individual destinations have different strengths and weaknesses. Chen (2006) used the analytical hierarchy process (AHP) to provide a general understanding of decision factors and determine the relative weight of critical attributes affecting site selection. Table 1 shows destination attributes determined by previous research.

Table 1: Destination Attributes

Category	Destination attributes
Accessibility	Availability of air service
	Cost of air service
	Convenience of local transportation
Local support	Destination service
	Government support
	Convention personnel
	Personnel efficiency
Extra-conference opportunities	Promotional appeal of city
	Climate
	Sightseeing
	Shopping
	Cultural attractions
Accommodation facilities	Hotel room availability
	Hotel room rates
	Helpfulness of service people
	Meeting space availability
	Cost of food and beverage
Costs	Transportation expenses
	Lodging expenses
	Food and beverage expenses
	Commodity prices
Others	Friendliness of local people
	Safety of attendees

Sources: Oppermann (1996), Crouch and Ritchie (1997), Chacko and Fenich (2000), Baloglu and Love (2005), Chen (2006).

The variables operationalized for this study are determined by previous studies and mainly extracted from the Crouch and Ritchie (1997) taxonomy findings from the Chacko and Fenich's (2000) study and Ching-Fu Chen (2006). In addition, even previous studies show many destination attributes are very critical in making a decision on site selection, other attributes (i.e., safety of delegations) might provide significant impact on the rating of a city. For example, a destination such as Seoul may be perceived negatively due to a tension between South Korea and North Korea.

III. Research Questions

The review of literature introduces many convention destination attributes, but the importance of those attributes is not clearly measured in the context of convention and meeting planners' point of view toward the top convention cities in Asia. Based on the research questions from Chacko and Fenich (2000) and Chen (2006), the following questions were constructed and augmented by the author in collaboration with the research objectives of this study in order to determine the perception of meeting planners on the selection of cities and their attributes.

1. If you have an opportunity to hold a meeting (convention, conference, or exhibition) in Asia, which city is your first choice, second choice, and so on?
2. What factors impact on you to prioritize sites as a prospective convention destination?
3. Within the factor, which attributes are more important than other attributes?

Once the research questions were constructed, the variables (i.e., attributes) were operationalized and designed in a survey questionnaire form to explore the extent to determine relative importance of each attribute in a sample of CMP Conclave participants.

IV. Methodology

A. Participants

Meeting and convention planners (Certified Meeting Professionals: CMPs) play a major role in the meetings, conventions and exhibitions (MCE) industry as they are deeply involved in the convention committee of the association and corporate segment in choosing the meeting and convention destination. Their selection criteria contain city, hotel, convention center, and extra-conference events. Therefore, their perceptions on convention destination are very critical whether the city could host the meeting or not. A sample of CMPs was selected from the participants of the 2010 CMP Conclave held by the Convention Industry Council (CIC) from July 17-19, 2010, in Baltimore, MD. The CMP Conclave is the industry's only exclusive meeting of CMPs (CIC, 2010) and the site survey was used in order to enhance higher return rate. From this source, 150 survey forms were distributed to the CMP Conclave attendees and a sample of 61 (40.7 percent) was returned from the meeting planners.

B. Instrument and Procedure

The questionnaire requested the CMPs to prioritize the cities as an overall destination for meetings, conventions, and exhibitions. Six major convention cities in Asia were chosen: Seoul, Tokyo, Beijing, Shanghai, Hong Kong, and Singapore. In addition, it requested the CMPs to enlist other cities that he/she might consider as another convention destination alternative besides six cities.

The six major-destination attributes or factors were based on a taxonomy developed by Crouch and Ritchie (1997) and used in the work of Chacko and Fenich (2000): meeting facilities, accommodation facilities, costs, site environment, local support, extra-conference opportunities (shopping, entertainment, etc). I chose these particular attributes because these attributes were used and verified by others in terms of their reliability and validity. Furthermore, the questionnaire requested the CMPs to rank the sub-attributes in the context of their importance within the six major attributes or factors in determining and selecting the city as their best alternatives.

To determine the selection with its priority of the best city from six alternate cities, importance of destination attributes and sub-attributes the data were analyzed using descriptive

statistics as measured through central tendency such as mean, mode, median, and frequency distribution.

V. Results and Discussions

The descriptive statistics of the meeting planners indicate that the largest portion of the meeting planners belong to the association (n=27, 44 percent) followed by the corporation meeting profession (n=21, 34 percent). Most meeting planners were female (n=50, 82 percent) while male planners were counted as only 11 (18 percent).

A. Destination Preference

Of the six Asian convention destinations given, Tokyo (mean=2.67, mode=1) achieved the lowest mean score on a prioritization, therefore, it is ranked as the first preferred convention destination. It was followed by Hong Kong, Singapore, Beijing, Shanghai, and Seoul (see Table 2).

Table 2: Mean Performance Scores of CMPs Perceptions on Six Convention Destinations

	Tokyo [a]	Hong Kong	Singapore	Beijing	Shanghai	Seoul
Mode	1	2	2	4	5	6
Mean	2.67	3.08	3.10	3.61	3.90	4.77
Rank	1	2	3	4	5	6

[a]n=61 for each city where n is the number of observations.
Note: rank refers to results from prioritizing of preferences from 1st to 6th (1= the most prioritized, 6= the least prioritized). Used as Non-parametric rank (order) statistics.

According to the mean performance scores of attributes/factors for affecting convention site preference reported in Table 3, meeting facilities (2.82) and site (city) environment (2.92) were the two most important factors for selecting a convention destination site, followed by accommodation facilities (3.16), costs (3.23), local support (4.07), and extra-conference opportunities (e.g., shopping) (4.75). In contrast to the findings of Chen's (2006) study which shows *site (city) environment* is of the highest importance, this study's finding proves the meeting planners give more weight to meeting facilities.

Table 3: Mean Performance Scores of Attributes for Site Selection

	Rank	Mean
Meeting facilities	1	2.82
Site (city) environment	2	2.92
Accommodation facilities	3	3.16
Costs	4	3.23
Local support	5	4.07
Extra-conference opportunities	6	4.75

n=61.
Note: rank refers to results from prioritizing of preferences from 1st to 6th (see footnote of Table 2).

The mean performance scores of attributes for meeting facilities which was chosen as the most important factor affecting site selection showed that space size and its availability (2.93) and condition and quality (3.82) of meeting facilities were the two most contributing attributes followed by suitability of facilities (3.98) and variety of properties (4.52). In contrast, reputation (5.07) and rental rates (5.87) were the bottom two rankings (see Table 4). These findings that *space size (capacity) and its availability* and *condition and quality* confirm the findings of previous research (Oppermann, 1996; Baloglu and Love, 2005).

Table 4: Mean Performance Scores of Attributes for Meeting Facilities

	Rank	Mean
Space size and its availability	1	2.93
Condition and quality	2	3.82
Suitability of facilities	3	3.98
Variety of properties	4	4.52
Reputation (image)	5	5.07
Rental rates	6	5.87

n=61.
Note: rank refers to results from prioritizing of preferences from 1st to 6th (see footnote of Table 2).

In Table 5, the mean performance scores of attributes/factors for affecting convention site environment indicated that city accessibility (2.13) and site (city) image (2.23) were the two most important factors for selecting a convention destination site, followed by suitability (2.66), and infrastructure (2.98). In contrast to the findings of Chen's (2006) study which shows *suitability* is of the highest importance, this study finds that meeting planners give more weight to *city accessibility* and *city image*. These attributes seem to indicate promotional appeal is important to meeting planners since these make it easier to market the convention destination and site to prospective attendees as mentioned by Chacko and Fenich (2000).

Table 5: Mean Performance Scores of Attributes for Site (City) Environment

	Rank	Mean
City accessibility	1	2.13
Country (city) image	2	2.23
Suitability	3	2.66
Infrastructure	4	2.98

n=61.
Note: rank refers to results from prioritizing of preferences from 1st to 4th (see footnote of Table 2).

The mean performance scores of attributes for accommodation facilities selection show that meeting space size and its availability (4.57) and condition and quality (4.64) of meeting facilities were two of the most contributing attributes followed by hotel room rates (5.18) and suitability of facilities (6.07). These findings do not support for DiPietro *et al.'s* (2008) and Oppermann's (1996) findings that "safety and security were among the most important factors." In contrast, reputation (6.38) and helpfulness of service people (6.54) were the bottom two rankings (see Table 6).

Table 6: Mean Performance Scores of Attributes for Accommodation Facilities

	Rank	Mean
Meeting space size and its availability	1	4.57
Condition and quality	2	4.64
Hotel room rates	3	5.18
Suitability of facilities	4	6.07
Reputation (image)	5	6.38
Helpfulness of service people	6	6.54

n=61.
Note: rank refers to results from prioritizing of attributes (see footnote of Table 2).

In Table 7, the mean performance scores of attributes/factors for affecting costs indicated that hotel room rates (2.36) and airfare (2.49) were the two most important factors for costs factor, followed by room rental (2.79), food costs (3.38), technological equipment rental (4.15), and local transportation (4.80) respectively. The finding from *costs* attribute shows the *hotel cost (rate)* is the most important attribute while Chacko and Fenich (2000) observe that *airfare* is highly regarded among the costs attributes.

Table 7: Mean Performance Scores of Attributes for Costs

	Rank	Mean
Hotel	1	2.36
Airfare	2	2.49
Meeting room rental	3	2.79
Food	4	3.38
Technological equipment rental	5	4.15
Local transportation	6	4.80

n=61.
Note: rank refers to results from prioritizing of attributes (see footnote of Table 2).

According to the results reports in Table 8, the mean of the performance scores of attributes for local support indicates that government's (CVB) support (1.82) and quality of industry personnel (1.93) were two of the most contributing attributes followed by efficiency of industry personnel (2.26) and helpfulness of service people (6.54). In contrast to Chacko and Fenich's (2000) findings which show *helpfulness of service people* was the important attribute for *local support,* this study shows *CVB's support* is the most important attribute for *local support*.

Table 8: Mean Performance Scores of Attributes for Local Support

	Rank	Mean
Government (CVB) support	1	1.82
Quality of industry personnel	2	1.93
Efficiency of industry personnel	3	2.26
Helpfulness of service people	4	6.54

n=61.
Note: rank refers to results from prioritizing of attributes (see footnote of Table 2).

In Table 9, the mean performance scores of attributes/factors for affecting extra-conference opportunities costs indicated that sightseeing and cultural attractions (1.72) and shopping (2.26) were the two most important factors for extra-conference opportunities factor, followed by outside entertainment (2.67), and climate (3.34) respectively. As the samples of this study are limited to Asia, the meeting planners give heavy weight to the *sightseeing and cultural attractions*. This finding is in contrast to that of a previous research focused on the US domestic destinations (Oppermann, 1996; Chacko and Fenich, 2000; Baloglu and Love, 2005; DiPietro *et al.*, 2008).

Table 9: Mean Performance Scores of Attributes for Extra-Conference Opportunities

	Rank	Mean
Sightseeing and cultural attractions	1	1.72
Shopping	2	2.26
Outside entertainment	3	2.67
Climate	4	3.34

n=61.
Note: rank refers to results from prioritizing of attributes (see footnote of Table 2).

The mean performance scores of attributes for other factors in Table 10 that could impact on the site selection show that safety of attendees (1.36) was the most important factor, followed by communication skill (English) of local employee's (2.41), friendliness of local people (2.72), and interaction with local CMP (PCO) (3.49), respectively. Meeting planners depend heavily on the safety of attendees as the meeting is held in foreign cities and this finding is similar to that of Chen's (2006) research.

Table 10: Mean Performance Scores of Attributes for Other Factors

	Rank	Mean
Safety of attendees	1	1.36
Communication skill of employees	2	2.41
Friendliness of local people	3	2.72
Interaction with local CMP (PCO)	4	3.49

n=61.
Note: rank refers to results from prioritizing of attributes (see footnote of Table 2).

VI. Conclusions

This study is one of the first cross-national empirical tests of its kind to demonstrate the convention and meeting planner's destination choice based on its attributes and image. It is worthy to note that planners give high weight toward Japan among other cities in Asia. In addition, the most important finding from this study is that meeting facilities and site (city) environment are vital contributors to a convention destination. As many Asian cities compete for hosting the meeting, convention, and exhibition business and due to an economic turmoil, it is very critical for the cities to consider those attributes as a way to enhance and alter meeting planners' perceptions of their cities.

A number of studies on the meeting planners' perception focus on their own city attributes that are important in hosting convention delegations. The drawback of previous studies is that they simply listed and ranked the significant attributes without performing the competitive analysis. The significant contribution of this study is that by comparing six major competing cities it provides indicators in terms of convention destination alternatives. In other words, this study implies which attributes or factors are the most important to a meeting planner as well as to city convention stakeholders.

The results of this study indicate that attributes can play a valuable role in examining image and positioning strategies of a convention destinations image. As competition among Asian cities

increases, a well-conceived positioning strategy becomes vital. For example, Seoul would do well to promote city environment as well as safety of attendees.

This study only examined six Asian convention destinations and their attributes. Other Asian cities as well as other attributes could have been utilized in order to elaborate the findings. Furthermore, this study uses only descriptive analysis techniques using an ordinal measurement. Other statistical analysis tools should be used such as factor and regression analysis in order to determine the importance of destination attributes and to understand the co-relationship among destination attributes. In addition, meeting planners may consider other attributes as an important considering factor in site selection. It is hoped that this study contributes to the base of understanding the meeting planners' perception on Asian convention destinations and their attributes and utilizes the results to promote their destination positioning strategy. Therefore, the city will attract more meetings, conventions, and exhibitions to their cities in order to flourish the positive economical impact of conventions to their communities.

As this study contains all meeting planners' segments, it would be better to classify or allocate them into different segments such as corporate, association, social, military, education, religion, and fraternal (SMERF), independent/individual, and professional convention organization (PCO) in order to observe each segment's meeting planners' perception specifically.

References

Baloglu, Seyhmus, and Curtis Love. 2005. "Association Meeting Planners' Perceptions and Intentions for Five Major US Convention Cities: The Structured and Unstructured Image." *Tourism Management,* 26(5): 743-52.

Beerli, Asuncíon, and Josefa D. Martin. 2004. "Tourists' Characteristics and the Perceived Image of Tourist Destinations: A Quantitative Analysis—A Case Study of Lanzarote, Spain." *Tourism Management*, 25(5): 623-36.

Bregoli, Ilenia. 2013. "Effects of DMO Coordination on Destination Brand Identity: A Mixed-Method Study on the City of Edinburgh." *Journal of Travel Research,* 52(2): 212-24.

Chacko, Harsha E., and George G. Fenich. 2000. "Determining the Importance of US Convention Destination Attributes." *Journal of Vacation Marketing*, 6(3): 211-20.

Chen, Ching-Fu. 2006. "Applying the Analytical Hierarchy Process (AHP) Approach to Convention Site Selection." *Journal of Travel Research*, 45(2): 167-74.

Chiang, Che-Chao, Ying-Chieh Chen, Lu-Feng Huang, and Kai-Feng Hsueh. 2012. "Destination Image and Marketing Strategy: An Investigation of MICE Travelers to Taiwan." *Journal of American Academy of Business, Cambridge*, 18(1): 224-31.

Clark, J. Dana, Michael R. Evans, and Bonnie J. Knutson. 1998. "Selecting a Site for an Association Convention: An Exploratory Look at the Types of Power Used by Committee Members to Influence Decisions." *Journal of Hospitality and Leisure Marketing*, 5(1): 81-93.

Crouch, Geoffrey I., and Jordan J. Louviere. 2004. "The Determinants of Convention Site Selection: A Logistic Choice Model from Experimental Data." *Journal of Travel Research*, 43(2): 118-30.

Crouch, Geoffrey I., and J. R. Brent Ritchie. 1997. "Convention Site Selection Research: A Review, Conceptual Model, and Propositional Framework." *Journal of Convention and Exhibition Management*, 1(1): 49-69.

DiPietro, Robin B., Deborah Breiter, Paul Rompf, and Marta Godlewska. 2008. "An Exploratory Study of Differences among Meeting and Exhibition Planners in their Destination Selection Criteria." *Journal of Convention and Event Tourism*, 9(4): 258-76.

García, Juan A., Mar Gómez, and Arturo Molina. 2012. "A Destination-Branding Model: An Empirical Analysis Based on Stakeholders." *Tourism Management*, 33(3): 646-61.

Hong Kong Exhibition Convention Industry Association Report 2008. HKECIA Economic Impact Study Report 2011. http://www.exhibitions.org.hk/english/industry_detail.php?id=266

Kim, Woo Gon and Hyeon-Cheol Kim. 2003. "The Analysis of Seoul as an International Convention Destination." *Journal of Convention and Exhibition Management*, 5(2): 69-87.

Lee, Myong Jae, and Ki-Joon Back. 2005. "A Review of Convention and Meeting Management Research 1990-2003." *Journal of Convention and Event Tourism*, 7(2): 1-20.

Mair, Judith, and Karen Thompson. 2009. "The UK Association Conference Attendance Decision-Making Process." *Tourism Management*, 30(3): 400-09.

Oppermann, Martin. 1996. "Convention Destination Images: Analysis of Association Meeting Planners' Perceptions." *Tourism Management*, 17(3): 175-82.

Weber, Karin, and Adele Ladkin. 2003. "The Convention Industry in Australia and the United Kingdom: Key Issues and Competitive Forces." *Journal of Travel Research*, 42(2): 125-32.

Wu, Ann, and Karin Weber. 2005. "Convention Center Facilities, Attributes and Services: The Delegates' Perspective." *Asia Pacific Journal of Tourism Research*, 10(4): 399-410.

The Journal of Business Inquiry 2014, 13, Issue 2, 81-100
http:www.uvu.edu/woodbury/articles/
ISSN 2155-4072

The Relation Between Private Ownership of Equity and Executive Compensation

By DANIEL AMES*

This study compares the executive compensation practices of firms with private equity and public debt to the compensation practices of firms with public equity and public debt. From 1992-2005, in a sample of 77 firms, it was found that privately-owned companies pay less equity compensation and less total compensation. Three explanations for observed differences in equity compensation are proposed and tested, and it was found that results consistent with an explanation that privately-held firms pay executives less equity due to inherent difficulties in valuing and/or liquidating equity.

Keywords: Private and Public Firms, Executive Compensation, Equity-Based Compensation

JEL Classification: G32, J31, J33, J41

I. Introduction

The forms and levels of executive compensation in a firm are varied and are influenced by several factors, including firm size, executive tenure and rank and, ideally, firm performance. However, one factor that has not been adequately researched is the relation of equity ownership to the type and amount of executive compensation in a firm. In this paper, I examine differences in executive compensation between public equity firms and private equity firms, holding public debt constant. Specifically, I test for differences in both aggregate compensation and its components: equity-based pay, other incentive-based pay (including bonuses and perquisites) and fixed compensation, such as salary. My intent is to identify and explain differences in compensation that occur as a result of the private ownership of equity and subsequently to identify the underlying causes of any existing differences.

Prior research on differences in executive compensation between public and private firms provides mixed results. Ke *et al.* (1999) find weak support for a difference in total CEO (Chief Executive Officer) compensation among insurance companies. Givoly *et al.* (2010) find that equity-based compensation increases in dollar value after an initial public offering, though the percentage of CEOs receiving equity-based compensation remains unchanged before and after an initial public offering (IPO).

Private equity firms and family-founded firms share a number of characteristics including a more concentrated ownership group and higher ownership levels by executives. Ali *et al.* (2006) test for compensation differences between family-founded firms and non-family-founded firms in a sample of S&P 500 firms. He finds that CEOs of family-founded firms receive less equity-based compensation and less total compensation.

---

*Assistant Professor of Accounting, Illinois State University, Department of Accounting-5520, Normal, IL 61761. E-mail: dames@ilstu.edu.

I employ a sample similar to Givoly *et al.* (2010) and extend their results by analyzing data from firms whose equity is privately traded but whose debt is publicly traded.
Specifically, I use a sample of 77 firms, containing 445 firm-year observations from 1992 to 2005, whose equity is privately held but whose debt is publicly owned. I then compare the executive compensation practices of these firms to the compensation practices of a corresponding sample of companies whose debt and equity are publicly owned.

Consistent with Ali *et al.* (2006) and Givoly *et al.* (2010) but in contrast to Ke *et al*. (1999), I find that private equity firms pay their executives less equity compensation than do their public equity counterparts, in dollars and as a percentage of total compensation. This result is robust when the sample is restricted to CEOs. I find that private equity firms pay less incentive-based compensation in some specifications, but more in others. I find that private equity firms offer less total compensation.

I also test for three possible determinants of the differences in equity compensation:

1. *Ownership differences*. Private firm managers own more of the firms they manage than their public counterparts. Increased ownership leads to better incentive alignment and decreases the need for compensation, especially equity-based compensation.

2. *Liquidity and valuation differences*. Accurately valuing and/or liquidating equity holdings in firms with no publicly-traded equity shares are comparatively difficult. As a result, private ownership of equity induces a firm to reduce the use of this form of compensation.

3. *Monitoring differences*. Private firms generally have few owners. This concentrated ownership yields superior monitoring that in turn reduces the need for incentive alignment mechanisms in compensation contracts, such as payment in the form of equity.

To test for differences in ownership between publicly- and privately-traded firms, I measure the number of shares owned by the executives in my sample. I find that private equity executives own more shares of the firms they manage than public equity executives but not larger percentages of total shares outstanding. This suggests that ownership differences are *not* a primary determinant in the difference in equity compensation in my sample.

To test the second explanation, that the difficulty in valuing and/or liquidating equity compensation drives differences, I collect a sample of firms who have either 'gone public' or 'gone private' while maintaining public debt. I find evidence that, prior to an initial public offering or after going private, firms offer less equity compensation than when a public market exists for their equity shares. This result is consistent with the second explanation. However, a change in equity status (from public to private or vice versa) not only affects liquidity and valuation but also monitoring as ownership changes. Thus, I perform an additional test designed to test monitoring specifically.

Because most of the firms in my sample do not file a proxy statement, many of the traditionally employed monitoring proxies are unavailable. In their place, I use earnings management. Following Givoly *et al.* (2010) and Burgstahler and Dichev (1997), I test for significant discontinuities in frequency distributions of firms' reported earnings around zero. I find no evidence of differences in monitoring using this test. In further monitoring tests, I compare differences in the number of large shareholders in public and private firms, as significant owners

in a company may be able to monitor a firm more closely than atomistic shareholders. The number of large shareholders (owners of at least 5 percent of outstanding common stock) does not differ significantly between public and private firms. Finally, I test for differences in the number of board meetings of public and private firms and find that private firms have fewer meetings. Thus, my tests do not support the explanation that superior monitoring drives the differences in equity compensation that I observe. Taken together, these findings are consistent with the argument that private equity firms compensate their executives differently because of liquidity and/or valuation concerns.

My study is based on a sample of US private equity firms from non-regulated industries, which improves the generalizability of findings relative to Ke *et al.* (1999). This study expands upon Givoly *et al.*'s (2010) description of CEO compensation differences and incorporates compensation data for all of the top five executives in each firm. In addition, I compare private equity firm compensation practices to public equity firm compensation practices during periods that do not immediately precede an initial public offering. Finally, I test for determinants of differences in executive compensation between public and private equity firms.

This paper improves our understanding of the relation between equity ownership and executive compensation. According to Sanders and Hambrick (2007), lower levels of certain types of equity compensation reduce the likelihood of risky decisions by managers and reduce the likelihood of big losses for shareholders (see also Sanders, 2001). Thus, my findings imply that managers of companies in which equity is privately held may be less likely to engage in risky behavior and are less likely to deliver large losses for owners of the corporations that they manage.

The rest of this paper is organized as follows: Section II includes a brief literature review and develops the hypotheses, Section III describes sample selection, Section IV explains results and Section V concludes.

II. Literature Review and Hypothesis Development

A. Literature review

Even assuming reasonably accurate and complete financial reports, relatively little is known of privately-owned firms, largely because, except in the case of firms with publicly-traded debt, they are not subject to the same reporting requirements as publicly-traded companies. Researchers have attempted to identify key differences between publicly-traded and privately-traded firms in several instances. Most of that research focuses on earnings management (Beatty *et al.*, 2002; Beatty and Harris; 1999, Burgstahler *et al.*, 2006; Penno and Simon, 1986), conservatism (Ball and Shivakumar, 2005), or earnings quality (Givoly *et al.*, 2010).

Within the realm of public and private firm executive compensation, prior research provides conflicting results. Ke *et al.* (1999) examine 43 privately-held and 18 publicly-held property liability insurers. They find that privately-held firms exhibit a weaker pay-for-performance relationship than the corresponding publicly-held firms. In levels, they find no significant difference in total compensation between these two groups. In changes, they do find that publicly-held firms offer greater total compensation.

Givoly *et al.* (2010) study the effect of equity ownership on accruals quality. Their sample of firms with private equity but public debt spans 1978 through 2003 and includes 531 distinct firms and observations on a total of 2,519 firm-years. They compare these to a sample of firms with public debt and public equity (3,954 distinct firms and 30,696 total firm-year observations).

They find that private equity firms are more conservative with respect to their reporting practices than public equity firms but that private equity firms are more likely to manage their earnings relative to public equity firms. They compare CEO compensation in the years immediately preceding and following an IPO and compare the compensation after the IPO. They conclude that CEOs were just as likely to receive stock options before an IPO as after the IPO, though the value of the options included in the annual compensation package was greater after the IPO.

Finally, Ali *et al.* (2006) study the compensation of managers of family-founded firms relative to the compensation of non-family-founded companies that share a number of characteristics with private firms, including concentrated ownership and better monitoring (Demsetz and Lehn, 1985). Using a 2002 sample of 177 S&P 500 firms, Ali *et al.* (2006) find that family-founded firms are less likely to grant equity-based compensation to their professional CEOs and that they pay less total compensation to their CEOs. Furthermore, he finds that family-founded firms use fewer, though more financial-based, performance measures in compensation contracts and use more discretion in determining CEO bonuses.

Taken together, the extant literature leaves an incomplete understanding of the compensation practices of privately-owned firms and how they may differ from those of publicly-owned companies. My purpose, in addition to addressing this question in a more generalizable setting, is to identify the determinants of extant differences.

B. Hypothesis Development

I expect the value of total compensation to differ between executives of privately- and publicly-owned firms. Employment as a manager of a private firm differs fundamentally from employment as a manager of a publicly-traded firm in some important ways. For example, shareholders in privately-owned firms hold their shares for longer periods than in publicly-owned firms, so managers are able to focus on the long term (Beatty and Harris, 1999). Managers of public firms are often pressured by the market to focus on short-run success, potentially at the cost of reduced long-term growth (James, 1999; Kwak, 2003; Stein, 1988 and 1989). This difference in focus by ownership may constitute a superior working environment or an implicit form of compensation.

This long-term focus by private ownership may also lead to greater job security for managers in privately-held firms. In other words, managers of privately-held firms may be less likely to be fired than managers of a publicly-traded firm. This reduced risk, if present, is another form of implicit compensation. Furthermore, to the extent that managers of private firms are already personally tied to the company through a large equity stake or through emotional attachment, as in some family-founded firms (Ali *et al.*, 2007; Anderson and Reeb, 2003; Shleifer and Vishny, 1997), less annual compensation is required to motivate managers to work.

It is important to note, however, that firms backed by financial sponsors may not enjoy the same long-horizon benefits. That is, the tenure of executives in firms backed by financial sponsors is actually very low (Kaplan and Strömberg, 2009 and Givoly *et al.*, 2010). Thus, to the extent that firms in this sample have a financial sponsor, I do not expect to observe this difference.

Given these two competing forces, I remain agnostic in expectation and state the hypothesis in the null form:

H1. No difference exists in the value of total compensation paid to executives in firms with privately-traded equity and firms with publicly-traded equity.

In firms with privately-owned equity, I expect that many managers are large stakeholders of the firms they manage, as is often the case with family-founded firms (Anderson and Reeb, 2003; Ali *et al.*, 2007). If a manager's wealth is strongly linked to firm wealth, additional equity-based compensation is not required (Hall and Murphy, 2003). Another reason to expect equity compensation to differ in firms with privately-traded equity is the difficulty inherent without a public market—liquidity and valuation of the shares. Equity-based compensation in a private firm is difficult to value and is likely to be difficult (or impossible) to liquidate. While independent valuations of stock price are available and required in some cases, this price is derived from a single source, offered at a specific point in time. This stands in stark contrast to the stock price of a publicly-traded firm, which is based on the valuation estimates of many, perhaps millions, updated almost constantly. Even if the stock price were readily and accurately estimable, the liquidity of a privately-held share of stock is unclear. In a publicly-held company, trading a share of stock for cash is a relatively straightforward exchange. This is unlikely to be the case in a privately-traded firm.

However, some privately-held firms may actually *emphasize* equity-based compensation. Many privately-held firms have a long term goal to become publicly traded. While this intention is impossible to observe, to the extent that it exists, executives may actually *prefer* to receive equity-based compensation in anticipation of a time when equity stakes in the firm become liquid and the opportunity arises to 'cash in'. It is also the case that privately-owned firms of this type may prefer to offer equity-based compensation because cash is scarce. Given these two competing forces, I remain agnostic in expectation and state the hypothesis in the null form:

H2. No difference exists in the value of equity compensation paid to executives in firms with privately-traded equity and firms with publicly-traded equity.

III. Sample Selection

A. Primary Sample

My sample consists of Security and Exchange Commission (SEC) registrants whose debt is publicly traded but whose equity is privately held. In order to compare the compensation practices of these privately-owned companies to those of publicly-traded firms, I compare my sample to a subsample of the Execucomp database, which contains compensation data for publicly-owned companies.

The private equity portion of my sample is based on a subsample of firms from the Compustat database whose equity status has been verified as private (per the 10-K filing). I begin with the entire Compustat industrial annual database for the years 1992-2005[1]. I then eliminate firms that have a stock price at fiscal year-end. I also exclude firms that have less than 1 million dollars in debt. As with Givoly *et al.* (2010), the sample excludes SIC (Standard Industrial Classification)

[1] The sample focuses on 1992 (the first year 10-k filings became available electronically) through 2005 (the last year prior to the implementation of the Summary of Financial Accounting Statement (SFAS) 123R This sample also avoids the impact of the financial crisis in 2008 and after.

codes 6000-6999 (financial institutions) and SIC codes 4800-4900 (regulated industries). Heavily-regulated industries face different incentives and may have different compensation contract designs that adversely affect the analysis. The sample excludes foreign firms, either listed as American Depository Receipts or with a state incorporation code of 99. I further eliminate firms with a stock ownership code of 1 (subsidiary of public firm) and firms with public equity. Of these, I randomly sample 2,500 firm-year observations for verification that each firm's equity is privately owned.

I eliminate firms that were publicly traded, even on small exchanges or over-the-counter markets, subsidiaries of public firms, and firms in bankruptcy. Because compensation contracts may change in the periods immediately prior to a "going public" transaction, I also exclude the two years prior to an initial public offering. Additionally, I eliminate firm years in which the executive compensation information is not included in the 10-K or proxy statement. Additionally, I eliminate observations that did not contain executive compensation information. The resulting sample of 445 firm-year observations from 77 firms[2] represents approximately 20 percent of the population of firms having these characteristics during the period 1992-2005. The other 2,055 firm-year observations sampled were excluded from further analysis.

Data for my sample was collected from the 10-K and proxy statements of each firm. The data includes names and titles of each executive, as well as relevant financial information such as salary, bonus, other annual compensation, long-term incentive payouts, and all other compensation. Equity compensation information is also available through these filings including stock awards, restricted stock awards, number of options granted, grant date, and, where available, grant date present value.

In many instances, an estimate of stock price is not included in the reports of privately-owned firms in the sample. As a result, an estimation process is required. In order to assess the value of private equity shares and options, I use a "pseudo stock price" and an estimated volatility measure as variables in the Black-Scholes pricing model. Calculation of these variables is described below.

As a control group, I employ firms from the Execucomp database, which contains corresponding data for executive compensation from public equity firms. Because each of the privately-owned firms in my sample has public debt, I eliminate Execucomp firms without public debt, in order to isolate the differences in compensation due to equity status. I eliminate firms without a debt rating and rated firms without current public debt. The final control group contains 1,994 firm-year observations from 1992-2005.

In order to calculate the "pseudo stock price", I match each private equity observation with a public equity counterpart from the control group based on industry (four digit SIC code), year and total assets. For each match, I calculate the price to sales ratio and multiply this value by the private equity firm's sales to obtain a "pseudo stock price". This practice is substantiated by findings from Liu *et al.* (2002), Alford (1992), Beaver and Morse (1978), Nissim and Thomas (2002) and Hines (2011).[3] [4]

Table 1 Panel A contains univariate statistics for private and public equity firms associated with my primary sample. Six variables are components of compensation and (in log form) serve

[2] For each firm-year observation, compensation information for up to five executives is available for each firm year. I excluded compensation information for executives joining or leaving the company mid-period. The remaining executives are clustered so that each firm year includes a single observation in the regression tests.

[3] The "pseudo stock price" is an estimate. However, this practice does not introduce any systematic estimation errors in favor of finding results consistent with my hypotheses. In fact, any noise in the measure renders the detection of differences more difficult.

[4] Using the price-to-sales ratio is required in order to ensure that all "pseudo stock prices" have positive values.

as dependent variables in later regressions. Equity compensation includes stock options, restricted stock awards, and stock awards. Equity valuation is based on a pseudo stock price, derived using the price-to-sales ratio of a public company and matched on size, year, and industry. For option valuation, I use the Black-Scholes model, using the pseudo stock price and volatility[5] ratings from Execucomp as inputs. Other income components are reported herein as they appear in the ExecuComp Database for public equity firms and SEC filings for private equity firms.

Table 1: Descriptive Statistics for Firms with Privately-Owned Equity and Public Debt and the Corresponding Control Group of Firms with Publicly-Owned Equity and Public Debt

Panel A. Primary Sample

	Private Equity Descriptive Statistics					Public Equity Descriptive Statistics			
n=445	Salary	Bonus	Other	Equity	n=1,994	Salary	Bonus	Other	Equity
Mean	333.68	185.42	115.56	1,392.5	Mean	384.15	524.4	276.88	835.34
Stdev	419.49	711.51	375.3	16,779.09	Stdev	243.57	1,536.55	1,851.37	2,342.12
Q3	350	164.5	38.44	0	Q3	468.15	433	130.41	683.61
Median	236.42	57.5	9.47	0	Median	310	198.61	37.74	224.08
Q1	165.96	0	2.14	0	Q1	225	85.63	9.68	0

	Total Comp	Assets	Net Income	Leverage		Total Comp	Assets	Net Income	Leverage
Mean	2,027.17	743.64	-1.71	26.14	Mean	2,020.77	6,473.84	352.64	2.18
Stdev	16,868.46	748.01	66.05	519.93	Stdev	4,098.55	14,659.16	1,150.26	7.82
Q3	682.35	870	15	5.4	Q3	1,870.6	5,702.52	271	2.57
Median	365.9	490	2.07	-1.57	Median	912.41	2,386.77	93.99	1.58
Q1	239.52	290	-9.3	-3.94	Q1	501.4	1,102.79	22.82	1.03

Variable definitions:
Salary and Bonus = As reported by 10-K filings or Execucomp.
Other = Other compensation which includes Long Term Incentive Payouts and Other Compensation, generally perquisites.
Equity = Stock awards + restricted stock awards + option awards as reported by 10-K filings or Compustat.
Total Comp = Total annual compensation.
Assets and Net Income = As reported by Compustat.
Leverage = Total liabilities scaled by book value of equity, as reported by Compustat.
Note: All compensation data are reported in thousands of dollars. Net Income and Assets are in millions of dollars.

[5] The use of the same volatility measure for privately-owned firms helps to eliminate the possibility that my results are driven by changes in compensation contracts as a result of Sarbanes-Oxley (Narayanan and Seyhun, 2006; Cohen *et al.*, 2013).

Table 1: Descriptive Statistics for Firms with Privately-Owned Equity and Public Debt and the Corresponding Control Group of Firms with Publicly-Owned Equity and Public Debt: Continues

Panel B. Secondary Sample of Firms During Periods of Private and Public Equity with Public Debt

	Private Equity Descriptive Statistics					Public Equity Descriptive Statistics			
n=162	Salary	Bonus	Other	Equity	n=155	Salary	Bonus	Other	Equity
Mean	255.69	126.16	76.83	300.96	Mean	266.88	189.8	78.78	79.13
Stdev	160.02	406.87	428.49	7,653.61	Stdev	165.79	582.76	640.31	607.6
Q3	306.2	135	340	0	Q3	337	175	17.74	0.000018
Median	233.57	67.5	286.9	0	Median	229.52	81.15	5.5	0
Q1	160.94	0	233.8	0	Q1	165.96	20	0	0

	Total Comp	Assets	Net Income	Leverage		Total Comp	Assets	Net Income	Leverage
Mean	760.04	659.31	-4.16	81.81	Mean	535.52	25,673.7	153.58	-15.38
Stdev	7672	1,125.83	67.76	943.91	Stdev	952.43	98,009.92	411.93	96.93
Q3	479.08	536.48	11.42	3.77	Q3	531.48	1,834.79	136.52	1.41
Median	318.83	358.15	0.29	-1.95	Median	338.71	82.44	1.77	0.57
Q1	213.75	203.2	-26.56	-5.39	Q1	217.9	4.47	-1.3	0.12

Contrary to expectations, the Equity Compensation is smaller for public equity firms than for private equity firms ($0.84 million versus $1.39 million, respectively). Examination of the distributions indicates that the mean of equity compensation for private equity firms is driven by a few extreme observations. These are Winsorized at the 1st and 99th percentiles prior to performing regression analysis. The distribution of total compensation is also skewed as a result. Other univariate statistics are generally in line with expectations.

Table 2: Correlation Matrix

N=2,439	Private	Equity	Bonus	Salary	Other	Total Comp	Assets	ROA	Leverage	Loss
Private		-0.48	-0.2	-0.17	-0.21	-0.32	-0.45	-0.23	0.05	0.25
Equity	-0.43		0.22	0.24	0.23	0.6	0.36	0.18	-0.01	-0.17
Bonus	-0.28	0.44		0.22	0.18	0.37	0.24	0.27	0	-0.27
Salary	-0.23	0.48	0.58		0.2	0.52	0.32	0.07	0	-0.07
Other	-0.23	0.31	0.34	0.43		0.42	0.33	0.07	0	-0.08
Total Comp	-0.36	0.76	0.72	0.8	0.54		0.52	0.17	-0.02	-0.15
Assets	-0.46	0.43	0.46	0.53	0.41	0.57		0.1	-0.02	-0.13

Table 2: Correlation Matrix: Continues

ROA	-0.32	0.24	0.32	0.1	0.09	0.24	0.12		-0.29	-0.65
Leverage	0.1	-0.07	-0.05	-0.01	0.01	-0.05	0.09	0		-0.01
Loss	0.25	-0.17	-0.25	-0.08	-0.07	-0.16	-0.13	0.37	0.06	

Scores reported above the diagonal are Pearson correlation coefficients and scores below the diagonal are Spearman rank correlation coefficients.
Variable definitions:
Private: Indicator variable equal to 1 if equity is privately owned, else 0.
ROA: Return on assets. Calculated as Net income divided by total assets as reported by Compustat.
Loss: Indicator variable equal to 1 if net income is negative, else 0.
Other variables defined in Table 1.

Table 2 presents the correlation matrix for the relevant variables as they appear in the subsequent multivariate regressions. Scores reported above the diagonal are Pearson correlation coefficients while scores below the diagonal reflect Spearman rank correlation coefficients. Several of the compensation-related variables have high correlations, such as the log of equity compensation and the log of total compensation (0.76 Pearson, 0.60 Spearman), probably due to the fact that equity compensation is a component in total compensation. In assessing *H1* and *H2*, assessment of the correlations show that private ownership is negatively associated with total compensation (-.32 Pearson, -.36 Spearman) and with equity compensation (-.48 Pearson, -.43 Spearman), suggesting that private firms offer less total compensation and less equity compensation. The log of assets and private equity are negatively correlated (-0.45 Pearson, -0.46 Spearman), advising that private equity firms in the sample are smaller. Non-reported variance inflation scores were low enough (generally between one and two) to dismiss concerns about multicollinearity.

B. Secondary Samples

I employ three distinct sets of secondary tests of explanations for differences in compensation between public equity firms and private equity firms. To test the first, I use the ownership information collected with my primary sample. To test the second and third explanations (whether equity valuation and liquidity drives differences in compensation between publicly-owned and privately-owned firms and whether superior monitoring in privately-owned firms reduces equity compensation), I collect a sample of firms that either 'went public' or 'went private' while maintaining public debt and thus SEC registration. This "gone public/private" sample consists of 43 firms that either "went public" or "went private" between 1992 and 2007 while maintaining public debt. This sample permits firms to serve as their own controls. The sample has 317 firm years, each containing the compensation information for approximately five executives for a total of 162 firm-year private equity observations and 155 firm-year public equity observations. Firms meeting this description are not common and this sample represents approximately 50 percent of the population of firms of this specific type. Table 1 Panel B shows the descriptive statistics for this secondary sample. These statistics are grouped by equity status of the firms—periods of private equity and periods of public equity. As with the primary sample, the two years prior to an IPO are excluded.

The descriptive statistics of these 317 firm-year observations show the averages of relevant variables for firm years corresponding to privately-owned equity and when ownership of equity is publicly owned. As with the primary sample, observing the non-zero equity compensation at the

75th percentile shows that a greater percentage of public equity firms employ equity-based compensation and the high mean for privately-traded firms is likely due to extreme observations. As with the primary tests, observations are Winsorized at the 1st and 99th percentiles. In other respects, the differences are similar generally to those observed using the primary sample.

The majority of firms with privately-traded equity do not file a proxy statement, which eliminates many of the traditionally available monitoring proxies. In the absence of these measures, I use earnings management as a proxy for monitoring—I interpret a high incidence of earnings management as evidence of poor monitoring. Thus, following Givoly *et al.* (2010) and Burgstahler and Dichev (1997)[6], I plot frequency distributions using my primary sample and its corresponding control group to test for earnings management by measuring the number of unexpected observations just above and below zero. I interpret a trend indicating aggregate avoidance of reporting small losses or a small negative change in earnings as evidence of earnings management. For this test, firms with missing Compustat values for net income and book value of equity are also eliminated.

To evaluate monitoring, I use two additional measures. From the gone public/private sample, I obtain the number of owners holding more than five percent of the total shares outstanding from 10-K filings. This measure assumes that investors holding at least five percent of outstanding shares are likely to more closely monitor the activities of managers than are atomistic shareholders. Finally, where available, I collect the number of board meetings held for each firm each year from the proxy statements of firms in the gone public/private sample. This measure assumes that a greater number of meetings is positively associated with superior monitoring.[7]

IV. Results

A. Primary Tests: Tests of Differences in Compensation

In order to test my hypotheses, I employ the following regression equation beginning in Table 3:

$$Log(x_i) = \beta_0 + \beta_1 Private_{jt} + \beta_2 Assets_{jt} + \beta_3 ROA_{jt} + \beta_4 Loss_{jt} + \beta_5 Private*ROA_{jt} + \beta_6 Private*Loss_{jt} + \beta_7 Leverage_{jt} + \beta_{8\text{-}36} Industry_{jt} + \beta_{37\text{-}49} Year_{jt} + \varepsilon \quad (1)$$

x_i = a component of compensation and subscripts *j* and *t* indicate firm and year.
Private = 1 if the firm's equity is privately traded and is equal to 0 otherwise.
Assets = a firm's assets, which is inserted in log form into the regression, *ROA* is return on assets.
Loss = 1 if the firm's reported net income in year *t* is negative and is equal to 0 otherwise.
Leverage = measured as total liabilities divided by book value of equity.

[6] This measure is not without limitations (see Beaver *et al.*, 2007). However, the limitations of other potential measures of earnings management, such as the various forms of the Jones model, are also well documented (see Dechow *et al.*, 1995, Kothari *et al.*, 2005).
[7] One limitation of this measure is that it fails to capture informal meetings that may be held by board members.

Table 3: Tests for Differences in Levels of Executive Compensation in Firms Whose Equity is Privately Owned Versus Firms Whose Equity is Publicly Owned, Both with Public Debt

Panel A. All Executives Listed in the 10-K

$$Log(x_i) = \beta_0 + \beta_1 Private_{jt} + \beta 2 Assetsjt + \beta_3 ROA_{jt} + \beta_4 Loss_{jt} + \beta_5 Private*ROA_{jt} + \beta_6 Private*Loss_{jt} + \beta_7 Leverage_{jt} + \beta_{8\text{-}36} Industry_{jt} + \beta_{37\text{-}49} Year_{jt} + \varepsilon$$

N=2,439	Intercept	Private	Assets	ROA	Loss	Private*ROA	Private*Loss	Leverage	Adj. R^2
X1=Salary	9.36	-0.1	0.24	354.91	0.02	1042.41	-0.06	0	47.56
t-stat	29.63	-0.82	11.46***	2.1**	0.56	1.21	-0.65	-1.76*	
X2=Bonus	6.24	-1.69	0.44	4205	-1.1	7840.33	0.24	0	20.61
t-stat	3.17	-2.3**	0.327	2.59***	-3.91***	1.83*	0.39	0.96	
X3=Other	1.94	-0.38	0.6	1289.22	0.46	1182.65	-0.38	0	28.79
t-stat	1.55	-0.8	7.33***	1.49	2.74***	0.39	-1.08	-0.28	
X4=Equity	-1.58	-8.79	1.06	1044.47	-0.93	4467.06	0.56	0	58.54
t-stat	-0.57	-8.57***	5.71***	0.43	-2.25**	0.76	0.88	-1.08	
X5= Total Comp	1.84	-4.46	0.83	1594.85	-0.43	2322.33	0.23	0	62.09
t-stat	1.33	-8.6***	9.03***	1.35	-2.04**	0.77	0.68	-1.91*	

See Table 1 and Table 2 for variable definitions.

Panel B. Chief Executive Officers

$$Log(x_i) = \beta_0 + \beta_1 Private_{jt} + \beta 2 Assetsjt + \beta_3 ROA_{jt} + \beta_4 Loss_{jt} + \beta_5 Private*ROA_{jt} + \beta_6 Private*Loss_{jt} + \beta_7 Leverage_{jt} + \beta_{8\text{-}36} Industry_{jt} + \beta_{37\text{-}49} Year_{jt} + \varepsilon$$

N=530	Intercept	Private	Assets	ROA	Loss	Private*ROA	Private*Loss	Leverage	Adj. R^2
X1=Salary	9.68	0.01	0.23	665.19	0.1	-58.02	-0.09	0	72.76
t-stat	18.81	0.05	7.87***	1.57	1.21	-0.07	-0.91	-0.92	
X2=Bonus	13.72	-4.18	0.39	-2083	1.04	26,219	-1.69	0	36.84
t-stat	2.05	-2.24**	1.05	-0.38	1	2.52**	-1.38	0.68	
X3=Other	10.61	1.91	0.05	3,652.38	1.13	-12,816	-1.9	0	35.27
t-stat	3.86	2.48**	0.31	1.62	1.65*	-3***	-3.79***	1.72*	
X4=Equity	-16.24	-10.5	1.72	-18,590	-1.36	25,995	1.6	0	63.58
t-stat	-1.85	-4.27***	3.48***	-2.57**	-1	1.9*	0.14	-3.12***	
X5= Total Comp	-5.22	-5.37	1.21	-7,733.3	-0.79	8,421.5	0.02	0	64.57
t-stat	-1.07	-3.93***	4.4***	-1.92*	-1.04	1.11	0.02	-2.59***	

See Table 1 and Table 2 for variable definitions

Table 3 Panel A shows the results of this test using this regression equation for all executives in the primary sample.[8] The primary variable of interest is *Private*. In this specification, we are able to test both *H1*, that total compensation is the same for managers in both privately- and publicly-owned firms, and *H2*, that equity compensation is the same for managers in both privately- and publicly-owned firms. These results show that executives in privately-owned firms receive significantly less bonus compensation (*t*-stat = 2.3, *p*-value < .05), less equity compensation (*t*-stat -8.57, *p*-value < .01) and less total compensation (*t*-stat -4.46, *p*-value <.01). Table 3 Panel B tests the same hypotheses while restricting the sample to CEOs. The inferences are identical with one exception—that private firm CEOs receive more perquisite compensation (labeled other) than public firm CEOs. These results suggest that *H1* and *H2* should be rejected.

Table 4: Tests for Relative Differences in Executive Compensation in Firms Whose Equity is Privately Owned Versus Firms Whose Equity is Publicly Owned, Both with Public Debt

Panel A. All Executives Listed in the 10-K

$$Log(x_i) = \beta_0 + \beta_1 Private_{jt} + \beta 2 Assetsjt + \beta_3 ROA_{jt} + \beta_4 Loss_{jt} + \beta_5 Private*ROA_{jt} + \beta_6 Private*Loss_{jt} + \beta_7 Leverage_{jt} + \beta_{8\text{-}36} Industry_{jt} + \beta_{37\text{-}49} Year_{jt} + \varepsilon$$

N=2,439	Intercept	Private	Assets	ROA	Loss	Private*ROA	Private*Loss	Leverage	Adj. R^2
X1=Salary/Total Comp	0.76	0.39	-0.05	-217.45	0.01	48.66	-0.01	0	37.9
t-stat	4.41	7***	-3.76***	-1.49	0.45	0.13	-0.21	0.69	
X2=Bonus/Total Comp	0.11	0.1	-0.01	47	-0.01	104.13	-0.02	0	10.98
t-stat	1.04	2.49**	-0.79	0.55	-0.91	0.42	-0.88	-0.34	
X3=Other /Total Comp	0.03	0.04	0	-393.15	-0.03	-7.26	0.04	0	6.89
t-stat	0.16	0.6	0.13	-2.61***	-1.14	-0.02	0.89	-0.66	
X4=Equity/Total Comp	0.21	-0.55	0.05	173.71	-0.01	-87.73	0.01	0	32.61
t-stat	0.78	-7.06***	2.56**	0.73	-0.2	-0.19	0.18	-0.31	

See Table 1 and Table 2 for variable definitions.

[8] The reader should exercise caution in interpreting parameter estimates due to the log transformation of the dependent variable in each specification.

Table 4: Tests for Relative Differences in Executive Compensatio in Firms Whose Equity is Privately Owned Versus Firms Whose Equity is Publicly Owned, Both with Public Debt: Continues

Panel B. Chief Executive Officers

$$Log(x_i) = \beta_0 + \beta_1 Private_{jt} + \beta 2 Assetsjt + \beta_3 ROA_{jt} + \beta_4 Loss_{jt} + \beta_5 Private*ROA_{jt} + \beta_6 Private*Loss_{jt} + \beta_7 Leverage_{jt} + \beta_{8\text{-}36} Industry_{jt} + \beta_{37\text{-}49} Year_{jt} + \varepsilon$$

N=530	Intercept	Private	Assets	ROA	Loss	Private*ROA	Private*Loss	Leverage	Adj. R^2
X1=Salary/Total Comp	0.31	0.46	-0.03	1,167.92	0.12	-757.62	0.57	0	58.01
t-stat	0.62	3.32***	-1.24	2.85***	1.57	-0.98	0.63	2.48**	
X2=Bonus/Total Comp	0.19	0.38	0	1,551.65	0.15	-1,500.53	-0.18	0	47.56
t-stat	0.47	3.43***	0.04	4.71***	2.39**	-2.41**	-2.51**	1.63	
X3=Other /Total Comp	0.49	0.13	-0.03	-38.34	0.04	-703.23	-0.05	0	49.04
t-stat	2.38	2.24**	-2.46**	-0.23	1.27	-2.2**	-1.36	2.22**	
X4=Equity/Total Comp	0.01	-0.98	0.06	-2,681.53	-0.31	2,969.91	0.18	0	65.35
t-stat	0.02	-5.52***	1.75*	-5.15***	-3.16***	3.02***	1.54	-3.71***	

See Table 1 and Table 2 for variable definitions.

Table 4 repeats the above tests with each compensation component scaled by total compensation. This specification allows for a comparison of compensation practices as a percentage of total. This controls for the possibility that differences observed in Table 3 may be driven by a significant difference in total compensation between public and private firms in the sample. Hypothesis 1 is not testable in this specification. In Panel A, with all executives included, we observe that private firm executives receive more salary as a percentage of total income (*t*-stat 7, *p*-value < .01), more bonus as a percentage of total income (*t*-stat 2.49, *p*-value < .05) and less equity compensation (*t*-stat -7.06, *p*-value < .01). Panel B restricts the sample to CEOs. Again, the only qualitative change from Panel A is that private CEOs appear to earn more perquisite compensation (*t*-stat 2.24, *p*-value < .05).

In sum, these results show that both *H1 and H2* should be rejected. That is, private firms offer less total compensation primarily as a result of offering less equity based compensation to their executives. The remainder of this paper is dedicated to tests of the underlying determinants of this difference in equity-based compensation.

B. Secondary Tests: Tests of Determinants of Equity Compensation Differences

Explanation 1: Private equity firm managers are more vested in the firm they manage.

Consistent with the previous discussion, although annual equity compensation is lower for private equity firms, I expect that executives of private equity firms will be more heavily invested in the firms they manage. In the absence of data on managers' total wealth, I use total number of shares owned and total number of shares owned as a percentage of total shares outstanding for the firm. Thus, I expect that executives of privately-held firms are similar to executives of family-founded firms in that I expect them to own more shares of the firms they manage than managers of public equity firms.

Table 5: Share Ownership Comparison Between Public Equity and Private Equity Firms, Both with Public Debt

Panel A. All Executives Listed in the 10-K

$$Log(x_i) = \beta_0 + \beta_1 Private_{jt} + \beta 2 Assetsjt + \beta_3 ROA_{jt} + \beta_4 Lagged\ ROA_{jt} + \beta_5 Leverage_{jt} + Industry_{jt} + Year_{jt} + \varepsilon$$

N=2,439	Intercept	Private	Log(Assets)	ROA	Lagged ROA	Leverage
Shares Owned =	3,833,906	1,625,177	416,489	1,028,162	-1,028,162	19.82
t-stat	-3.58***	4.86***	3.86***	-0.51	0.54	0.06
Shares owned/shares outstanding	27.81	7.69	-2.57	9.23	22.33	0
t-stat	1.51	1.27	-1.26	0.3	0.75	-0.43

Panel B. Chief Executive Officers

N=530	Intercept	Private	Log(Assets)	ROA	Lagged ROA	Leverage
Shares Owned =	-801,375	476,874	102,326	-695,933	295,161	-4.74
t-stat	-2.81***	5.08***	3.51***	-1.13	0.53	-0.01
Shares owned/shares outstanding	33.88	6.3	-4.59	48.91	5.27	0
t-stat	1.36	0.61	-1.7*	0.3	0.91	-0.69

Variable Definitions:
Shares owned = Total shares owned of any type as reported in 10-K filings and Execucomp.
Shares owned/Shares outstanding = Shares owned scaled by total shares outstanding as reported in 10-K filings and Execucomp.
Lagged ROA = Calculated as Net income for firm *j* in year *t* divided by total assets for firm *j* in year *t*-1 as reported by Compustat.
See Table 1 and Table 2 for other variable definitions.

Table 5 tests *Explanation 1* using the primary sample executives. Data were obtained using the stock ownership information found in the relevant 10-K filing for private equity firms and as reported in Execucomp for public equity firms. The variable of interest continues to be the *Private* indicator. Table 5 demonstrates that private equity executives hold more shares of the firm they manage (*t*-stat 4.86, *p*-value < .01). However, as a percentage of total shares outstanding, the difference is no longer statistically significant (*t*-stat 1.17, *p*-value > .1). The same is true for CEOs. These findings suggest that private equity firm executives do not own a significantly different percentage of outstanding equity than do public equity executives. These results do not support

the first Explanation for compensation differences - that managers of private equity firms are more economically vested in the firm they manage than public equity executives.

Explanation 2: Inherent difficulties associated with valuing and/or liquidating privately held stock causes private firm managers to receive less equity-based compensation.

The second Explanation for differences in executive compensation between publicly-traded and privately-held firms is the difficulty in valuing or liquidating shares of stock. This leads private equity firms to substitute other forms of compensation for equity-based compensation. In terms of liquidity, some private equity firms may offer to buy back shares of stock owned by employees, creating an outlet for current and departing employees to trade stock and stock options for cash. However, the frequency of such agreements is unobservable and low in expectation. I predict that the challenges associated with valuing and liquidating private firms' equity is the primary force behind differences in equity compensation.

Table 6: Tests for Differences in Executive Compensation During Periods of Private Equity Ownership Versus Periods of Public Equity Ownership for Firms with Public Debt

Panel A. All Executives Listed in the 10-K, in Levels

$$Log(x_i) = \beta_0 + \beta_1 Private_{jt} + \beta 2 Assetsjt + \beta_3 ROA_{jt} + \beta_4 Loss_{jt} + \beta_5 Private*ROA_{jt} + \beta_6 Private*Loss_{jt} + \beta_7 Leverage_{jt} + Industry_{jt} + Year_{jt} + \varepsilon$$

n=317	Intercept	Private	Assets	ROA	Loss	Private*ROA	Private*Loss	Leverage	Adj. R^2
Equity	-1.58	-0.75	-0.24	0.25	0.51	0	-0.77	0.51	60.47
t-stat	-5.51	-1.71*	8.64***	1.56	0.03	-2.28**	-3.17***	3.64***	

Panel B. Chief Executive Officers, in Levels

n=232	Intercept	Private	Assets	ROA	Loss	Private*ROA	Private*Loss	Leverage	Adj. R^2
Equity	1.32	-1.88	0.26	-0.44	-0.1	-0.18	-0.9	-0.16	51.57
t-stat	2.88	-0.88	3.5***	-0.13	-0.7	-0.67	-0.4	-0.34	

Panel C. All Executives Listed in the 10-K, Scaled by Total Compensation

n=317	Intercept	Private	Assets	ROA	Loss	Private*ROA	Private*Loss	Leverage	Adj. R^2
X4=Equity/Total Comp	-0.1	-0.02	0.02	0.09	0.01	-0.24	-0.03	0	11.45
t-stat	-3.75	-2.03**	5.21***	1.84*	1.12	-3.01***	-2.7**	-0.53	

Panel D. Chief Executive Officers, Scaled by Total Compensation

n=232	Intercept	Private	Assets	ROA	Loss	Private*ROA	Private*Loss	Leverage	Adj. R^2
X4=Equity/Total Comp	-0.05	-0.03	0.02	0.02	-0.02	-0.58	-0.05	0	8.15
t-stat	-2.06	-1.62	4.74***	0.37	-1.5	-1.91*	-1.75*	-1.66*	

See Table 1 and Table 2 for variable definitions.

Using my 'gone public/private' sample, I test to see whether equity compensation is significantly greater during periods of public equity. I use the same multiple regression model as with the primary sample, with one difference.[9] The results are found in Table 6. In levels, the amount of equity compensation is lower for private equity firms, significant at the five percent level in a one-tailed test. As a percentage of total compensation, equity compensation is again significantly lower for private equity firms. Restricting the sample to CEOs, the difference again becomes insignificant, presumably due to sample size.

A change in equity status, either from private to public or vice versa, captures the effect of a change in liquidity and valuation for equity shares. However, such a change in equity status also captures a change in monitoring to the extent that differences exist. That is, when a firm "goes public" or "goes private", not only does the ease of valuing and/or liquidating stock change, so may the quality of monitoring as ownership changes. As a result, I employ a third set of tests designed to specifically test for differences in monitoring between these two groups.

Explanation 3: Private equity firms are superior monitors

The third Explanation posits that, due to more concentrated ownership in firms with privately-held equity, superior monitoring substitutes for other incentive-aligning mechanisms, such as equity compensation. In order to determine if this effect drives the results found in Table 6, I conduct tests to determine whether private equity firms appear to have superior monitoring. Because many traditional measures of monitoring are unobservable, I use earnings management, which is observable using my data. Earnings management is one potential result of inferior monitoring. Using earnings management as a test of monitoring is predicated on the assumption that superior monitors will exert influence on managers to use their reporting discretion to maximize the reliability and transparency of financial statements and subsequently, earnings. This approach further assumes that managers, in the absence of monitoring, will always manage earnings in the presence of an opportunity to do so. Using the same secondary sample of 43 firms used to test *Explanation 2*, I test for earnings management following Givoly *et al.* (2010). In the absence of earnings management, a distribution of firms' reported earnings should be approximately normal. However, in the presence of earnings management, the distribution may be distorted such that an unexpectedly low number of firms report small losses and an unexpectedly high number of firms report small gains (see Burgstahler and Dichev, 1997).

[9] In order to avoid over fitting with a small sample, I exclude control variables that were statistically insignificant. The associated F-change statistic was insignificant.

Table 7: Tests for Differences in Monitoring During Periods of Private Equity Ownership Versus Periods of Public Equity Ownership for Firms with Public Debt

Panel A. Frequency Distribution of Earnings Around Zero[a]

	N	Number "Just Below Zero"[b]	Expected[c]	Standardized Difference[d]	Number "Just Above Zero"[b]	Expected[c]	Standardized Difference[d]
Private Equity Firms	162	10	22.5	-2.18	38	18.5	3.34
Public Equity Firms	155	10	12.5	-0.49	49	19	4.92

Panel B. Number of Owners Holding at Least 5 Percent of Outstanding Equity

	N	Mean	Std Dev	t-stat
Private Equity Firms	162	4.34	2.69	1.45
Public Equity Firms	155	3.81	3.01	

Panel C. Number of Board Meetings Per Year

	N	Mean	Std Dev	t-stat
Private Equity Firms	6	4.3	3.07	4.38*
Public Equity Firms	100	6.3	0.82	

*Satterthwaite adjustment

[a]In Panel A, the distribution of net income in year *t* scaled by total assets at the end of year *t*-1 (Income/Assets) is measured to assess earnings management around zero.

[b]"just below zero" and "just above zero" refer to intervals. Intervals, or bin widths, are calculated following Degeorge *et al.* (1999), as 2*2(IQR)n^(1/3), where IQR is the sample inter-quartile range and *n* is the number of observations. The resulting bin widths for the distribution of Income/Assets are .052 for private equity firms and .058 for public equity firms.

[c]The expected frequency in the interval is calculated as the average of the number of observations observed in the adjacent intervals on each side.

[d]The standardized difference is the difference between the actual frequency and expected frequency, divided by the standard deviation of the difference. The standard deviation of the difference is computed, following Burgstahler and Dichev (1997), as $[n*P_i*(1- P_i) + .25*n*(P_{i-1} + P_{i+1}) *(1-P_{i-1} + 1- P_{i+1})]^{.5}$, where *n* is the number of observations and P_i is the probability that an observation will fall into interval *i*.

The results of this test of differences in earnings management are reported in Table 7, Panel A. For private firms, the number of firms reporting earnings "just below zero" is significantly lower than expected, while for public firms, the number of firms reporting earnings "just below zero" was not significantly lower than expected. For the "just above zero" bin, both public and private equity firms exhibit frequencies that were significantly higher than expected. Together, this provides evidence of earnings management during periods in which equity is privately owned as well as periods of public ownership. That private firms do not manage earnings less than public firms is substantiated by Burgstahler *et al.* (2006). These results do not support *Explanation 3*.

I use two other tests designed to detect differences in monitoring using this sample. According to Hill and Jones (1992), larger stockholders are likely to be better monitors than atomistic ones. Based on this theory, if private firms are superior monitors, they are likely to have

more large stakeholders than public firms. Thus, I test for differences in the number of large shareholders during periods of private equity versus periods of public equity. For purposes of this test, I define a large shareholder as one owning at least five percent of total shares outstanding, regardless of class. The results of this test are reported in Table 7 Panel B. These results provide no evidence of a difference in the number of large shareholders between publicly- and privately-owned firms.

Superior monitoring also seems likely to be positively correlated with the number of board meetings occurring during the year. Thus, I also test for differences in the number of board meetings during periods of private ownership and periods of public ownership in Table 7 Panel B. Though of limited generalizability due to data availability, the results of this test suggest that during periods of private ownership, firms hold significantly fewer board meetings than during periods of public ownership. Together, these results do not provide support for *Explanation 3*.

V. Conclusion

In summary, I find support for my first hypothesis that privately-owned companies pay their executives less total compensation. I also find strong support for my second hypothesis—that executives of privately-owned corporations receive less equity-based compensation. As this difference in equity compensation appears to be the primary difference in compensation practices between publicly- and privately-owned firms, I test for the determinants of these differences. In subsequent tests, I find that the difference in equity compensation is due to difficulty in value/liquidity associated with equity compensation in a privately-held corporation. Based on prior research, this difference in the form of compensation may result in less risky behavior on the part of private firm managers (Sanders, 2001; Sanders and Hambrick, 2007).

Future research may explore the differences in incentive-based compensation between publicly- and privately-owned firms, especially perquisites. Future researchers could improve generalizability further by examining the compensation arrangements of firms whose equity is privately owned and whose debt is also owned privately. Additionally, researchers might compare the long-term performance and decision making of private firm managers compared to public ones.

REFERENCES

Alford, Andrew W. 1992. "The Effect of the Set of Comparable Firms on the Accuracy of the Price-Earnings Valuation Method." *Journal of Accounting Research,* 30(1): 94-108.

Ali, Ashiq, Tai-Yuan Chen, and Suresh Radhakrishnan. 2007. "Corporate Disclosures by Family Firms." *Journal of Accounting and Economics,* 44(1-2): 238-86.

Ali, Ashiq, Tai-Yuan Chen, and Ramachandran Natarajan. 2006. "Compensation Contracts of Professional CEOs of Family Firms." Working Paper, The University of Texas at Dallas**.**

Anderson, Ronald C., and David M. Reeb. 2003. "Founding-Family Ownership and Firm Performance: Evidence from the S&P 500." *Journal of Finance,* 58(3): 1301-28.

Ball, Ray and Lakshmanan Shivakumar. 2005. "Earnings Quality in UK Private Firms: Comparative Loss Recognition Timeliness." *Journal of Accounting and Economics,* 39(1): 83-128.

Beatty, Anne, and David G. Harris. 1999. "The Effects of Taxes, Agency Costs and Information Asymmetry on Earnings Management: A Comparison of Public and Private Firms." *Review of Accounting Studies,* 4(3): 299-326.

Beatty, Anne L., Bin Ke, and Kathy R. Petroni. 2002. "Earnings Management to Avoid Earnings Declines Across Publicly and Privately Held Banks." *The Accounting Review,* 77(3): 547-70.

Beaver, William, and Dale Morse. 1978. "What Determines Price-Earnings Ratios?" *Financial Analysts Journal*, 34(4): 65-76.

Beaver, William, Maureen F. McNichols, and Karen K. Nelson. 2007. "An Alternative Interpretation of the Discontinuity in Earnings Distributions." *Review of Accounting Studies*, 12(4): 525-56.

Burgstahler, David, and Ilia Dichev. 1997. "Earnings Management to Avoid Earnings Decreases and Losses*." Journal of Accounting and Economics,* 24 (1): 99-126.

Burgstahler, David C., Luzi Hail and Christian Leuz. 2006. "The Importance of Reporting Incentives: Earnings Management in European Private and Public Firms." *The Accounting Review*, 81(5): 983-1016.

Cohen, Daniel A., Aiyesha Dey, and Thomas Z. Lys. 2013. "Corporate Governance Reform and Executive Incentives: Implications for Investments and Risk Taking." *Contemporary Accounting Research*, 30(4), 1296-1332.

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney. 1995. "Detecting Earnings Management." *The Accounting Review*, 70(2): 193-225.

Demsetz, Harold, and Kenneth Lehn. 1985. "The Structure of Corporate Ownership: Causes and Consequences." *Journal of Political Economy*, 93(6): 1155-77.

Degeorge, François, Jayendu Patel, and Richard Zeckhauser. 1999. "Earnings Management to Exceed Thresholds." *The Journal of Business*, 72(1):1-33.

Givoly, Dan, Carla K. Hayn, and Sharon P. Katz. 2010. "Does Public Ownership of Equity Improve Earnings Quality?" *The Accounting Review*, 85(1): 195-225.

Hall, Brian J., and Kevin J. Murphy. 2003. "The Trouble With Stock Options." *Journal of Economic Perspectives*, 17(3): 49-70.

Hill, Charles W. L., and Thomas M. Jones. 1992. "Stakeholder-Agency Theory." *Journal of Management Studies*, 29(2): 131-54.

Hines, Christopher. 2011. "Determinants and Consequences of Risk Management Committee Formation." Working paper, University of Arkansas.

James, Harvey S. 1999. "Owner as a Manager, Extended Horizons and the Family Firm." *International Journal of the Economics of Business*, 6(1): 41-55.

Kaplan, Steven N., and Per Strömberg. 2009. "Leveraged Buyouts and Private Equity." *Journal of Economic Perspectives*, 23(1): 121-46.

Ke, Bin, Kathy Petroni, and A. Safieddine. 1999. "Ownership Concentration and Sensitivity of Executive Pay to Accounting Performance Measures: Evidence from Publicly and Privately-Held Insurance Companies." *Journal of Accounting and Economics*, 28(2): 185-209.

Kothari, S. P., Andrew J. Leone, and Charles E. Wasley. 2005. "Performance Matched Discretionary Accrual Measures." *Journal of Accounting and Economics*, 39(1): 163-97.

Kwak, Mary. 2003. "The Advantages of Family Ownership." *MIT Sloan Management Review*, 44(2): 12.

Liu, Jing, Doron Nissim, and Jacob Thomas. 2002. "Equity Valuation Using Multiples." *Journal of Accounting Research*, 40(1): 135-72.

Narayanan, M. P., and H. Nejat Seyhun. 2006. "Effect of Sarbanes-Oxley Act on the Influencing of Executive Compensation.*"* Working Paper, University of Michigan.

Penno, Mark, and Daniel T. Simon. 1986. "Accounting Choices: Public Versus Private Firms." *Journal of Business Finance and Accounting*, 13(4): 561-69.

Sanders, William Gerard. 2001. "Behavioral Responses of CEOs to Stock Ownership and Stock Option Pay." *Academy of Management Journal*, 44(3): 477-92.

Sanders, William Gerard, and Donald C. Hambrick. 2007. "Swinging for the Fences: The Effects of CEO Stock Options on Company Risk Taking and Performance." *Academy of Management Journal*, 50(5): 1055-78.

Shleifer, Andrei, and Robert W. Vishny. 1997. "A Survey of Corporate Governance." *Journal of Finance*, 52(2): 737-83.

Stein, Jeremy C. 1988. "Takeover Threats and Managerial Myopia." *Journal of Political Economy*, 96(1): 61-80.

Stein, Jeremy C. 1989. "Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior." *Quarterly Journal of Economics*, 104(4): 655-69.

The Journal of Business Inquiry 2014, 13, Issue 2, 101-127
http:www.uvu.edu/woodbury/jbi/articles/
ISSN 2155-4072

Labor Income Tax and Output: A Structural VAR Analysis

By SANJIB SARKER, BASUDEB BISWAS, AND PETER J. SAUNDERS*

This paper analyzes two channels through which a change in labor income tax may affect output. First, a tax cut provides higher work incentives, thereby increasing the aggregate output through an increase in the aggregate labor supply. Second, a tax-cut increases disposable income and the aggregate demand. An increase in the aggregate demand leads to a higher level of aggregate output. The first channel is believed to have a permanent effect on output movements, while the latter has only a temporary effect. This paper captures these two effects by defining two disturbances on the basis of the existing economic theory.

Keywords: Labor Income Tax, Output Growth, Structural VAR Model, Blanchard-Quah Decomposition, Supply-Side Effect, Impulse Response Functions

JEL Classification: C32, E01, E62

I. Introduction

The U.S. economy experienced a deep economic recession that began in 2008. To date, the economic recovery from this recession has been unusually slow as evidenced from the relatively low growth rate of the real GDP and a persistently high rate of unemployment. This anemic economic growth has renewed interest in analyzing the effectiveness of fiscal policy in restoring economic growth. Fiscal policy can either be focused on stimulating the aggregate demand, or it can be designed to affect primarily the aggregate supply of the economy subjected to a fiscal stimulus. In the U.S., the objective of the unprecedentedly large fiscal stimulus that began in 2009 has been on increasing the U.S. economy's aggregate demand. Yet in spite of all massive recent fiscal expenditures, the U.S. real GDP growth rate averaging 2.5 percent in 2011 and 1.9 percent in the first quarter of 2012 leaves many doubts about the effectiveness of fiscal policy that targets the aggregate demand in restoring economic growth. The key objective of this paper is to provide empirical evidence on the effectiveness of fiscal policy in promoting economic growth. This objective is achieved by developing a structural vector autoregressive (SVAR) model and subjecting it to empirical tests. Our paper analyses both the aggregate demand and the aggregate supply channels by which fiscal policy affects economic growth.

A large number of studies have analyzed the role of fiscal policy on aggregate economic activity. Empirical studies have been increasingly using SVAR models in tracking the dynamics of output response to unanticipated fiscal policy shocks. For example, Galí *et al.* (2007) study the effects of

* Sanjib Sarker: Risk Modeling Analyst, JPMorgan Chase**, Columbus, OH. E-mail: sanjsar@gmail.com. Basudeb Biswas: Professor, Far East University, Gamgok-nyeon, South Korea and Emeritus Professor of Economics and Finance, Huntsman School of Business, Utah State University. E-mail: basudeb.biswas10@gmail.com. Peter J. Saunders: Professor, Department of Economics, College of Business, Central Washington University, 400 E. University Way, Ellensburg, WA 98926-7486. E-mail: saunders@cwu.edu.

** This paper is an outgrowth of the Ph. D. dissertation by Sanjib Sarker completed at Utah State University. None of the views, opinions, data, facts and figures expressed in this paper are related to the bank.

government spending shocks to several macroeconomic variables in a new Keynesian framework. Ramey and Shapiro (1997) investigate the effects of military buildups on a variety of macroeconomic variables in a multi-sector neoclassical framework. Edelberg *et al.* (1999) examine the effects of exogenous shocks to real government spending on the U.S. output. Clarida and Prendergast (1999) present some empirical findings on the dynamic effects of fiscal policy on real exchange rates in the G3 countries. Blanchard and Perotti (2002) estimate the dynamic effects of shocks to government spending and taxes on the postwar U.S. output. Fatas and Mihov (2000) analyze the role of exogenous shocks to government spending on U.S. output using an identified VAR system. Perotti (2002) investigates the effects of fiscal policy on GDP, interest rate, and prices in five OECD (Organization for Economic Cooperation and Development) countries. Although numerous studies have examined the effects of aggregate government spending on output, few studies have attempted to examine empirically the impact of other fiscal instruments, namely labor income taxation, on output. This paper provides new empirical evidence on this key economic issue by examining the dynamic response of output to exogenous shocks of labor income tax policy innovations in the U.S. economy during 1979-2006. Providing new empirical evidence on effects of income taxes on economic growth is of particular importance in the current world-wide economic decline.

In Keynesian framework, a reduction in labor income tax increases the aggregate consumption demand through higher disposable income. An increase in consumption expenditures raises the aggregate demand, and through this channel an economy's output expands. However, the supply-side economic theory maintains that a reduction in the labor income tax affects the aggregate output through a fundamentally different channel. Lower tax rates increase the work incentive of laborers by increasing their after-tax return. Hence the aggregate labor supply increases, and so does the aggregate output.

The U.S. experienced a large scale tax restructuring under the leadership of President Regan during the 1980s. Federal personal income tax rates were reduced drastically and the tax structure was simplified considerably. Prior to the enactment of the Economic Recovery Tax Act (ERTA) in 1981, the U.S. income tax structure comprised of 15 rates ranging from 14 to 70 percent. The ERTA lowered tax rates across the board by more than 20 percent, with lowered spread of rates ranging from 11 to 50 percent. Income tax rates were reduced further under the Tax Reform Act (TRA86) in 1986. The tax structure was simplified considerably to a two-rate schedule of 15 and 28 percent. The design of ERTA and TRA86 was primarily motivated by the idea of the supply-side stimulus to economic growth. The U.S. personal income tax structure was further modified in 2001 during President Bush's first term in the White House. However, unlike the previous Reagan tax cuts, the basic motivation behind the 2001 tax legislation was to provide a demand-side stimulus to a recessionary economy. President Obama also adopted the Keynesian demand-side approach to combating the economic recession that began in 2008. This policy included, among others, the record high fiscal stimulus of $837 billion in 2009.

Clearly, as stated above, the reasons for lowering taxes can be different. There are two distinct channels through which taxes impact economic growth. In general, during sound economic environment, a lower marginal income tax rate is supposed to motivate workers to work more (supply-side stimulus), while during sluggish economic environment, a lower income tax rate is targeted to stimulate spending (demand-side stimulus). This paper explicitly identifies those two channels, and measures the relative contributions of the supply-side and the demand-side effects of unanticipated changes in labor income tax on the real output of the U.S. economy during 1979-2006. Relative lack of empirical research of this subject to date as well as the recent 2009 economic

recession make it imperative to gain a better understanding of the supply-side and demand-side effects of labor income tax changes on an economy's output.

The present paper develops a SVAR model with two variables, namely the real output growth rate and the labor income tax rate. These two variables are used to isolate the supply and the demand-side effects. We define two structural disturbances on the basis of the nature of their impact on output. The mechanism of work incentives is believed to have a "permanent effect" on output. It is captured in the supply disturbance. The mechanism of higher consumption spending is believed to have only a "temporary effect" on output. It is captured in the demand disturbance. The portion of the growth rate of the actual output due to the supply disturbance is called the supply component of the growth rate of output. The portion of the growth rate of the actual output due to the demand disturbance is called the demand component of the growth rate of output. We recover the time series of output in *level* from the time series of growth rate of output, given an initial value of output. The portion of the actual output due to the supply disturbance is called the supply component of output, while the portion of the actual output due to the demand disturbance is called the demand component of output. Time series of actual output, its supply component and its demand component, are *not* in the linear relationship because the latter two time series are recovered from their respective growth rate series. The movement in the supply component of output is regarded as the long-run trend in the actual output[10] under fully flexible prices, while demand disturbance causes short-run deviations of the actual output from its long-run trend. However, under imperfectly flexible prices this assumption is unwarranted. In that event, deviations from trend arise not only due to demand disturbance, but also due to supply disturbance. The conventional view of fluctuations in output being temporary deviations from the trend does not hold (Campbell and Mankiw, 1987). Accordingly, we also investigate whether the long run trend is stochastic.[11]

Additionally, it is also important to address the issue of the functional dependence of tax rates on output. Fiscal policy decisions are largely governed by the current or prior state of the economy. Therefore, the tax rate is endogenously influenced by the growth rate of the real output. Blanchard and Perotti (2002) use similar specifications where it is assumed that the exogenous changes in the tax rate are due to the unpredictable component of tax rates. Following the resolution of this issue, the rest of the paper is organized as follows. Section II explains the fundamentals of the supply-side and the demand-side effects of changes in the labor income tax on output. Methodological issues and the data used to analyze these effects are outlined in sections III and IV. In Section V all test results are reported and analyzed. Final conclusions on the impact of labor taxes on the U.S. economy are reached in Section VI.

II. Effects of a Change in the Labor Income Tax

A. Supply-Side Effects: Substitution Effect and Income Effect

The labor-leisure analysis is often used to describe the effect of a reduction in the labor income tax on individuals' labor supply decisions (Gwartney and Stroup, 1983; Bohanon and Cott, 1986). A reduction in the labor income tax generates two opposing impacts. First, it leads to a

[10] Generally, the actual output is characterized by a unit-root process. A natural implication is that it can be decomposed into the supply (permanent) and demand (temporary) components.

[11] If the long-run trend is stochastic, it may generate short-run fluctuations in the actual output. Consequently, a conventional view that fluctuations in output are temporary deviations from the trend does not hold (Campbell and Mankiw, 1987).

higher real wage. This makes the consumption of leisure expensive, as the opportunity cost of leisure increases. Since leisure is a normal good, individuals substitute away from leisure. This constitutes a substitution effect. Given the similar preference function, and *ceteris paribus*, the substitution effect increases the total labor supply in the economy. Second, lower taxes increase real income. Higher real income induces individuals to increase consumption of all normal goods, including leisure. This is referred to as an income effect. Given the similar preference function, and *ceteris paribus*, the income effect reduces the total labor supply in the economy. Thus, the net effect of a reduction in the labor income tax on total labor supply depends on the relative strengths of the substitution and income effects.

B. Demand-Side Effects

The focus of the Keynesian economic theory is on the output determination in the short run. In the short run, the aggregate output is primarily determined by the aggregate demand. Tax cuts raise the consumption demand through a higher disposable income and, thereby, the aggregate demand. Over time, the effects of the aggregate demand shocks die out. The aggregate demand changes cannot influence the aggregate output in the long run. The long-run effects of a tax cut are reflected in higher prices and wages through a dynamic adjustment mechanism. The long-run adjustment in output takes place through an upward revision of an expected wage rate and consequent changes in the price level. Output eventually returns to its natural level of output. When an economy is above the natural level of output, the price level goes up. The higher price level causes a decrease in the demand and output. When economy is below the natural level of output, the price level decreases. The lower price level causes an increase in the demand and output. Thus the demand-side forces do not have a permanent effect on the aggregate output. They can only cause short-run cyclical fluctuations in output around the long-run trend (Blanchard, 2006).

III. Methodology

Our modeling of the time-series data is based upon the methodology pioneered by King *et al.* (1991), Galí (1992), Enders and Lee (1997), and Claus (1999), among others. We develop a SVAR model with long-run identifying restrictions proposed by Blanchard and Quah (1989). First, we construct a two-variable VAR model where output and tax affect each other. Effects of a tax cut on output are realized through the supply and demand channels. Effects of output on tax rates are due to the fact that fiscal policy decisions are influenced by the aggregate state of the economy. Tax policy decisions are contingent on a government's prevailing budgetary circumstances. The feedback of output and tax rate is inherent in the dynamic analysis of the Laffer curve. Followers of the supply-side economics claim that higher economic growth resulting from a tax cut can be large enough to make tax rate even lower.[12] This assumption is used by Blanchard and Perotti (2002) who attribute the unexpected movements in output to the unexpected movements in tax rates and *vice versa*.

[12] Mankiw and Weinzierl (2005) examine the extent to which tax cuts are capable of generating higher revenue through economic growth.

A. Identification

The VAR approach is often criticized for having little economic content in its results. That is why numerous empirical studies in recent years impose a theoretical structure on the joint process of the constituent variables. This paper uses the *a priori* restriction that the demand disturbance does not affect the output in the long run. This restriction follows from the natural rate hypothesis developed in the mid-1950s by Friedman (1968).[13] Only one restriction is required to identify a structural model with two endogenous variables. No *a priori* assumption is made about the effects of the two disturbances on the tax rate, and the effect of the supply disturbance on output. We further assume that these two disturbances are uncorrelated at all leads and lags.

Let y_t and z_t denote the logarithm of the real GDP and the first difference of tax rates, respectively. Since y_t is the logarithm of the real GDP, Δy_t is the growth rate of real GDP. Our data suggests that both the growth rate in real GDP, Δy_t, and the first difference of tax rates, z_t, are stationary. This result is a necessary condition for constructing a VAR model. We consider a bivariate system where $\{\Delta y_t\}$ is affected by the current and the past realizations of $\{z_t\}$ along with its own past realizations, and likewise $\{z_t\}$ is affected by the current and the past realizations of $\{\Delta y_t\}$ along with its own past realizations. Structural equations are written as

$$\Delta y_t = \gamma_{z0} z_t + \gamma_{y1}\Delta y_{t-1} + \gamma_{z1} z_{t-1} + ... + \gamma_{yk}\Delta y_{t-k} + \gamma_{zk} z_{t-k} + v_{1t} \quad (1)$$

$$z_t = \beta_{y0}\Delta y_t + \beta_{y1}\Delta y_{t-1} + \beta_{z1} z_{t-1} + ... + \beta_{yk}\Delta y_{t-k} + \beta_{zk} z_{t-k} + v_{2t} \quad (2)$$

where v_{1t} and v_{2t} are uncorrelated white noise structural disturbances, γs are structural coefficients in Δy_t equation, and βs are structural coefficients in z_t equation. Both structural equations are considered without an intercept and have a finite lag order k. A shock to either of the structural disturbances affects both $\{\Delta y_t\}$ and $\{z_t\}$ simultaneously. Using matrix algebra, the above bivariate system can be written as

$$A_0 X_t = A_1 X_{t-1} + ... + A_k X_{t-k} + v_t \quad (3)$$

where X is the column vector $(\Delta y, z)'$, v is the column vector of unobserved structural disturbances $(v_1, v_2)'$, and

$$A_0 = \begin{bmatrix} 1 & -\gamma_{z0} \\ -\beta_{y0} & 1 \end{bmatrix},\ A_1 = \begin{bmatrix} \gamma_{y1} & \gamma_{z1} \\ \beta_{y1} & \beta_{z1} \end{bmatrix},\ \ldots,\ A_k = \begin{bmatrix} \gamma_{yk} & \gamma_{zk} \\ \beta_{yk} & \beta_{zk} \end{bmatrix}$$

Using lag operator on X_t, Equation (3) can be rewritten as

$$A_0 X_t = A_1 L X_t + ... + A_k L^k X_t + v_t \quad (4)$$

Alternatively,

$$A(L) X_t = v_t \quad (5)$$

where $A(L) = \left(A_0 - A_1 L - .... - A_k L^k\right)$.

Therefore,

$$X_t = A(L)^{-1} v_t \quad (6)$$

or

[13] For detailed outline of the natural rate hypothesis, see Friedman (1968).

$$X_t = S(L)v_t \tag{7}$$

where $S(L) = A(L)^{-1}$ is a matrix polynomial of infinite order and for which we assume that the bivariate invertibility conditions hold. Equation (7) is a bivariate moving average representation of structural equations of (3). Each equation in (7) can then be written as

$$\Delta y_t = \sum_{p=0}^{\infty} s_{11}(p)v_{1t-p} + \sum_{p=0}^{\infty} s_{12}(p)v_{2t-p} \tag{8}$$

$$z_t = \sum_{p=0}^{\infty} s_{21}(p)v_{1t-p} + \sum_{p=0}^{\infty} s_{22}(p)v_{2t-p} \tag{9}$$

Equations (8) and (9) express $\{\Delta y_t\}$ and $\{z_t\}$ as linear combinations of the current and past structural shocks.

In a more compact form, equations (8) and (9) can be written as

$$\begin{bmatrix} \Delta y_t \\ z_t \end{bmatrix} = \begin{bmatrix} S_{11}(L) & S_{12}(L) \\ S_{21}(L) & S_{22}(L) \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} \tag{10}$$

$S_{ij}(L)$ in Equation (10) are polynomials in the lag operator, where individual coefficients are denoted by $s_{ij}(p)$ in equations (8) and (9), and p is the lag length of infinite order.

We further assume unit variance for each of the disturbances (normalization assumption). This along with the assumption of uncorrelated white noise structural disturbances gives a diagonal variance-covariance matrix of the structural disturbance:

$$E(v_t v_t') = \begin{bmatrix} \text{var}(v_{1t}) & \text{cov}(v_{1t}, v_{2t}) \\ \text{cov}(v_{2t}, v_{1t}) & \text{var}(v_{2t}) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{11}$$

v_{1t} is regarded as the aggregate supply disturbances and v_{2t} is regarded as the aggregate demand disturbances, and $s_{ij}(p)$ s' are impulse responses of aggregate shocks. For instance, $s_{11}(1)$, $s_{11}(2)$, $s_{11}(3)$,… etc. are separate impulse responses of $\{\Delta y_t\}$ to an aggregate supply shock on $\{v_{1t}\}$, and $S_{11}(L)$ is the corresponding cumulative impulse response. Specifications of (8) and (9) do not assume that supply component (permanent component) of output follows a random walk.[14]

We impose an identifying restriction that the aggregate demand disturbance v_2 has no effect on the *level* of output (logarithmic scale) y_t in the long run. This indicates that the cumulative effects of v_{2t} on $\{\Delta y_t\}$ must be equal to zero:

$$\sum_{p=0}^{\infty} s_{12}(p) = 0 \tag{12}$$

[14] Lippi and Reichlin (1994) argue that the assumption of the permanent component of output being stationary is inconsistent with the true nature of technological adoption. For example, the random-walk characterization of a permanent component of the output precludes the possibility of learning by doing at the firm level. Moreover, a false random walk characterization of a permanent component of output, when in fact it is not, may mislead policy makers.

It is important to understand how Equation (12) ensures that demand shock v_2 has no effect on the *level* of output in logarithmic scale (y), hence on the *level* of output (Y). The proof of this assertion is presented in the following part of this paper in a simple example.

Let the sequence of growth rate of real GDP $\{\Delta y_t\}$ be governed by shocks on v_2 only. The corresponding moving average representation of $\{\Delta y_t\}$ is $\Delta y_t = \sum_{p=0}^{\infty} s_{12}(p) v_{2t-p}$. For expositional purpose, we set upper limit of p arbitrarily at 1. Then the sequence of $\{\Delta y_t\}$ follows $\Delta y_t = s_{12}(0)v_{2t} + s_{12}(1)v_{2t-1}$, where $s_{12}(p)$ is the effect of v_{2t-p} for $p = 0,1$ on Δy_t. We can write

$$y_t = y_{t-1} + s_{12}(0)v_{2t} + s_{12}(1)v_{2t-1} \tag{13}$$

Notice that the left-hand side of Equation (13) is the *level* of output. Successive substitutions of expressions y_{t-j}, $j = 1,2,...,\infty$ in Equation (13) by moving backward through time after setting initial shock value v_{20} at zero yields

$$y_t = y_0 + s_{12}(0)v_{2t} + \left[\sum_{p=0}^{1} s_{12}(p)v_{2t-1} + \sum_{p=0}^{1} s_{12}(p)v_{2t-2} + ...\right] \tag{14}$$

The term within the parentheses on the right-hand side of Equation (14) captures the long-run effect of past shocks in v_2 on the logarithm of output (y), and $s_{12}(0)$ captures the contemporaneous effect of shocks in v_2 on y. Clearly, for v_2 to have no effect on the level of output in log scale (y), and hence on the level of output (Y) in the long run, we must have $\sum_{p=0}^{1} s_{12}(p) = 0$. Thus for $p = 0,1,...,\infty$, the long-run restriction of demand shock on output level is equivalent to Equation (12).

In Equation (10), structural shocks $\{v_t\}$ are unobservable. So, $S(L)$ is not directly estimable. In order to recover the series of $\{v_t\}$, we construct the reduced form VAR from Equation (3), then estimate it in its unrestricted form. From Equation (3), we can write:

$$X_t = A_0^{-1}A_1X_{t-1} + ... + A_0^{-1}A_kX_{t-k} + A_0^{-1}v_t \tag{15}$$

Using lag operator on X_{t-1}, Equation (15) can be written as

$$X_t = A_0^{-1}A_1X_{t-1} + ... + A_0^{-1}A_kL^{k-1}X_{t-1} + A_0^{-1}v_t \tag{16}$$

or

$$X_t = A_0^{-1}\left(A_1 + ... + A_kL^{k-1}\right)X_{t-1} + A_0^{-1}v_t \tag{17}$$

or

$$X_t = \Phi(L)X_{t-1} + e_t \tag{18}$$

where $\Phi(L) = A_0^{-1}\left(A_1 + ... + A_kL^{k-1}\right)$ and $e_t = A_0^{-1}v_t$.

Alternatively,

$$\begin{bmatrix}\Delta y_t \\ z_t\end{bmatrix} = \begin{bmatrix}\Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L)\end{bmatrix}\begin{bmatrix}\Delta y_{t-1} \\ z_{t-1}\end{bmatrix} + \begin{bmatrix}e_{1t} \\ e_{2t}\end{bmatrix} \tag{19}$$

In Equation (19), e is the vector of reduced form disturbances $(e_1, e_2)'$, $\Phi(L)$ is the $2 \cdot 2$ matrix whose elements are the polynomials $\Phi_{ij}(L)$ in Equation (19), for example, $\Phi_{11}(L) = \phi_{11}(0) + \phi_{11}(1)L + \phi_{11}(2)L^2 + ...$, where $\phi_{ij}(p)$ are coefficients of $\Phi_{ij}(L)$, and p is the lag length. The residuals of a reduced form VAR, e_{1t} and e_{2t} are the composites of structural disturbances v_{1t} and v_{2t}. Hence, they are correlated, and an exogenous shock to one structural disturbance affects both variables simultaneously. Since v_{1t} and v_{2t} are white-noise innovations, both e_{1t} and e_{2t} have zero means, constant variances, and are individually serially uncorrelated (Enders, 2003, pp. 264-266). Estimation of Equation (18) is preceded by choosing an optimal number of lags by applying a lag-length selection criterion, such as Akaike Information Criterion (AIC) or Schwarz's Bayesian Criterion (SBC). Appropriately selected lag length eliminates serial correlation from reduced form residuals. Since the right-hand side of Equation (18) contains only predetermined variables, each error term has constant variance and error terms are serially uncorrelated. Therefore, each equation in the system can be estimated using OLS. The estimated unrestricted reduced form VAR can then be inverted to the vector moving average (VMA) representation using the Wold decomposition theorem (Hamilton, 1994, pp. 108-109).

$$\begin{bmatrix} \Delta y_t \\ z_t \end{bmatrix} = \begin{bmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \tag{20}$$

or

$$X_t = C(L)e_t \tag{21}$$

where $C(L) = (I - \Phi(L)L)^{-1}$.

Now we establish the relationship between the reduced form disturbances e and the structural disturbances v. One-step ahead forecast error of Δy_t in Equation (19) is $e_{1t} = \Delta y_t - E_{t-1}\Delta y_{t-1}$. Equivalent expression in Equation (8) is $s_{11}(0)v_{1t} + s_{12}(0)v_{2t}$. Thus,

$$e_{1t} = s_{11}(0)v_{1t} + s_{12}(0)v_{2t} \tag{22}$$

Similarly by comparing one-step ahead forecast errors of z_t in equations (20) and (9),

$$e_{2t} = s_{21}(0)v_{1t} + s_{22}(0)v_{2t} \tag{23}$$

equations (22) and (23) can be represented in a more compact form as

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} s_{11}(0) & s_{12}(0) \\ s_{21}(0) & s_{22}(0) \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} \tag{24}$$

or

$$e_t = S(0)v_t \tag{25}$$

Equation (25) forms the crucial relationship between reduced form disturbances e and structural disturbances v that will help recover unobservable structural disturbances v. It is clear from Equation (25) that a recovery of structural disturbances v requires coefficient estimates of $S(0)$,[15] which is the matrix of a contemporaneous effect of the structural disturbances v_t on X_t. In order to estimate four coefficients of $S(0)$ viz, $s_{11}(0)$, $s_{12}(0)$, $s_{21}(0)$ and $s_{22}(0)$, we use equations (22) and (23) along with the assumptions in Equation (11) and obtain the following three equations:

$$\text{var}(e_{1t}) = s_{11}(0)^2 + s_{12}(0)^2 \tag{26}$$

$$\text{var}(e_{2t}) = s_{21}(0)^2 + s_{22}(0)^2 \tag{27}$$

$$\text{cov}(e_{1t}, e_{2t}) = s_{11}(0)s_{21}(0) + s_{12}(0)s_{22}(0) \tag{28}$$

Equations (26), (27) and (28) can be viewed as three equations in four unknowns. We need one more equation to identify $S(0)$. The long-run restriction of Equation (12) provides that additional equation.

Using equations (10), (20) and (24),

$$\begin{bmatrix} S_{11}(L) & S_{12}(L) \\ S_{21}(L) & S_{22}(L) \end{bmatrix} = \begin{bmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} s_{11}(0) & s_{12}(0) \\ s_{21}(0) & s_{22}(0) \end{bmatrix} \tag{29}$$

or

$$S(L) = C(L)S(0) \tag{30}$$

Application of the long-run restriction of Equation (12) replaces $S_{12}(L)$ in the left-hand side of Equation (29) by zero and makes $S(L)$ lower triangular. Consequently, we obtain an additional equation:

$$C_{11}(L)s_{12}(0) + C_{12}(L)s_{22}(0) = 0 \tag{31}$$

Thus, equations (26), (27), (28), and (31) comprise the set of four equations that can be used to identify four coefficients of $S(0)$. Once $S(0)$ is estimated, the entire $\{v_{1t}\}$ and $\{v_{2t}\}$ sequences can be identified using Equation (25), $v_t = e_t S(0)^{-1}$, hence $v_{t-i} = e_{t-i} S(0)^{-1}$. Also, the elements of $S(L)$, namely $S_{11}(L)$, $S_{21}(L)$ and $S_{22}(L)$ can be recovered using Equation (29). For instance, $S_{11}(L) = C_{11}(L)s_{11}(0) + C_{12}(L)s_{21}(0)$.

Upon estimation of $S(0)$ and $S(L)$, we conduct historical decomposition of $\{\Delta y_t\}$. In order to construct a series that reflects only the effects of supply disturbances, we set all realizations of $\{v_{2t}\}$ to zero. Accordingly, the supply component of the growth rate in $\{y_t\}$ is given by

---

[15] $S(0)$ has a straight forward interpretation as the Cholesky factor of the variance-covariance matrix of the vector of reduced form disturbances (Lucas, 1990). Let the variance-covariance matrix of the vector of reduced form disturbances be Ω. Then, $\Omega = E(e_t e_t') = S(0)E(v_t v_t')S'(0) = S(0)S'(0)$, since $E(v_t v_t') = I$. Hence $S(0)$ is identified.

$$\Delta y_t^{Supply} = S_{11}(L)v_{1t} \tag{32}$$

Similarly a series reflecting only the effects of demand disturbances is obtained by setting all realizations of $\{v_{1t}\}$ to zero. Accordingly, the demand component of the growth rate in $\{y_t\}$ is given by

$$\Delta y_t^{Demand} = S_{12}(L)v_{2t} \tag{33}$$

From Equation (32), the level of output due to the supply disturbance $\{Y_t^{Supply}\}$ is generated by an appropriate transformation of $\{\Delta y_t^{Supply}\}$. First, $\{y_t^{Supply}\}$ is generated by taking a starting value of $\{y_{t-1}^{Actual}\}$. Then, $\{Y_t^{Supply}\}$ is obtained by taking antilog of $\{y_t^{Supply}\}$. The level of output due to the demand disturbance $\{Y_t^{Demand}\}$ is generated in either of the two ways: first, $\{\Delta y_t^{Demand}\}$ is computed using Equation (33), then $\{y_t^{Demand}\}$ is obtained by initiating the series with $\{y_{t-1}^{Actual}\}$, correspondingly $\{Y_t^{Demand}\}$ is obtained by taking antilog of $\{y_t^{Demand}\}$; alternatively, by $\{\Delta y_t^{Demand}\} = \{\Delta y_t^{Actual}\} - \{\Delta y_t^{Supply}\}$. Then $\{y_t^{Demand}\}$ and $\{Y_t^{Demand}\}$ are obtained successively. Choice of $\{y_{t-1}^{Actual}\}$ as an initial point is somewhat arbitrary.

In the same way, we generate the level of tax rates due to the supply disturbance $\{\tau_t^{Supply}\}$ and the level of tax rates due to the demand disturbance $\{\tau_t^{Demand}\}$ after computing

$$z_t^{Supply} = S_{21}(L)v_{1t} \tag{34}$$

and

$$z_t^{Demand} = S_{22}(L)v_{2t} \tag{35}$$

IV. Data

The data are obtained from the National Income and Product Accounts (NIPA) collected by the Bureau of Economic Analysis for the period from 1978:I to 2006:III. Focusing our research on this particular time period enables us to analyze fully the impact of recent tax changes on the U.S. economy's output. Average tax rates on the labor income (τ_t) are calculated following Jones (2002), and Mendoza *et al.* (1994). The data are the quarterly U.S. observations on the Real Gross Domestic Product (Y_t), Personal Current Taxes of Federal Government in billions of dollars (FIT_t), Personal Current Taxes of State and Local Government in billions of dollars (SIT_t), Wage and Salary Disbursements in billions of dollars (W_t), Proprietor's Income with Inventory Valuation and Capital Consumption Adjustment in billions of dollars (PRI_t), Rental Income of Persons with Capital Consumption Adjustment in billions of dollars (RI_t), Personal Interest Income (PII_t), and Personal Dividend Income (PDI_t). Real Gross Domestic Product (Y_t) is obtained as Seasonally Adjusted Quantity Indexes measured at the base year 2000. Personal Current Taxes of Federal Government (FIT_t) include the dividend tax for 1933-34, and the automobile use tax for 1942-46. All other series are expressed in billions of dollars and are seasonally adjusted at annual rates.

Average tax rates on labor income are calculated using the $\tau_t = \frac{FIT_t + SIT_t}{W_t + PRI_t + RI_t + PII_t + PDI_t}$ expression. The denominator of this expression comprises Labor Income (LI_t) and Capital Income (CI_t), where $LI_t = W_t + PRI_t / 2$ and $CI_t = PRI_t / 2 + RI_t + PII_t + PDI_t$. The division of Proprietor's Income into labor and capital income is somewhat arbitrary (Joines, 1981).

For the purpose of cross-verification, the data on Personal Income (PI_t) and Personal Current Taxes (PCT_t) are also obtained. An *ad hoc* measure of the labor income tax rates (τ_t^a) is calculated taking the ratio of PCT_t and PI_t. The correlation coefficient between τ_t and τ_t^a is found to be more than 98 percent.[16] In the following section, Real GDP in logarithmic scale is denoted by y i.e., $y_t = \log(Y_t)$, and the first difference of labor income tax rates is denoted by z i.e. $z_t = \tau_t - \tau_{t-1}$.

V. Results and Analysis

A. Unit-Root Test

The initial step in analyzing any time-series data necessitates stationarity testing of each individual time-series. The objective of stationarity tests is to determine the degree of integration of all time-series data used in any subsequent econometric modeling. This determination is made upon establishing the number of unit roots in all data series under empirical investigation. Only stationary time-series data can be used in any subsequent econometric modeling. Numerous unit-root tests are outlined in econometric literature. The most commonly used unit-root tests are the Augmented Dickey-Fuller (Fuller, 1976), Dickey and Fuller (1979) (ADF) test and the Phillips-Perron (1988) (PP) test. We used initially the ADF test to examine the presence of unit roots (non-stationarity) in the series of real GDP (Y_t), the natural logarithm of real GDP (y_t), the first difference of natural logarithm of real GDP (Δy_t), the average tax rate (τ_t) and the first difference of the average tax rate (z_t). The first difference of natural logarithm of real GDP (Δy_t) is the growth rate of real GDP.

Table 1 reports the ADF test results at the 5 percent significance level. The Schwarz's information criterion is used to determine the lag length p in each series. The first row of the table indicates the selected lag lengths of each series with 'no trend' and 'trend' specifications. The critical value (t-critical) corresponding to a test with 5 percent level of significance changes depending on model specifications of each series (even though the number of observations remains the same).[17]

---

[16] The NIPA data source on each variable stated above is as follows, Y - Table 1.1.3: line 1, FIT - Table 3.2: line 3, SIT - Table 3.3: line 3, W - Table 2.1: line 3, PRI - Table 2.1: line 9, RI - Table 2.1: line 12, PII - Table 2.1: line 14, PDI - Table 2.1: line 15, PI - Table 2.1: line 1, PCT - Table 2.1: line 25.

[17] Appropriate critical values depend on both the model specification and the sample size.

Table 1: Augmented Dickey-Fuller (ADF) Test Results[a] of Output Series

	Y		y		Δy	
	No Trend	Trend	No Trend	Trend	No Trend	Trend
p	2	2	2	2	1	1
a_0	-0.0285 (-0.1578)[b]	1.8649 (2.5830)	0.0003 (0.0269)	0.3909 (3.5159)	0.0041 (4.0876)	0.0034 (2.3289)
a_2	---	0.0266 (2.7038)	---	0.0008 (3.5290)	---	0.00001 (0.6317)
γ	0.0045 (1.8456)	-0.0397 (-2.4032)	0.0009 (0.3485)	-0.1000 (-3.4853)	-0.5757 (-5.4324)	-0.5798 (-5.4461)
$\gamma+1$	1.0045 (1.8456)	0.9603 (-2.4032)	1.0009 (0.3485)	0.9000 (-3.4853)	0.4243 (-5.4324)	0.4202 (-5.4461)
ADF Test Statistic	1.8456	-2.4032	0.3485	-3.4853*	-5.4324*	-5.4460*
t-critical (5% level)	-2.8874	-3.4504	-2.8874	-3.4504	-2.8874	-3.4504
Observations	112	112	112	112	112	112

Note: Y, y and Δy represent real GDP, the natural logarithm of real GDP and the first difference of real GDP (i.e. the growth rate of real GDP), respectively.

[a] All test regressors include a constant.

[b] t-statistics are in parentheses.

* Reject the null hypothesis of the presence of unit root at 5 percent significance level.

ADF test statistics show that we fail to reject the null of the presence of a unit root for the series of real GDP (Y_t) with both 'no trend' and 'trend' specifications. Additionally, test results reject the null hypothesis of the presence of a unit root in the series of the natural logarithm of real GDP (y_t) with 'trend', and the first difference of the natural logarithm of real GDP (Δy_t) under both 'no trend' and 'trend' assumptions. Therefore, real GDP (Y_t) is non-stationary and the growth rate of real GDP between two consecutive quarters (Δy_t) is trend stationary.

We also deployed the ADF analysis to test for the presence of a unit root in tax rates in *level* (τ_t), and the first difference of tax rates (z_t). Although these tabulated results are not reported, they are available upon request. They indicate that average tax rate (τ_t) is non-stationary, but its first difference (z_t) is stationary. In order to test further the robustness of our unit-root tests, we subjected all our time-series data to two additional tests, the Phillips-Perron (1988) (PP) test and

the Zivot-Andrews (1992) (ZA) test. The PP tests yielded the same stationarity conclusions for our time-series data as those obtained by using the ADF tests. The ZA test supported stationarity conclusions reached by both the ADF and the PP tests. Due to space constraint, the individual PP and AZ test results are not reported. However, they will be made available upon request to interested readers.

B. Estimation

Having identified two stationary processes, $\{\Delta y_t\}$ and $\{z_t\}$, we use the AIC and the SBC methods to select the lag order p in the reduced form representations of the VAR system (corresponding to Equation (19) in Section III). We obtain the AIC and the SBC numbers of each series for a lag length of 2 through 8 quarters. Table 2 shows that the minimum AIC occurs at a lag length of 3 for both $\{\Delta y_t\}$ and $\{z_t\}$. The minimum SBC occurs at a lag length of 3 for $\{\Delta y_t\}$ and a lag length of 1 for $\{z_t\}$. We choose the lag length of 3.

Table 2: Lag Length Selection

Lag Length	AIC		SBC	
	Δy	z	Δy	z
2	-7.0012	-7.7582	-6.9041	-7.6611*
3	-7.0974*	-7.0974*	-6.9266**	-7.5976
4	-6.9586	-7.7462	-6.7622	-7.5499
5	-6.9231	-7.7016	-6.6762	-7.4546
6	-6.9052	-7.6945	-6.6072	-7.3964
7	-6.8732	-7.6556	-6.5235	-7.3059
8	-6.8498	-7.6429	-6.4478	-7.2409

* AIC and SC are minimum.

Since both equations in the reduced form VAR have the same regressors, and each regressor is independent of disturbances, then each equation can be estimated separately using OLS. Coefficient estimates of an unrestricted reduced form VAR are given in Table 3. Reduced form coefficients, also known as *impact multipliers*, measure the response of endogenous variables to changes in the predetermined (lagged) variables. All three coefficients of lagged z in Δy equation are negative, indicating that a tax cut in the past causes higher output growth. The coefficient of Δy_{t-1} in z equation is also negative, indicating that a higher output-growth rate in the last quarter causes reduced current tax rate.

Table 3: Coefficient Estimates of Unrestricted Reduced Form VAR

Regressors	(Δy_t)	(z_t)
Δy_{t-1}	0.4326	-0.0647
	(4.4896)[a]	(-0.9979)
Δy_{t-2}	0.2750	0.0544
	(2.7369)	(0.8045)
Δy_{t-3}	0.0839	0.0505
	(0.9265)	(0.8299)
z_{t-1}	-0.0418	-0.4330
	(-0.2971)	(-4.5736)
z_{t-2}	-0.0715	0.1740
	(-0.4784)	(1.7302)
z_{t-3}	-0.1931	0.2276
	(-1.3777)	(2.4132)
$Var(e_{it})$	0.0005	0.0002

[a] t-statistics are in parentheses.

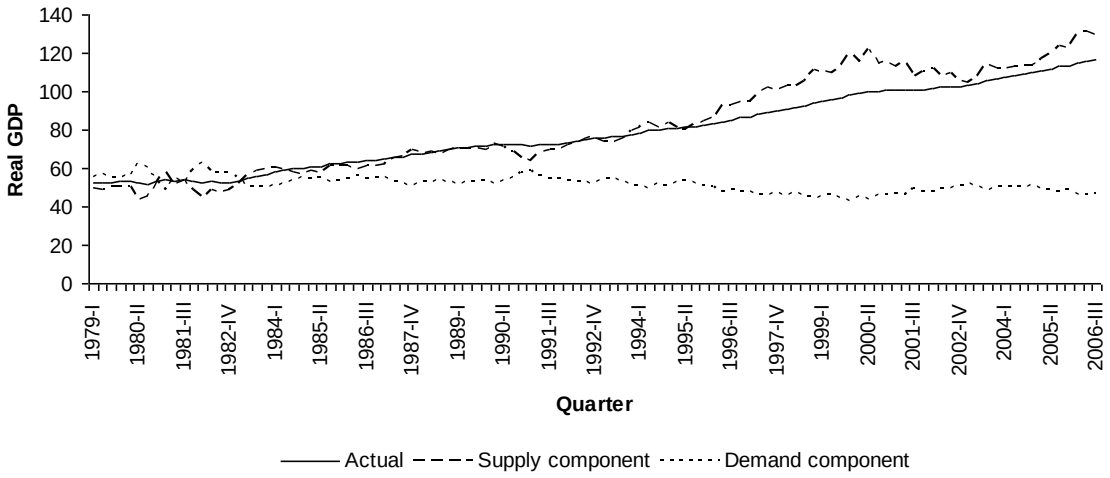
C. Decomposition of Output

Estimation of the joint process of $\{\Delta y_t\}$ and $\{z_t\}$ along with the long-run identifying restriction entails a recovery of the unobserved supply and demand disturbances (corresponding to Equation (25) in Section III). The 'supply component' and the 'demand component' of real GDP and the tax rates in level are recovered by an appropriate transformation of the series generated by Equation (32) to Equation (35). Figure 1 presents the decomposition of the actual real GDP into its supply and demand components in level.[18] The supply component of the real GDP is the time

[18] Our sample ranges from 1978:I to 2006:III. Blanchard-Quah technique yields the series $\{\Delta y_t^{Supply}\}$. From this term we generate the series $\{y_t^{Supply}\}$ by taking y_{1978IV}^{Actual} as the initial value (since the first four observations are lost due to taking the first difference of $\{y_t\}$ and choosing the lag length of 3. By appropriate scaling (taking antilog) of $\{y_t^{Supply}\}$, we recover the series $\{Y_t^{Supply}\}$. In order to recover the series $\{Y_t^{Demand}\}$, we proceed by generating the series $\{\Delta y_t^{Demand}\} = \{\Delta y_t^{Actual}\} - \{\Delta y_t^{Supply}\}$. From this equation we generate $\{y_t^{Demand}\}$ in the same way as we do for $\{y_t^{Supply}\}$. An appropriate scaling of $\{y_t^{Demand}\}$ yields $\{Y_t^{Demand}\}$. Figure 2 depicts the time series plots for $\{Y_t^{Actual}\}$, $\{Y_t^{Supply}\}$ and $\{Y_t^{Demand}\}$. A visual inspection indicates that $\{Y_t^{Supply}\}$ is non-stationary and $\{Y_t^{Demand}\}$ is stationary.

path of the real GDP that would have been obtained in the absence of a demand disturbance. The supply component is obtained by setting the demand innovations at zero. By the same token, the demand component of the real GDP is the time path of the real GDP that would have been obtained in the absence of a supply disturbance. The latter can be achieved in two ways: (a) by setting the supply innovations at zero, or (b) by taking the difference between $\{\Delta y_t^{Actual}\}$ and $\{\Delta y_t^{Supply}\}$. Figure 1 follows (b). However, either approach yields almost identical results. The time path of the supply component and the demand component of the real GDP are consistent with the identifying restriction that the demand disturbance has no long-run effect on real GDP. The demand component of the real GDP in level $\left(Y_t^{Demand}\right)$ is mean reverting (stationary) whereas the supply component of the real GDP $\left(Y_t^{Supply}\right)$ exhibits a trend (non-stationary). A close look at the supply component of output reveals that the trend is not deterministic.[19] It exhibits a higher growth in the 1990s compared to the growth in the 1980s. Periods for which the actual real GDP falls short of its supply component are characterized by the lack of sufficient demand. Thus the supply component of output can also be viewed as the level of the 'potential output'. The opposite interpretation holds when the actual output is above the supply component. These are the periods of an overheated economy with an increasing demand pressure. Figure 1 indicates that after the mid-1990s, the U.S. economy has operated below its potential. This is the period when either the demand-side stimulus due to a tax cut is not significant or there is a negative demand effect due to a tax hike. In fact, the U.S. experienced a tax increase under the Deficit Reduction Act of 1993.

Figure 1: Decomposition of Actual Real GDP into Supply and Demand Components



[19] Appendix A reports the ADF test results.

D. Business Cycles

Commonly the supply component of output is considered as 'trend', that part of an output that would be realized under perfectly flexible prices. All temporary deviations of the actual output from its trend due to demand disturbances are 'business cycles'. Under the assumption of perfectly flexible prices, the trend is deterministic. Nelson and Plosser (1982) challenged the assumption of a constant trend over time. In real life, prices are imperfectly flexible. The presence of nominal rigidities in prices may change the long-run adjustment mechanism in the output. A time-varying trend is called the stochastic trend. Results in our sample indicate that the supply component of the output exhibits a stochastic trend. Therefore, both the supply and the demand disturbances contribute to business cycles.

Figure 2: Deviations of Actual Real GDP from Supply Component of Real GDP

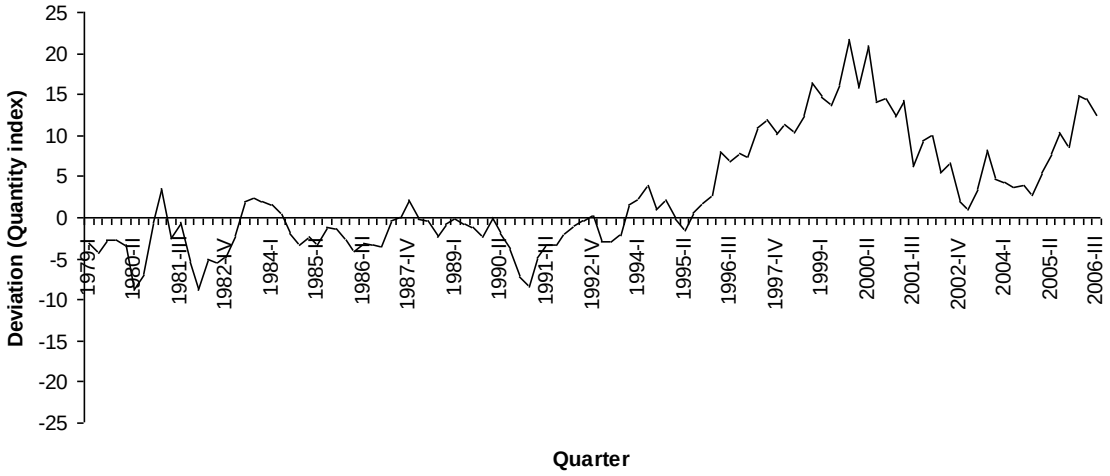


Figure 2 presents deviations of the actual output from its supply component *in levels* $\left(Y_t^{Supply} - Y_t^{Actual}\right)$. However, identifying separately the effects of a stochastic trend and business cycles on these deviations is difficult. Tax cuts contribute to output deviations from the long-run trend through both the supply and the demand channels. Our sample indicates a marked difference in deviations in two different phases. Deviations became more volatile after mid-1990s. This is mainly due to the volatility in the supply component of output.

It is difficult to identify business cycles and the trend separately because of the stochastic nature of the supply component of output. However, it is important to analyze the movements in the demand component of output over time because business cycles are primarily driven by the demand side factors, such as the consumption effect of a tax cut.

Figure 3: Output Fluctuations Due to Demand

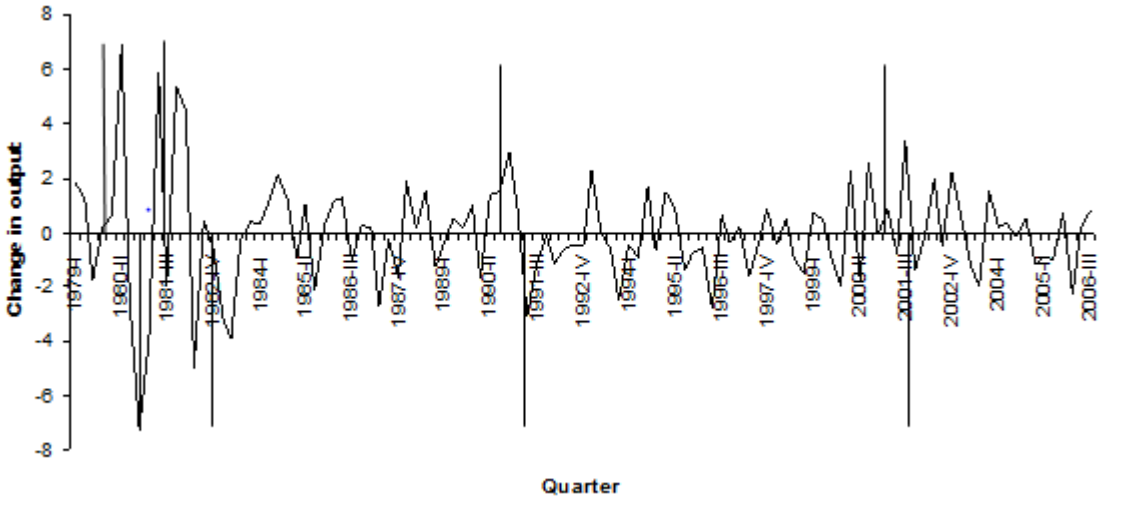


Figure 3 magnifies output fluctuations in the short run by taking the difference in the demand component of output in two consecutive quarters. The peaks and troughs of the demand component of output match closely with the NBER peaks and troughs.[20] The NBER peaks and troughs are marked by vertical lines. The recession of 1980 deserves a special mention. Results of our study indicate that historically large fluctuations in the U.S. output are mainly demand driven.

E. Decomposition of Tax Rates

Figure 4 shows the decomposition of the actual tax rate *in level* into its supply and demand components. The time path of the supply component of the tax rate is obtained by setting all demand disturbances at zero. The time path of the demand component of the tax rate $\left(\tau_t^{Demand}\right)$ is obtained by generating the series $\left\{z_t^{Demand}\right\} = \left\{z_t^{Actual}\right\} - \left\{z_t^{Supply}\right\}$, then $\left\{\tau_t^{Demand}\right\}$ is obtained by taking τ_{1978IV}^{Actual} as the initial value. Since there is no restriction on the short-run and long-run effects of the supply and the demand disturbances on the tax rate, some implications of this relationship can be derived informally.

[20] The NBER peaks and troughs as reported in http://www.nber.org/cycles/cyclesmain.html after 1979 are as follows: Peak – 1980:1, 1981:3, 1990:3, and 2001:1; Trough – 1980:3, 1982:4, 1991:1, and 2001:4.

Figure 4: Decomposition of Actual Tax Rate into Supply and Demand Components

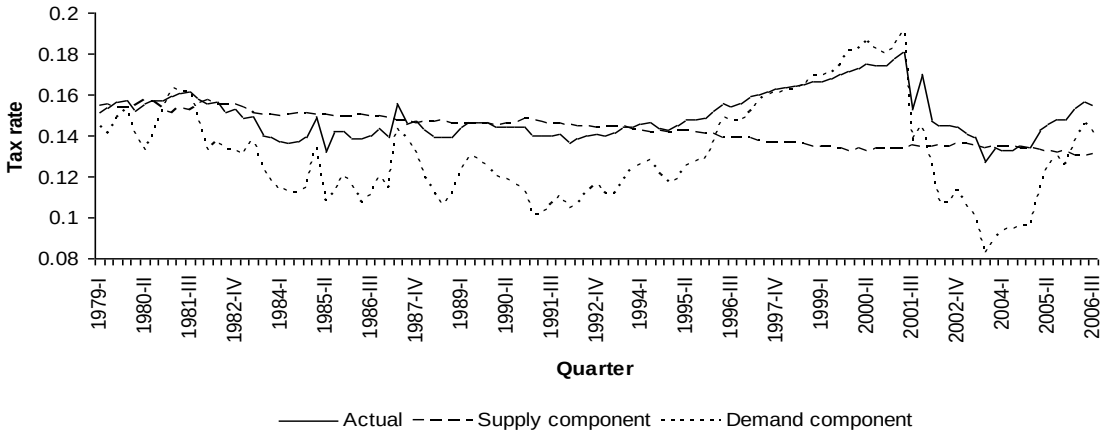


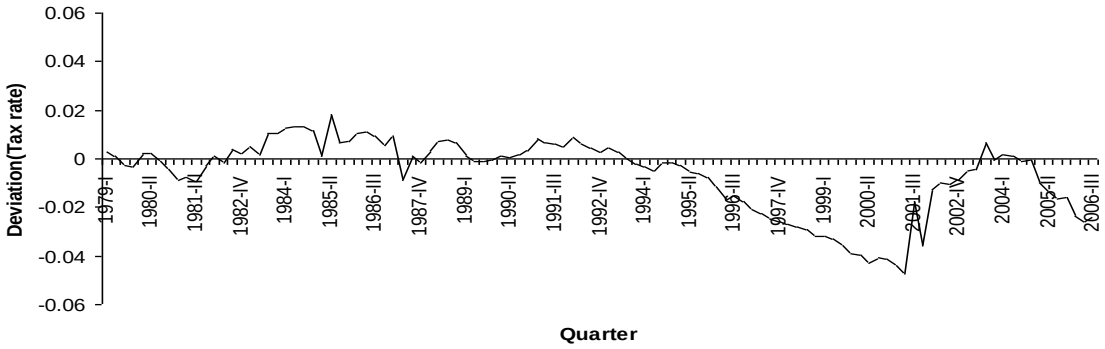
Figure 4 indicates that the time series of the actual tax rate and its demand component seem to move together. The correlation coefficient between these variables is 0.87. This result suggests that tax policies are mainly influenced by the demand disturbances. Table 4 reports the standard deviation and the mean of the times series of the actual tax rate, the supply component of the tax rate and the demand component of the tax rate, respectively. Clearly, the variation in the demand component is higher than the variation in the supply component by more than 97 percent. At the same time, it is clear that the variations in the demand component account for almost all variation in the actual tax rates. The supply component of the tax rate exhibits a slightly declining trend over time, averaging approximately 15 percent. This trend is deterministic. The declining trend along with a low standard deviation (0.0077) indicates that the demand disturbances of the tax rate changes are reduced slowly over time. Therefore, tax policies are not effective. Since the demand component of tax rate has a unit root, and it is difference stationary (not reported), any change in the tax policy due to demand disturbance seems to have a long-run effect on future tax rates.

Table 4: Contribution of Supply and Demand Disturbances in Tax Policy

	Actual	Supply	Demand
Standard Deviation	0.0133	0.0077	0.0152
Mean	0.1495	0.1430	0.1593

Figure 5 shows the time path of the deviations of the actual tax rate from its supply component $\left(\tau_t^{Supply} - \tau_t^{Actual}\right)$. Since, $\left\{\tau_t^{Supply}\right\}$ exhibits a deterministic trend, any such variations are due to the demand disturbance only. Also when figures 2 and 5 are compared, it is evident that the demand components of the output and the tax rate are near mirror images of each other. This result is consistent with the pattern of the relationship between the two variables in the short run.

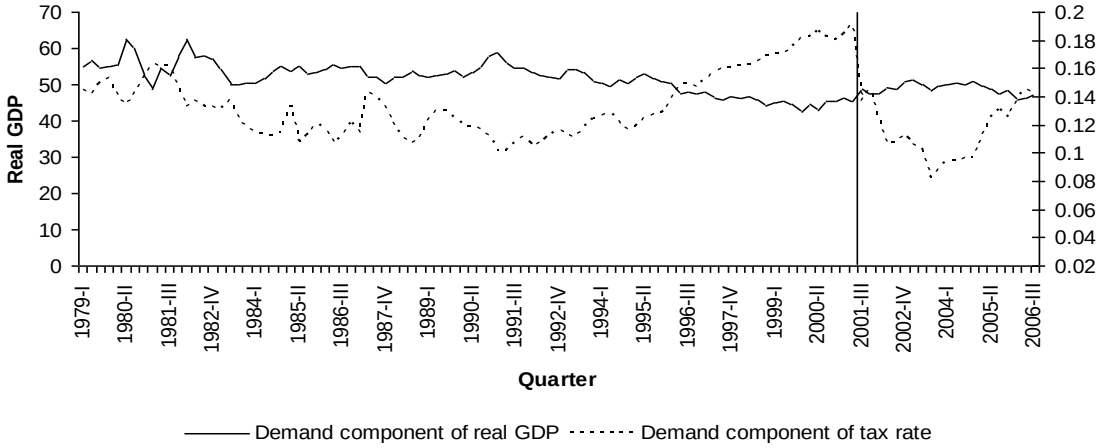
Figure 5: Deviations of Actual Tax Rate from Supply Component of Tax Rate



F. Demand Components of the Output and the Tax Rate

Since it is clear from the above analysis that the demand disturbance has a dominant influence on tax rates, it would certainly be of interest to examine the patterns of the short-run movements in the demand components of output and the tax rates. This objective is accomplished in Figure 6. This figure presents the time paths of the demand components of the real GDP (measured along the vertical axis on the left) and the tax rate (measured along the vertical axis on the right) for the time period 1979:I-2006:III. The correlation coefficient of -0.78 indicates a significant negative relationship between the two time series. However, the negative strength of this relationship declines after the third quarter of 2001 (marked by a vertical line at 2001:III). The correlation coefficient between the demand component of output and the tax rates during 2001:III-2004:IV is -0.43. This coefficient is -0.79 during 1979:I-2001:II. This result indicates a behavioral change in the pattern of the relationship between the tax rate and output in the short run after the third quarter of 2001. This result also provides a plausible explanation for the relative ineffectiveness of the first 2001 Bush's tax cut and the need to reduce taxes further in 2003.

Figure 6: Movements of Demand Components of Real GDP and Tax Rate



G. Relative Contributions of Supply and Demand Disturbances

G.1. Variance Decomposition

The next obvious step in analyzing the impact of tax changes on the real GDP in the U.S. necessitates undertaking a statistical assessment of the relative contributions of the supply and the demand disturbances on the U.S. output. This objective can be accomplished by computing the variance decompositions of the forecast error for the growth rate of output (Δy_t) and the change in tax rates (z_t) at various time horizons. The forecast error variance decompositions provide estimates of proportions of movements in $\{\Delta y_t\}$ due to the supply shocks in $\{v_{1t}\}$ and the demand shocks in $\{v_{2t}\}$ at various time horizons. The proportions of movements in $\{z_t\}$ due to each of these two shocks at different time horizons can also be measured using this method. Our long-run identification restriction on the demand disturbance has a connotation for variance decompositions, namely the contributions of the supply disturbance to the variance of the output movements tend to hundred percent as the horizon increases.

Variance decompositions of the two endogenous variables are given in tables 5 and 6. Numbers are computed as follows. First, the k-quarter, $k = 1,2,...,40$, ahead of forecast errors in Δy_t and z_t are calculated by the difference between the observed value of the variable and its forecast. A reduced form VAR of Equation (19) is used for these computations. The resulting forecast error is due to both the supply and the demand disturbances because the reduced form disturbances are composites of structural disturbances. Second, structural disturbances are identified in the forecast error variance using Equation (25). Third, the percentages of the forecast error variance due to the supply and demand disturbances are obtained against each k. For instance, the percentage of one-step ahead forecast error variance due to supply disturbance in the growth rate of output is 96.0972. In both tables 5 and 6, the numbers under the second and the third columns for each k add up to hundred.

Table 5: Variance Decomposition of Growth Rate of Output

Percentage of Variance Due to Supply and Demand Disturbances		
Quarters	Supply	Demand
1	96.0972	3.9028
5	97.1001	2.8999
10	97.2622	2.7378
15	97.2769	2.7231
20	97.2788	2.7212
25	97.2791	2.7209
40	97.2792	2.7208

Table 6: Variance Decomposition of Change in Tax Rates

Percentage of Variance Due to Supply and Demand Disturbances		
Quarters	Supply	Demand
1	2.2875	97.7125
5	1.8541	98.1459
10	2.0766	97.9234
15	2.1413	97.8587
20	2.1538	97.8462
25	2.1561	97.8439
40	2.1566	97.8434

Several important conclusions about the relative importance of the supply and demand disturbances on the U.S. economy's output emerge from tables 5 and 6. First, the relative contribution of the supply disturbance to output is very significant even in the shorter run. It amounts to 96 percent at one quarter horizon. Second, the effects of the demand disturbance on output die out at a faster pace than the effects of the supply disturbance increase over time. For instance, the proportion of the forecast error variance due to the demand disturbance decreases by thirty percent from one to forty quarter horizon whereas the proportion of the forecast error variance due to supply disturbance increases by one percent during the same time span. Third, the effect of the demand disturbance on tax rates starts declining gradually after the fifth quarter. However, it remains high at all horizons.

G.2 Impulse Response Functions

The dynamic effects of structural disturbances on the growth rate of output (Δy_t) and the changes in tax rate (z_t) can be analyzed most effectively by impulse response functions. These functions are illustrated in figures 7 and 8. The vertical axis in figures 7 and 8 denote the growth rate of output while the horizontal axes denote time in quarters. Figure 7 depicts the time path Δy_t due to a one standard deviation shock on the supply disturbance v_{1t}. The growth rate of output (also level of output) cumulates steadily over time. The peak response is approximately four times the initial effect, and it takes place after twenty two quarters. This effect dies out and stabilizes eventually at a growth rate of 3.2 percent. This result indicates that the supply disturbance has a permanent effect on the output in the long run. Figure 8 depicts the time path Δy_t due to a one standard deviation shock on the demand disturbance v_{2t}. For expositional purpose, Figure 8 has an amplified vertical axis. The demand disturbance has a hump shaped effect on Δy_t. This effect peaks during the third quarter. It decays slightly between the third and fourth quarters. It rises again during the fifth quarter, and gradually declines thereafter. Dynamic effects of output changes to demand disturbances are consistent with the traditional adjustment view that assumes that the

initial demand disturbance is followed by dynamic adjustments in prices and wages. These adjustments lead the economy back to its original steady-state value.

Figure 7: Response of Growth Rate of Output to Supply Shock

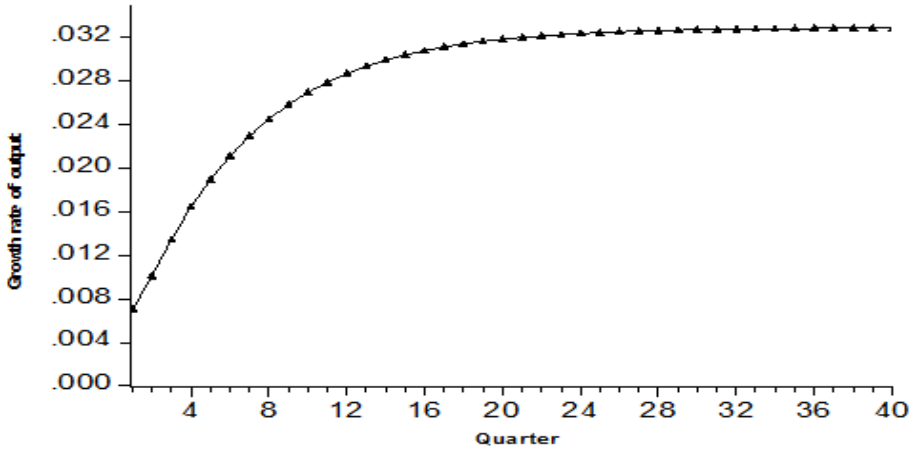
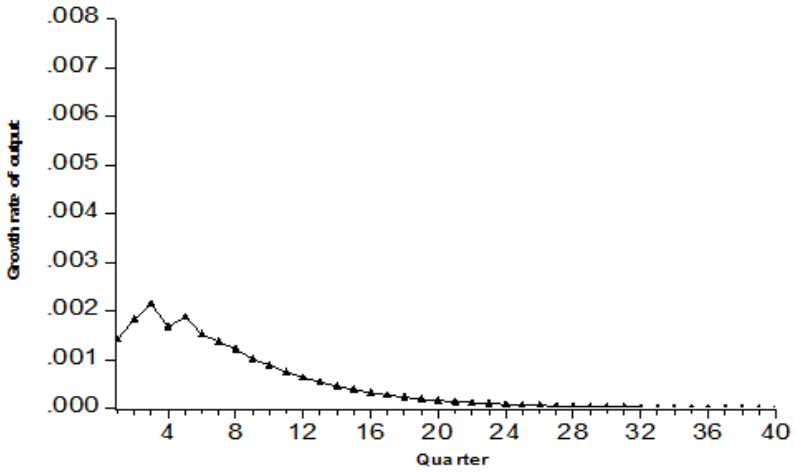


Figure 8: Response of Growth Rate of Output to Demand Shock



Figures 9 and 10 present dynamic effects of the supply and the demand disturbances respectively, on changes in the tax rate (z_t). In Figure 9, a positive supply disturbance decreases the tax rate slightly initially. This effect peaks up after the second quarter and it stabilizes approximately the same time as the supply effect of output stabilizes (after twenty two quarters).

Figure 9: Response of Change in Tax Rate to Supply Shock

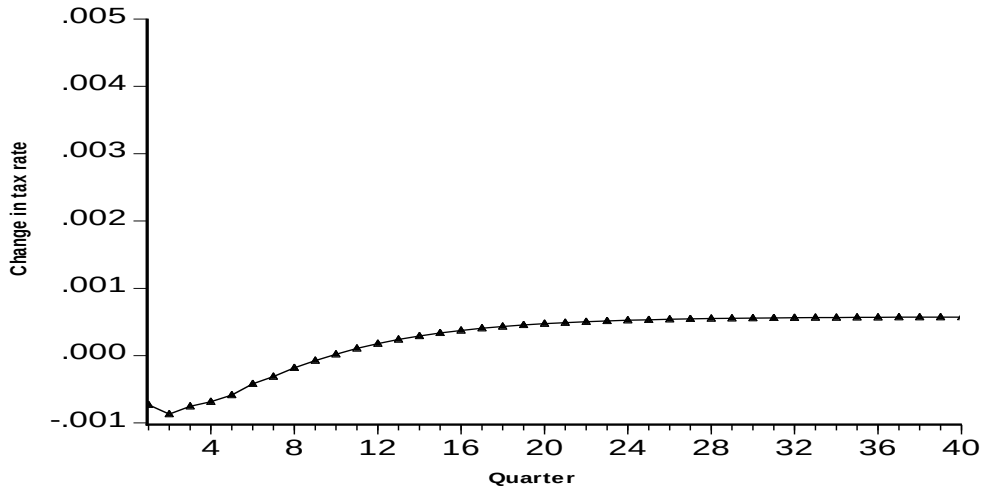


Figure 10: Response of Change in Tax Rate to Demand Shock

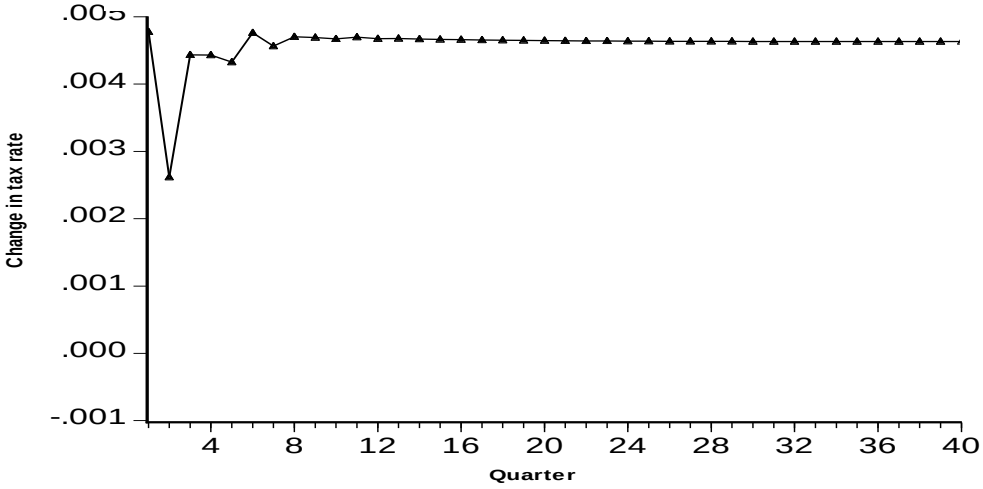


Figure 10 suggests that an exogenous shock to the demand disturbance results in a steep drop in the tax rate in the second quarter, followed by a steep rise in the third quarter. The tax rate returns to its original level after eight quarters.

Conclusion

This study provides new empirical evidence on the effects of a change in the labor income tax on the U.S. economy's output. A SVAR model comprising the real output growth rate and the labor income tax rate is developed to achieve this objective. The novelty of the present research is its extensive analysis of both the supply-side and the demand-side channels through which labor income tax affects real output. The first channel, namely the supply-side effect, is based on the premise that a tax cut provides higher work incentives, thereby increasing the aggregate labor supply in the economy. This increase in the labor supply leads to a higher aggregate output. Thus

the supply-side effect operates through the aggregate production function. Theoretically, the supply-side impact of a labor income tax-cut is permanent. It can be viewed as the long run, permanent effect on an economy's output. However, the assertion that the labor income tax-cut contributes permanently to economic growth, although undoubtedly extremely important, is thus far only a theoretical possibility. As such, it is not universally accepted in economic literature. Empirical evidence on this key economic issue is also mixed and far from satisfactory. The present research provides new empirical evidence on this issue.

The second channel through which a labor income tax cut affects the aggregate output, namely the demand-side effect, is based on the Keynesian theory of the aggregate demand. A tax cut results in higher disposable income, thereby increasing the aggregate consumption and the aggregate demand. The demand-side effect of a tax cut is based on the premise that demand changes determine output only in the short run. This channel of the tax-cut impact on the aggregate output is well documented in economic literature. The impact of a tax cut on consumption expenditures may be lesser if consumers maximize their utility subject to their life time budget constraint. If consumers anticipate that taxes will have to be increased to finance current budget deficit, then the current tax cut may not cause an increase in the current consumption and output (Ricardian equivalence). A tax cut will also not increase consumption unless it is perceived to be a permanent tax cut. However, most empirical studies do not support the Ricardian equivalence hypothesis. Given the preceding theoretical controversies and regardless of the channel of transmission, tax cuts may have an ambiguous effect on output.[21]

Empirical research can help to resolve the above outlined theoretical controversies. The present paper provides new empirical evidence on this issue. As mentioned above, this objective is accomplished by constructing a two-variable (output growth and labor tax rate) SVAR model that investigates the effects of an exogenous shock of the tax policy on the aggregate output. The Blanchard-Quah decomposition technique is used in our data analyses. We define the demand and the supply disturbances according to their assumed theoretical impact on the output dynamics. The demand disturbance is believed to have a temporary effect on the output, whereas the supply disturbance affects output permanently. Both endogenous variables in the VAR system are affected by the two disturbances along with their own current and lagged values. An exogenous shock to either of the disturbances affects both endogenous variables simultaneously.

Given the above stated conditions, we generate a non-stationary permanent component and a stationary temporary component of output. Variance decomposition and impulse response functions techniques are deployed to analyze the supply and the demand effects of labor income tax cuts on the aggregate output. These analyses provide new startling evidence on the impact of tax cuts on the U.S. economy. Variance decomposition and impulse response tests indicate that the contribution of the supply disturbances to output is very significant even in the short run. The supply-side effect on the output reaches its maximum after approximately five to six years. We also conclude that demand disturbances cause a substantial contribution to output fluctuations in the short run. The demand-side effect disappears after approximately the same time as the supply-side effect reaches its peak. Consequently, it would appear that labor income tax cuts impact positively the U.S. economy's output not only in the long run, but also in the short run. When analyzing the effects of the supply and the demand disturbances on the tax rates, our research indicates that most of the fluctuations in the tax rate are due to demand disturbances.

[21] On the contrary, most economists are in agreement on the effects of capital gains taxation on an economy. Harberger (1966), Chamley (1981), Jorgenson and Yun (1990), and Lucas (1990) argue strongly against taxing any form of income from capital.

The results of our research have important implications for the use and the effectiveness of the aggregate demand and aggregate supply fiscal policies in the U.S. Our research indicates that the supply-side fiscal policy is more effective in promoting economic growth than the fiscal policy that focuses on stimulating the aggregate demand. It appears that reducing labor income taxes affects the U.S. economy's output not only in the long run, but also in the short run. At the same time, it is evident that the traditional Keynesian side effect of tax policies primarily impacts the U.S. economy in the short run. Given these results, it is fair to conclude that reducing the labor income tax may be the most appropriate economic policy to implement for achieving economic growth in the U.S. One additional general advantage of relying on the supply-side economic policies for the purposes of economic stabilization is the fact that such policies, unlike the traditional Keynesian aggregate demand policies, do not have an inflationary bias.

References

Blanchard, Olivier, and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *The Quarterly Journal of Economics*, 117(4): 1329-68.

Blanchard, Olivier Jean, and Danny Quah. 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review*, 79(4): 655-73.

Blanchard, Olivier. 2006. *Macroeconomics*, 4th ed., New York: Prentice Hall.

Bohanon, Cecil E., and T. Norman Van Cott. 1986. "Labor Supply and Tax Rates: Comment." *American Economic Review*, 76(1): 277-9.

Bureau of Economic Analysis (http://www.bea.gov/national/nipaweb/)

Campbell, John Y., and N. Gregory Mankiw. 1987. "Are Output Fluctuations Transitory?" *The Quarterly Journal of Economics*, 102(4): 857-80.

Chamley, Christophe. 1981. "The Welfare Cost of Capital Income Taxation in a Growing Economy." *Journal of Political Economy*, 89(3): 468-96.

Clarida, Richard, and Joe Prendergast. 1999. "Fiscal Stance and the Real Exchange: Some Empirical Estimates." NBER Working Paper No. 7077.

Claus, Iris. 1999. "Estimating Potential Output for New Zealand: A Structural VAR Approach." Federal Reserve Bank of New Zealand Discussion Paper 2000/03.

Dickey, David A., and Wayne A. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series with Unit Root." *Journal of American Statistical Association*, 74(366a): 427-31.

Edelberg, Wendy, Martin Eichenbaum, and Jonas D. M. Fisher. 1999. "Understanding the Effects of a Shock to Government Purchases." *Review of Economic Dynamics*, 2(1): 166-206.

Enders, Walter. 2003. *Applied Econometric Time Series*, 2nd ed. New York: John Wiley and Sons.

Enders, Walter, and Bong-Soo Lee. 1997. "Accounting for Real and Nominal Exchange Rate Movements in the Post-Bretton Woods Period." *Journal of International Money and Finance*, 16(2): 233-54.

Fatás, Antonio, and Ilian Mihov. 2000. "The Macroeconomic Effects of Fiscal Policy." Working Paper.

Friedman, Milton. 1968. "The Role of Monetary Policy." *American Economic Review*, 58(1): 1-17.

Fuller, Wayne A. 1976. *Introduction to Statistical Time Series*. New York: John Willey and Sons.

Galí, Jordi. 1992. "How Well Does the IS-LM Model Fit Post-War U.S. Data." *The Quarterly Journal of Economics*, 107(2): 709-38.

Galí, Jordi, J. David López-Salido, and Javier Vallés. 2007. "Understanding the Effects of Government Spending on Consumption." *Journal of the European Economics Association*, 5(1): 227-70.

Gwartney, James D., and Richard Stroup. 1983. "Labor Supply and Tax Rates: A Correction of the Record," *American Economic Review*, 73(3): 446-51.

Hamilton, James D. 1994. *Time Series Analysis*. Princeton University Press.

Harberger, Arnold C. 1966. "Efficiency Effects of Taxes on Income from Capital" in *Effects of the Corporation Income Tax*, ed. Marian Krzyzaniak. Detroit: Wayne State University Press, 107-17.

Joines, Douglas H. 1981. "Estimates of Effective Marginal Tax Rates on Factor Incomes." *Journal of Business*, 54(2): 191-226.

Jones, John Bailey. 2002. "Has Fiscal Policy Helped Stabilize the Postwar U.S. Economy?" *Journal of Monetary Economics*," 49(4): 709-46.

Jorgenson, Dale W., and Kun-Young Yun. 1990. "Tax Reform and U.S. Economic Growth," *Journal of Political Economy*, 98(5): S151-93.

King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson. 1991. "Stochastic Trends and Economic Fluctuations." *American Economic Review*, 81(4): 819-40.

Lippi, Marco, and Lucrezia Reichlin. 1994. "Diffusion of Technical Change and the Decomposition of Output into Trend and Cycle." *Review of Economic Studies*, 61(1): 19-30.

Lucas Jr., Robert E. 1990. "Supply-Side Economics: An Analytical Review." *Oxford Economic Papers*, 42: 293-316.

Mankiw, N. Gregory., and Matthew Weinzierl. 2005. "Dynamic Scoring: A Back-of-the-Envelope Guide." Harvard Institute of Economic Research, Discussion Paper No. 2057.

Mendoza, Enrique G., Assaf Razin, and Linda L. Tesar. 1994. "Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption." *Journal of Monetary Economics*, 34(3): 297-323.

Nelson, Charles R., and Charles R. Plosser. 1982. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications." *Journal of Monetary Economics*, 10(2): 130-62.

Perotti, Roberto. 2002. "Estimating the Effects of Fiscal Policy in OECD Countries." European Central Bank Working Paper No. 168.

Phillips, Peter C.B., and Pierre Perron. 1988. "Testing for a Unit Root in Time Series Regression." *Biometrika*, 75(2): 335-46.

Ramey, Valerie A., and Matthew D. Shapiro. 1997. "Costly Capital Reallocation and the Effects of Government Spending." NBER Working Paper No. 6283.

Zivot, Eric, and Donald W. K. Andrews. 1992. "Further Evidence on the Great Crush, the Oil-Price Shock, and the Unit-Root Hypothesis." *Journal of Business and Economic Statistics*, 10(3): 251-70.

Appendix A

The ADF test results conducted on the supply component of the output under the 'no trend' and the 'trend' (deterministic) assumptions as well as on the demand components of output under the random walk and the random walk with drift assumptions are reported in Table 7. We fail to reject the null of the presence of a unit root for both specifications of the supply component of output at 5 percent level. The presence of a unit root in the supply component of output under the 'trend' (deterministic) specification is indicative of a stochastic trend. Also, we reject the null hypothesis of the presence of a unit root for both specifications of the demand component of output at 5 percent level.

Table 7: Augmented Dickey-Fuller (ADF) Test Results of the Permanent and the Cyclical Components

	Y_t^{Supply}		Y_t^{Demand}	
	No Trend	Trend	Without Constant	With Constant
p	1	0	0	0
a_0	0.5936	5.6067	---	6.1323
	(0.5936) [a]	(2.8827)		(2.5879)
a_2	---	0.1081	---	---
		(2.9620)		
γ	0.0034	-0.1339	-0.0022	-0.1210
	(0.2766)	(-2.8487)	(-0.5925)	(-2.6273)
$\gamma+1$	1.0034	0.8661	0.9978	0.8790
	(0.2766)	(-2.8487)	(-0.5925)	(-2.6273)
ADF Test Statistic	0.2766	-2.8487	-0.5925*	-2.6273*
t-critical (5% level)	-2.8882	-3.4512	-1.9438	-2.8879
Observations	109	110	110	110

Note: Y_t^{Supply} and Y_t^{Demand} represent the supply component of real GDP and the demand component of real GDP respectively.

[a] t-statistics are in parentheses.

* Reject null hypothesis of the presence of unit root at 5 percent significance level.

The Journal of Business Inquiry 2014, 13, Issue 2, 128-141
http:www.uvu.edu/woodbury/articles
ISSN 2155-4072

The Formation of the Desire for Retribution

By JEFF PETERSON*

In this paper I examine how aspects of a person who commits an organizational violation affect a third-party observer's desire that the person be punished. Specifically I look at how the third party's desire for retribution is affected by the offending party's past behavior and offer of an apology. I further propose a model of desire for retribution in which observers rely on aspects of the violation (such as severity) and aspects of the violator (such as a previous history of the violation) to determine the strength of this desire. Using repeated measures ANOVAs I found that apologies reduce the desire for punishment (F=8.55, p<.01, eta=.09), while a history of the offense increases it (F=11.08, p<.00, eta=.12). Also, desire for punishment is highest when there is no apology with a previous history of the offense and is lowest when there is an apology and no history (F=12.95, p<.00, eta=.13). Violation severity has a main effect on desire for retribution (F=24.48, p<.00, eta=.20) and it also interacts with apology and history, with history making a difference regardless of severity, but apologies having no effect with severe violations (F=12.95, p<.00, eta=.13).

Keywords: Punishment, Justice, Retribution, Attitudes

JEL Classification: M12

I. Introduction

People in organizations make mistakes, perform poorly, steal, cheat and do other things that are contrary to the interests of their organization (Vardi and Wiener, 1996). One of the responsibilities of managers in modern organizations is dealing with subordinates who engage in these behaviors. This often takes the form of administering some type of punishment (Butterfield *et al.*, 2005).

Most of the research on organizational justice has focused on the employer-employee dyad. For example, a number of punishment studies have been conducted looking at the manager's response to a subordinate's poor performance (Ashkanasy and Gallois, 1994; Crant and Bateman, 1993; Green and Mitchell, 1979; Kipnis and Cosentino, 1969; Klaas and Wheeler, 1990; Miner, 1976; Mitchell *et al.*, 1981; Mitchell and O'Reilly, 1983) or on the effects of punishment on specific subordinate behaviors that influence a punishment's effectiveness (Arvey and Ivancevich, 1980; Arvey and Jones, 1985). And a few studies have examined the experience of the recipients (Atwater *et al.*, 2001).

*Assistant Professor: Department of Business Management, Utah Valley University, 800 West University Parkway, Orem, UT 84058-5999. E-mail: jeff.peterson@uvu.edu. Phone: (425) 773-6262.

Many of these studies have found negative consequences associated with punishment episodes. For example, a number of these studies have found that while punishment may be effective in changing an employee's behavior in the short term, recipients often experience resentment, hostility and can even engage in sabotaging behaviors in response to being punished (Arvey and Ivancevich, 1980; Arvey and Jones, 1985; Butterfield *et al.*, 1996). There also have been serious challenges raised about the effectiveness of punishment as a motivator of performance (Atwater *et al.*, 2001). These studies imply that perhaps managers should avoid punishment and look to other methods to influence employees to not engage in misconduct.

However, managers have more to be concerned about than just the direct effectiveness of punishment on the recipient's behavior. Researchers have extended findings beyond the recipient-centered approach to include looking at the experience of the manager (Butterfield *et al.*, 1996) and at punishment as a social experience, by including the reactions of observers (Treviño, 1992). For example, it may be that punishing an individual does indeed have the negative consequences that have been suggested by some and found in a few studies. However, it may be that avoiding punishment has negative outcomes for observers. If this is the case, then it would be the classic situation of the needs of the many outweighing the need of the few, or at a minimum, that managers would need to consider both perspectives in order to find a solution that minimizes the impact on all parties.

There is a growing amount of research on the importance of third parties in organizations. Darley and Pittman (2003) describe the psychological processes that cause third parties to take an interest in whether organizational members are treated fairly. Other studies have found that third parties actively assign blame to violators for actions that had no impact on the third party (Alicke, 1992; Shaver, 1970; Walster, 1966). And research shows that third parties will sanction violators even when the sanctions are costly to them and the violation itself had no affect on the third party (Fehr and Fischbacher, 2004; Turillo *et al.*, 2002). Research on observers' reactions has grown to include areas such as reactions to layoffs (Skarlicki *et al.*, 2008; Skarlicki *et al.*, 1998), and mistreatment of coworkers by managers (O'Reilly and Aquino, 2011; Skarlicki and Kulik, 2005). However, there is still little known about how third parties react to the punishment of other organizational members. This paper seeks to begin to fill that gap.

II. Theory and Hypotheses

A. Third-Party Observers

In this paper I define third-party observers as individuals who become aware that an employee has committed an organizational infraction. This information may come from direct observation of the offense, or through other means such as hearing rumors or becoming aware that someone has been punished. Regardless of the source, the third party has become aware that the employee has violated the organization's rules. Third-party observers can be coworkers, part of the violator's social network, other organizational members, and even people outside of the organization (Skarlicki and Kulik, 2005).

This paper, however, is primarily concerned with third-party observers who are also coworkers. While many people might take an interest in an employee's violation, how the observer is related to the violator makes a difference in how they react. For example, Chi and Lo (2003) found that coworker's perceptions of justice were affected by the quality of their relationship to the violator. Simply liking a person more affected how they perceived the punishment. This suggests that, for

example, being a third party in a different department might result in different reactions than you would see from a coworker. Therefore, in this paper I will focus primarily on the coworker as the third party simply to eliminate the moderating effect of the observer's general relationship to the violator.

B. Retributive Justice – The Desire for Punishment

Third-party reactions are important because of the fact that as social beings, we are highly concerned about social norms, and the violation of important norms has been shown to influence the behavior of third-party observers (Fehr *et al.*, 2002). A key theory in examining the reactions of third parties is Folger's (2001) concept of deontic justice. The term deontic justice refers to the psychological state where people react emotionally to actions that violate social norms of conduct. This specifically includes the notion of how people should or ought to be treated. Therefore, these feelings of a desire for fairness come from a moral conviction about appropriate behavior. This is what provides the intensity behind third-party observer's reactions to events that do not impact them directly.

Fehr and Fischbacher (2004) offer an evolutionary explanation for why a person is so concerned with enforcing social norms even when the violations do not affect the individual directly. Essentially, for a given society to function, there must be a way for norms to be enforced, otherwise, interactions become chaotic and unpredictable. In any given interaction, it is possible for one party to take advantage of the other. If it were left only to the second party to ensure that things were fair, it becomes much easier for any given individual to take advantage. We would become subject to the whims of the powerful or deceitful. Societies where interested parties ensured that all members were treated fairly were more successful than those where certain individuals were able to manipulate other individuals. Over time, we have simply been shaped to care about maintaining the structure of societal norms. This is best accomplished by taking an interest in the situation of our fellow beings. While Fehr and Fischbacher (2004) argue that their research shows that people are not simply self-interested, one can also argue that it is actually in one's self-interest to prevent people from abusing societal norms. While the desire for retribution may have arisen from evolutionary pressures, what we now experience as human beings is a moral outrage of a certain degree when someone violates a norm and gets away with it, even if we were not directly harmed.

Because people care about how other people are treated even when the actions do not impact them directly, it is logical that witnessing someone commit a violation and seeing a manager punish the violator will result in emotional reactions from observers that will eventually result in important behaviors. However, to this point, little research has examined the ways that observers react and what factors moderate their reactions. One likely set of moderators are the various aspects of the violator. For example, how is the third party related to the violator? Did the violator express remorse? Does the violator have a history? Does it appear that the violation was done intentionally? Presumably, these variables will impact the strength of the third party's moral reaction to the violation. And a second likely set of moderators are the aspects of the violation. How severe was it? What was the outcome? Who was impacted? These also should impact the observer's moral reaction.

This paper proposes a basic model for the formation of the desire for retribution. When an observer sees a violation, they look at the aspects of the violation (how severe, the outcome, etc.) and aspects of the violator (is there a history of the violation, did the violator apologize) and then makes a judgment about how much punishment the violator deserves. First, this basic model needs to be explored to see if that is indeed how observers behave, and secondly the many aspects of violations and violators need to be explored to see which matter, which do not, and how they interact

with each other. To date we know very little about this area, and this paper seeks to begin the process of exploring this model. While the model is not particularly complex, it is possible that all of the many aspects could have very complex relationships and interdependencies.

C. Presence of an Apology

One important aspect of a violator is whether they are remorseful and specifically whether they offer an apology. Apology has been defined as an utterance intended to remedy a social disruption (Scher and Darley, 1997). Research has shown that when a violator apologizes, those they have offended behave less aggressively towards them (Ohbuchi *et al.*, 1989). Bisel and Messersmith (2012) examined how employees reacted to apologies from supervisors and from organizational representatives apologizing on behalf of an organization and found that employees were more forgiving when an apology was offered. However, while there is a strong consensus that apologies work in both private and organizational settings between the violator and the person who was wronged, the question remains as to whether an apology between two parties has an effect on the third-party observer. It makes sense however, that if an apology is meant to remedy a social disruption that it might also work on an observer. The presence of an apology should serve as a reducing factor in an observer's feelings about a violation. Therefore, I propose:

> *H1: When a violator offers an apology, coworkers will have less desire for retribution than they will when the coworker does not offer an apology.*

D. History of Previous Violations

One issue related to retributive justice is the extent to which a violator can be held responsible for their behavior (Kidd and Utne, 1978). Niehoff *et al.* (1998) looked at a violator's past performance and found that when the violator was a good performer, observers had less desire for retribution than they did when the violator was a poor performer. However, they did not examine the case where the violator had committed the same violation before. Previous work has shown that a violator's history of committing a specific violation resulted in the observer making an attribution that the violator willingly engaged in the behavior (Klaas and Wheeler, 1990). A previous violation of the extract nature should indicate a greater level of culpability than situations where workers had not committed prior transgressions. If the violator has a previous history of committing the violation, that fact should be an additional aspect that should influence third-party evaluations. Therefore, I propose:

> *H2: When a violator has a history of the violation, coworkers will have more desire for retribution than they will when the coworker does not have a history of the violation*

These two aspects of the violator, apology and previous history, should each contribute to the desire for retribution such that they will have an additive effect, meaning that the presence of an apology and the absence of a previous history should result in the least negative feelings towards the violator and the absence of an apology and the presence of a previous history should result in the most negative feelings, with the other two combinations being between the two extremes. This assumes that these two are relatively equal in strength. Therefore, I propose:

H3: Desire for retribution will be highest when the violator does not apologize and has a history of the violation, while desire for retribution will be lowest when the violator apologizes and has no history of the violation, with the other two combinations having intermediate values.

E. Violation Severity

While aspects of the violator are likely to be quite important in the context of an organizational violation, different violations carry with them differing levels of moral outrage regardless of who commits them. Research has shown that people are naturally able to assess the seriousness of different violations and that there is considerable consistency in the rank ordering of common violations (Warr *et al.*, 1983). Carlsmith *et al.* (2002) argue that observers have a desire to make punishment proportional to the violation. In other words, the punishment should fit the crime. If this is the case, more severe violations should result in a greater desire for retribution. Therefore I propose:

H4: Severe violations will result in a stronger desire for retribution than will mild violations.

While severity should be a strong factor in the desire for retribution, there is also the possibility that the violation's severity will interact with the aspects of the violator. For example, a mild violation might be amenable to an apology, where a sufficiently severe violation might render an apology meaningless. Therefore, I propose:

H5: For severe violations, desire for retribution will not be affected by the violator's apology or history of the violation.

H6: For mild violations, desire for retribution will be highest when the violator does not apologize and has a history of the violation, while desire for retribution will be lowest when the violator apologizes and has no history of the violation, with the other two combinations having intermediate values.

III. Methods

A. Study Design

This study used a nested repeated-measures design where each subject was measured on the dependent variable multiple times. The subjects were initially divided into two main groups based on the severity of the violation. They were then presented with multiple scenarios and given instructions to vividly imagine that the scenarios referred to a coworker. I chose this design because it required a smaller number of subjects as compared to alternate designs. The advantages of a within-subjects design are two fold. First, it increases power by requiring fewer subjects. Since the various treatments are applied to each subject, I needed many fewer subjects than I would have using a completely between-subjects design. A second reason for using a within-subjects design is that it reduces the error variance that is associated with individual differences. This makes it clearer that the difference between the treatments is caused by the treatment, rather than some difference that exists between characteristics of the subject.

There are also some inherent weaknesses of the within-subject design, namely what are known as "*carryover effects*." Since a subject receives multiple treatments, it is more difficult to conclude that each treatment is completely independent. It is possible that the previous treatment has had an influence on the subject and therefore I cannot be as confident that any given treatment would be equally effective in isolation. However, to minimize this effect subjects had the four permutations of the vignette presented in a random order, which should reduce any order effect, but cannot completely compensate for the fact that each subject ultimately was presented with all four variations. Repeated-measures ANOVA attempts to compensate for this statistically.

B. Sample

I obtained subjects from the StudyResponse Project at Syracuse University. StudyResponse maintains a large panel of subjects who are interested in being research subjects for a modest payment. With their large base it is easy to request subjects who meet certain criteria, such as working adults, equally divided between men and women. I selected StudyResponse in order to quickly obtain subjects who were not students and had substantial work experience. StudyResponse sent an invitation via email to members of their database inviting them to participate. The invitation went to an equal number of male and female subjects. The invitation directed the subjects to a website which contained the survey. A total of 107 subjects completed the survey: 14 were discarded because of incomplete data and 9 were discarded for not following instructions properly. This resulted in 84 complete and usable responses.

Of the respondents 54.8% were male, 70.6% were Caucasian, 11.8% were Asian/Pacific Islander, 3.5% were African American, 2.4% were Hispanic, 2.4% were Multiethnic, 1.2 were Native American and 1.2% identified themselves as other. Age ranged from 19 to 72 years with an average of 38.3 years. Work experience ranged from 1 to 50 years with an average of 21.4 years. 84% reported having had some supervisory experience. Data on non-respondents was available for gender, race and age. Gender and race between responders and non-responders were non-significant, however age was significant (F=12.56, p<.01) with respondents being significantly older (34.2 years for non respondents and 38.3 for respondents).

C. Procedures

When subjects received the email invitation from StudyResponse, they clicked on a link to the survey page where they are assigned randomly to one of the two conditions. Subjects were told to imagine that person they are thinking of has been caught selling a stolen company laptop containing sensitive employee data. For the four apology and history conditions subjects were given four different scenarios in random order where the violator either did or did not apologize and did or did not have a history of doing the behavior previously. After each scenario the subjects were asked questions about their desire for retribution and then told to disregard all previous information for the next scenario, in essence asking them to "now imagine that instead your coworker did the following." In the other condition, subjects were told to imagine that the person they are thinking of had called in sick when in reality they were taking a fishing trip, causing extra work for the remaining employees. Again the scenarios were manipulated with the four variations of apology and history.

D. Measures

D.1 Desire for Retribution

I measured the coworker's desire for retribution using a three-item measure consisting of two items developed by Niehoff *et al.* (1998), with a third item (question #3) added for this study. The two-item measure had a coefficient Alpha of .91 in the Neihoff *et al.* study. For the current study the modified measure had a coefficient Alpha of .90. Responses were measured using Likert scales ranging from "strongly agree" (1) to "strongly disagree" (7). The items were:

1) The actions of the worker should have been punished
2) The employee deserved to be disciplined
3) It would bother me if this person was not punished

The three questions were consolidated with simple averaging after confirming their relatedness with Cronbach's Alpha. All statistics were conducted using the consolidated measure.

IV. Results

Hypothesis 1 stated that when a violator offers an apology, coworkers will have less desire for retribution than they will when the coworker does not offer an apology. To test all the hypotheses simultaneously I ran a repeated-measures ANOVA with apology and history as within subjects factors and violation severity as a between-subjects factor. This allows me to see all of the various interactions in addition to the main effects of each of the variables. The main effect for apology was significant ($F(1, 83)=8.55$, $p<.01$). The estimated marginal means showed that desire for retribution was lower when the violator apologized ($M=5.14$) than when violator did not apologize ($M=5.34$). This shows that apologizing for a violation does reduce an observer's desire for retribution. Therefore, hypothesis 1 was fully supported. Analysis results are presented in Table 1, with the descriptive measures presented in Table 2.

Table 1: Effect of Apology on Desire for Retribution

Effect	df	Mean Square	F	Sig.	Partial Eta Squared
Apology	1	3.26	8.55	.01	.09

Table 2: Descriptives for Apology

Apology	Mean	Std. Error
Apology	5.14	.140
No Apology	5.34	.146

Hypothesis 2 stated that when a violator has a history of the violation, coworkers will have more desire for retribution than they will when the coworker does not have a history of the violation. Using the same repeated-measures ANOVA described above I found that the main effect of history was significant ($F(1, 83)=11.08$, $p<.01$). When the violator had a history of the violation, subjects reported more desire for retribution ($M=5.38$) than when the violator had no previous history ($M=5.09$). This shows that a history of the violation increases the desire for retribution. Thus hypothesis 2 was fully supported.

Table 3: Effect of History on Desire for Retribution

Effect	df	Mean Square	F	Sig.	Partial Eta Squared
History	1	7.30	11.08	.00	.12

Table 4: Descriptives for History

History	Mean	Std. Error
No History	5.09	.140
History	5.38	.146

Hypothesis 3 stated that desire for retribution will be highest when the violator did not apologize and has a history of the violation, while desire for retribution will be lowest when the violator apologized and had no history of the violation, and when a violator apologizes but has a history of the violation or when a violator did not apologize but had no history of the violation, desire for retribution will greater than when there is an apology and no history, but lower than when there is no apology and a history. This relationship is more difficult to test since it represents gradations of values. It is unlikely that each cell would be significantly different than the other cells. Therefore, I ran a repeated-measures ANOVA between the condition with no apology and a history of the violation and the condition with an apology and no history (the two combinations representing the extremes). The difference between the means for desire for retribution was significant ($F(1, 83)=12.95$, $p<.01$) When the violator had a history of the violation and did not apologize, subjects reported more desire for retribution ($M=5.46$) than when the violator had no previous history and did apologize ($M=4.96$). The two intermediate conditions returned intermediate values ($M=5.32$) and ($M=5.22$). This relationship suggests that there is a simple additive effect taking place rather than a more complex moderating relationship. Therefore, hypothesis 3 was fully supported.

Table 5: Mean Difference Between Extreme Conditions

Effect	df	Mean Square	F	Sig.	Partial Eta Squared
Apology and No History vs. No Apology and History	1	10.05	12.95	.00	.13

Table 6: Descriptives for Permutations of Apology and History

History	Mean	Std. Error
Apology and No History	4.97	.177
Apology and History	5.32	.149
No Apology and No History	5.22	.161
No Apology and History	5.46	.160

Hypothesis 4 stated that severe violations will result in a stronger desire for retribution than will mild violations. To test this I ran the same repeated-measures ANOVA as in hypotheses 1 and 2 using the four scenarios made up of the permutations of history and apology (e.g. history and apology, history and no apology, etc.) with severity (severe or mild) as a between-subjects factor. This resulted in 42 subjects for the mild condition and 42 subjects for the severe condition. The between-subjects factor of severity was significant ($F(1, 83)=24.48$, $p<.01$). When the violation was severe, subjects reported more desire for retribution *(*$M=5.93$) than when the violation was mild ($M=4.55$). Therefore, hypothesis 4 was fully supported. The statistical results are presented in Table 7, and the descriptive results are presented in Table 8.

Table 7: Effect of Severity on Desire for Retribution

Effect	df	Mean Square	F	Sig.	Partial Eta Squared
Severity	1	161.01	24.48	.00	.23

Table 8: Descriptives for Severity

Severity	Mean	Std. Error
Illness	4.55	.198
Laptop	5.93	.196

Hypothesis 5 stated that for severe violations, desire for retribution will not be affected by the violator's apology or history of the violation. Proposing the null is not a common practice, but combining this with hypothesis 6 provides theoretical justification for this procedure. To test this I selected only subjects who were presented the severe scenarios and then ran the same repeated-measures ANOVA as in hypotheses 1 and 2. This results in 42 subjects. Apology was not significant ($F(1, 41)=.48$, $p=.49$). However, history was significant ($F(1, 42)=4.95$, $p<.05$). When the violation was severe and the violator had a history, subjects reported more desire for retribution *(*$M=6.05$) than when the violator did not have a history ($M=5.81$). This means that it didn't make a difference whether the violator apologized, but it did make a difference if the violator had a history. Therefore, hypothesis 5 was partially supported with apologies not influencing desire for retribution, but history still making a difference. Statistical and descriptive results of these analyses are presented in tables 9, 10, and 11.

Table 9: Effect of Apology on Desire for Retribution for Severe Violations

Effect	df	Mean Square	F	Sig.	Partial Eta Squared
Apology	1	.13	.48	.49	.011

Table 10: Effect of History on Desire for Retribution for Severe Violations

Effect	df	Mean Square	F	Sig.	Partial Eta Squared
History	1	2.48	4.95	.03	.11

Table 11: Descriptives for History for Severe Violations

History	Mean	Std. Error
No History	5.81	.178
History	6.05	.164

Hypothesis 6 stated that for mild violations, desire for retribution will be highest when the violator did not apologize and has a history of the violation, while desire for retribution will be lowest when the violator apologized and had no history of the violation, and when a violator apologizes but has a history of the violation or when a violator does not apologize but has no history of the violation, desire for retribution will greater than when there is an apology and no history, but lower than when there is no apology and a history. To test this I ran the same repeated-measures ANOVA that I ran in hypothesis 3 after selecting only those subjects who had a mild violation. This resulted in 42 subjects. Since hypothesis 3 was confirmed and hypothesis 5 was partially confirmed it is not surprising that the variance is coming mostly from the mild condition. The difference between the means for desire for retribution was significant ($F(1, 41)=10.11$, $p<.01$) When the violator had a history of the violation and did not apologize, subjects reported more desire for retribution ($M=4.88$) than when the violator had no previous history and did apologize ($M=4.20$). The two intermediate conditions both returned intermediate values ($M=4.56$). This relationship again suggests that there is a simple additive effect taking place rather than a more complex moderating relationship. Therefore, hypothesis 6 was fully supported. The pertinent analytic results are presented in tables 12 and 13.

Table 12: Mean Difference Between Extreme Conditions for Mild Violations

Effect	df	Mean Square	F	Sig.	Partial Eta Squared
Apology and No History vs. No Apology and History	1	9.78	10.11	.00	.20

Table 13: Descriptives for Permutations of Apology and History for Mild Violations

History	Mean	Std. Error
Apology and No History	4.20	.243
Apology and History	4.56	.240
No Apology and No History	4.56	.267
No Apology and History	4.88	.246

V. Discussion

The first thing that this study does is to reconfirm that employees do indeed care about the violations of their coworkers. I also found support for the proposition that the desire for retribution comes from assessments of the aspects of the violation and the violator. As they are given information about the violation and the violator, their desire to see that justice takes place and their feelings towards the violator are significantly impacted. Not surprisingly, there is a strong link between the seriousness of a violation and how much coworkers want to see the violator punished. This was true when there was no other information than the basic details of the violation, as well as when they were given information about different aspects of the violator. Severity matters on top of other variables, showing that this is a key part of the evaluation of observers. Serious violations result in more desire for retribution and more negative feelings toward the violator regardless of other factors. This would suggest that a manager should first take into account the overall seriousness of the violation before considering other factors. It is also appears that severity may be prone to anchoring and adjustment bias, where any subsequent information is used to adjust the initial anchor of severity. So, observers first decide on a level of desire for retribution based on the severity and then make minor adjustment based on other relevant factors.

A second and somewhat surprising result was that severity interacts differently with different aspects of the violator. It seems logical to suppose that severe violations might have a different structure than do mild violations. For example, one might assume that with severe violations, the characteristics of the violator would be swamped by the magnitude of the violation. For example, when someone steals a laptop with sensitive employee data on it, it would not matter if it is a first violation or a repeat violation, this act is so serious that the outrage makes other details irrelevant. However, I did not find that here. Apologies do not make any difference when a violation is serious, but a history still does. It is likely that each aspect of the violator interacts with severity in an idiosyncratic factor, meaning that we need to understand which factors when combined with severity will add to an observer's desire for retribution, and which will be discarded as irrelevant. And while it was not observed here, it may be that for mild violations some things matter and some do not.

The model proposed that the formation of the desire for retribution takes place by examining both aspects of the violator and the violation. The results support this assertion. The results also support the idea that these aspects seem to work in a complicated way. Sometimes they are simply additive, where each contributes a small amount of variance and in other cases they exert large amounts of influence and other times no influence. The mental calculus of desire for retribution may present some interesting twists and turns, even if it is completely invisible even to the person who seemingly without effort comes up with a final judgment of how much the violator is deserving of punishment.

VI. Implications for Practice

If an observers' desire for retribution is driven both by the seriousness of the violation and characteristics of the violator, then managers would need to take this into account when deciding on what punishment is appropriate. We can assume that if an observer has a strong desire for retribution and a manager does not punish or give a slap on the wrist to the violator, then the observer will experience feelings of injustice that could have negative implications. Managers may also need to manage the perceptions of their employees. Because of the strength of the effect, managers should clearly elucidate the seriousness of a violation. If observers assume a mild violation and see a strong

response, or they assume a severe violation and see a weak response, they will likely feel that the punishment is unjust. And severity seems to matter regardless of what other information is obtained.

Second, after punishment has taken place, if a manager perceives that observers have concerns about the response, providing additional details could alter the observer's perceptions. For example, if an observer concludes that a punishment was too severe, a manager could explain that the violator had a history of the violation. Or, on the other hand, if a coworker felt the punishment was too lenient, the manager could explain that an apology had been offered, or that the violation did not result in a negative outcome.

And third, is simply a confirmation that observers do indeed have significant reactions to the violations of coworkers. Managers would be well served to keep in mind that a punishment episode is not just between the manager and the violator, but if the violation is publicly known, then the manager does need to think about how the observers will react. At the very least they will have more negative attitudes about the violator, and if there is high interdependence in a group or team these attitudes could negatively affect performance. This presents a bit of a quandary at this point. Most HR departments would not react well to the idea of sharing with coworkers the details of a violation and punishment of another employee. However, studies like this can help to demonstrate the need to make others aware of the punishment and the aspects involved in a violation when the violation itself is already widely known.

VII. Limitations

Because I was asking subjects to imagine, they may not have been able to accurately assess their true desires and feelings. For example, it may be difficult for them to actively imagine a close coworker committing a violation if that person is of high integrity. Additionally, this study only examined the formation of the desire for retribution and not what happens after the fact. Presumably, upon administration of a punishment the observer's attitudes would be adjusted to take into account the new information. Additional research may reveal that some of these aspects are not relevant once punishment has taken place, or they could continue their effect even after punishment.

Another issue is the potential for demand characteristics. This issue often occurs in a repeated-measures design. Each time the subjects get a new scenario, they see a variation on what they previously saw. It is not hard to imagine that subjects could infer the point of the study by observing what was being changed at each stage. This may have influenced their responses so that they answered as they thought the question should be answered, according to their implicit theories of punishment (or according to the researcher's theory), rather than exactly how they felt. The only way to avoid this would have been to give each subject only one situation, which would have required many more subjects. And lastly, since this was done using a survey panel, it is possible that subjects did not take the study as seriously as would employees in their actual workplace.

REFERENCES

Alicke, Mark D. 1992. "Culpable Causation." *Journal of Personality and Social Psychology,* 63(3): 368-78.

Arvey, Richard D., and John M. Ivancevich. 1980. "Punishment in Organizations: A Review, Propositions, and Research Suggestions." *Academy of Management Review,* 5(1): 123-32.

Arvey, Richard D., and Allen P. Jones. 1985. "The Use of Discipline in Organizational Settings: A Framework for Future Research." In *Research in Organizational Behavior*, ed. Barry M. Staw and Larry L. Cummings, Vol. 7: 367-408, Greenwich, CT: JAI Press.

Ashkanasy, Neal M., and Cynthia Gallois. 1994. "Leader Attributions and Evaluations - Effects of Locus of Control, Supervisory Control, and Task Control." *Organizational Behavior and Human Decision Processes,* 59(1): 27-50.

Atwater, Leanne E., David A. Waldman, James A. Carey, and Priscilla Cartier. 2001. "Recipient and Observer Reactions to Discipline: Are Managers Experiencing Wishful Thinking?" *Journal of Organizational Behavior,* 22(3): 249-70.

Bisel, Ryan S., and Amber S. Messersmith. 2012. "Organizational and Supervisory Apology Effectiveness: Apology Giving in Work Settings." *Business Communication Quarterly,* 75(4): 425-48.

Butterfield, Kenneth D., Linda Klebe Treviño, and Gary A. Ball. 1996. "Punishment From the Manager's Perspective: A Grounded Investigation and Inductive Model." *Academy of Management Journal,* 39(6): 1479-512.

Butterfield, Kenneth D., Linda Klebe Treviño, Kim J. Wade, and Gary A. Ball. 2005. "Organizational Punishment From the Manager's Perspective: An Exploratory Study." *Journal of Managerial Issues,* 17(3): 363-83.

Carlsmith, Kevin M., John M. Darley, and Paul H. Robinson. 2002. "Why Do We Punish? Deterrence and Just Deserts as Motives for Punishment." *Journal of Personality and Social Psychology,* 83(2): 284-99.

Chi, Shu-Cheng, and Hsin-Hsin Lo. 2003. "Taiwanese Employees' Justice Perceptions of Co--workers' Punitive Events." *The Journal of Social Psychology,* 143(1): 27-42.

Crant, J. Michael, and Thomas S. Bateman. 1993. "Assignment of Credit and Blame for Performance Outcomes." *Academy of Management Journal,* 36(1): 7-27.

Darley, John M., and Thane S. Pittman. 2003. "The Psychology of Compensatory and Retributive Justice." *Personality and Social Psychology Review,* 7(4): 324-36.

Fehr, Ernst, and Urs Fischbacher. 2004. "Third-Party Punishment and Social Norms." *Evolution and Human Behavior,* 25(2): 63-87.

Fehr, Ernst, Urs Fischbacher, and Simon Gächter. 2002. "Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms." *Human Nature,* 13(1): 1-25.

Folger, Robert. 2001. "Fairness as Donance." In *Research in Social Issues in Management,* ed. Stephen W. Gilliland, Dirk D. Steiner and Daniel P. Skarlicki, 3-31. Greenwich, CT: Information Age.

Green, Stephen G., and Terence R. Mitchell. 1979. "Attributional Processes of Leaders in Leader-Member Interactions." *Organizational Behavior and Human Performance,* 23(3): 429-58.

Kidd, Robert F., and Mary K. Utne. 1978. "Reactions to Inequity: A Prospective on the Role of Attributions." *Law and Human Behavior,* 2(4): 301-11.

Kipnis, David, and Joseph Cosentino. 1969. "Use of Leadership Powers in Industry." *Journal of Applied Psychology,* 53(6): 460-66.

Klaas, Brian S., and Hoyt N. Wheeler. 1990. "Managerial Decision-Making About Employee Discipline: A Policy-Capturing Approach." *Personnel Psychology,* 43(1): 117-34.

Miner, John B. 1976. "The Management of Ineffective Performance." In *The Handbook of Industrial and Organizational Psychology,* ed. Marvin D. Dunnette, 995-1029, Chicago: Rand McNally.

Mitchell, Terence R., Stephen G. Green, and Robert E. Wood. 1981. "An Attributional Model of Leadership and the Poor Performing Subordinate: Development and Validation." In *Research in Organizational Behavior,* ed. Barry M. Staw and Larry L. Cummings, 197-234. Greenwich, CT: JAI Press.

Mitchell, Terence R., and Charles A. O'Reilly. 1983. "Managing Poor Performance and Productivity in Organizations." *Research in Personnel and Human Resources Management,* 1, 201-34.

Niehoff, Brian P., Robert J. Paul, and John F. Bunch. 1998. "The Social Effects of Punishment Events: The Influence of Violator Past Performance Record and Severity of the Punishment on Observers' Justice Perceptions and Attitudes." *Journal of Organizational Behavior,* 19(6): 589-602.

O'Reilly, Jane, and Karl Aquino. 2011. "A Model of Third Parties' Morally Motivated Responses to Mistreatment in Organizations." *Academy of Management Review,* 36(3): 526-43.

Ohbuchi, Ken-ichi, Masuyo Kameda, and Nariyuki Agarie. 1989. "Apology as Aggression Control - Its Role in Mediating Appraisal of and Response to Harm." *Journal of Personality and Social Psychology,* 56(2): 219-27.

Scher, Steven J., and John M. Darley. 1997. "How Effective Are the Things People Say to Apologize? Effects of the Realization of the Apology Speech Act." *Journal of Psycholinguistic Research,* 26(1): 127-40.

Shaver, Kelly G. 1970. "Defensive Attribution: Effects of Severity and Relevance on the Responsibility Assigned for an Accident." *Journal of Personality and Social Psychology,* 14(2): 101-13.

Skarlicki, Daniel P., Laurie J. Barclay, and S. Douglas Pugh. 2008. "When Explanations for Layoffs Are Not Enough: Employer's Integrity as a Moderator of the Relationship Between Informational Justice and Retaliation." *Journal of Occupational and Organizational Psychology,* 81(Part 1): 123-46.

Skarlicki, Daniel P., John H. Ellard, and Brad R. C. Kelln. 1998. "Third-Party Perceptions of a Layoff: Procedural, Derogation, and Retributive Aspects of Justice." *Journal of Applied Psychology,* 83(1): 119-27.

Skarlicki, Daniel P., and Carol T. Kulik. 2005. "Third-Party Reactions to Employee (Mis)Treatment: A Justice Perspective." *Research in Organizational Behavior: An Annual Series of Analytical Essays and Critical Reviews,* 26: 183-229.

Treviño, Linda K. 1992. "The Social Effects of Punishment in Organizations: A Justice Perspective." *Academy of Management Review,* 17(4): 647-76.

Turillo, Carmelo J., Robert Folger, James J. Lavelle, Elizabeth E. Umphress, and Julie O. Gee. 2002. "Is Virtue Its Own Reward? Self-Sacrificial Decisions for the Sake of Fairness." *Organizational Behavior and Human Decision Processes,* 89(1): 839-65.

Vardi, Yoav, and Yoash Wiener. 1996. "Misbehavior in Organizations: A Motivational Framework." *Organization Science,* 7(2): 151-65.

Walster, Elaine. 1966. "Assignment of Responsibility for an Accident". *Journal of Personality and Social Psychology, 3*(1), 73-9.

Warr, Mark, Robert F. Meier, and Maynard L. Erickson. 1983. "Norms, Theories of Punishment, and Publicly Preferred Penalties for Crimes." *Sociological Quarterly,* 24(1): 75-91.

The Journal of Business Inquiry 2014, 13, 2, 142-157
http:www.uvu.edu/woodbury/articles
ISSN 2155-4072

How do Chinese and Saudi Customers Perceive Online Service Quality? A Comparative Study

By HABIB ULLAH KHAN, MAHMOOD A. AWAN AND HAN CHIANG HO*

This study has two main objectives. The first one aims to examine the service quality of websites offering services using E-S-QUAL (Online Service Quality) and E-RecS-QUAL (Online Recovery of Service Quality) models in two markets: China and Saudi Arabia (KSA). The second one is to compare the online service quality of websites offering services in China with that of KSA in order to identify cultural variations. For the purpose of this study, a survey was conducted in China and KSA. A total of 550 customers who were using online services like websites, portals to carry out online purchases, are identified in both countries and are asked to fill the questionnaires.

Keywords: Service Quality Factors, Perceived Online Service Quality, Online Customer Satisfaction, E-S-QUAL, E-RecS-QUAL

JEL Classification: O1, M3, Z00

I. Introduction

The internet and web-based technologies have a strong impact on today's business and the business organizations which have their operations primarily online are increasing their market share rapidly (Kamhawi and Gunasekaran, 2009; Ye *et al*., 2012; Parthasarathy, 2012; Khan 2013b; Khan and Faisal, 2015). This shift in the market place is also encouraging conventionally-operated companies to adopt the internet for their operations. As a result, the internet is an important channel for selling and buying products/services online. According to Chen and Hitt (2000), the competitors utilize three main types of strategies in the businesses to compete. These strategies include: geographical differentiations, service quality, and modest cost of switching over to other service providers. The role of geographical location is reduced to a bare minimum due to online services which are now being made available to the customers at a convenient location. Among these three strategies, the service quality is vital for the companies to compete in the market.

During this digital dawn, the e-service quality has also become an essential strategy for the companies to persuade new customers and to retain the current customers (Sar and Garg, 2012; Liu and Yang, 2012; Khan, 2013a). Companies not only realize the importance and quality of online portals and services for their business, but also face problems in understanding the perception of quality from the view of customers (Awan *et al*., 2012; Khan, 2012; Hirmukhe, 2013). A wider evidence of the demerits because of not having adequate service quality through

* Habib Ullah Khan: Department of Accounting and Information Systems, College of Business and Economics, Qatar University, Doha, Qatar. Email: habib.khan@qu.edu.qa. Mahmood A. Awan: SolBridge International School of International Business, Daejeon, South Korea. Email: mawan@solbridge.ac.kr. Han Chiang Ho: Solbridge International School of International Business, Daejeon, South Korea. Email: hchiang@solbridge.ac.kr.

the electronic business is available in the market (Ahmad, 2002; Lennon and Harris, 2002; Khan *et al.,* 2013; Shatat and Udin, 2013). Research in this angle of service delivery by online means indicates how the service quality and service recovery can contribute to customer's perceptions of overall quality (Sirdeshmukh *et al.,* 2002; Cronin Jr. *et al*., 2000; Chen and Chen, 2014).

This study focuses on how customers perceive service quality in different sectors of China and Saudi Arabia. Saudi Arabia is rich in oil and natural gas products across the Gulf region. China is one of the largest leading producers in the Far East. Their significance in their respective regions makes them favorites for their selection in the current study.

II. Literature Review

The challenges in adopting internet-based services necessitate the service providers to analyze the attributes contributing to the evaluation of quality service by the consumers. Parasuraman *et al.* (1988) have shown that the perceived quality is the global judgment or attribute evaluated by the customer related to the superiority of a service among other alternatives. Thus, customers are of the opinion that "Online Service Quality" is a crucial factor for the success of internet-based services (Yang *et al*., 2004; Hirmukhe, 2013). It was found that the service quality brings about a greater degree of perceived satisfaction resulting in its adoption (Zhang and Prybutok, 2005). Since organizations are totally dependent on the internet for providing services over online portals and hence websites act as a pivotal media for communication (Merwe and Bekker, 2003; El Halabi *et al.,* 2014; Hassan *et al.,* 2016; Khan and Ahmed, 2013), online service quality becomes even more crucial.

Across all ages, various researchers have studied the characteristics which have contributed very significantly in assessing the quality of services. (Papaioannou *et al.,* 2013; Agrawal *et al*., 2012; Gronroos, 1983). Table 1 provides a summary of the literature on identification of factors related to service quality (face-to-face and on-line). Parasuraman *et al.* (1988) have identified five attributes constituting the global measurement device of service quality, SERVQUAL. As per Zeithemal *et al.* (2002) SERVQUAL is inadequate to calibrate the service quality for online dependent organizations. Cai and Jun (2003) identified that 'SERVQUAL' is mainly based on customer-employee interactions and hence cannot comprehend on interaction between consumers and online portals.

Table 1: Summary of the Literature

S. No.	Factor	Supported by	Relevance
F1	The relationships between service quality, consumer satisfaction, and purchase intentions are studied and service quality measured	Cronin Jr. *et al*. (2000)	Concluded that it is possible to measure the performance-based service quality
F2	Study on effects of quality satisfaction on consumers' behavior intentions	Cronin Jr. *et al.* (2000)	The study confirmed that service quality, service value, and satisfaction may all be directly related to behavioral intentions

Table 1: Summary of the Literature: Continues

S. No.	Factor	Supported by	Relevance
F3	Research on model for measuring Service Quality	Parasuraman *et al.* (1988), Hirmukhe (2013)	Developed SERVQUAL model
F4	Framework for understanding the behavior of service providers to convert customer trust to loyalty	Sirdeshmukh *et al.* (2002)	Evidence of asymmetric relationship between trustworthiness dimensions and consumer trust
F5	Perception of service quality among both internet purchasers and non-purchasers	Yang and Jun (2002)	Identification of five dimensions of internet purchasers: reliability, ease of use, personalization, security, and credibility
F6	Key dimensions of Online Service Quality	Cai and Jun (2003)	Two groups of internet users were identified: online buyers and information searchers
F7	A possible Online Service Quality Model	Zeithemal *et al.* (2002)	Recommended an Online Service Quality model
F8	Systematic scale development for measuring Online Service Quality	Parasuraman *et al.* (2005)	Developed a service quality and recovery service scale which is also utilized in the current study
F9	Relationship between Online Service Quality and perceived satisfaction	Zhang and Prybutok (2005)	It was concluded that Service Quality brings a greater degree of perceived satisfaction which leads to its adoption
F10	Quality of Online Service in Banking Sector of Saudi Arabia	Sohail and Shaikh (2008)	Identification of three factors including efficiency and security, fulfillment, and responsiveness as factors of Service Quality among Saudis also utilized in the current study

A. Perceived Online Service Quality Measurement

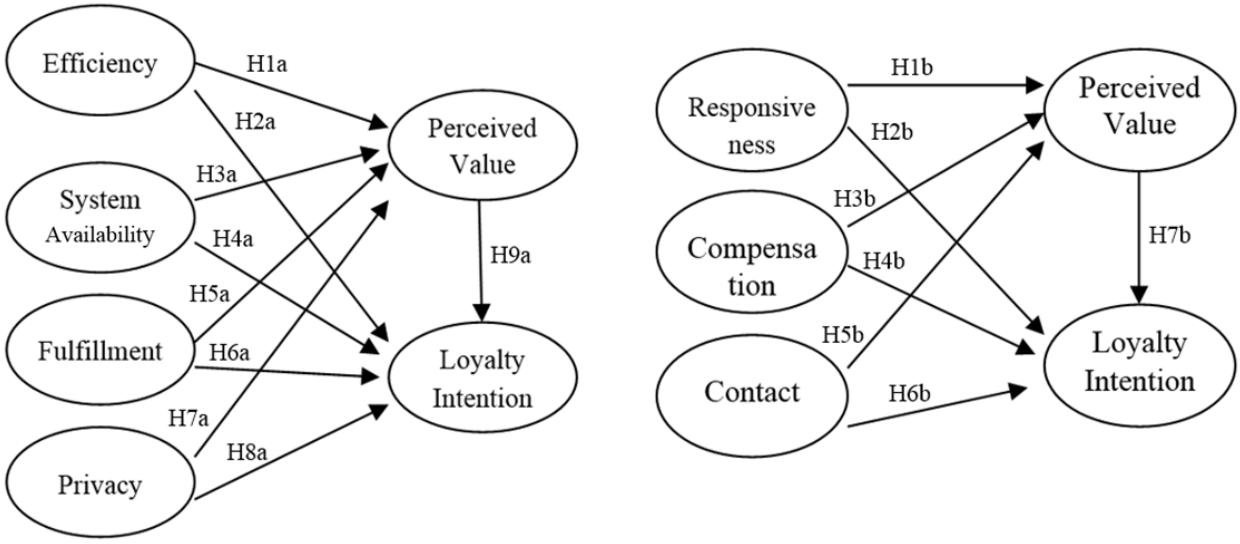
Yang and Jun (2002) reported five quality components which are used extensively by service providers: reliability, ease of use, personalization, security and credibility. This conclusion is based on the sample obtained from online users and conventional users. Cai and Jun (2003) identified two groups of users, the online buyers and the information searchers. The study also pointed out that website design/content, trustworthiness; prompt/reliable service and communication are the main dimensions of online service quality.

Additional research was conducted in order to provide insights about the criteria that are relevant for evaluating online service quality resulting in the E-S-QUAL – a multiple parameter

scale to measure service quality. This is developed by Parasuraman *et al.* (2005). E-RecS-QUAL is a subscale of E-S-QUAL focused on service issues. The focus is on the consumers who are not much at ease with online services. The reliability and validity tests showed that E-RecS-QUAL and E-S-QUAL demonstrated fairly adequately psychometric attributes.

Parasuraman *et al.* (2005) developed an E-S-QUAL systematic scale with four dimensions. They used these models to evaluate websites of various major online organizations. They empirically tested that the model is fit to data through structural equation modeling. The goodness of fit statistics also validated the model structure. The E-S-QUAL model also has an e-recovery quality scale called E-RecS-QUAL as a subscale for problem resolution. The E-RecS-QUAL has dimensions of responsiveness, compensations and contact, which are shown in Figure 1.

Figure 1: Path diagram for E-S-QUAL and E-RecS-QUAL



The definitions of the key terms used for E-S-QUAL and E-RecS-QUAL are provided in Table 2. The first dimension of E-S-QUAL, 'efficiency', is defined by Parasuraman *et al.* (2005, p. 220) as 'the ease and speed of accessing and using the site'. Ranganathan and Ganapathy (2002) reported that preference for online shopping is due to the consumer's convenience of location and most significant is the time saved unlike conventional shopping.

Table 2: Definitions of Key Terms Used

Term	**Definition**
Customer Perceived Quality	Global judgment or attribute relating to superiority of a service relative to competing offering (Parasuraman *et al.*, 1988)
SERVQUAL	Parasuraman *et al.* (1988) have also identified five attributes which constitute the base of a global measurement devise of service quality, namely, SERQUAL

Table 2: Definitions of Key Terms Used: Continues

Term	Definition
E-S-QUAL	Multiple-item scale for measuring the service quality delivered by websites and developed by Parasuraman *et al.* (2005)
E-RecS-QUAL	Subscale which contains how to handle service problems and inquiries, and how to handle non-routine encounters with the customers
Efficiency	Ease and speed of accessing and using the site (Parasuraman *et al.*, 2005, p. 220)
System Availability	Correct technical functioning of the site (Parasuraman *et al.*, 2005, p. 220)
Fulfillment	Extent to which the site's promises about the order delivery and item availability are fulfilled (Parasuraman *et al.*, 2005, p. 220)
Privacy	The degree to which the site is safe and protects customer information
Responsiveness	Effective handling of problems and returns through the site (Parasuraman *et al.*, 2005, p. 220)
Compensation	Dimension related to refunding shipping and handling costs etc. in case of problems
Contact	Availability of assistance through telephone even online representatives; it is important that, in case the customer has a problem, the customer service agent is available

Evaluation and management of quality of service are the key challenges of web-based service providers and they directly influence customer satisfaction (Chen and Chen, 2014). Kim *et al.* (2006) have shown that privacy has a strong impact on the intention to make any purchase. In the E-RecS-QUAL model, the 'responsiveness' dimension is defined as effective handling of problems and service-related responses and returns through the site (Parasuraman *et al.*, 2005, p. 220). The dimension 'compensation' is related to the refund for the services of shipping and operational costs in case of any hindrance in delivering the goods and services. Finally, the "contact" dimension of E-RecS-QUAL relates to the availability of assistance through telephone and accessibility of online representatives.

Companies have to understand which components might affect and attract customers to use online service. For the sake of sustainable operation and development, companies should cultivate long-term relationships with present customers. Morgan and Hunt (1994) proclaim the commitment-trust theory of relationship is suitable for varied relational exchanges such as customers, employees or suppliers. Developing a long-term relationship with customers could be one of the most important strategies for national and international companies so that they could maintain their competitiveness globally. In addition, companies should understand how they could attract various customers around the world to use their service online. Thus, for the purpose of exploring how companies can keep a successful relationship with their customers online and providing quality service, this study applies the commitment-trust theory as its foundation.

In situations where in the customer encounters certain challenges from the service provider over distant locations, there is always a need for customer executives to be available as and when required. Parasuraman *et al.* (2005) defined e-SQ as 'the extent to which a web site facilitates efficient and effective shopping, purchasing and delivery of products and services.' In their leading work, Parasuraman *et al.* (2005) provided an efficient system to measure the e-service quality

perceived by the consumer. However, the study focused on websites that sell products in contrast to services such as those offering financial and informational services. Hence there arises every need for further research to be carried out to validate the E-RecS-QUAL scale, in the context of diverse web sites, especially those involving unusually challenging situations from customers. Thus, one of the current research objectives is to validate online service quality of websites offering services, using E-S-QUAL and E-RecS-QUAL models in different cultures.

The Kingdom of Saudi Arabia (KSA) is one of the largest countries in the Middle East. The internet was introduced in the KSA in 1998 and the internet infrastructure is in the high state of readiness. AME Info (2007) has reported that the KSA showed internet users a growth rate of around 1,170 percent. According to the Internet World Stats report (2011) by December 2010, 11.4 million Saudi internet users went online. With the increasing online services, institutions offering online services in the KSA face challenges in identifying and addressing consumer concerns. Sohail and Shaikh (2008) have also studied internet banking and the quality of service in the KSA. The study identified efficiency and security, fulfillment, and responsiveness as factors that influence users' evaluation of service quality of internet banking services.

Similarly, Internet World Stats (2013) notes that by the end of December 2013, internet users in China reached 620 million, which is around 49 percent of users in Asia. Thus, it is the largest market in the region. BCG (2010) reports that 8 percent of the Chinese population shopped online in 2009, compared with only 3 percent in 2006. Presently, consumer to consumer marketing (C2C) accounts for the largest segment in the Chinese ecommerce industry. Currently, China's online shopping markets are dominated by C2C marketing that accounts for 93.2 percent of total online sales (Su, 2009). However, business-to-consumer marketing (B2C) is growing (Backaler, 2010). For China, however, there are very limited studies focusing on consumer perceptions of online service quality.

B. Hypotheses Development

Based on the dimensions of online service quality identified in the previous section and defined in Table 2, the following hypotheses are developed regarding the expected differences among the online consumers of China and the KSA across various dimensions.

The proposed hypotheses of E-S-QUAL are:

H1a: Online shoppers perceive that the efficiency of a website affects their perceived value toward the website.

H2a: Online shoppers perceive that the efficiency of a website affects their loyalty intention toward the website.

H3a: Online shoppers perceive that the system availability of a website affects their perceived value toward the website.

H4a: Online shoppers perceive that the system availability of a website affects their loyalty intention toward the website.

H5a: Online shoppers perceive that the fulfillment of a website affects their perceived value toward the website.

H6a: Online shoppers perceive that the fulfillment of a website affects their loyalty intention toward the website.

H7a: Online shoppers perceive that the privacy of a website affects their perceived value toward the website.
H8a: Online shoppers perceive that the privacy of a website affects their loyalty intention toward the website.
H9a: Online shoppers perceive that their perceived value affects their loyalty intention.

The proposed hypotheses of E-RecS-QUAL are:
H1b: Online shoppers perceive that the responsiveness of a website affects their perceived value toward the website.
H2b: Online shoppers perceive that the responsiveness of a website affects their loyalty intention toward the website.
H3b: Online shoppers perceive that the compensation of a website affects their perceived value toward the website.
H4b: Online shoppers perceive that the compensation of a website affects their loyalty intention toward the website.
H5b: Online shoppers perceive that the contact of a website affects their perceived value toward the website.
H6b: Online shoppers perceive that the contact of a website affects their loyalty intention toward the website.
H7b: Online shoppers perceive that their perceived value affects their loyalty intention.

III. Methodology

A survey is conducted to measure service quality of websites offering online services in both China and the KSA. This is done with the help of the scales developed by Parasuraman *et al.* (2005). They developed a 22-item scale (E-S-QUAL) measuring service quality and an 11-item E-RecS-QUAL scale measuring e-recovery. In this study, parameters from both scales were utilized without any modifications. The components for E-S-QUAL and E-RecS-QUAL are Likert-scale questions with responses ranging from "Strongly agree" (5) to "Strongly disagree" (1). The reliabilities are calculated for each dimension, and most of the results are above 0.81. In addition to the service quality of perception scales, consumer's online shopping experience, value assessment and likelihood of online shopping are also measured using the scales developed by Parasuraman *et al.* (2005).

A. Data Collection and Sample Characteristics

Data is collected in such a way that the sample constitutes of 250 Chinese and 301 KSA customers who had prior experience in online service transactions. The researchers enlisted the help of 20 executives in Saudi Arabia to explain the questionnaires. The possibilities for data collection are very limited in the KSA and hence very less secondary data are available in the market. To get reliable data, executives with varied expertise from stock markets and financial institutions like banks are requested to collect the data. The demographic information of the respondents is collected through the final section of the questionnaire. The questionnaire is designed in English by the principal researchers. The questionnaire is translated into Chinese and Arabic by the nationals.

IV. Analysis

AMOS™ software is used to perform multi-group analyses to test the conceptual model. Results revealed a good model fit. The values of goodness of fit (GFI), comparative fit index (CFI) and incremental fit index (IFI) are above 0.9 (Bagozzi and Yi, 1988). The root mean square error of approximation (RMSEA) which is an important index of measurement of fit also had a value of less than 0.5, representing a good model fit (Baumgartner and Homburg, 1996). Lastly, the *p*-value of χ^2 (chi-square test) was larger than 0.05 ($\chi^2(1) = 2.01$, $p = 0.15$; GFI = 0.99; CFI = 0.99; IFI = 0.99; RMSEA = 0.04). After sequentially fixing the non-significant parameters in each sample to zero, the models are shown in Figure 2.

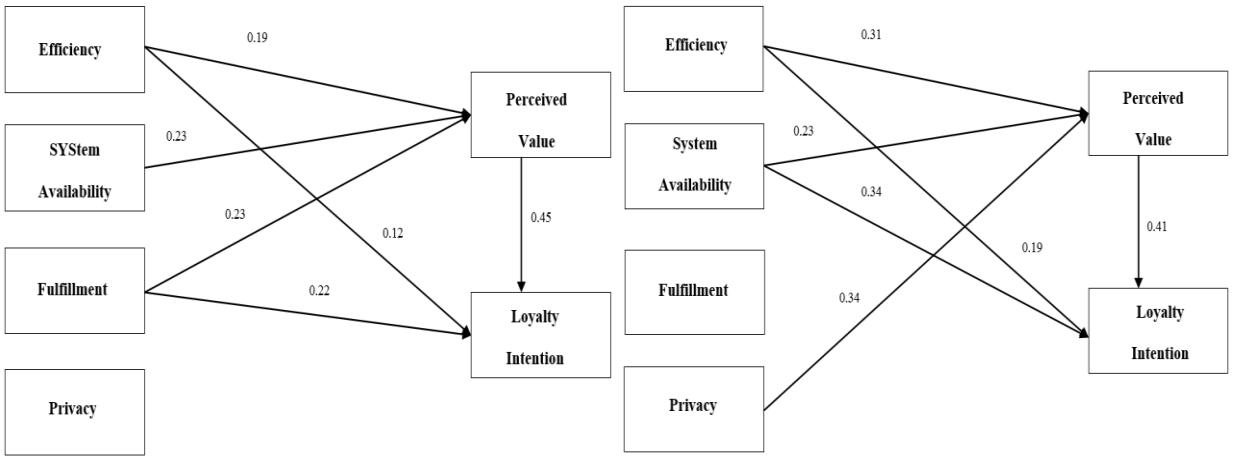
Table 3 shows that the relationships between *Efficiency* and *Perceived Value* ($EFF_{PV\ (C)} = 0.19$, $p < 0.05$; $EFF_{PV\ (SA)} = 0.31$, $p < 0.01$), *Efficiency* and *Loyalty Intention* ($EFF_{LI\ (C)} = 0.12$, $p < 0.10$; $EFF_{LI\ (SA)} = 0.19$, $p < 0.01$) and *System Availability* influences *Perceived Value* ($SYS_{PV\ (C)} = 0.23$, $p < 0.01$; $SYS_{PV\ (SA)} = 0.23$, $p < 0.01$) in both countries. Therefore, H1a, H2a and H3a are supported. *System Availability* affects *Loyalty Intention* only in Saudi Arabia ($SYS_{LI\ (C)} = 0.03$, $p = 0.660$; $SYS_{LI\ (SA)} = 0.34$, $p < 0.01$); H4a is not supported. *Fulfillment* influences *Perceived Value* ($FUL_{PV\ (C)} = 0.21$, $p < 0.05$; $FUL_{PV\ (SA)} = -0.02$, $p = 0.723$) and *Loyalty Intention* ($FUL_{LI\ (C)} = 0.22$, $p < 0.01$; $FUL_{LI\ (SA)} = 0.03$, $p = 0.594$) only happened in China. Hence, H5a and H6a are not supported. In Saudi Arabia, *Privacy* affects *Perceived Value* ($PRI_{PV\ (SA)} = 0.34$, $p < 0.01$), but it does not happen in China ($PRI_{PV\ (C)} = 0.08$, $p = 0.264$). H7a is not supported. H8a is not supported because *Privacy* does not influence *Perceived Value* ($PRI_{LI\ (C)} = 0.08$, $p = 0.139$; $PRI_{LI\ (SA)} = -0.02$, $p = 0.669$) in both countries. *Perceived Value* affects *Loyalty Intention* ($PV_{LI\ (C)} = 0.45$, $p < 0.01$; $PV_{LI\ (SA)} = 0.41$, $p < 0.01$) in both countries, supporting H9a.

Table 3: Results of E-S-QUAL Model

Hypotheses		China			Saudi Arabia			Critical Ratios for Coeff. Differences	
	Proposed Path	Coeff.	*p*		Coeff.	*p*		|z|	
H1a	EFF → PV	0.19	0.029	*	0.31	0.000	*	0.85	
H2a	EFF → LI	0.12	0.094	**	0.19	0.000	*	0.68	
H3a	SYS → PV	0.23	0.009	*	0.23	0.000	*	-0.30	
H4a	SYS → LI	0.03	0.660		0.34	0.000	*	3.35	††
H5a	FUL → PV	0.21	0.015	*	-0.02	0.723		-2.32	††
H6a	FUL → LI	0.22	0.002	*	0.03	0.594		-2.47	††
H7a	PRI → PV	0.08	0.264		0.34	0.000	*	3.08	††
H8a	PRI → LI	0.08	0.139		-0.02	0.669		-1.41	
H9a	PV → LI	0.45	0.000	*	0.41	0.000	*	-0.32	

* $p < 0.05$; ** $p < 0.10$; ††: $|z| > 1.645$, $p < 0.10$

Figure 2: E-S-QUAL Models of China and Saudi Arabia*



*only significant relationships are presented

China Saudi Arabia

Table 4 describes the results of the E-RecS-QUAL model. Only in China, *Responsiveness* influences *Perceived Value* ($RES_{PV\ (C)}$ = 0.34, $p < 0.01$; $RES_{PV\ (SA)}$ = 0.07, $p < 0.318$), not supporting H1b. H2b is supported because the relationships between *Responsiveness* and *Loyalty Intention* are significant ($RES_{LI\ (C)}$ = 0.28, $p < 0.01$; $RES_{LI\ (SA)}$ = 0.11, $p < 0.10$) in both countries. *Perceived Value* is only affected by *Contact* in China ($CON_{PV\ (C)}$ = 0.27, $p < 0.01$; $CON_{PV\ (SA)}$ = 0.10, $p = 0.120$). H3b is not supported. *Contact* influences *Loyalty Intention* ($CON_{LI\ (C)}$ = 0.13, $p < 0.05$; $CON_{LI\ (SA)}$ = -0.13, $p < 0.01$) in both countries, supporting H4b. H5b and H6b are not supported since only in Saudi Arabia *Compensation* affects *Perceived Value* ($COM_{PV\ (C)}$ = 0.09, $p = 0.247$; $COM_{PV\ (SA)}$ = 0.19, $p < 0.05$) and *Loyalty Intention* ($COM_{LI\ (C)}$ = -0.05, $p = 0.447$; $COM_{LI\ (SA)}$ = 0.24, $p < 0.01$). Lastly, in both countries *Perceived Value* ($PV_{LI\ (C)}$ = 0.47, $p < 0.01$; $PV_{LI\ (SA)}$ = 0.55, $p < 0.01$) affects *Loyalty Intention*. These results support H7b. Figure 3 shows the results of E-RecS-QUAL Model.

Furthermore, concerning the two entire models, the E-S-QUAL model explains a similar percentage of variance of *Perceived Value* ($R^2_{PV(C)}$ = 0.38; $R^2_{PV(SA)}$ = 0.33) and *Loyalty Intention* ($R^2_{LI(C)}$ = 0.58; $R^2_{LI(SA)}$ = 0.52), but the E-RecS-QUAL model explains a distinct percentage of variance of *Perceived Value* ($R^2_{PV(C)}$ = 0.36; $R^2_{PV(SA)}$ = 0.09) and *Loyalty Intention* ($R^2_{LI(C)}$ = 0.52; $R^2_{LI(SA)}$ = 0.46). According to the results of Tables 3 and 4, consumers in both countries do share similar attitudes toward *Perceived Value* and *Loyalty Intention*, e.g., in the E-S-QUAL model, *Efficiency* has a direct effect on *Perceived Value* and *Loyalty Intention. System availability* affects *Perceived Value*, and *Perceived Value* plays a mediator role between *Efficiency* and *Loyalty Intention.* In the E-RecS-QUAL model, *Loyalty Intention* is affected by *Responsiveness*, *Contact*, and *Perceived Value*. Further insight about consumers' *Perceived Value* and *Loyalty Intention* is obtained by computing the direct, indirect and total effects (Tables 5 and 6). In both countries, *Perceived Value* plays a major role in affecting *Loyalty Intention*. In the E-S-QUAL model, *System Availability* for China and *Privacy* for Saudi Arabia are the most important roles to affect

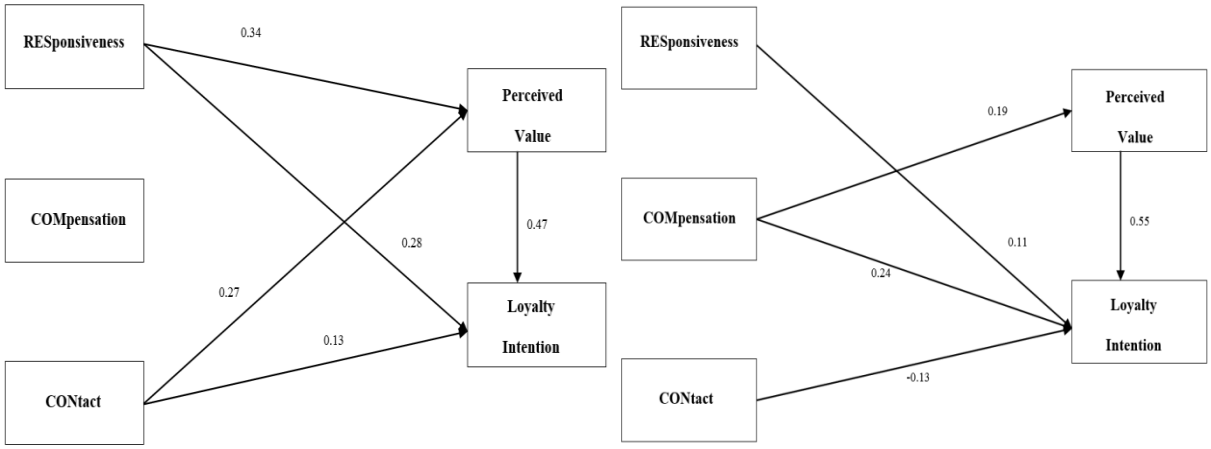
Perceived Value. In the E-RecS-QUAL model, *Responsiveness* for China and *Compensation* for Saudi Arabia are important to *Perceived Value*.

Table 4: Results of E-RecS-QUAL Model

Hypotheses		China			Saudi Arabia			Critical Ratios for Coeff. Differences	
	Proposed path	Coeff.	*p*		Coeff.	*p*		|z|	
H1b	RES → PV	0.34	0.000	*	0.07	0.318		-2.65	††
H2b	RES → LI	0.28	0.000	*	0.11	0.059	**	-2.26	††
H3b	CON → PV	0.27	0.000	*	0.10	0.120		-2.54	††
H4b	CON → LI	0.13	0.016	*	-0.13	0.007	*	-3.49	††
H5b	COM → PV	0.09	0.247		0.19	0.010	*	0.65	
H6b	COM → LI	-0.05	0.447		0.24	0.000	*	2.91	††
H7b	PV → LI	0.47	0.000	*	0.55	0.000	*	0.60	

* $p < 0.05$; ** $p < 0.10$; ††: $|z| > 1.645$, $p < 0.10$

Figure 3: E-RecS-QUAL Models of China and Saudi Arabia*



*only significant relationships are presented

China Saudi Arabia

Table 5: Standardized Total, Indirect and Direct Effects of E-S-QUAL Model on Perceived Value and Loyalty Intention

China			
(1) Effect on *Perceived Value* from:	Total Effect	Direct Effect	Indirect Effect
Privacy	0.075	0.075	0.000
Efficiency	0.190	0.190	0.000
Fulfillment	0.208	0.208	0.000
System Availability	0.229	0.229	0.000
(2) Effect on *Loyalty Intention* from:	Total Effect	Direct Effect	Indirect Effect
Privacy	0.116	0.082	0.033
Efficiency	0.206	0.121	0.085
Fulfillment	0.315	0.222	0.093
System Availability	0.134	0.032	0.102
Perceived Value	0.445	0.445	0.000
Saudi Arabia			
(1) Effect on *Perceived Value* from:	Total Effect	Direct Effect	Indirect Effect
Privacy	0.342	0.342	0.000
Efficiency	0.306	0.306	0.000
Fulfillment	-0.021	-0.021	0.000
System Availability	0.226	0.226	0.000
(2) Effect on *Loyalty Intention* from:	Total Effect	Direct Effect	Indirect Effect
Privacy	0.120	-0.020	0.140
Efficiency	0.319	0.194	0.125
Fulfillment	0.018	0.027	-0.009
System Availability	0.434	0.341	0.093
Perceived Value	0.409	0.409	0.000

Note:
(1) Total effect = Direct effect + Indirect effect
(2) Indirect effect = Indirect effect via affective components + Indirect effect via cognitive components

Table 6: Standardized Total, Indirect and Direct Effects of E-RecS-QUAL on Perceived Value and Loyalty Intention

China			
(1) Effect on *Perceived Value* from:	Total Effect	Direct Effect	Indirect Effect
Compensation	0.090	0.090	0.000
Responsiveness	0.336	0.336	0.000
Contact	0.274	0.274	0.000
(2) Effect on *Loyalty Intention* from:	Total Effect	Direct Effect	Indirect Effect
Compensation	-0.009	-0.052	0.043
Responsiveness	0.440	0.280	0.159
Contact	0.261	0.131	0.130
Perceived Value	0.474	0.474	0.000
Saudi Arabia			
(1) Effect on *Perceived Value* from:	Total Effect	Direct Effect	Indirect Effect
Compensation	0.188	0.188	0.000
Responsiveness	0.073	0.073	0.000
Contact	0.097	0.097	0.000
(2) Effect on *Loyalty Intention* from:	Total Effect	Direct Effect	Indirect Effect
Compensation	0.344	0.239	0.104
Responsiveness	0.147	0.107	0.040
Contact	-0.076	-0.129	0.053
Perceived Value	0.554	0.554	0.000

Note:
(1) Total effect = Direct effect + Indirect effect
(2) Indirect effect = Indirect effect via affective components + Indirect effect via cognitive components

V. Discussion and Future Research Recommendations

Efficiency is widely recognized as the ratio of outputs to inputs. Customers who prefer online services weigh their ordered products or services, time, money and effort they invest for the online transaction. Consumers from both China and Saudi Arabia believe that a reasonable tradeoff between the outputs and inputs is very crucial and will determine the value of the online service. The convenience or pleasure obtained from high efficiency service helps to foster a long-term customer service-provider relationship. System availability is another key aspect that influences consumers' perceived value in both countries.

However, system availability has a different impact on the loyalty intention of consumers from both countries. The system availability could not significantly increase or decrease the loyalty intention for Chinese consumers, while it could for consumers from Saudi Arabia. The reason for this is attributed to the Chinese companies that provide their online services approachable to consumers anywhere and anytime. Hence by nature, consumers prefer the system availability as an integral part of the whole service. For example, Chinese consumers would be happy if an online

service provider would have good system availability but would not be very surprised so as to limit themselves to this company.

Apart from the above, fulfillment is proved to have an important role for perceived value and loyalty intention among Chinese consumers. However, it is insignificant in Saudi Arabia for either perceived value or loyalty intention. There are many factors that could affect the process order fulfillment to order completion. Among these factors, logistics is one of the most critical elements. Since China is a huge country with a huge population, it might still take longer to deliver to the place located elsewhere from the service provider, especially in remote areas. Hence maintaining the logistics service quality can be very different. With a lot of comparisons, consumers in China would prefer those who can deliver the products or services in time. On the other hand, this is not the case with Saudi Arabia as the country is relatively small and as a result the order fulfillment quality is not very heterogeneous. This in turn makes the order fulfillment a success.

Privacy also exhibits an obvious discrepancy between two countries. It does not have a significant effect on both perceived value and loyalty intention in China. But, in Saudi Arabia, it is significant while discussing the relationship of privacy and perceived value. In China, people are confident that their privacy information will be protected by sophisticated and secure online system. According to the E-RecS-QUAL model, Chinese companies should watch out for the responsiveness of the customers in order to enhance consumer value and loyalty. A possible reason for that is China has a very big market size due to its large population. The competition is fiercer in attracting the consumers than retaining the consumers. Therefore, Chinese consumers appreciate a quick response and accessibility after sale. Compensation is not a big deal for Chinese consumers, since many online business platforms serve as moderators and protectors of consumers. On the other hand, compensation is essential in Saudi Arabia. It has significant influence on both perceived value and loyalty intention, while responsiveness and contact are only influential to loyalty intention, but not to the perceived value.

The overall outcomes of the research are providing a good base to the managers of the organizations in the region to explore the perception of the customer related to the products and services they offer. This can give them ideas of enhancing loyalty of customer by improving the quality of their online services. The outcome of the current research also provides support for the E-S-QUAL model and E-RecS-QUAL model implementation in Asian country-based researches, which will ultimately help to improve the theoretical aspect of the models. Any future research done in this domain can observe these findings and hence can achieve the required goals.

REFERENCES

Agrawal, Vivek, Vikas Tripathi, and Sucheta Agrawal. 2012. "Measuring Relation Among E-Service Quality, Customer's Satisfaction, and Intention to Purchase." *International Journal of Marketing and Technology*, 2(12): 32-42.

Ahmad, Sohel. 2002. "Service Failures and Customer Defection: A Closer Look at Online Shopping Experiences." *Managing Service Quality: An International Journal*, 12(1): 19-29.

AME Info. 2007. "E-commerce reaches highs in Saudi Arabia." www.ameinfo.com/61990.html (Accessed October 2014).

Awan, Mahmood A., Habib Ullah Khan, Wenqing Zhang. 2012. "A Comparative Study on Online Service Quality Perception of Two Major Regional Economies." *International Journal of e-Education, e-Business, e-Management and e-Learning*, 2(6): 529-33.

Backaler, Joel. 2010. "Chinese E-commerce Tops $38.5 Billion; What Comes Next." http://readwrite.com/2010/04/19/chinese_e-commerce_tops_385_billion_what_comes_next (Accessed April 2014).

Bagozzi, Richard P., and Youjae Yi. 1988. "On the Evaluation of Structural Equation Models." *Journal of the Academy of Marketing Science*, 16(1): 74-94.

BCG. 2010. "Explosive Growth in Internet Use Is Fundamentally Changing China's Economy and Society." http://www.marketwire.com/press-release/Explosive-Growth-Internet-Use-Is-Fundamentally-Changing-Chinas-Economy-Society-Says-1160501.htm (Accessed May 2014).

Baumgartner, Hans, and Christian Homburg. 1996. "Applications of Structural Equation Modeling in Marketing and Consumer Research: A Review." *International Journal of Research in Marketing*, 13(2): 139-61.

Cai, Shaohan, and Minjoon Jun. 2003. "Internet Users' Perceptions of Online Service Quality: A Comparison of Online Buyers and Information Searchers." *Managing Service Quality: An International Journal*, 13(6): 504-19.

Chen, Fei-Yun, and Shun-Hsing Chen. 2014. "Application of Importance and Satisfaction Indicators for Service Quality Improvement of Customer Satisfaction." *International Journal of Services Technology and Management*, 20(1/2/3): 108–22.

Chen, Sharon, and Lorin M. Hitt. 2000. "Switching Cost and Brand Loyalty in Electronic Markets: Evidence from On-Line Retail Brokers." *Proceedings of 21st Annual International Conference on Information Systems*, Brisbane, Australia.

Cronin Jr., Joseph, Michael K. Brady, and G. Tomas M. Hult. 2000. "Assessing the Effects of Quality, Value, and Customer Satisfaction on Consumer Behavioral Intentions in Service Environments." *Journal of Retailing*, 76(2): 193-218.

El Halabi, Ammar, Ali Hachem, Louay Al-Akhrass, Hassan Artail, and Habib Ullah Khan. 2014. "Identifying the Linkability Between Web Servers for Enhanced Internet Computing." *17th IEEE Mediterranean Electrotechnical Conference (MELECON) 2014*, Beirut, Lebanon.

Gronroos, Christian. 1983. "Strategic Management and Marketing in the Service Sector." *Marketing Science Institute*, Cambridge, MA, 83-104.

Hassan Inuwa Musa, Habib Ullah Khan, and Madhavi Lalitha. 2016. "Pedagogical Potentials of IEEE 802.11 WLAN to Nigerian Universities: A Case Study of the University of Uyo." *International Journal of Information and Education Technology*, 6(4): 256-61.

Hirmukhe, Jyotsna. 2013. "Measuring the Customer's Perception of Service Quality Using SERVQUAL in Public Services." *International Journal of Services Technology and Management*, 19(4/5/6): 267-77.

Internet World Stats. 2013. China. http://www.internetworldstats.com/stats3.htm (Accessed February 2014).

Internet World Stats. 2011. Saudi Arabia. http://www.internetworldstats.com/stats5.htm (Accessed February 2014).

Jain, Pooja, and Narender Kumar. 2011. "Factors Affecting Evaluation of Online Service Quality." *International Journal of Multidisciplinary Research*, 1(3): 26-33.

Kamhawi, Emad M., and Angappa Gunasekaran. 2009. "ERP Systems Implementation Success Factors: IS and Non-IS Managers Perceptions." *International Journal of Business Information Systems*, 4(6): 688-704.

Khan, Habib Ullah. 2013a. "Use of E-Learning Tools to Solve Group Work Problems in Higher Education: A Case Study of Gulf Countries." *The Advances in Computer Science: An International Journal*, 2(3): 90-6.

Khan, Habib Ullah. 2013b. "Role of Computer Mediated Communication in Affect Empowerment and Performance Improvement." *International Journal of Computing*, 3(3): 165-71.

Khan, Habib Ullah. 2012. "Computer Mediated Communication, Quality of Learning, and Performance." *Journal of GSTF Business Review*, 1(3): 81-8.

Khan, Habib Ullah, and Shafiq Ahmed. 2013. "The Possible Challenges of Information Technology Adoption in the Framework of Supply Chain Management: A Case Study of Leading Private and Public Based Service Organizations Of UAE." *44th Annual Conference of Southwest Decision Sciences Institute (SWDSI)*, Albuquerque, New Mexico, USA.

Khan, Habib Ullah, and Mohammed Nishat Faisal. 2015. "A Grey Based Approach for ERP Vendor Selection in Small and Medium Enterprises (SMEs) in Qatar." *International Journal of Business Information Systems*.

Khan, Habib Ullah, **Shafiq Ahmad, and Mali Abdollahian**. 2013. "Supply Chain Technology Acceptance, Adoption, and Possible Challenges: A Case Study of Service Organizations of Saudi Arabia." *10th International Conference on Information Technology: New Generations (ITNG), 2013*, Las Vegas, Nevada, USA.

Kim, Minjeong, Jung-Hwan Kim, and Sharron J. Lennon. 2006. "Online Service Attributes Available on Apparel Retail Web Sites: An E-S-QUAL Approach." *Managing Service Quality: An International Journal*, 16(1): 51-77.

Lennon, Ron, and Judy Harris. 2002. "Customer Service on the Web: A Cross-Industry Investigation." *Journal of Targeting, Measurement and Analysis for Marketing*, 10(4): 325-38.

Liu, Xiao, and Jianmei Yang. 2012. "Social Buying Metanetwork Modelling and Analysis." *International Journal of Services Technology and Management*, 18(1/2): 46-60.

Merwe, Van, and Bekker J. 2003. "A Framework and Methodology of Evaluating E-Commerce Websites." *Internet Research: Electronic Networking Applications and Policy*, 13(5): 330-41.

Morgan, Robert, and Shelby Hunt. 1994. "The Commitment-Trust Theory of Relationship Marketing." *Journal of Marketing*, 58(3): 20-38.

Papaioannou, Eugenia, Costas Assimakopoulos, Christos Sarmaniotis, and Christos K. Georgiadis. 2013. "Investigating Customer Satisfaction Dimensions with Service Quality of Online Auctions: An Empirical Investigation of E-Bay." *Information Systems and e-Business Management*, 11(2): 313-30.

Parasuraman, A., Valarie A. Zeithaml, and Arvind Malhotra. 2005. "E-S-QUAL – A Multiple-Item Scale for Assessing Electronic Service Quality." *Journal of Service Research*, 7(3): 213-33.

Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry. 1988. "SERVQUAL – A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality." *Journal of Retailing*, 64(1): 12-40.

Parthasarathy, S. 2012. "Research Directions for Enterprise Resource Planning (ERP) Projects." *International Journal of Business Information systems*, 9(2): 202-21.

Ranganathan, C., and Shobha Ganapathy. 2002. "Key Dimensions of Business-to-Consumer Web Sites." *Information and Management*, 39(6): 457-65.

Sar, Archana, and Poonam Garg. 2012. "Analysis of Critical Failure Factors in ERP Implementation: An Indian Experience." *International Journal of Business information Systems*, 11(4): 359-86.

Shatat, Ahmad, and Zulkfli M. Udin. 2013. "Factors Affecting ERP System Effectiveness in Post-Implementation Stage Within Malaysian Manufacturing Companies." *International Journal of Business Information Systems*, 14(3): 348-92.

Sirdeshmukh, Deepak, Jagdip Singh, and Barry Sabol. 2002. "Consumer Trust, Value, and Loyalty in Relational Exchanges." *Journal of Marketing*, January, 66(1): 15-37.

Sohail, M. Sadiq, and Nassar M. Shaikh. 2008. "Internet Banking and Quality of Service: Perspectives From a Developing Nation in the Middle East." *Online Information Review*, 32(1): 58-72.

Su, Hui. 2009. "*iResearch: 09Q3 Online Shopping Market Size Reaching 65.85 Billion Yuan.*" https://researcharchive.lincoln.ac.nz/bitstream/10182/3939/3/Zhang_MCM.pdf (Accessed April 2014).

Yang, Zhilin, Minjoon Jun, and Robin T Peterson. 2004. "Measuring Customer Perceived Online Service Quality: Scale Development and Managerial Implications." *International Journal of Operations and Production Management*, 24(11): 1149-74.

Yang, Zhilin, and Minjoon Jun. 2002. "Consumer Perception of Service Quality: From Internet Purchaser and Non-Purchaser Perspectives." *Journal of Business Strategies*, 19(1): 19-41.

Ye, Lili, Zhihao Chen, and Xueni Li. 2012. "A Study of Impacts on Online Purchase Behaviour: Information Richness and Website Interactivity Perspectives." *International Journal of Services Technology and Management*, 18(1/2): 3-15.

Zeithemal, Valarie, A. Parasuraman, and Arvind Malhotra. 2002. "A Conceptual Framework for Understanding E-Service Quality: Implications for Future Research and Managerial Practice." *MSI Working Paper Series*, Report 00-115. Cambridge, MA.

Zhang, Xiaoni, and Victor R Prybutok. 2005. "A Consumer Perspective of E-Service Quality." *IEEE transactions on Engineering Management*, 52(4): 461-77