THE UNIVERSITY OF TENNESSEE MARTIN
COLLEGE OF BUSINESS AND GLOBAL AFFAIRS

JOURNAL OF INTERNATIONAL BUSINESS DISCIPLINES

www.jibd.org

The International Academy of Business Disciplines

Volume 16, Number 2 November 2021

Published By:
University of Tennessee at Martin and the International Academy of Business Disciplines

ISSN 1934-1822 WWW.JIBD.ORG

Journal of International Business Disciplines

Volume 16, Number 2 November 2021

Chief Editor:
Ahmad Tootoonchi
College of Business and Global Affairs
University of Tennessee at Martin
100 Business Administration Building
231 Lovelace Avenue
Martin, TN 38237
Tel: 731-881-7227
Tootoonchi@utm.edu

Editor:
Michele A. Krugh
University of Pittsburgh
4200 Fifth Avenue
Pittsburgh, PA 15260
Tel: 412-526-4271
michele.krugh@pitt.edu

Published By:
University of Tennessee at Martin and the International Academy of Business Disciplines

ISSN 1934-1822 WWW.JIBD.ORG

JOURNAL OF INTERNATIONAL BUSINESS DISCIPLINES
www.jibd.org

OFFICERS

Journal of International Business Disciplines

Volume 16, Number 2 November 2021

Editorial Note

The November 2021 issue of the *Journal of International Business Disciplines (JIBD)* has been the result of a rigorous process of blind reviews, and in the end, six articles were recommended for publication in this issue of *JIBD*.

JIBD is committed to maintaining high standards of quality in all of its publications.

Ahmad Tootoonchi, Chief Editor
Journal of International Business Disciplines

EFFECT OF COVID-19 ON THE UNITED STATES STOCK MARKET: AN EMPIRICAL STUDY

Morsheda Hassan, Grambling State University
Morshedat@yahoo.com

Raja Nassar, Louisiana Tech University
nassar@latech.edu

ABSTRACT

In this study, we investigate the effects of the COVID-19 pandemic on the daily returns, volatility, and trading volume for the DOW, S&P 500 and NASDAQ on the US stock market. In addition, we examine the pandemic effect on the relationships between market volatility and market returns and trading volume. The daily data used was over the period January 1, 2019 to December 31, 2020. Results from the time series intervention analysis showed a significant positive impact of the pandemic on volatility and trading volume for the DOW, S&P 500, and NASDAQ. Furthermore, the onset of the pandemic caused a significant sharp drop in the returns over a two-week period. After this period, the markets recovered and the returns during the pandemic were significantly higher than the returns before the pandemic. Also, the pandemic affected the relationships that existed between volatility and returns for the DOW and SP 500. In addition, it had an effect on the relationship between volatility and volume for the NASDAQ and S&P 500.

INTRODUCTION

The coronavirus (COVID-19) pandemic has created drastic negative impacts on the economic activity in the US and in countries around the world. The COVID-19 crisis has affected health, employment, GDP, the hospitality industry, stock markets and other sectors of the economy here in the U.S and globally. The pandemic has reduced consumer spending, particularly in hotels and restaurants, and disrupted the supply chain. Also, there was a substantial decline in consumer demands for goods and services. The pandemic has affected adversely the airline industry and tourism and created a global economic crisis. Public health measures to contain the spread of the coronavirus have caused severe economic downturns and a substantial increase in unemployment.

While there is ample research on the effect of the pandemic on the economy, few studies have dealt with the effect of the pandemic on the stock markets. In this study, we employ time series analysis to investigate the effect of the COVID-19 pandemic on stock market volatility, trading volume, and returns for the NASDAQ, S&P 500, and the DOW Jones Industrial Average.

LITERATURE REVIEW

Chernick et al. (2020) evaluated the fiscal effect of the COVID-19 pandemic on revenue in 150 US cities. The authors provided estimates of shortfalls in revenue for fiscal year 2021 as compared to the forecast prior to the COVID-19 recession. The study considered revenue forecasts for two scenarios, less severe and more severe. It predicted a shortfall in revenue of 5.5% and 9% for the less severe and more severe scenarios. Some cities that were hardest hit faced a shortfall of 15% or more.

He et al. (2020) investigated the effect of COVID-19 on spill-overs in stock markets. The analysis, using t-tests and Mann-Whitney tests, on daily returns in China, Italy, South Korea, France, Spain, Germany, Japan, and the United States showed that the COVID-19 pandemic had negative short-time effects on the stock market returns. In addition, COVID-19 had spill-over effects on stock markets of other countries.

Shapiro (2020), in a research study from the Federal Reserve Bank of San Francisco, investigated the effect of COVID-19 on inflation in the US. The author reported that personal consumption expenditures (PCE) were about a percentage point below the 2% target set by the Federal Reserve. Data showed that the drop in inflation was due to a substantial decline in consumer demands for goods and services.

Sattar et al. (2020) investigated the effect of COVID-19 on stock market indexes of different countries. The authors used a log-log regression model relating log market index to log number of COVID-19 cases. Data obtained were from January 1, 2020 to March 31, 2020. The countries studied included China, USA, UK, Japan, Germany, Hong Kong, Russia, and India. For all indices, there was a negative relationship between the log number of COVID-19 cases and log market index.

Geert et al. (2020) studied the effect of COVID-19 on the US economy. Their investigation revealed that two thirds of the decline in the GDP during the first quarter of 2020 was caused by a decline in aggregate demand. On the other hand, two thirds of the decline in the GDP in the second quarter of 2020 was caused by a decline in aggregate supply. The authors' analysis predicted a slow recovery due to the supply shock.

Andrea et al. (2020) studied the effects of the COVID-19 pandemic on macroeconomic and financial uncertainty and the consequence of uncertainty on key economic variables. The model used for the analysis was the heteroskedastic vector autoregression (VAR). Results of the analysis indicated that COVID-19 contributed to an increase in macroeconomic and financial uncertainty, which caused a downturn in the economic and financial conditions. However, it was determined that the contribution of uncertainty on the economic downturn was small compared to the effects of other factors attributable to the COVID-19 pandemic.

Gibson and Olivia (2020) investigated the effect of the COVID-19 pandemic on life expectancy in Indonesia. The analysis was based on the mortality table, which gives for each age x the probability of death before age x+1. Results of the analysis showed that the indirect effect on life expectancy, as a result of loss in future income, was larger by at least five orders of magnitude than the direct effect due to COVID-19 related deaths. The effect of COVID-19 on poverty was spatially heterogeneous. As expected, the increase in poverty was higher in provinces that had lower initial poverty rates.

Jelilov et al. (2020) examined the relationship between stock market returns and inflation in Nigeria, while controlling for the effect of the COVID-19 pandemic. The authors used the generalized conditional heteroscedasticity (GARCH (1,1)) model in their analysis. Results of the analysis on daily data from February 27, 2020 to April 30, 2020 revealed that Covid-19 increased market volatility and had a negative effect on returns. In addition, inflation, in the presence of COVID-19 had a negative effect on returns.

Schmitz et al. (2020) estimated that the cost to producers of ethanol, corn, gasoline, and oil in the US, due to the COVID-19 pandemic, was 176.8 billion dollars in 2020. The cost for oil producers alone was 151 billion dollars in the US and 1055.8 billion dollars worldwide. When the unemployment effect is taken into consideration, the cost estimate was 1266.9 billion dollars.

Rababah et al. (2020) examined the effect of the COVID-19 pandemic on the financial performance of Chinese listed companies. Regression analysis was performed on industries over time (panel regression). Financial performance (ROA) was the dependent variable and the independent variables were the natural log of assets of the company (size), leverage, growth rate, the natural log of total revenues, industry, year, and pandemic period (dummy variable). Results from the analysis showed that the COVID-19 pandemic had a negative impact on financial performance., predominantly for small and medium size companies. Also, the negative effect was industry specific.

Singh and Neog (2020) discussed the economic impact of the COVID-19 pandemic on India’s travel and tourism, transportation, stock market, macro-economy, human capital, and trade. The assessment was that India could experience significant economic contraction in the near future as a result of COVID-19. The authors offered policy suggestions to mitigate the health and economic crises.

Moen et al. (2020) examined the effect of the COVID-19 pandemic on unemployment in the United States over the period January through April 2020. There was evidence of an increase in unemployment for all ages and across gender, educational attainment, and race/ethnicity. Young adults and women were especially at risk of being unemployed. Black men with college degrees experienced an increase of 12.4% in unemployment. Individuals in their 50s and 60s with less than a college degree were less likely to become employed, regardless of race.

Von Wachter (2020) studied the long-term effects of the COVID-19 pandemic on unemployment with regard to job losers and labor market entrants in the United States. Measures of the scale of job loss during the pandemic, compared to those from past recessions, indicated that the cost of job loss can reduce earnings for several decades. The predicted losses in lifetime earnings were about two trillion dollars. Furthermore, losses in unemployment could mean a lasting reduction in the employment to population ratio.

Gursoy and Chi (2020) examined the effect of the coronavirus pandemic on the hospitality industry. Study showed that over 50 % of individuals were not willing to dine at a restaurant or stay at a hotel. Also, results of the study indicated that 25% of individuals would feel comfortable with dining at a restaurant and 18% staying at a hotel when testing, tracing and isolating cases in a community improves. A smaller percentage (14% to 17%) will feel comfortable dining at a restaurant or staying at a hotel only when vaccine becomes available. These results suggest that research efforts in behavioral and operational hospitality marketing and management must be directed at determining what factors will make customers return.

Ngwakwe (2020) examined the effects of the COVID-19 pandemic on the SSE Composite Index in China, the Euronext 100 in Europe, and the Dow Jones Industrial Average and the S&P 500 in the United States. Data on stock market performance were gathered 50 days before and 50 days within the COVID-19 pandemic. The before and after stock market mean values were compared using a paired sample t-test. Results of the analysis showed that the Dow Industrial Average had a significant reduction in mean value during the pandemic. However, the Chinese stock market had a significant increase in mean value during the pandemic. There was no change in mean value for the European and S&P 500 markets.

Gruenwald (2020) reported on the impact of the coronavirus pandemic on the world economy in 2020. Forecasts were presented showing a drastic decrease in GDP growth in 2020 for the U.S., Eurozone, China, India, Japan, Russia, and Brazil. In addition, global growth forecast was just 0.4% in 2020 with an increase to 4.9% in 2021.

DATA

Daily market data for the DOW, S&P 500, and NASDAQ were obtained, for the period January 1, 2019 to December 31, 2020, from Yahoo Finance. The time period encompassed the periods before and during the COVID-19 pandemic. The data obtained included daily stock market closing index, trading volume in millions, and the high and low for the day. The range, as a measure of daily volatility, was equal to high – low for the day. Considering the timeline of COVID-19 developments in 2020, the intervention model time T (Equation (1)) for the onset of the effect of the pandemic in the US was taken to be March 2, 2020, the first day in March when the stock market opened.

METHODS

The statistical analysis utilized for this study included time series intervention analysis, time series regression, using the transfer function approach, and the autoregressive conditional heteroscedasticity (ARCH) model. SAS was used in the analysis.

Intervention Analysis

Intervention analysis (Box & Tiao, 1975) is used to study the effect of an intervention on a time series response variable when the time (T) of the intervention is known. The intervention in this case is the onset of the coronavirus pandemic, taken to be March 2, 2020, and is entered in the model as a step function (0 before March 2, 2020 and 1 afterwards). The impact of the intervention on the response variable can be specified in general as

$$wB^bS_t^T \,, \tag{1}$$

where, B is the backshift operator, b is the time shift or delay until the impact of the intervention is felt, and w is the impact coefficient and

$$S_t^T = \begin{cases} 0, & t < T \\ 1, & t \geq T \end{cases}$$

If the response due to the impact is gradual, the impact can be specified as

$$(wB^b/(1-\delta))S_t^T \tag{2}$$

Where δ is between 0 and 1 (Wei, 2006).

For this analysis, the intervention model that gave the best fit to the data was

$$y_t = x_t + wB^bS_t^T \tag{3}$$

where y_t is the observed series and x_t is the series before the intervention, expressed as an ARIMA process, $\frac{\varphi(B)}{\theta(B)}$ x_t.

ARCH Model

Market volatility, as measured by the error variance, is rarely constant over time. A model incorporating a non-constant error variance over time is called autoregressive conditional heteroscedasticity (ARCH), Engel (1982). The model expresses the residual (η_t) from regression as

$$\eta_t = \sigma_t e_t \tag{4}$$

where σ_t is the standard error at time t, e_t is the independent error, and

$$\sigma^2_t = \Theta_0 + \Theta_1 \eta^2_{t-1} + \Theta_2 \eta^2_{t-2} + . . + \Theta_s \eta^2_{t-s} \tag{5}$$

An extension of the ARCH(s) model is the generalized autoregressive conditional heteroscedasticity (GARCH (r, s)) model (Bollerslev, 1986) where the error variance is expressed as a function of its own lags and the lags of the residuals.

$$\sigma^2_t = \Theta_0 + ø_1\sigma^2_{t-1} + ø_2 \sigma^2_{t-2} + . . . + ø_r\sigma^2_{t-r} + \Theta_1 \eta^2_{t-1} + \Theta_2 \eta^2_{t-2} + . . + \Theta_s \eta^2_{t-s} \tag{6}$$

In general, if the residuals ($€_t$) from regression are auto-correlated, the GARCH model can be expressed as

$$€_t = \delta_1€_{t-1} + . . . + \delta_p €_{t-p} + \eta_t$$
$$\eta_t = \sigma_t e_t$$
$$\sigma^2_t = \Theta_0 + ø_1\sigma^2_{t-1} + ø_2 \sigma^2_{t-2} + . . . + ø_r\sigma^2_{t-r} + \Theta_1 \eta^2_{t-1} + \Theta_2 \eta^2_{t-2} + . . + \Theta_s \eta^2_{t-s} \tag{7}$$

The portmanteau Q statistics (McLeod and Li, 1983) was used for determining the order of the GARCH or ARCH model. The model that gave the best fit to the data and was used in this study for estimating the daily standard (SE) as a measure of volatility, was the ARCH (1) model. Also, in the ARCH regression, the regression model used included the first difference (return) of the DOW index, the S&P 500 index, or the NASDAQ index as the stationary dependent variable and the first lag of the difference, if significant, as an independent variable.

Time Series Regression

The standard error (SE_t), as an estimate of volatility at time t in days, for the DOW, S&P 500 (SP), and NASDAQ was calculated from the ARCH (1) model, for a given data set. We then studied the effect of SE and range (defined as high – low for the day) on the DOW index and volume, the S&P 500 index and volume, and the NASDAQ index and volume for the periods before and during the Pandemic.

The linear time series regression (transfer function approach) was used to determine the effect of a stationary input series on a stationary output series. It should be noted that the first difference, Δ, was used on the DOW, SP, and NASDAQ indexes in order to render them stationary for the time series analysis.

The model relating a stationary output series y_t to a stationary input series x_i is expressed as

$$y_t = v(B)\, x_t + a_t, \tag{8}$$

where a_t, is the residual and

$v(B) = w(B)B^c/d(B)$.

Here, $w(B) = w_0 – w_1B - \ldots -w_sB^s$

$d(B) = 1-d_1B- \ldots -d_rB^r$.
and c represents the time delay (or lag) until the input variable x_t produces an effect on the output variable y_t.

We assume that the input series follows an ARMA process, $\frac{\varphi(B)}{\theta(B)}$ x_t. The function v(B) with its lags is determined from the cross correlations between the white noise input series $\frac{\varphi(B)}{\theta(B)}$ x_t and the filtered output series $\frac{\varphi(B)}{\theta(B)}$ y_t (Wei, 2006).
Once v(B) is identified, one can express a_t in Eq. (8) as

$$a_t = y_t – v(B)\, x_t \tag{9}$$

and identify the appropriate time series model for Eq. (9). With a_t known, one can determine the final model in Eq. (8).

For this analysis, all the time series variables were tested for stationarity using the Phillips-Perron test. Where a variable was not stationary, we used its first difference, which was stationary. Thus, all variables that entered the model were stationary.

RESULTS AND DISCUSSION

Table 1 presents the estimates of the impact of the COVID-19 pandemic on the daily DOW returns, daily trading volume, standard error(SE), and range. The standard error was computed from the ARCH (1) model. Both standard error and range were used as a measure of market volatility. It is seen from Table 1 that the COVID-19 impact (w-estimate from Equation (3)) was positive and significant at the 10% level for the DOW returns and highly significant and positive for volume, SE, and range. The impact of COVID-19 was felt 15 days after March 2, 2020. March 2,2020 or day 293 was taken as time T in Equation (1), the onset of the step function.

Figure 1 shows the sharp drop in the DOW index at about day 293 or March 2, 2020 which lasted about 15 days, after which the DOW started its ascent. The delay or time shift of 15 days agrees with the observation in Figure 1. Also, Figures 2, 3 and 4 clearly show that there was an increase in the series mean for volume, SE, and range during the COVID-19 pandemic. This is in agreement with results in Table 1, which shows that the increase in mean was significant. These results indicate that during the COVID-19 pandemic, the daily market returns, trading volume, and volatility increased significantly relative to their values before the pandemic.

TABLE 1. ESTIMATES OF THE PANDEMIC EFFECTS (W), FROM THE INTERVENTION MODEL IN EQUATION (3)

Variable	w -estimate	Standard Error	t -value	p-value	Time-shift in days
ΔDOW	83.66	48.77	1.72	0.0869	15
Volume	119.81	26.48	4.52	<.0001	0
SE	380.23	24.32	15.63	<.0001	0
Range	279.27	68.86	4.06	<.0001	0

The symbol Δ indicates first difference for stationarity

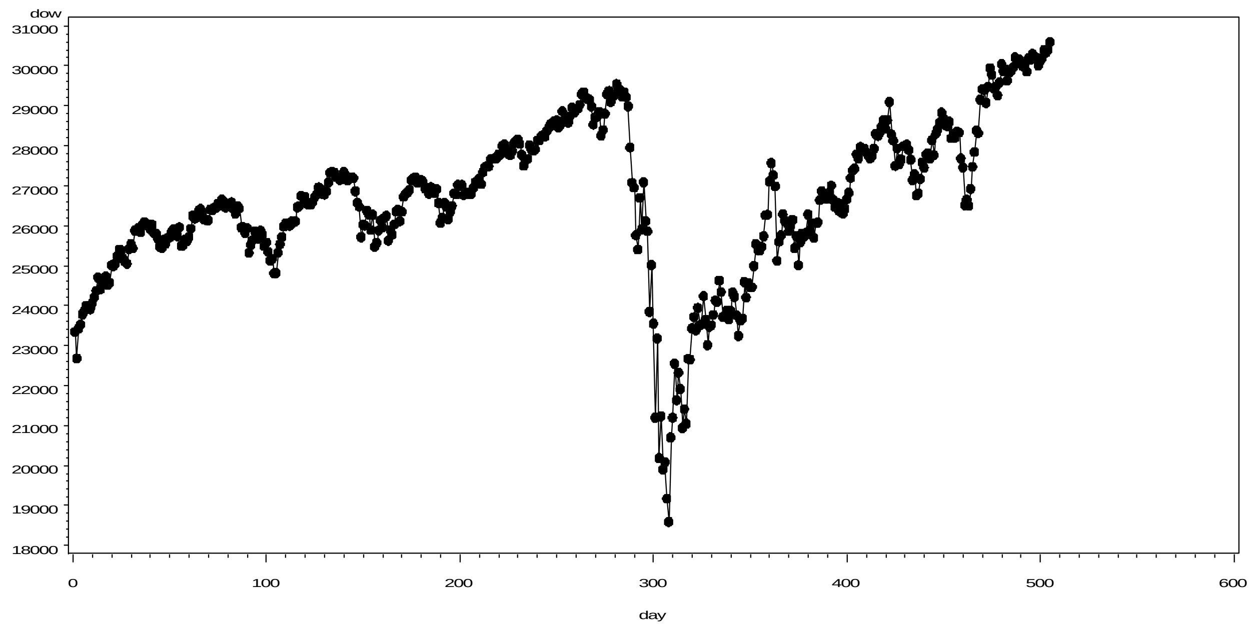

FIGURE 1. PLOT OF THE DOW INDEX OVER DAYS FOR THE PERIOD, JANUARY 1,2019 – DECEMBER 31, 2020.

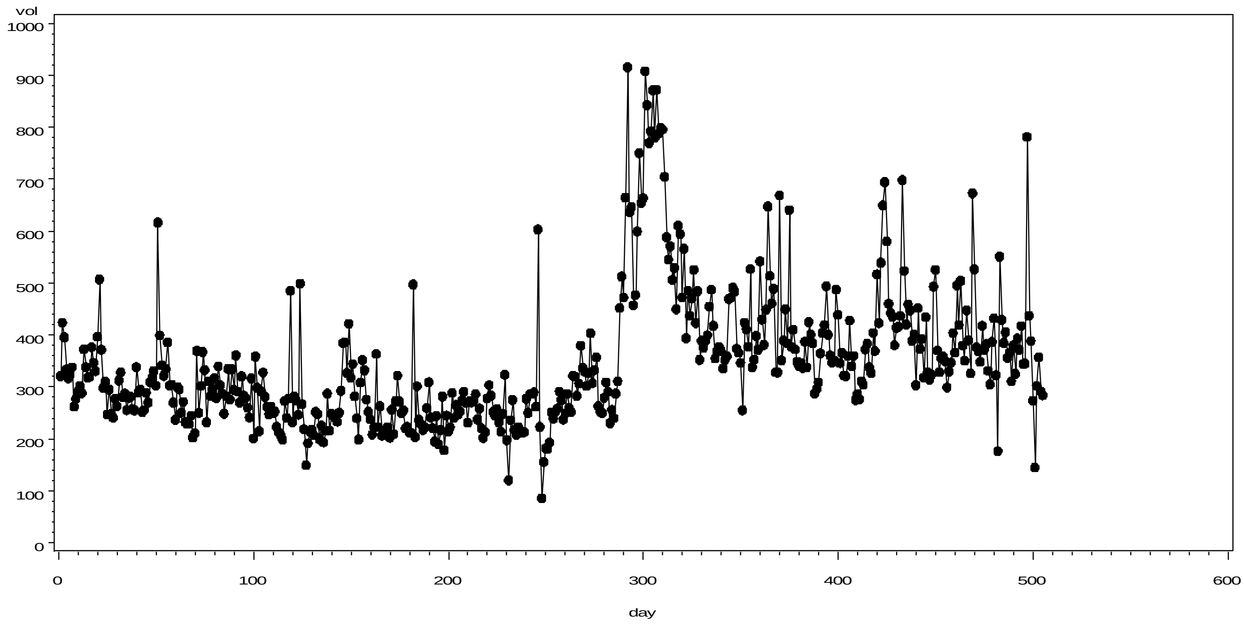

FIGURE 2. PLOT OF THE DOW TRADING VOLUME OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020.

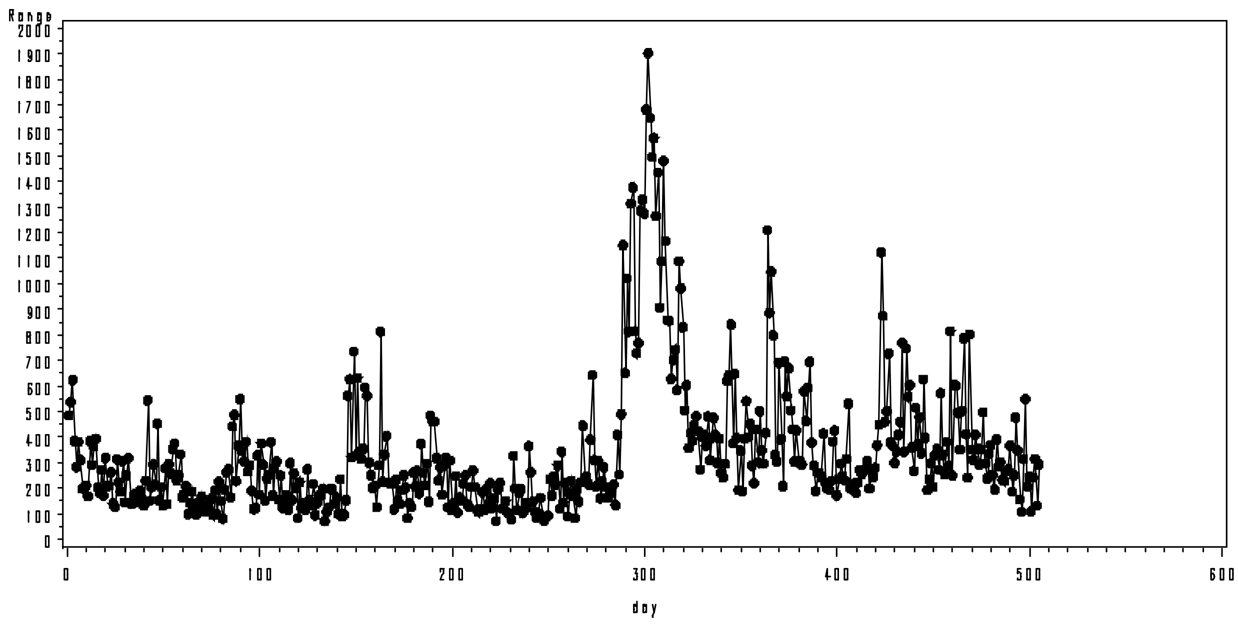

FIGURE 3 PLOT OF THE DOW RANGE OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020.

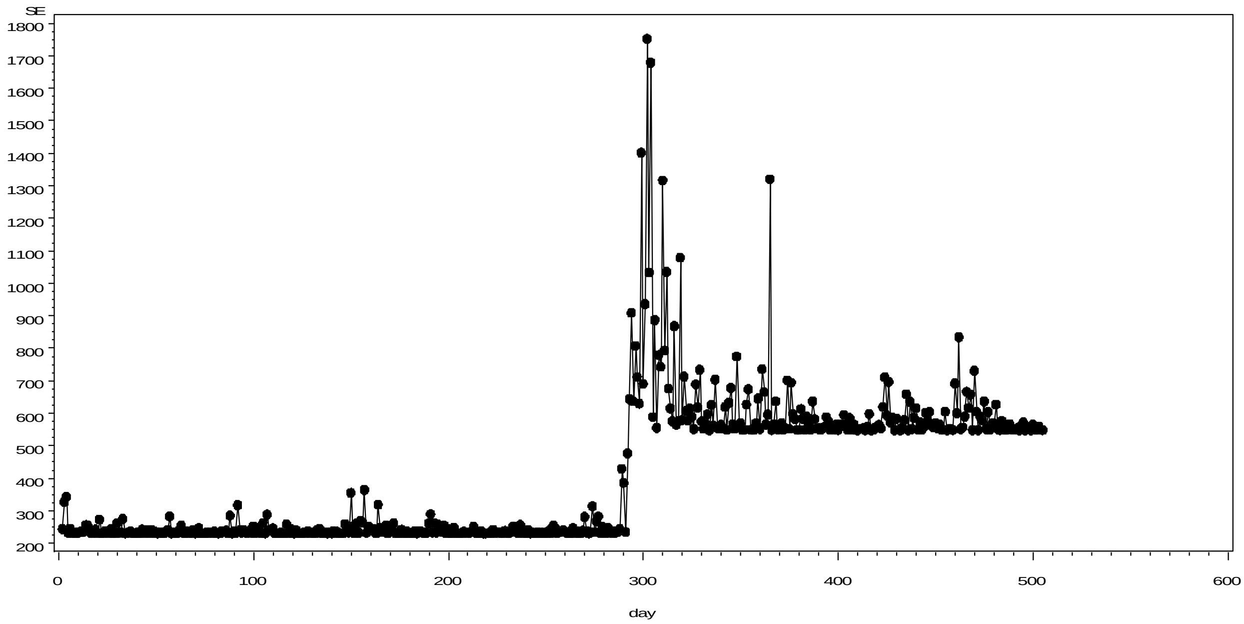

FIGURE 4. PLOT OF THE DOW STANDARD ERROR OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020.

Table 2 shows the effects of the pandemic on the daily NASDAQ returns, trading volume, and volatility. It is seen that the means of returns, volume, SE, and range increased significantly during the pandemic. As in the DOW case, the impact on the NASDAQ returns was felt 15 days after the onset of the Pandemic. These results are in agreement with the trends over days in Figures 5,6,7 and 8. Figure 5 shows an increase in the rate of growth of the NASDAQ after the drop as compared to the period before the pandemic. Also, the means of the series in figures 6, 7, and 8 were larger during the pandemic period as compared to the period before the pandemic.

TABLE 2. ESTIMATES OF THE PANDEMIC EFFECT (W), FROM THE INTERVENTION ANALYSIS MODEL IN EQUATION (3)

Variable	w -estimate	Standard Error	t -value	p-value	Time-shift in days
ΔNASDAQ	30.58	10.50	2.91	0.0038	15
Volume	1971.70	103.96	18.96	<.0001	0
SE	124.10	3.17	39.11	<.0001	0
Range	110.69	19.95	5.55	<.0001	0

The symbol Δ indicates first difference for stationarity.

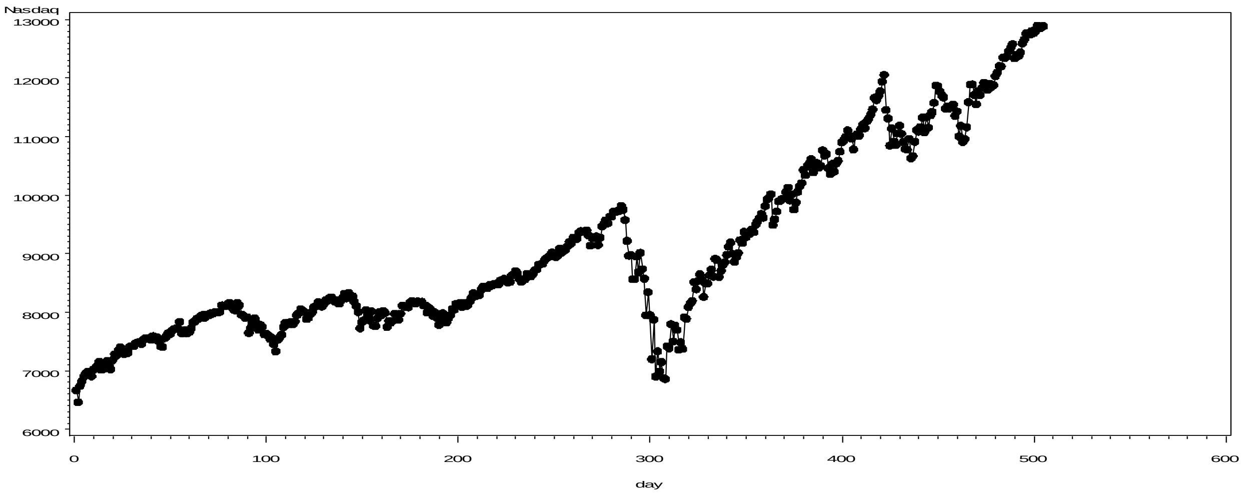

FIGURE 5. PLOT OF THE NASDAQ INDEX OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020.

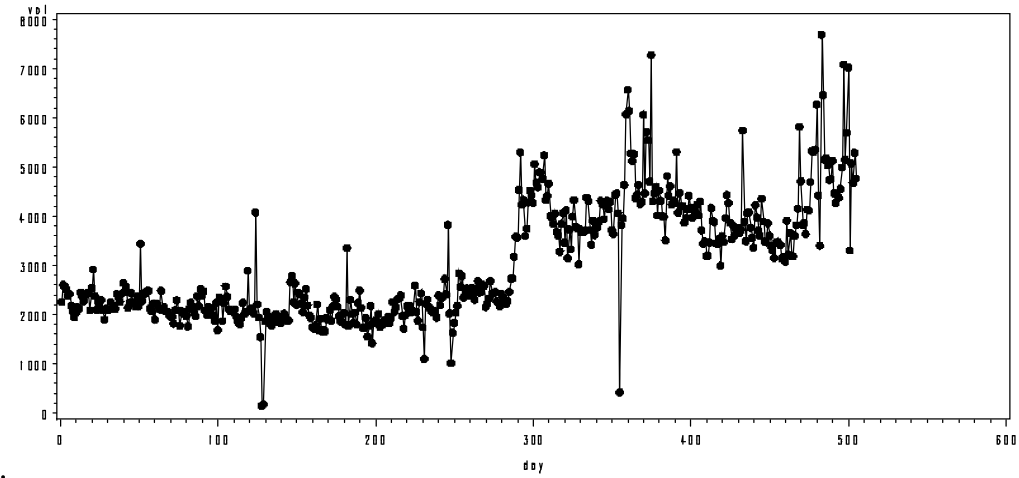

FIGURE 6. PLOT OF THE NASDAQ TRADING VOLUME OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020

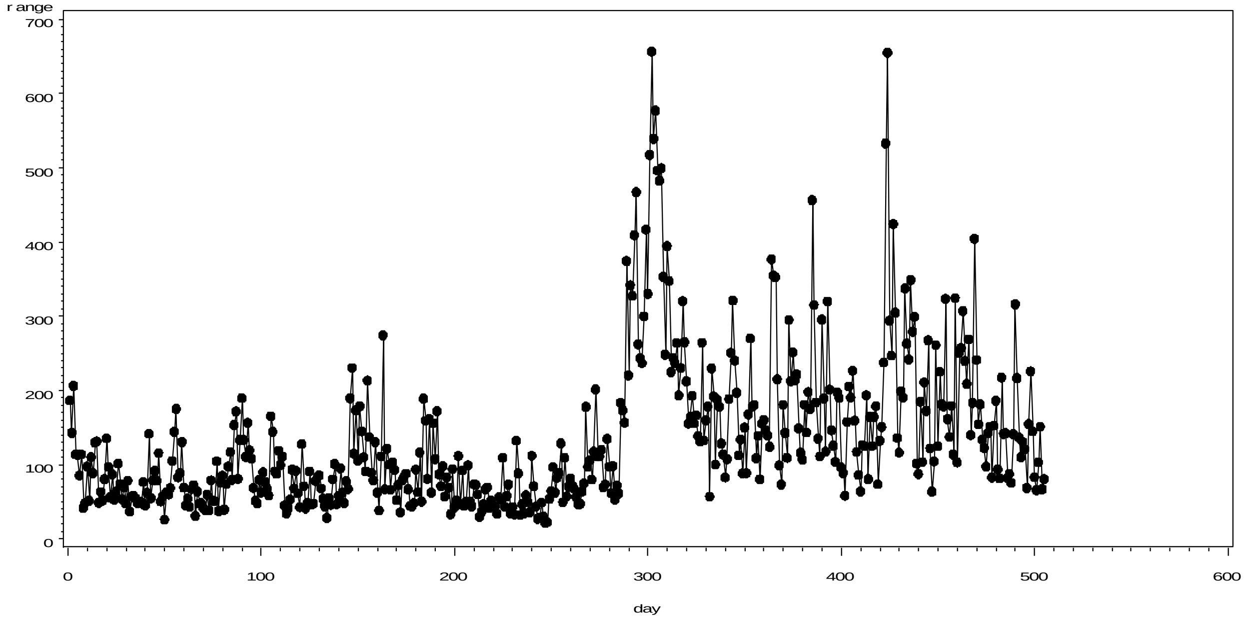

FIGURE 7. PLOT OF THE NASDAQ RANGE OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020.

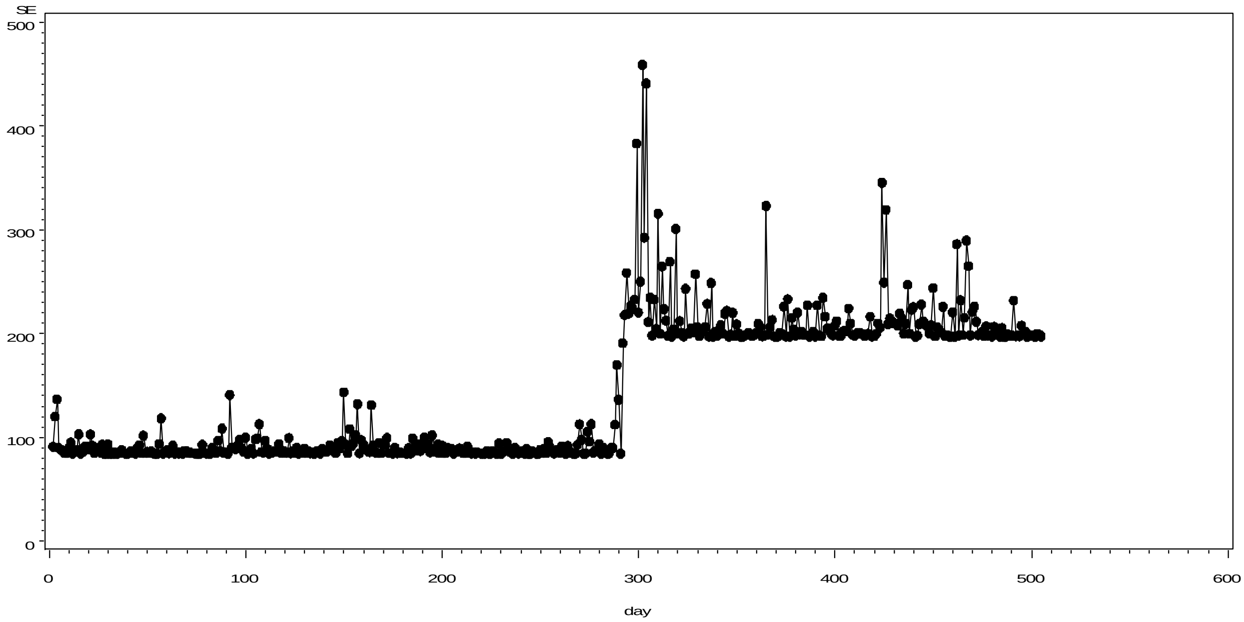

FIGURE 8. PLOT OF THE NASDAQ STANDARD ERROR OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020.

Results in Table 3 indicate that the impact of the pandemic, as seen from the w-estimates, was positive and significant for the S&P 500 returns, trading volume, SE, and range. The impact of the pandemic, as in the case of the DOW and NASDAQ, was felt 15 days after the onset of the pandemic. The graphs in Figures 9 to 12 are in agreement with the intervention analysis results in Table 3. They show a larger rate of increase, after the drop in the SP index, than is seen before the pandemic. In addition, the means of trading volume, SE, and range were larger during the pandemic as compared to the means before the pandemic.

Figures 1, 5 and 9 were similar in that they show that the DOW, NASDAQ, and S&P 500 dropped in March after which they recovered relatively fast. Table 4 shows that this drop for the period March 2, 2020 to March 20. 2020, (days 293-307) was highly significant in that the w-estimates were all negative with a p-value less than 0.0001. For this intervention, the step function in Equation (1) was taken to be 1 for the days 293 to 307 and 0 elsewhere.

TABLE 3. ESTIMATES OF THE PANDEMIC EFFECT (W), FROM THE INTERVENTION ANALYSIS MODEL IN EQUATION (3)

Variable	w estimate	Standard Error	t -value	p-value	Time-shift in days
ΔSP	8.41	3.16	2.66	0.0081	15
Volume	1309.70	195.62	6.70	<.0001	0
SE	40.84	3.18	12.84	<.0001	0
Range	30.34	8.39	3.61	0.0003	0

The symbol Δ indicates first difference for stationarity.

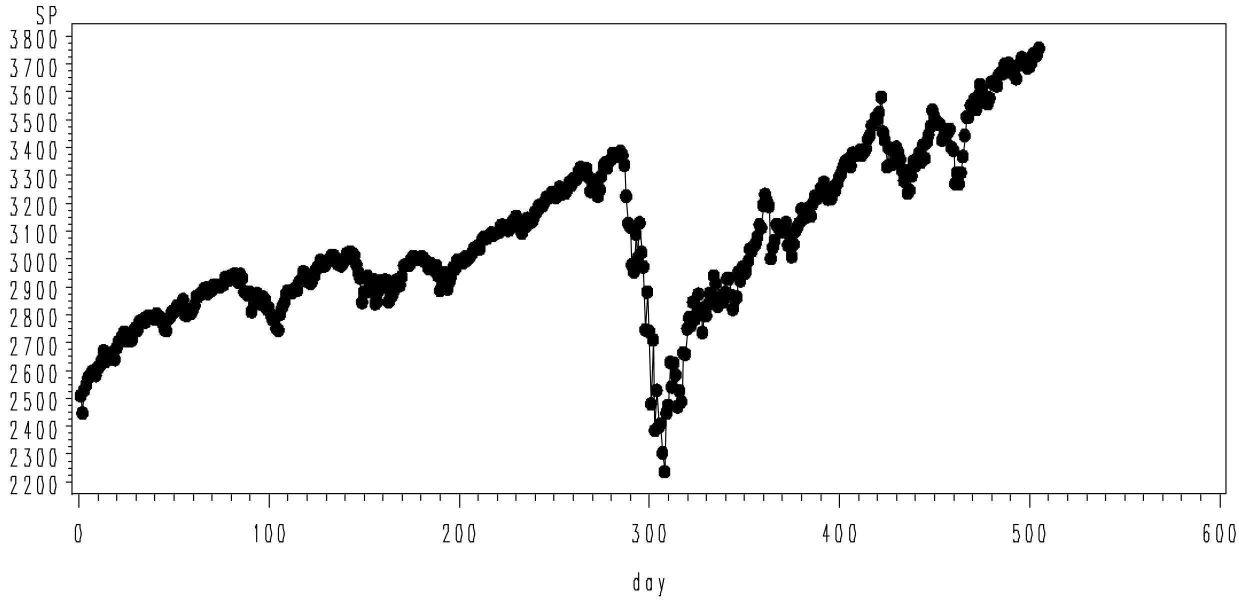

FIGURE 9. PLOT OF THE S&P 500 INDEX OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020

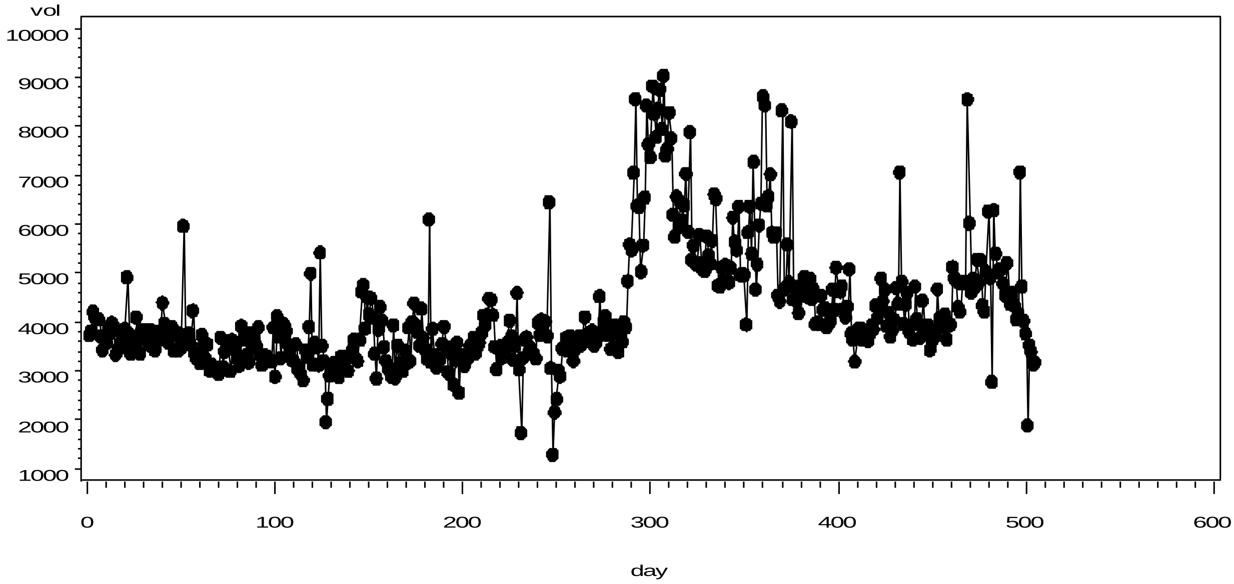

FIGURE 10. PLOT OF THE S&P 500 TRADING VOLUME OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 –DECEMBER 31, 2020.

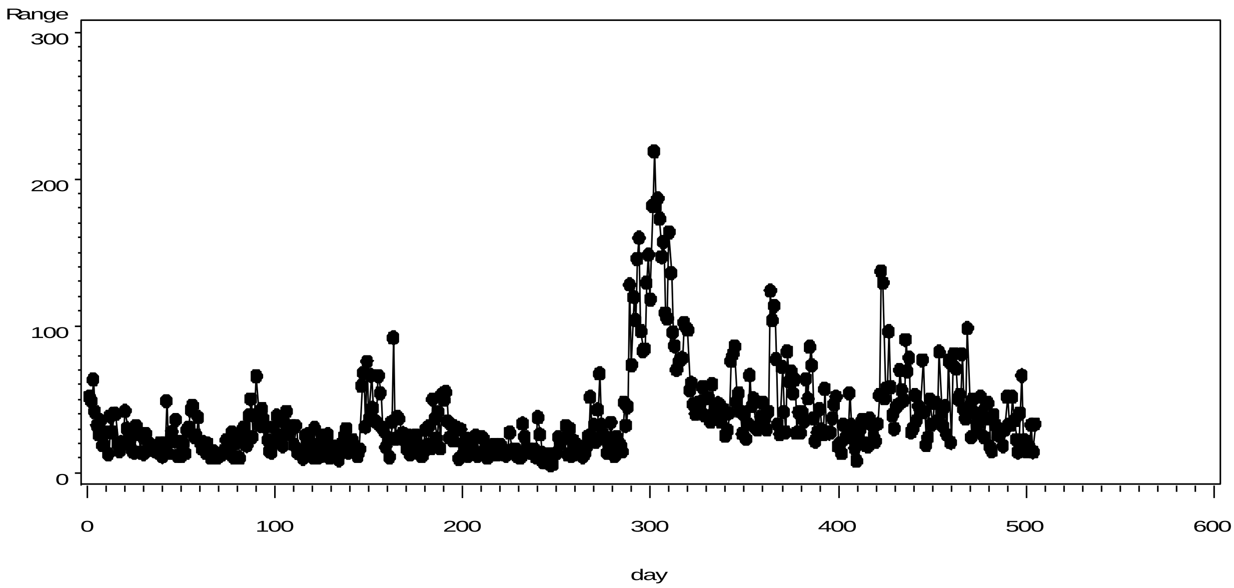

FIGURE 11. PLOT OF THE S&P 500 RANGE OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020.

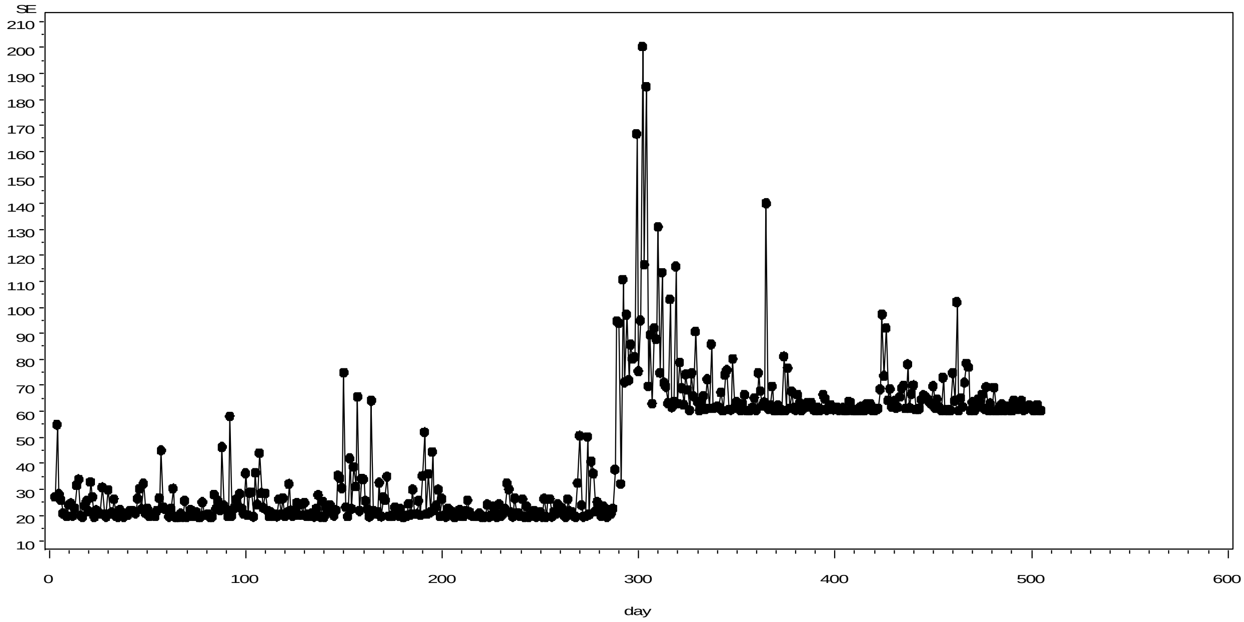

FIGURE 12. PLOT OF THE S&P 500 STANDARD ERROR OVER DAYS FOR THE PERIOD, JANUARY 1, 2019 – DECEMBER 31, 2020.

TABLE 4. ESTIMATES OF THE PANDEMIC EFFECT (W), FROM THE INTERVENTION ANALYSIS MODEL IN EQUATION (3) FOR THE PERIOD, MARCH 2, 2020 TO MARCH 20. 2020, (DAYS 293-307).

Variable	w-estimate	Standard Error	t-value	p-value
ΔDOW	-455.44	87.61	-5.20	<.0001
ΔNASDAQ	-133.55	29.25	-4.57	<.0001
ΔSP	-48.61	8.33	-5.83	<.0001

The symbol Δ indicates first difference for stationarity.

Table 5 presents results, from the model in Equation (8), on the effect of volatility (as measured by the standard error or SE) on the Dow, S&P 500, and NASDAQ daily returns as well as their daily trading volumes before and during the pandemic. As seen from the w_0 estimate in Table 5, the relationship between SE and the DOW returns was negative, but not significant, before the pandemic and changed to being positive and significant during the pandemic. The SE volatility effect on volume was positive and significant before and during the pandemic.

TABLE 5. EFFECTS OF DAILY STANDARD ERROR OF RETURNS (SE), (V(B) IN EQUATION (8) = W_0 or $W_0 - W_1B$) FOR THE PERIODS BEFORE AND DURING THE PANDEMIC

Period: January 1, 2019 to March 1, 2020					March 21, 2020 to December 31, 2020			
Variable	Estimate	Std. Error	t-value	p-value	Estimate	Std. Error	t-value	p-value
ΔDOW	w_0 = - 0.353	0.514	-0.69	0.4930	w_0 = 0.731 shift = 1	0.322	2.27	0.0245
Volume	w_0 = 0.400	0.167	2.39	0.0177	w_0 = 0.199	0.076	2.60	0.0099
ΔNASDAQ	w_0 = - 0.028	.479	-0.06	0.954	w_0 = -1.76	7.30	-0.24	0.809
Volume	w_0 = 7.29 shift = 2 w_1B = -7.26 shift =2	2.23 2.20	3.27 -3.29	0.0012 0.0011	w_0 = -59.97 shift = 3 w_1B = 63.21 shift = 3	32.96 32.89	-1.82 1.92	.0705 0.0562
ΔSP	w_0 = -0.105	0.211	-0.50	.6182	w_0 = 1.171 shift = 1	0.494	2.37	0.0189
Volume	w_0 = 9.65	2.344	4.12	<.0001	w_0 = 4.82	10.12	0.48	0.6345

The symbol Δ indicates first difference for stationarity

The effect of volatility on the NASDAQ returns was negative, but not significant, for both periods, before and after the pandemic. On the other hand, the effect of volatility on the NASADQ trading volume was positive before the pandemic and negative during the pandemic.

The effect of volatility on the S&P 500 returns was negative, but not significant, before the pandemic and changed to being significantly positive during the pandemic. The effect of volatility on the daily trading volume was positive and significant before the pandemic and positive, but not significant, during the pandemic. Clearly, there was an effect of the pandemic with regard to the relationships between volatility and the DOW returns, the NASDAQ trading volume, S&P 500 returns, and its trading volume.

Table 6 presents the relationships between volatility, as measured by the daily range, and the daily returns on the DOW, S&P 500, and NASDAQ indexes and their trading volumes for the periods before and during the pandemic. It is seen from Table 6 that the relationship between volatility and the DOW returns was negative before the onset of the pandemic and positive during the pandemic. There was no change in the positive relationship between volatility and the trading volume as a result of the pandemic.

The effect of volatility on the NASADQ returns was negative for both periods. Also, the effect on volume was positive for both periods. The relationship between volatility and the returns on the S&P 500 was negative for both periods but became insignificant during the pandemic. The relationship of volatility on the SP trading volume was positive for both periods.

There is some evidence in the literature of a negative relationship between market returns and volatility in the case of the S&P 500 (Crestmont Research, 2011). This is seen to be the case, in this analysis, for the DOW, S&P 500, and NASDAQ before the pandemic. This relationship changed for the DOW and the S&P 500 as a result of the pandemic.

TABLE 6. EFFECTS OF DAILY RANGE (V(B) IN EQUATION (8) = W_0 or $W_0 - W_1B$) FOR THE PERIODS BEFORE AND DURING THE PANDEMIC

Period: January 1, 2019 to March 1,2020					March 21, 2020 to December 31, 2020			
Variable	Estimate	Std. Error	t-value	p-value	Estimate	Std. Error	t-value	p-value
ΔDOW	w_0 = -0.651	0.082	-7.91	<.0001	w_0 = 0.340 shift = 1	0.123	2.76	0.0063
Volume	w_0 = 0.244	0.029	8.39	<.0001	w_0 = 0.336	0.0232	14.46	<.0001
ΔNASDAQ	w_0 = -0.588	0.090	-6.51	<.0001	w_0 = -0.589	0.106	-5.51	<.0001
Volume	w_0 = 3.485	0.468	7.43	<.0001	w_0 = 2.304	0.588	3.91	0.0001
ΔSP	w_0 = -0.649	0.080	-8.03	<.0001	w_0 = -0.093	0.111	-0.83	0.4049
Volume	w_0 = 18.95	2.22	8.55	<.0001	w_0 = 17.35	2.54	6.82	<.0001

The symbol Δ indicates first difference for stationarity.

CONCLUSION

This study investigated the effect of the COVID-19 pandemic on the daily returns, trading volume, and volatility with regard to the DOW, S&P 500, and NASDAQ on the US stock market.

In addition, the study examined the effect of the pandemic on the relationships between volatility and returns as well as volatility and trading volume. The time series analysis used was the intervention analysis, the transfer function time series regression approach, and the autoregressive conditional heteroscedasticity (ARCH) model. Volatility was measured by the daily range (high - low) of an index and by the daily standard error of returns from the ARCH (!) model.

Results from the intervention analysis showed a significant positive impact of the pandemic on trading volume, and volatility for the three indexes. Also, the onset of the pandemic caused first a steep drop in returns for the three indexes, which was followed by an increase in returns as compared to the period before the pandemic. Furthermore, the pandemic affected the relationships that existed between volatility and returns in the case of the DOW and S&P 500. Also, it had an effect on the relationship between volatility and volume for the NASDAQ and S&P 500.

REFERENCES

Andrea, C., Clark, T. E., Marcellino, M., & Mertens, E. (2020). *Measuring uncertainty and its effects in the COVID-19 era* (Working Paper No. 2032). Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202032

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics, 31*, 307-327.

Box, G. E. P. & Tiao; G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association, 70,* 70-79.

Chernick, H., Copeland, D., & Reschovsky, A. (2020). The fiscal effects of the Covid-19 pandemic on cities: An initial assessment. *National Tax Journal, 73,* 699–732

Crestmont Research (2011). *Market volatility report.* https//www.crestmontresearch.com/stock-market/

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50,* 987-1007

Geert, B., Engstrom, E. & Ermolov, A. (2020). *Aggregate demand and aggregate supply effects of COVID-19: A real-time analysis* (Finance and Economics Discussion Series 2020-049). Board of Governors of the Federal Reserve System. https://doi.org/10.17016/FEDS.2020.049.

Gibson, J. & Olivia, S. (2020) Direct and indirect effects of COVID-19 on life expectancy and poverty in Indonesia. *Bulletin of Indonesian Economic Studies, 56,* 325-344.

Gursoy, D. &. Chi, C. G. (2020) Effects of COVID-19 pandemic on hospitality industry: Review of the current situations and a research agenda. *Journal of Hospitality Marketing & Management*, *29,* 527-529.

He, Q., Liu, J. , Wang, S., & Yu, J. (2020). The impact of COVID-19 on stock markets. *Economic & Political Studies, 8,* 275-288.

Jelilov, G., Iorember, P. T., Usman, O., & Yua, P.M. (2020) Testing the nexus between stock market returns and inflation in Nigeria: Does the effect of COVID-19 pandemic matter? *Journal of Public Affairs, 20,* 1-9.

Moen, P., Pedtke, J. H., & Flood, S. (2020). Disparate disruptions: Intersectional COVID-19 employment effects by age, gender, education, and race/ethnicity. *Work, Aging and Retirement*, *6*(4), 207–228.

Ngwakwe, C. C. (2020). Effect of COVID-19 pandemic on global stock market values: A differential analysis. *AUDŒ, 16,* 255-269.

Gruenwald, P., F. (2020). *Economic research: The escalating coronavirus shock is pushing 2020 global growth toward zero* (15). S&P Global Ratings. https://www.spglobal.com/ratings/en/research/articles/200330-economic-research-the-escalating-coronavirus-shock-is-pushing-2020-global-growth-toward-zero-11413969

Rababah, A., Al-Haddad, L., Sial, M. S., Chunmei, Z., & Cherian, J. (2020). Analyzing the effects of COVID-19 pandemic on the financial performance of Chinese listed companies. *Journal of Public Affairs, 20,* 1-6. https://doi.org/ 10.1002/pa.2440.

Sattar, M. A., Arcilla Jr, F. E., & Sattar, M. F. (2020). The response of financial market indices to COVID-19 pandemic. *Financial Studies*, *3,* 83-92.

Schmitz, A., Moss, C. B., & Schmitz, T. G. (2020). The economic effects of COVID-19 on the producers of ethanol, corn, gasoline, and oil. *Journal of Agricultural & Food Industrial Organization*, *18*, 1-18.

Shapiro, A. H. (2020). Monitoring the inflationary effects of COVID-19. *FRBSF Economic Letter, 2020-24,* 1-5.

Singh, M. K., & Neog, Y. (2020). Contagion effect of COVID-19 outbreak: Another recipe for disaster on Indian economy. *Journal of Public Affairs, 20,* 1-8.

Von Wachter, T. (2020). Lost generations: Long-term effects of the COVID-19 crisis on job losers and labor market entrants, and options for policy. *Fiscal Studies, 41,* 549–590.

Wei, W. W. S. (2006). *Time series analysis: Univariate and multivariate methods* (2^{nd} ed.). Pearson Addison Wesley.

EFFECTS OF ANCHORING PARADIGM ON YOUNG CONSUMERS' PURCHASING DECISION-MAKING

C. Christopher Lee, Central Connecticut State University
christopher.lee@ccsu.edu

Teodor Panaitisor, Central Connecticut State University
tradu@my.ccsu.edu

Ramadan Hemaida, University of Southern Indiana
rhemaida@usi.edu

ABSTRACT

This paper examined how young consumers responded to the anchoring paradigm. This research proposed five hypotheses: Anchoring was positively related to the consumer's willingness to pay; Anchoring was positively related to price perception; Anchoring was positively related to product quality perception; Gender would have a moderating effect; Income would have a moderating effect. We collected data via in-person survey of young consumers in a public university in America. For data analysis, we used ANOVA and regression models. Anchoring was statistically significant in most ANOVA and regression models. Price Perception, Product Quality, and Gender variables showed strong correlation, but they had no statistical significance in regression models. Data showed no income effects on the anchoring paradigm. Thus, evidence supported Hypothesis 1 but did not support Hypothesis 5, while Hypothesis 2, 3, & 4 were inconclusive. This study provided empirical evidence to shed light on the anchoring paradigm effect on young consumers.

INTRODUCTION

Consumer behavior is an area of study that still offers plenty of secrets left to be uncovered. Consumers have become increasingly price sensitive, leading to further research into the anchoring, imprinting, and priming paradigm that may offer new insight into the thought patterns of consumers when purchasing a product.

The topic of consumer behavior can hardly be considered fully fleshed out, and despite the numerous amounts of studies done on the topic, there is still much that we do not fully understand. The anchoring paradigm is something that has become more popular recently, existing in the subconsciousness of the consumer. Many of the studies done recently tend to focus on the psychological aspect of the theory, namely the facet of consumer perception. According to Poundstone (2010), anchoring is essentially the starting point when a consumer is unfamiliar with a product. Reference prices help the consumer to perceive the benefits of a product, while priming and imprinting do the same for a brand. However, it is important to look beyond merely the psychological factors at play, as demographic characteristics may also play a role in the anchoring paradigm. It is important that businesses, specifically management, understand this topic, so they can potentially utilize their knowledge to better manage how their customer perceives their business. Much research has done on the anchoring paradigm, but few studies investigated the effect of anchoring paradigm on Gen Z consumers. Such a lack of literature motivated this research.

This research will attempt to gather further information about the different effects of the anchoring paradigm across various different demographics. The paper will be studying any differences that arise between gender and income level, in an effort to determine if different demographics experience anchoring differently.

Data was collected from students at a New England public university. The hypothesis testing will be conducted through a multivariate statistical analysis.

The following section will include the review of prior studies performed on the subject. Section 3 will flesh out our methodology, followed by the results of our testing in Section 4. Section 5 will include any implications from a managerial perspective, and Section 6 will conclude the study.

LITERATURE REVIEW

According to Poundstone (2010), anchoring serves as a mental benchmark or starting point for estimating an unknown product. To further explain how anchoring works, Poundstone (2010) refers to an experiment done by Tversky and Kahneman. The researcher's goals were to illustrate human decision making and how human judgment drives decisions (Lagnado, 2007). Tversky and Kahneman used one piece of apparatus that resembled a carnival-style wheel of fortune with numbers up to 100. The wheel was spun giving the researchers to a random number that was presented to the group of college students. This is where the independent variable comes in; the truth was that the wheel was actually rigged. The wheel was set up in a way where only two numbers would show up, 10 or 65. After the number was presented the participants were asked to answer a set of two questions. The questions where: "a) Is the percentage of African nations in the United Nations higher or lower than 65 or 10 [the number that just came up on the wheel]?" And "b) What is the percentage of African nations in the United Nations?" When the wheel stopped at 10 the average estimate of the proportion of African nations in the United Nations was 25 percent. But when the wheel stopped at 65 the estimate of the proportion of African nations in the United Nations was 45 percent. This is a classic example of how anchoring is able to influence an individual's decisions and how in unfamiliar circumstances anchoring could impact people's decisions.

Imprinting, as described by Dan Ariely, professor of psychology and behavior economics, is much like comparing the human brain to that of a gosling (Ariely, 2008). Konrad Lorenz, an Austrian zoologist, ethologist [i.e. the study of animal behavior], and ornithologist, describes imprinting as the effect that takes place when goslings hatch and attach themselves to the first organism they see (Nisbett, 1976). Lorenz realized that the first person that goslings see after they hatch the goslings get imprinted to. What Lorenz found remarkable is the fact that they can get imprinted to anything, even humans. This psychological effect that takes place between organisms can also be generalized from organism (i.e., human) to inanimate object (i.e., product or service). In marketing, consumers get "imprinted" to organizations and companies all the time. As a result, this effect causes low survival rate for new firms and companies that are trying to enter the market (Dobrev & Gotsopoulus, 2010). Consumers, when dealing with purchasing decisions, try to use the most efficient means when dealing with their purchase choices. Rarely do consumers take time out their shopping experience to read and examine the labels that the product is being provided with. Therefore, consumers rely on what marketers call familiarity of the brand or product (Solomon et al., 2009).

Priming increases the likeliness of an outcome due to prior exposure. It works much like imprinting where exposure to a stimulus will increase the likeliness of a decision. What distinguishes imprinting from priming is the fact that imprinting utilizes explicit memory while priming utilizes implicit memory. According to Dias (2009), priming may also utilize explicit memory. Dias concludes that priming occurs in two separate ways: through explicit (direct) or implicit (indirect) ways (Dias, 2009). Some examples of priming through explicit ways are through the use of print ads in magazines, newspapers, radio, and television commercials. Priming through implicit ways are through the use of word of mouth. Explicit measures produce more than a 4:1 ratio of sales to initial cost, where the ratio for implicit cost is unknown.

Does brand name imprinting improve a consumer's ability to recall information specific to that brand? Baker (2003) seems to think so. The end goal of brand imprinting is to have consumers associate those brands with positive thoughts and experiences. For example, the McDonald's brand wants the consumer to associate that brand with their products, like happy meals or Big Macs. Baker (2003) completed a study that intends to demonstrate how brand imprinting works with regards to product recollection. When an understanding of that process is gained, companies can then use that information to help design their own method of imprinting their brand, thus allowing them to instill positive thoughts about their brand into their consumer's head. In a survey, Baker gathered the opinions of 200 college students, using laundry detergent and batteries as the baseline products to test. Four brands were used for each product, two brands that implied high benefits and two brands that offered neutral benefits. The volunteers were exposed to each brand, performed a task to learn about their benefits, and then given a memory test to determine which brand they most easily recalled. The results indicated that imprinting did indeed have an effect on brand recollection, as those surveyed were able to recall brands they were imprinted on much more easily than non-imprinted brands.

Willingness to pay is a concept designed to determine the maximum amount a consumer is going to pay for a specific product. Koçaş & Dogerlioglu-Demir (2013) attempted to determine what the cumulative effect that willingness to pay (WTP) has when attempting to sell a product. The ultimate purpose of their study was to potentially identify what behaviors lead to the final maximum price a consumer is willing to pay. For example, a seller may price a product at $80, so the consumer will use that as a base point to attempt to haggle the price from there. The two researchers conducted a series of studies to collect as many observations as they possibly could be regarding the WTP topic. These studies intended to observe the effects of priming on WTP, as well as demonstrate that hypothetical WTP is higher than real WTP (people will theoretically pay more than they actually will). After organizing that data into a diagnostic model, it was determined that anchoring did play a major role in overall WTP. Interestingly, the results demonstrated that each individual was affecting by anchoring, not just the extremes outside of the mean range.

Gwebu et al. (2011) continued to examine the effects of anchoring, this time pertaining to name your own price auctions. Websites like eBay offer name your own price auctions, and this study attempted to examine how anchoring can affect the price that the consumer is willing to pay. The researchers posited that there are certain tiers of pricing that auctioneers can aim for; prices that will allow sellers to sell a product for the best possible price. This research can be utilized in the future by sellers on websites like eBay, where those sellers can find a price point that best fits the product they are attempting to sell. In a lab experiment, a total of 140 data points, from which the researchers could analyze their proposed hypotheses, were organized into groups, and then run through an ANOVA model to either prove or disprove their various hypotheses. The results were interesting, in that it found that consumers tended to be swayed more by the low-end pricing rather than the high-end pricing. Although it appeared the high-end pricing had some sway as to what a consumer would end up paying, it was the low-end pricing that consumers used most to gauge what they should pay.

In a study about how consumers perceive reference, or anchor, prices that seem to be unlikely to be encountered, Suter and Burton (1996) further study the anchoring paradigm. A survey was conducted in which the anchor price was manipulated between each participant. Each participant was then asked to evaluate the perceived value of the prices they were presented with. The ensuing data was then run through a regression model, with the results demonstrating that implausible pricing actually does influence customer behavior. This study shows how powerful the anchoring effect is, when even a ludicrous price for a product can be set, and still consumer's use that price as a reference point. This of course could present an ethical issue, especially if sellers begin pricing products at absurd amounts, in which case it could be construed as an ethical and perhaps legal wrongdoing.

Wansink et al. (1998) also studied the effects of anchoring, in this instance applying it to how much a consumer is going to purchase. By demonstrating the effect that anchoring has on purchase quantity, Wansink et al. (1998) hopes to give retailers and manufacturers valuable information about consumer behavior. The sellers can utilize the information from this study to potentially increase customer satisfaction, as well as customer loyalty. Using an ANOVA model, the researchers hoped to determine if anchoring did in fact play a role in a consumer's decision regarding purchase quantity. Their research indicates that anchoring does in fact play a role in how much a consumer is going to purchase. Retailers can offer anchor-based promotions, hyping up a bundle of a given product to improve the consumer's perception of purchasing that quantity. Although it is not determined if this translates into improved sales, it certainly presents an interesting area for further research to be conducted.

Grewal et al. (1998) ran a similar study, researching the effects of price-comparison advertising on consumer perception. The researchers were attempting to further examine the perception of price from a consumer perspective, which is a valuable subject that can benefits retailers and sellers immensely. Using a regression model, the researchers found that while there was no perception difference between advertised selling price or advertised reference price, there was an influence on perception for the consumer's internal reference price and advertised selling price. This is interesting because it implies that a consumer has a preconceived notion of what they should be paying for a product. Further study into how this internal price was arrived at would be a particularly interesting, albeit difficult, topic to research, but it could provide useful information about the average consumer.

Blair and Landon (1981) conducted a study focused on reference prices within advertisements of the retail environment. The purpose of the study was to determine what effect reference prices had on consumer perception within the retail arena. An ANOVA model was developed to determine if reference prices presented within an advertisement would have any effect on the consumer's perception of savings. The researchers concluded that the effect was somewhat apparent, in that consumers typically did not accept a reference price as advertised but felt a greater sense of savings having been presented with a reference price. This is interesting because even if the reference price isn't completely accepted, it still has some effect on consumer perception. This could be useful to retailers who want their consumers to perceive they are getting big savings on products, although it can also be abused, manipulating consumers into purchasing products they may not necessarily need.

The aforementioned studies about anchoring seem to indicate that the effect is very difficult for consumers to avoid. However, Smith and Windschitl (2015) seem to think that if the consumer is knowledgeable enough, they can avoid the effects of anchoring. Using psychology students as subjects, the two researchers conducted experiments to determine the knowledge the students had about the specific tasks they were given. Using an ANOVA model, the researchers found that the increased knowledge about each task did not reduce the effects of anchoring present in the participants. The effect of the anchoring bias is incredibly powerful, so much so that not even improved knowledge on a given subject can reduce its implicit bias. Perhaps the correct type of knowledge is needed to help reduce any anchoring bias, but the fact that these students could not completely eliminate the effects from their perception demonstrates how powerful a bias of that nature can be.

Information about priming is not limited to gaining additional insight into consumer behavior. Pacheco (2005) implies that priming can be used by rival brands as a means of gaining some competitive advantage. Start-ups would find this information extremely useful, as it is very difficult for a new entrant to establish themselves in a given market. In a 2x2x2 study, Pacheco used 283 students to determine if priming could play a factor in the consumer's evaluation of the brand. After running his data through a multivariate analysis, Pacheco found that the students did indeed offer higher appraisals of products they were primed with versus products they weren't. New or relatively unknown brands can utilize this information in an effort to prime their target market to their products. They can attempt to endear their market to the product in an effort to gather a sustainable portion of the market share. If the brand can perform it well, they can gain a competitive advantage over other brands that might not necessarily be capable of priming their markets at the same level.

Priming can also be used to help build brand names or advertising slogans. Tseng (2013) offers research into this topic, theorizing that the priming effect would have greater impact on slogans than brand names. Many companies are built around their brand name, Coca-Cola for example, but, at least in recent times, few have come up with an advertising slogan that sticks in the consumer's brain. To test his theory, Tseng used an ANOVA model to test the priming effects of slogans under various conditions, both positive and negative. Interestingly, the results indicated that priming effects were more significant when paired with brand names, but only with the caveat of a savings appeal. This means that when offered some level of savings when purchasing a product, consumers were primed by the brand name rather than the slogan. However, the priming effect of slogans were more powerful when a level of savings was not implied in the transaction. For all intents and purposes, the priming effects of slogans and brand names are exact opposites.

METHODOLOGY

Variables and Hypotheses

According to various researchers, priming has an incredibly powerful effect on a customer's willingness to pay (Koçaş & Dogerlioglu-Demir, 2013). Willingness to pay is effectively how much a customer is going to pay for a particular product. Priming can alter that factor by manipulating the consumer's perception as to what the value of a particular product actually is. Therefore, our first hypothesis is:

Hypothesis 1: If anchoring is present, then the consumer's willingness to pay will become more likely.

When considering the factor of anchoring, one major point to consider is the idea of reference prices. When given a reference price, consumers are anchored to that point, fluctuating their perception of value around that price. Several studies have indicated that these reference, or anchor, prices demonstrated a significant impact on a consumer's perception (Grewal et al., 1998; Suter & Burton, 1996).

It can be concluded that willingness to pay is positively correlated with the anchored group ($100, $200, $300, and $400) or amount an individual is anchored with. When the amount an individual is anchored with increases, willingness to pay for the product also increases. The item used to anchor individuals were a set of four relatively similar surveys questionnaires, all four survey questioners/groups received the same products; the only difference is the amount they were anchored with. Group 1 was anchored with $100; group 2 was anchored with $200; group 3 with $300; and group 4 with $400 (see Appendix A). The item used to measure willingness to pay was measured on a continuous scale where individuals were asked to indicate the maximum amount they would be willing to pay (in US dollars) for each product (Question 2 of survey).

Hypothesis 2: If anchoring is present, then the consumer will have a more favorable perception of price.

The next hypothesis will be derived from studies concerning priming and imprinting enhancing a consumer's perception of quality (Baker 2003). The theory behind this notion is that imprinting a consumer to a particular brand or product will enhance their perceptions of said brand or product. This will cause them to associate positive thoughts and attributes towards the brand, including quality. If the person is imprinted or primed to a given brand or product, they will associate that with positive benefits and higher quality. Additionally, anchoring that consumer to a reference price they find attractive will also enhance their perception of quality.

Therefore, it can be concluded that perception of price is positively correlated with the anchoring group. The anchoring groups are similar to hypothesis one. Similarly, willingness to pay is similar to hypothesis one. The variable which measures the perception of price is measured on an ordinal scale where individual's perception of "a deal" (Question 4 of survey) will increase as anchoring and willingness to pay will also increase.

Hypothesis 3: If anchoring/imprinting is present, then the consumer will perceive the product as being higher quality.

The hypothesis will focus on all three aspects of this paradigm: anchoring, priming, and imprinting. The hypothesis is looking to answer the question: Is the perception of quality impacted by the amount an individual is anchored?

It is hypothesized that when the amount an individual is anchored with increases, perception of quality also increases. The anchoring groups are similar to hypothesis one. The variable which measures perceived quality is measured on an ordinal scale where individuals' perception of "good deal" (Question 3 of survey) will increase as anchoring increase.

Hypothesis 4: Anchoring has an effect on gender

This hypothesis will focus on how anchoring affects people on a demographic level. It has been said that the three influences are extremely difficult to resist (Smith & Windschitl, 2015). Therefore, gender will not be a good indicator of measuring anchoring effect.

Hypothesis 5: If anchoring is present, then the consumer with the higher disposable income will be more willing to pay

Similarly, we are interested in seeing if a consumer's income level will have any effect on the anchoring paradigm. A person with a higher weekly income will more than likely be willing to spend more on a given product. We will be attempting to determine if that relates to anchoring, in that if the consumer is given a higher reference price, their average income level will lead them to be willing to pay that higher price.

ANOVA Models

To conduct our hypothesis testing, 4 one-way ANOVA models will be created to test 4 different dependent variables. Our first model will test anchoring's effect on "willingness to pay", where willingness to pay increases with the anchoring group ($100, $200, $300, and $400) The second model will test if anchoring has an effect on the customer's perception of price, where perception of price will increases with the anchoring group ($100, $200, $300, and $400) The third model will evaluate the customer's perception of quality as influenced by the anchoring paradigm, when the amount an individual is anchored with increases, perception of quality also increases. The last model tested will determine if gender and income levels are affected predictors of anchoring paradigm. Additionally, we will employ 4 multiple regression models in an effort to determine what variables are most significant when it comes to the anchoring paradigm. Each of the four models will test the relationship of the independent variables to each anchoring group. For example, model 1 will use the $100 anchoring group as it's dependent variable, while its independent variables will be perception of quality, perception of price, gender, and disposable income level. Three more models will be run for each anchoring group, which are the $200, $300, and $400 groups. These models are ultimately testing our first three hypotheses; demonstrating the relationship, if any, between anchoring and willingness to pay, anchoring and perception of price, and anchoring and perception of quality. The 4 ANOVA models will be used to test our last 2 hypotheses, pertaining to anchoring's effects on gender and various income levels.

In order to determine the effect that the independent variables may have on the dependent variable, we collected data from a survey conducted prior to this study. Our survey consisted of polling 180 students from a university on the east coast, where roughly 12,000 students attend. The survey consists of a selection of products, each priced within a specific range. These ranges are intended as the anchor, or reference, price with which we can draw our conclusions from. The administered survey can be found on Appendix A.

RESULTS

Sample Data

There was a total sample of 180 students where 81 (45%) were males and 99 (55%) were female (Mean age = 22.06 years, age range: 18-37 years, SD = 3.12). Most of the students that were tested were students from various university classes. The majority were students of upper division Marketing and Psychology courses. Also, their average weekly disposable income ranged from $101 to $150. Table 1 reported the sample data demographics. These participants were taking part in the study as part of a requirement that they must meet in order to pass their introductory psychology courses, while the rest participated for extra credit in their designated classes. The participants who took part due to class requirement, signed up using a computer system called SONA. SONA was designed in order for students, who are taking introductory psychology classes, to participate in various experiments and earn credit for doing so.

TABLE 1. DEMOGRAPHICS OF SAMPLE DATA

Age	22 ± 3.12 years old		
Weekly Disposal Income	$101 - $150		
Gender	Male	45%	81
	Female	55%	99
School Year	Freshman	3%	6
	Sophomore	21%	38
	Junior	28%	51
	Senior	46%	83
	Graduate Student	1%	2
Total			180

In order to examine the difference between groups that were anchored, imprinted and primed with values equaling $100, $200, $300 and $400 on participants' bidding decisions, we conducted four one-way ANOVA between participant's analysis of variance. Results indicated that all four products were statistically significant.

Across all four products, the independent variable (Group) is measured at a categorical level while the dependent variable is measured on a continuous level in dollar amount (willingness to pay). Results showed that Product 1 (Headphones) with a $F(3, 175) = 3.48$, $p < .05$, $\eta^2 = .06$ as shown in Table 2; Product 2 (Tablet) with a $F(3, 169) = 9.56$, $p < .01$, $\eta2 = .14$; Product 3

(TV) with a $F(3, 168) = 9.46$, $p < .01$, $\eta2 = .14$; Product 4 (Smartphone) with a $F(3, 173) = 7.44$, $p < .01$, $\eta2 = .11$. Table 2 also showed means and standard deviations.

TABLE 2. ANOVA MODEL RESULTS ON ANCHORING EFFECT

Product	Anchoring Group	N	Mean	SD	DF	F
Product 1 (Headphone)	Anchoring Value ($100)	51	61.62	55.44	3, 175	3.48*
	Anchoring Value ($200)	44	85.07	80.12		
	Anchoring Value ($300)	39	107.05	89.85		
	Anchoring Value ($400)	46	107.50	92.26		
	Total	180	89.07	81.53		
Product 2 (Tablet)	Anchoring Value ($100)	51	185.31	90.76	3, 169	9.56**
	Anchoring Value ($200)	44	191.40	90.20		
	Anchoring Value ($300)	39	259.45	121.24		
	Anchoring Value ($400)	46	278.87	106.34		
	Total	180	227.02	109.27		
Product 3 (TV)	Anchoring Value ($100)	51	191.22	106.67	3, 168	9.46**
	Anchoring Value ($200)	44	256.53	165.08		
	Anchoring Value ($300)	39	313.72	170.37		
	Anchoring Value ($400)	46	346.80	157.86		
	Total	180	272.99	160.74		
Product 4 (Smartphone)	Anchoring Value ($100)	51	122.55	63.57	3, 173	7.44*
	Anchoring Value ($200)	44	129.95	72.36		
	Anchoring Value ($300)	39	169.74	81.22		
	Anchoring Value ($400)	46	188.96	94.89		
	Total	180	151.67	82.70		

Note: *p < 0.05, **p < 0.01; Dependent Variable = Maximum $ amount willing to pay; SD = Standard Deviation; DF = Degree of Freedom

WillingtoPay1 variable is a 0-1 dummy variable with 0 being the respondent would be willing to buy and 1 being the respondent would not be willing to buy. Gender is also a 0-1 dummy variable with 0 being male and 1 being female. Table 3 showed the results of the Pearson correlation, as each variable pertains to Price Group 1.

TABLE 3. PEARSON CORRELATIONS FOR PRICE GROUP 1 ($100)

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Max Amount Willing to Pay	1					
(2) Disposable Income	.190*	1				
(3) Perceived Quality	.382**	.011	1			
(4) Is it a Deal?	.345**	.063	.319**	1		
(5) Gender	-.361*	-.081	-.133	-.036	1	
(6) Willing to Pay1	.683**	.101	.272**	.580**	-.295**	1

*p<0.05, **p<0.01

Models 1 through 4 show statistical significance ($p < 0.001$), and the Pearson correlation results demonstrate no real issue with multicollinearity. With maximum amount willing to pay as the dependent variable, we find that the most significant variable is in fact our anchoring variable. In each case the anchoring variable (Willing to Pay) is the most significant variable ($p < 0.001$). We find that none of the other variables are significant at any level, with the exception of gender being significant in two of the products at the $100 value level. Note that a model was not included for price group 1 for product 2, as the model did not demonstrate reliability at any alpha value. This validates the claim made in our first hypothesis, namely that since the respondent is anchored at the $100 price range, they are willing to pay more for each product. There seems to be no relation between quality and anchoring in this price group, nor does there seem to be any relation between disposable income and anchoring either. Finally, our fourth hypothesis regarding gender is neither confirmed nor denied, as gender was significant in 2 of the 4 models. Table 4 showed the regression results on Price Group 1.

TABLE 4. REGRESSION ANALYSIS RESULTS OF PRICE GROUP 1 ($100)

	Model 1	Model 2	Model 3	Model 4
Dependent Variable	Max $ willing to pay Product 1	Max $ willing to pay Product 2	Max $ willing to pay Product 3	Max $ willing to pay Product 4
Disposable Income	-1.755	-1.373	-5.459	-3.703
Perceived Quality	8.645	1.454	14.384	16.325
Is it a Deal?	4.865	13.716	-6.521	-1.938
Gender	-20.391*	-12.635	64.813*	-22.219
Willing to pay	85.820***	101.231***	112.806*	74.545***
Adjusted R^2	0.690	0.521	0.287	0.509
Degree of Freedom	5, 43	5, 43	5, 42	5, 42
F	22.326***	13.321***	4.789***	10.759***

*p<0.05, **p<0.01, ***p<0.001

WillingtoPay2 variable is a 0-1 dummy variable with 0 being the respondent would be willing to buy and 1 being the respondent would not be willing to buy. Gender is also a 0-1 dummy variable with 0 being male and 1 being female. Table 5 shows the results of the Pearson correlation, as each variable pertains to price group 2.

TABLE 5. PEARSON CORRELATIONS FOR PRICE GROUP 2 ($200)

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Max Amount Willing to Pay	1					
(2) Disposable Income	.136	1				
(3) Perceived Quality	.372**	-.036	1			
(4) Is it a Deal?	.241**	.005	.411**	1		
(5) Gender	.048	-.081	.242**	.022	1	
(6) Willing to Pay2	.390**	.098	.340**	.620**	-.024	1

*p<0.05, **p<0.01

Models 5 through 8 all show statistical significance at either the p<0.001 level or the p<0.01 level, and the Pearson correlation is very similar to our first model. Maintaining the maximum amount willing to pay as the dependent variable; the anchoring variable is still the most significant variable. Perception of price (Is it a deal?) also become significant with products 2 (p < 0.05) and 4 (p < 0.01) respectively. This is consistent with our first set of models. Once again hypothesis 1 is validated, while hypothesis 2 is neither confirmed nor denied. With this particular price group, our other hypotheses are rejected and the null is held true, namely, that there is no relationship between quality, gender, and income, with regards to anchoring. Table 6 showed the regression results on Price Group 2.

TABLE 6. REGRESSION ANALYSIS RESULTS OF PRICE GROUP 2 ($200)

	Model 5	Model 6	Model 7	Model 8
Dependent Variable	Max $ willing to pay Product 1	Max $ willing to pay Product 2	Max $ willing to pay Product 3	Max $ willing to pay Product 4
Disposable Income	1.631	-2.595	-5.82	-6.298*
Perceived Quality	7.428	2.554	-5.743	12.492
Is it a Deal?	9.555	23.815*	50.218	18.455**
Gender	-21.224	-12.908	36.490	-18.301
Willing to pay	133.211***	112.673***	111.523	109.946***
Adjusted R^2	0.690	0.649	0.249	0.827
Degree of Freedom	5, 38	5, 36	5, 37	5, 38
F	20.009***	16.188***	3.786**	42.199***

*p<0.05, **p<0.01, ***p<0.001

WillingtoPay3 variable is a 0-1 dummy variable with 0 being the respondent would be willing to buy and 1 being the respondent would not be willing to buy. Gender is also a 0-1 dummy variable with 0 being male and 1 being female. Table 7 shows the results of the Pearson correlation, as each variable pertains to price group 3.

TABLE 7. PEARSON CORRELATIONS FOR PRICE GROUP 3 ($300)

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Max Amount Willing to Pay	1					
(2) Disposable Income	-.076	1				
(3) Perceived Quality	.290**	-.124	1			
(4) Is it a Deal?	.405**	-.207**	.468**	1		
(5) Gender	.168	-.081	.262**	.099	1	
(6) WillingtoPay3	.421**	-.090	.392**	.685**	.096	1

*p<0.05, **p<0.01

Models 9 through 12 all show significance, with all models showing significance at the p<0.001 level. Similar to our prior models, the most significant variable when determining the maximum amount willing to pay (our dependent variable) is still the anchoring variable. In each model, it was significant at either the p<0.01 level or the p<0.05 level. Additionally, Quality became significant in this price group for product 4, but we cannot confirm nor deny the validity of our 3rd hypothesis based on this lone result. Once again, we accept our 1st hypothesis as being true, while rejecting all other hypotheses. The Pearson correlation for this particular group is similar to the correlation models for our prior groups. Table 8 showed the regression results on Price Group 3.

TABLE 8. REGRESSION ANALYSIS RESULTS OF PRICE GROUP 3 ($300)

	Model 9	**Model 10**	**Model 11**	**Model 12**
Dependent Variable	Max $ willing to pay Product 1	Max $ willing to pay Product 2	Max $ willing to pay Product 3	Max $ willing to pay Product 4
Disposable Income	5.607	-6.412	-12.475	8.154
Perceived Quality	12.609	19.500	14.445	23.983
Is it a Deal?	18.422	9.803	19.482	15.544
Gender	-21.769	34.550	-11.300	-.911
Willing to pay	178.941***	174.797***	184.097**	121.040***
Adjusted R^2	0.728	0.652	0.451	0.707
Degree of Freedom	5, 33	5, 31	5, 30	5, 33
F	21.327***	14.506***	6.740***	19.321***

*p<0.05, **p<0.01, ***p<0.001

WillingtoPay4 variable is a 0-1 dummy variable with 0 being the respondent would be willing to buy and 1 being the respondent would not be willing to buy. Gender is also a 0-1 dummy variable with 0 being male and 1 being female. Table 9 shows the results of the Pearson correlation, as each variable pertains to price group 4.

TABLE 9. PEARSON CORRELATIONS FOR PRICE GROUP 4 ($400)

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Max Amount Willing to Pay	1					
(2) Disposable Income	.054	1				
(3) Perceived Quality	.384**	-.055	1			
(4) Is it a Deal?	.316**	-.042	.394**	1		
(5) Gender	-.161*	-.081	-.013	-.093	1	
(6) Willing to Pay4	.413**	.006	.362**	.604**	.096	1

*p<0.05, **p<0.01

Models 13 through 16 continue testing the effects of anchoring, this time at the $400 level. With the maximum amount willing to pay continuing to be our dependent variable, all four models in price group 4 demonstrate significance at the p<0.001 level. Once again, the most significant variable is the willingness to pay, or anchoring, variable. Price perception becomes significant in Product 2 (p < 0.05), and Product 3 (p < 0.01). Also, gender becomes significant in Products 1 (p < 0.05) and 4 (p < 0.05) respectively. Therefore, we could conclude that data supported Hypothesis 1, while Hypotheses 2 and 4 were inconclusive. The other hypotheses were rejected. Table 10 reported the regression results on Price Group 4.

TABLE 10. REGRESSION ANALYSIS RESULTS OF PRICE GROUP 4 ($400)

	Model 13	Model 14	Model 15	Model 16
Dependent Variable	Max $ willing to pay Product 1	Max $ willing to pay Product 2	Max $ willing to pay Product 3	Max $ willing to pay Product 4
Disposable Income	12.676	10.669	4.290	14.822
Perceived Quality	9.606	21.810	-10.985	13.781
Is it a Deal?	-4.443	27.994*	85.220***	7.618
Gender	-57.531**	-28.796	26.296	-48.095*
Willing to pay	235.339***	120.197***	110.456**	259.666***
Adjusted R^2	0.728	0.649	0.759	0.542
Degree of Freedom	5, 33	5, 40	5, 38	5, 39
F	21.327***	17.610***	28.048***	11.4000***

*p<0.05, **p<0.01, ***p<0.001

Examining the model sets as a whole, we find that the anchoring variable is easily the most significant among all 16 models. Therefore, we can reasonably claim hypothesis 1 to be true across all price groups. However, variables such as gender, price perception, and quality perception seemingly become significant based on the product or price level being tested. Therefore, we cannot reasonably claim hypotheses 2, 3, and 4 to be true. However, we cannot flatly deny those hypotheses either. Logically, we must conclude that hypotheses 2, 3, and 4 are inconclusive across all price groups. Table 11 showed evidence on Hypothesis 2. Table 12 reported evidence on Hypothesis 3.

TABLE 11. DATA ON PRICE PERCEPTION

	Price Group				
	$100	$200	$300	$400	Total
	Mean	Mean	Mean	Mean	Mean
Is Product 1 a Deal?	2.46	1.84	1.82	1.50	1.92
Is Product 2 a Deal?	3.94	3.30	3.00	2.67	3.25
Is Product 3 a Deal?	3.94	3.41	3.15	2.98	3.39
Is Product 4 a Deal?	3.34	2.61	2.44	2.18	2.67

TABLE 12. DATA ON PRODUCT QUALITY PERCEPTION

	Price Group				
	$100	$200	$300	$400	Total
	Mean	Mean	Mean	Mean	Mean
Perceived Quality for Product 1	3.71	3.25	4.08	3.78	3.69
Perceived Quality for Product 2	4.12	3.61	4.23	4.11	4.02
Perceived Quality for Product 3	4.16	3.61	3.97	3.91	3.92
Perceived Quality for Product 4	3.86	3.36	3.64	3.60	3.63

The four ANOVA models indicate that there was no significant difference across the various groups of weekly disposable income. This means that regardless of income level, anchoring, imprinting, and priming will be effective. There were no consistent results when measuring the impact of anchoring, imprinting and priming on gender. Results were chaotic, where in some instances there was a significance in amount willing to bid on products and in other examples there was no significance. Therefore, we concluded that gender is not a reliable variable. Table 13 showed evidence on Hypothesis 4.

TABLE 13. DATA ON GENDER

	Price Group									
	$100		$200		$300		$400		Total	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Product	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Product 1	86.64	41.96	116.38	47.50	135.00	83.10	158.82	77.41	121.35	62.40
Product 2	178.57	190.56	188.26	195.00	229.41	285.00	315.62	258.59	221.17	231.71
Product 3	137.27	235.19	231.78	285.00	299.94	324.75	343.53	348.85	243.46	297.49
Product 4	140.45	107.96	146.42	110.20	177.78	162.86	214.71	173.32	166.10	139.50

Finally, our fifth hypothesis can be outright rejected, as disposable income is consistently insignificant across all of the products and price points. Table 14 reported evidence on Hypothesis 5.

TABLE 14. DATA ON DISPOSABLE INCOME

		Disposable Income				
		<$50	$51-$100	$101-$150	$151-$200	>$201
Price Group	Product	Mean	Mean	Mean	Mean	Mean
Group 1 ($100)	Product 1	68.45	66.82	38.89	45	66.67
	Product 2	188.5	179.55	183.33	125	208.33
	Product 3	213.75	186.25	163.89	85	202.5
	Product 4	120.75	125	110	150	133.33
Group 2 ($200)	Product 1	48.58	107.73	72.5	95	101.36
	Product 2	150	235	200	158.33	211.82
	Product 3	271.33	285	175	241.67	252.27
	Product 4	128.33	133.55	133.75	100	143.09
Group 3 ($300)	Product 1	95.91	72.78	143	62.5	157
	Product 2	250	255.56	277.78	250	260
	Product 3	364.5	305.56	305.44	266.67	270
	Product 4	154.55	152.22	185	187.5	190
Group 4 ($400)	Product 1	87.27	66.07	130	160	175
	Product 2	231.82	246.43	300	389.8	325
	Product 3	367	369.29	293.33	349.8	383.33
	Product 4	181.82	201.85	183.33	205.8	162.5
Total	Product 1	73.29	77.78	103.71	100.59	115.38
	Product 2	200.38	228.98	250	244.06	237.69
	Product 3	285.5	289	248.35	260.56	259.6
	Product 4	141.76	154.73	159.29	157.59	152.85

DISCUSSION

Similar to Ariely's (2008) study, we were able to support the claim that anchoring, imprinting, and priming occurs on individuals when presented with an initial unrelated value. Therefore, we can conclude that anchoring does serve as a mental benchmark or starting point for estimating an unknown product (Poundstone, 2010). Also, when presented with a product of an unknown value, the anchored amount serves as a reference price (Grewal et al., 1998; Suter & Burton, 1996).

Researchers have suggested that the reason why consumers engage in these mental shortcuts is due to the fact that they do not wish to take the time and learn about new products or engage in the mental process of determining the true value (Dobrev & Gotsopoulus, 2010; Solomon et al., 2009).
The study also concluded that consumers' willingness to pay the full amount for the product is irrelevant since the anchor is such a powerful influence in manipulating consumer's perception of what the actual value of a product really is (Koçaş & Dogerlioglu-Demir, 2013).

The study builds on previous research by adding a series of new dimensions. These dimensions are gender and income. The results for gender were inconsistent. Some groups/products showed no significant difference among gender where other groups did.

However, when looking at income, the amount of disposable capital had no significant difference on the effect of anchoring, imprinting, or priming. This means that anchoring affects individuals of all income levels. Additional research on the topic needs to be conducted in order to better understand the effects of anchoring, imprinting, and priming and its impact on consumers.

CONCLUSION

The results of this research show that regardless of income level, anchoring, imprinting, and priming will be effective. Also, there were no consistent results when we reviewed the impact of anchoring, imprinting and priming on gender. In terms of the effects of anchoring, priming, and imprinting; the results indicate that there was a significant presence of that paradigm amongst our survey respondents. It not only influenced their decision to purchase the product; it also influenced how much the person was willing to pay for the product.

Further studies on this subject matter should be investigated on the impacts of anchoring, imprinting and priming on gender. The limitations of this sample size could be also put in question and increased. This investigation could be improved on by increasing the sample size, changing the location, and using subjects that rely more on their income than on an allowance or a limited income. It should be noted that this study was exploratory in nature. All our data was collected using a computer system, SONA. We caution in generalizing the findings to the entire US population.

REFERENCES

Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions*. Harper Collins.

Arney, C. (2010). Predictably irrational: The hidden forces that shape our decisions. *Mathematics and Computer Education, 44*(1), 68-69.

Baker, W. (2003). Does brand name imprinting in memory increase brand information retention? *Psychology & Marketing, 20*(3), 1119-1135.

Blair, E. A. & Landon, E. L., Jr. (1981). The effects of reference prices in retail advertisements. *Journal of Marketing, 45*(2), 61-69.

Boush, D. (1993). How advertising slogans can prime evaluations of brand extensions. *Psychology & Marketing, 10*(1), 67-78.

Dias, M. (2009). *Ask your doctor: The direct-to-consumer (DTC) priming effect of pharmaceutical marketing on purchasing and health seeking behavior* [Unpublished master's thesis]. Central Connecticut State University.

Dobrev, S. D. & Gotsopoulos, A. (2010). Legitimacy vacuum, structural imprinting, and the first mover disadvantage. *Academy of Management Journal, 53*(5), 1153-1174.

Grewal, D., Monroe, K. B., & Krishnan, R. (1998). The effects of price-comparison advertising on buyers' perceptions of acquisition value, transaction value, and behavioral intentions. *Journal of Marketing, 62*(2), 46-59.

Gwebu, K. W., Wang, J., Hao, A. W., & Hu, M. Y. (2011). Effects of price recommendation in name-your-own-price auctions. *Journal of Electronic Commerce Research, 12*(1), 61-77. Kholekile L.

Hardesty, D. & Suter, T. A. (2005). E-tail and retail reference price effects. *Journal of Product and Brand Management, 14*(2), 129-136.

Koçaş, C. & Dogerlioglu-Demir, K. (2013). An empirical investigation of consumers' willingness-to-pay and the demand function: The cumulative effect of individual differences in anchored willingness-to-pay responses. *Marketing Letters, 25*(2), 139-152.

Lagnado, D. A. (2007). Perspectives on Daniel Kahneman. *Thinking & Reasoning*, *13*(1), 1-4.

Lee, A. (2002). Effects of implicit memory on memory-based versus stimulus-based brand choice. *Journal of Marketing Research, 39*(4), 440-454.

Nisbett, A. (1976). *Konrad Lorenz*. A Helen and Kurt Wolff Book, 43-54.

Pacheco, B. G. (2005). *Implicit priming as a competitive strategy for challenger brands* [Unpublished doctoral dissertation]. University of Colorado.

Poundstone, W. (2010). *Priceless: The myth of fair value (and how to take advantage of it)*. Hill and Wang.

Smith, A. R., & Windschitl, P. D. (2015). Resisting anchoring effects: The roles of metric and mapping knowledge. *Memory and Cognition*, *43,* 1071-1084.

Solomon, M. R., Cornell, D. L., & Nizan, A. (2009). *Lunch! Advertising and promotion in real time*. Flat World Knowledge.

Suter, T. A., & Burton, S. (1996). Believability and consumer perceptions of implausible reference prices in retail advertisements. *Psychology & Marketing, 13*(1), 37–54.

Tseng, C. (2013). The priming effect of brand names and slogans with various appeals. *International Conference on Enterprise Marketing and Globalization Proceedings*, 1-7.

Wansink, B., Kent, R. J., & Hoch, S. J. (1998). An anchoring and adjustment model of purchase quantity decisions. *Journal of Marketing Research, 35*(1), 71-81.

APPENDIX A.
SURVEY QUESTIONAIRE ON GROUP ANCHORED WITH $100 VALUE

1) Are you willing to pay $100 for this product? (a) Yes (b) No

2) Please indicate the maxim amount you would be willing to pay:
$_________________

3) Item number one's perceived quality: Poor 1 2 3 4 5 Good

4) I think item number one is a deal: Disagree 1 2 3 4 5 Agree

1) Are you willing to pay $100 for this product? (a) Yes (b) No

2) Please indicate the maxim amount you would be willing to pay:
$_________________

3) Item number two's perceived quality: Poor 1 2 3 4 5 Good

4) I think item number two is a deal: Disagree 1 2 3 4 5 Agree

1) Are you willing to pay $100 for this product? (a) Yes (b) No

2) Please indicate the maxim amount you would be willing to pay:
$_________________

3) Item number three's perceived quality: Poor 1 2 3 4 5 Good

4) I think item number three is a deal: Disagree 1 2 3 4 5 Agree

1) Are you willing to pay $100 for this product? (a) Yes (b) No

2) Please indicate the maxim amount you would be willing to pay: $________________

3) Item number four's perceived quality: Poor 1 2 3 4 5 Good

4) I think item number four is a deal: Disagree 1 2 3 4 5 Agree

1) Please indicate your age:__________

2) Gender (please circle one): (a) Male (b) Female

3) Please indicate your race (please circle one):
   a) White
   b) African American
   c) American Indian or Alaska Native
   d) Asian
   e) Native Hawaiian or other Pacific Islander
   f) Other; Specify:_____________

4) Please indicate your school year:
   a) Freshman b) Sophomore c) Junior d) Senior e) Graduate Student

5) Weekly disposable Income:
   a) < $50 b) 51 – 100 c) 101 – 150 d) 151 – 200 e) > 201

HOW BOARD GENDER AND KNOWLEDGE-BASED DIVERSITY INFLUENCE FIRM PROCESS INNOVATION

Pingying Zhang, University of North Florida
pingying.zhang@unf.edu

Sadi Bogac Kanadli, Republic of Turkey Ministry of Trade
sadi.kanadli@alumni.esade.edu

Nada Kakabadse, University of Reading
n.kakabadse@henley.ac.uk

ABSTRACT

The paper aims to concurrently examine effects of knowledge-based diversity and gender diversity on firm process innovation. It also investigates board chairpersons' gender effect on these relationships using the categorization–elaboration model (CEM). The paper benefits from the survey method and applies structural equation modeling (SEM) to examine responses of 462 CEOs of publicly listed firms and private firms. We find that knowledge-based diversity has a more substantial impact on process innovation than gender diversity. Meanwhile, it is more likely for female board chairpersons to utilize gender diversity to improve firm process innovation. For male chairpersons, they are more likely to use knowledge-based diversity to improve process innovation. The finding suggests that social categorization processes based on gender matter to some extent. Practitioners should pay particular attention to the skills and quality makeup of boards while being cognizant of the varied support to gender diversity between female and male chairpersons.

INTRODUCTION

Research shows that board gender diversity enhances firm innovation outcomes (Torchia et al., 2018). Board gender diversity is measured by a composition degree between male and female directors. Researchers also show that the knowledge-based diversity of a board can also improve firm innovation. Knowledge-based diversity is measured by the variety in directors' educational background, functional background, and industry experience (Forbes & Milliken, 1999). Independently, both diversities exert a desirable impact on innovation. Together, their implications are, however, seldom explored. The fruitfulness of examining multiple dimensions of group diversity underlies our research interests (e.g., Guillaume et al., 2017; Homan et al., 2007). We intend to investigate how board gender diversity and knowledge-based diversity affect firm process innovation.

Firm process innovation is one critical innovation outcome (Crossan & Apaydin, 2010; Galia & Zenou, 2012), which influences a firm's product innovation (Martinez-Ros, 2000) and overall firm performance (Adner & Levinthal, 2001). Process innovation describes the introduction of new process technologies that may comprise changes to the production process and adaptations of modern management practices to achieve lower costs and higher quality (Adner & Levinthal, 2001; Reichstein & Salter, 2006). For example, a lean production method is process innovation, involving new material-processing technologies, task designs, and management practices (Reichstein & Salter, 2006). It is reasonable to expect that board knowledge-based diversity and gender diversity can improve firm process innovation. However, the effects become complicated through a social categorization lens when there are distinctions between subgroups of “we vs. they” and “in-group vs. out-group” due to diversity (We apply a categorization–elaboration model (CEM) to understand the boardroom's social categorization effect. The CEM suggests that some demographic variables, such as gender, can create separation between males and females, leading to conflicts between subgroups and ultimately slowing down creative solutions in a group (Van Knippenberg et al., 2004; Van Knippenberg & Schippers, 2007). Unless certain conditions/mechanisms are in place, women directors' diverse talent may not be fully realized due to the negative consequences of social categorization processes (Eagly, 2016; Harrison & Klein, 2007; Kanadlı et al., 2018a).

The CEM also emphasizes contingency factors that can reduce the separation effect by facilitating information elaboration (Van Knippenberg et al., 2004). Board leadership is one such factor that promotes the use of directors’ diverse talent and reduces the adverse effects of social categorization (Åberg & Shen, 2020; Guerrero et al., 2015; Kakabadse et al. 2015; Kanadlı et al., 2018a). The chairperson’s gender is particularly interesting in this regard, which may influence the board leadership style. For example, a female chairperson might be more aware of and sensitive to the social categorization processes (Eagly, 2007; Huse & Solberg, 2006; Kakabadse et al., 2015). She might put more emphasis than a male chairperson on managing social categorization by acting as a mentor and helping women minorities fit into male-dominated boardrooms (Eagly, 2007; Huse & Solberg, 2006; Kakabadse et al., 2015). Hence, chairperson gender presents an interesting opportunity in understanding the leadership style. We present our research model in Figure 1.

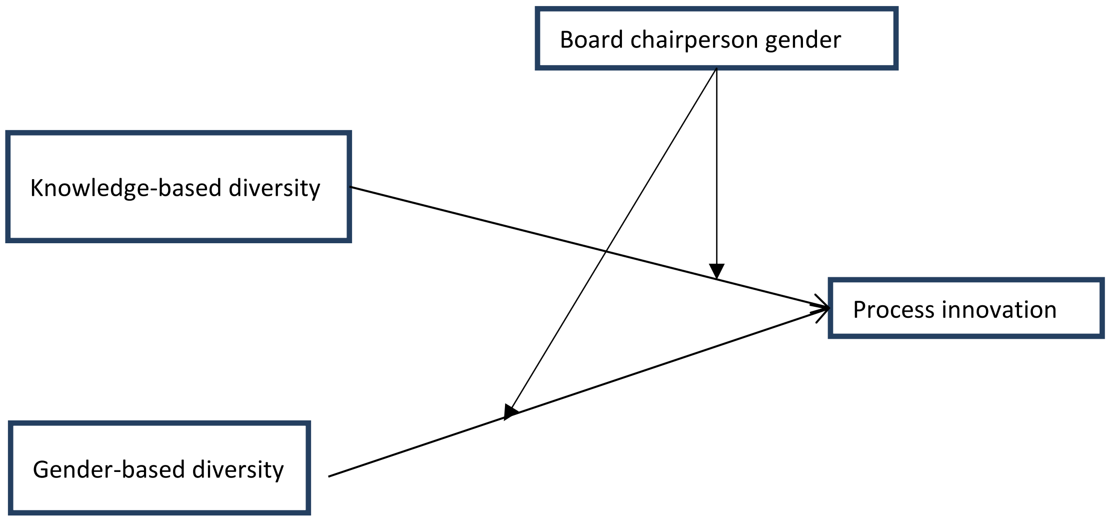

FIGURE 1. RESEARCH MODEL

The contribution of this study is three-fold. First, we introduce CEM as a novel approach to board diversity research by drawing scholarly attention to the interaction between social categorization processes and the use of directors’ knowledge and skills. Contrary to the argument that knowledge-based and gender-based diversity both have favorable influences on strategic decisions (Page, 2007), we hypothesize and demonstrate the two types of diversities can elicit different effects on process innovation. As a result, our study echoes the multifaced nature of board diversity (Harrison & Klein, 2007; Milliken & Martins, 1996). Second, the paper contributes to the field of board leadership research. A growing body of research has acknowledged the importance of board leadership on the utilization of diversity potential (e.g., Eagly, 2016; Gabrielsson et al., 2007; Guerrero et al., 2015; Kakabadse et al. 2015; Kanadlı et al., 2018a; Pugliese et al., 2015). We confirm the fruitfulness of this line of research by extending the analysis towards the gender of board chairpersons (Eagly, 2016; Guerrero et al., 2015; Kakabadse et al. 2015). Finally, the finding of the paper may enrich the discussion of corporate governance practices regarding controlling the negative impact of social categorization.

THEORETICAL BACKGROUND AND HYPOTHESES

Categorization–Elaboration Model

Integrating information processing and social categorization perspectives, Van Knippenberg et al. (2004) have introduced the categorization-elaboration model (CEM). The information-processing theory postulates that information creates value through individual contributions such as ideas, decisions, and judgments in the group information-processing space, such as in group discussions and meetings (De Dreu et al., 2008; Hinsz et al., 1997). The purpose of group-level information processing is to produce a "*coherent, feasible, sensible, and, if needed, creative group judgment or decision*" (De Dreu et al., 2008, p. 28). First, the information-processing view postulates that information creates value through individual contributions such as ideas, decisions, and judgments in a group setting (De Dreu et al., 2008). The quality of group-level information processing varies greatly, as group members "*differ in the depth with which information is searched and processed*" (De Dreu et al., 2008, p. 26).

In the context of boards of directors, group-level information-processing is exemplified by board meetings and discussions. Variety in directors' skills, knowledge, and experience enables them to contribute with a broader range of unique information and increases the comprehensiveness of decision-making at the group level (Gabaldon et al., 2018; Van Knippenberg et al., 2004). Informational diversity could also come from gender difference, where a gender-diverse board is more likely to take up various issues and initiate different approaches (Eagly, 2016; Huse & Solberg, 2006; Terjesen et al., 2009). Given the rich bundle of informational diversity, groups may use in-depth information to understand problems better. Such comprehensiveness improves group performance, especially group decision-making creativity and quality (Van Knippenberg et al., 2004).

Still, in a group of apparent informational diversity, not all members' information is added to the group information-processing space. Whereas some information is shared among all group members, other information remains unshared and individually possessed (Brodbeck et al., 2007; Scholten et al., 2007). Information bias occurs if the unshared information becomes critical to decision-making, resulting in groupthink and more inferior decision-making quality (Brodbeck et al., 2007).

To limit groupthink's potential damage, CEM posits that the elaboration of task-related information is the primary process underlying the positive effects of diversity on deep and deliberate information processing at the group level. Van Knippenberg et al. (2004, p. 1101) explain that "*elaboration is defined as the exchange of information and perspectives, individual-level processing of the information and perspectives, the process of feeding back the results of this individual-level processing into the group, and discussion and integration of its* implications." Furthermore, by disseminating and processing unshared task-related information, groups make superior quality decisions (Scholten et al., 2007).

Social categorization is the second pillar of the CEM. It calls attention to the impact of social barriers and intergroup biases on the information-elaboration process (Van Knippenberg et al., 2004). Social categorization offers explanations through social identity and self-categorization theories (Turner et al., 1987). Social identity theory explains why a group may be separated into "us" and "them" (in-groups/out-groups), and the self-categorization theory explains the consequences of such a categorization on minority members, such as female directors in the board.

When gender becomes a salient observable, social categorization starts to form, and intergroup biases emerge. Intergroup bias refers to more favorable perceptions of and attitudes and behavior toward, for example, a gender-based group category (e.g., woman vs. man) rather than individual merits. Moreover, negative attributes are exaggerated, and positive ones are discounted in intergroup biases, while the permissible behavior of the stereotyped person is constrained (Van Knippenberg et al., 2004). In this case, the information generated by an out-group member tends to be perceived as less relevant or credible than the information from an in-group member. As a result, in-group members are more inclined to cooperate with other in-group members than out-group members (Balliet et al., 2010). In-group members are also likely to resist out-group members' influence and dismiss or devalue their inputs in group decision-making (Tanford & Penrod, 1984).

It is interesting to point out that threats or challenges to the value of group identity between in- and out-group members can also trigger intergroup biases (Branscombe et al., 1999). Consequently, we observe social competition for status and prestige, outright derogation, discrimination of the group, and unequal status between groups (Van Knippenberg et al., 2004). In short, CEM enables us to address one key question in diversity research: "How differences between workgroup *members affect group process and performance, as well as group member attitudes and subjective well-being*" (Van Knippenberg & Schippers, 2007, p. 517). While we acknowledge that board diversity affects board processes and board task performance (e.g., Kakabadse et al., 2018; Kanadlı et al., 2018a; Zhang, 2010) as well as directors' attitudes towards minority directors (Zhu et al., 2014), we propose that social categorization in the board may be a crucial factor moderating the above processes.

Impacts of Board Diversity on Process Innovation

Corporate boards of directors are expected to be active players in facilitating innovation through board diversity (Adams et al., 2015). They are also identified as the primary determinant of firm process innovation (Crossan & Apaydin, 2010). It means directors are enablers of introducing new technologies, involving the use of new material-processing technologies, task designs, and management practices. Board diversity can play a role in spreading new ideas through its network ties. Information and knowledge from external partners through board networks are valuable in opportunity recognition (Reichstein & Salter, 2006). Specifically, directors form boundary-spanning activities that help focal firms explore opportunities from the external environment by bringing in information about new technologies, advanced processes, and best management practices adopted by other firms (Hillman & Dalziel, 2003). As articulated by Galaskiewicz and Wasserman (1989, p. 456), network ties between boundary-spanning personnel, such as board members, "*act as a conduit to disseminate ideas and innovation.*" It is essential to point out that, with board diversity, board members are in a stronger position to seek more unique information from their network ties (Gabaldon et al., 2018; Kakabadse et al., 2018; Shropshire, 2010; Zhang, 2010). Additionally, diverse boards act as a catalyst for strategic change and contribute to the adaptation of new processes, technologies, and practices (Haynes & Hillman, 2010).

Following the above discussions, we expect knowledge-based diversity to enhance information depth and the number of perspectives in the group information-processing space, facilitating firm process innovation. Similarly, gender-based diversity could also positively influence process innovation through social networks of female directors that bring in non-redundant knowledge from other firms (Hillman et al., 2002;), supporting innovative activities in focal firms (Torchia et al., 2018). However, there may be differences between gender-based diversity and knowledge-based diversity (Harrison & Klein, 2007).

Gender is a salient observable among upper echelons, and it can quickly create a social categorization divide between female and male directors (Zhu et al., 2014). Studies have shown that female directors influence not only the type of issues to be considered but also how they are discussed (Eagly, 2016; Huse & Solberg, 2006; Terjesen et al., 2009). For example, females may be more willing to bring new problems and perspectives to the table, start lively debates, and broaden the content of boardroom discussions. They are sensitive to the interests of others and usually consider the views of multiple parties. They may also ask critical questions of their peers or managers and pursue answers. It is noticed that, in an attempt to help to break the glass ceiling, female directors can be impatient to take leading roles or influence board decisions to show that women in the upper echelons have it all (Singh et al., 2008). The behaviors and attitudes of female directors, therefore, can be perceived as challenges or threats to the male directors' status and prestige, or to existing norms and distinctiveness of the dominant male group (Kakabadse et al., 2015). It can create friction between female and male directors, which results in a longer decision-making time in the boardroom.

Female directors, as a result, may face adverse consequences of social categorization from their male colleagues. In a male-dominated board, information from female directors about firm process innovation is likely to be overlooked or not taken seriously when the male directors perceive them as a threat. Male directors may thus neglect female directors' involvement, creating unshared information in the group information-processing space. Under such a circumstance, the CEM predicts that the elaboration of task-related information cannot reach its potential. On the other hand, knowledge-based diversity does not seem to lead to a similarly intense social categorization between female and male directors.

Hypothesis 1: Board knowledge-based diversity has a stronger impact on process innovation than gender-based diversity.

Impact of Chairperson's Gender

A board chairperson is an official leader in the board and positively influences how directors share and exchange information in board meetings and discussions (Kakabadse et al., 2015; Pugliese et al., 2015). Studies show that the board chairperson contributes to the creation of a safe (Guerrero et al., 2015) and an open (Gabrielsson et al., 2007) atmosphere in board meetings, enabling the elaboration of task-related information and exchange of different views, benefiting a board's contributions to strategy process (Kakabadse et al., 2015; Zhang, 2010).

There is evidence suggesting that gender may explain why some chairpersons make better use of directors' diverse talent than others (Eagly, 2016; Nielsen & Huse, 2010). For example, studies of gender have shown that female chairpersons tend to display a collaborative leadership style (Eagly, 2007, 2016; Huse & Solberg, 2006). These studies suggest that female leaders demonstrate a strong focus on listening, an inclination to gain social support, a desire to achieve a win-win solution, an attitude to bring in new perspectives, and a willingness to engage in various discussions. Indeed, such a leadership style underlies a general finding of women in the workforce. Women are more sensitive to other people's interests than their male colleagues, and they are more likely to consider multiple perspectives (Terjesen et al., 2009). The collaborative leadership style feature even discusses how a female chairperson is better than a male chairperson when the utilization of directors' unique knowledge and perspectives becomes the concern (Brodbeck et al., 2007; Nielsen & Huse, 2010). As process innovation calls for informational diversity, a female chairperson can take advantage of knowledge-based diversity and gender diversity through her collaborative leadership style.

Nevertheless, when social categorization is considered, CEM suggests that a chairperson's influence might be complicated, where a female chairperson's attention could differ between knowledge-based diversity and gender diversity. The unequal attention could affect the relationship between gender diversity and process innovation as well as knowledge-based diversity and process innovation. We present two arguments below.

First, women directors have long faced struggles in the corporate boardroom, such as unfavorable social status (Huse & Solberg, 2006; Kakabadse et al., 2015). Support in the form of legalizing a gender quota system, for example, is called on to help women break away from these obstacles. However, even with the progress made by implementing a gender quota system, female directors' involvement is still limited (Labelle et al., 2015). It seems that women directors need additional assistance beyond the quota system. A female chairperson might be considered helpful for women directors. A female chairperson could systematically divert more attention to women directors, helping them break away from the weak and out-group members (Eagly, 2007; Kakabadse et al., 2015; Nielsen & Huse, 2010; Tate & Yang, 2015). As such, in a board chaired by a female, there could be more exchange of gender-related issues than task-related issues. Everything else equal, a female chairperson is thus likely to favor the value-creation from gender diversity more than knowledge-based diversity.

Second, from an operational perspective, a female chairperson can limit potential resentment from the male directors when the chairperson helps women directors advance in the boardroom (Kakabadse et al., 2006, 2015). As mentioned earlier, when women are aware of the adverse impact of social categorization, they become more sensitive to these issues than their male colleagues (Terjesen et al., 2009). It is reflected in their actions of mentoring. Researchers have shown that a female chairperson is careful in selecting and implementing mentoring practices, which help women directors interact with their male counterparts without rocking the boat (Kakabadse et al., 2015). Under a female chairperson's leadership, women directors may "sit back, observe others, and learn. Once [they] understand the issues, question [the male directors] in the contextually acceptable manner" to limit unwanted personal conflicts (Kakabadse et al., 2015, p. 273). The practices could reduce hostilities against women directors and create an open and safe boardroom atmosphere for both female and male directors (Kakabadse et al., 2006). Consequently, mentoring practices encourage sharing unique information, increase the communication of diverse views, and promote the value-creation from gender diversity.

To conclude, female chairpersons could become sympathetic towards female directors, and they may favor a greater gender diversity when male directors dominate a typical board. As a result, female chairpersons could focus more on tapping the benefits of gender diversity to improve firm process innovation. It leads to the following hypothesis:

Hypothesis 2a: Gender diversity will have a greater positive impact on process innovation when the board is led by a female chairperson.

Male chairpersons, on the other hand, may not be equally enthusiastic about gender-related issues. They may not prioritize the active participation and contribution of women directors in board discussions. In particular, when value differences surface between male and female directors, male directors may perceive threats from the participation of women directors (Branscombe et al., 1999). In this case, a male chairperson could downplay the importance of gender diversity and focus more on knowledge-based diversity instead. As a result, we have the following hypothesis:

Hypothesis 2b: Knowledge-based diversity will have a greater positive impact on process innovation when the board is led by a male chairperson.

METHOD AND ANALYSIS

Data

We have used the survey method in this study. We obtained survey data from a public survey database, the Value Creating Board from Norway, collected in 2005 (Sellevoll et al., 2007). There are several benefits of using a public database, such as the Value Creating Board. First, the data collection process is rigorous, providing quality data in statistical analysis (Huse, 2018; Huse et al., 2011; Sellevoll et al., 2007). Publications using this database have covered a broad area such as board knowledge, power, trust, and board task performance (Huse et al., 2011; Zhang, 2010). Second, the database contains information on board diversity and firm process innovation, which are the key variables of our investigation. In particular, the gender issue during the data collection period had received great attention, resulting in a 40% female director quota system to be enforced shortly after (Sweigart, 2012). Therefore, data collected during the period provides rich information to examine the mechanism of social categorization due to gender. Third, the development of the boards of directors in Norway has been greatly influenced by corporate governance systems and practices of other countries such as the U.S.A. and the U.K. (Huse, 2007; Oxelheim & Randøy, 2003). Studies that have used Norwegian governance data may thus contribute to interesting discussions for the international audience (Oxelheim & Randøy, 2003).

2,954 CEOs participated in the survey with a response rate of 33%. For this study, companies selected are publicly listed firms and private firms with the number of employees at or more than 50. Among these firms, 42% are in manufacturing, 30% in service, 6% in banking and finance, and the rest in other industries. We have excluded data without information on firm process innovation and board chairperson gender. The final size of the data is 462. We have not found selection bias when checking the data against sales revenue. The descriptive statistics and data correlations are presented in Table 1.

In Table 1, 60% of the boards have women directors. The average size of the board is six, and the mean of women directors is one. Eight percent of all cases have a female board chairperson. The data suggests that male directors dominated the Norwegian boards at the time of data collection. The correlation between board size and the number of female directors is significant at 0.61. The correlation between board size and gender-based diversity (Blau Index, 1977) is also significantly positive. It indicates the women directors are more likely to serve on a big board. The correlation between chairperson tenure and the board size is significant but negative at -0.17, suggesting the bigger the board size, the newer the board chairperson. Past product innovation and past technology innovation are positively correlated with each other at 0.44.

Constructs and Variables

We have applied Structural Equation Modeling (SEM) in the analysis. SEM is a powerful tool that combines Explanatory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) and simultaneously runs linear equations between observed variables and latent constructs, and among latent constructs themselves (Nunnally & Bernstein, 1994). Constructs are measured by survey items, where a 7-point Likert scale is used to measure the perception. A short description of the constructs and variables are presented below.

Knowledge-based diversity. This is a three-item construct measuring board knowledge-based diversity. The items capture the perception of CEOs regarding the degree of variation of their board members in 1) functional knowledge, 2) industry knowledge, and 3) educational background.

Gender-based diversity. We approach gender-based diversity by using the Blau Index (1977), and it is calculated as 1 – ΣPi^2, where Pi is the proportion of the group members in the category i. The value of the Blau Index of gender diversity varies between 0 and 0.5. The higher the number, the more diverse the board is. The mean value of the Blau Index for this study is 0.21.

Firm process innovation. This construct measures the perception of respondents regarding the degree of the process innovation of the firm. It includes three items: 1) the firm invests heavily in R&D to create innovative process technology; 2) The firm is the first in the industry to develop and introduce completely new process technology; 3) The firm is the pioneer in creating new process technology.

TABLE 1. DESCRIPTIVE STATISTICS AND CORRELATION

	Minimum Statistic	Maximum Statistic	Mean Statistic	Std Error									
Number of employees	50	25,000	636	103.66	1								
Revenues ($ million)	0.67	10,000	181	172.88	-.05	1							
Board size	3	14	6	.08	-.07	.33**	1						
No. of female directors	0	6	1	.05	-.04	.34**	.61**	1					
Blau Index	0	.5	.23	.00	.001	.18**	.39**	.87**	1				
Board chair tenure	0	32	5.10	.22	-.18**	-.07	-.17**	-.07	-.03	1			
Age of chairperson	25	82	54	.39	-.08	.01	.07	.01	.001	.29**	1		
Past technology innovation	1	7	4.31	.08	-.06	.09	.08	.07	10*	-.05	.02	1	
Past product innovation	1	7	4.62	.07	-.01	.08	.05	.03	.06	-.13**	.03	.44**	1

Number of cases: 426
**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

In this study, we control board chair tenure, which is likely to influence a diverse board (Buchanan, 1974). Board chairperson tenure may signal a degree of complexity in the understanding of the board's work, which may influence how effectively the board chairperson handles gender-related and task-related issues. We also control board size, which is positively associated with firm process innovation (Zahra et al., 2000). Last, we control the degree of product innovation and technology innovation in the past three years. Studies have found a strong complementary effect between technology innovation, product innovation, and process innovation (Adner & Levinthal, 2001; Martinez-Ros, 2000). We thus expect that technology innovation and product innovation in the past can influence process innovation.

Measurement of Constructs

The results of the two constructs—knowledge-based diversity and firm process innovation—are strong (see Table 2 below). The values of five indices for the measurement model fit are presented in 1A of the Appendix. The five indices are also used in the assessment of the structural model fit later on. The index CMIN/df measures the minimum discrepancy, and a value between 1 and 3 indicates an acceptable fit (Carmines & McIver, 1981). The other four indices also suggest a good fit. The Comparative Fit Index (CFI) has a value higher than 0.95 (Joreskog & Sorbom, 2009). Root Mean Square Error of Approximation (RMSEA) is lower than 0.06 (MacCallum et al., 1996) or 0.08 (Diamantopoulos & Siguaw, 2000). PCLOSE (p of Close Fit) is higher than 0.05 (Hair et al., 2010). Standardized Root Mean Square Residual (SRMR) is lower than 0.08 (Hair et al., 2010). All indices have met the requirement of a good model fit.

TABLE 2. CONFIRMATORY FACTORY ANALYSIS

Items (N=462)	Factor loading
Firm process innovation: Cronbach's α=0.87, CR=0.87, AVE=0.68, MSV=0.25	
The firm invests heavily in the R&D to create innovative process technology.	0.81
The firm is the first in the industry to develop and introduce completely new technology.	0.84
The firm is the pioneer in creating new process technology.	0.84
Board diversity: Cronbach's α=0.76, CR=0.74, AVE=0.49, MSV=0.04	
There is a high level of board diversity regarding: – functional knowledge	0.75
– industry knowledge	0.64
– educational background	0.70

It is vital to check the reliability and validity of the constructs. The reliability of a construct measures the consistency of items in a construct. Factor loadings suggest the item consistency in Table 2, where all loadings are higher than 0.60 (Bollen, 1989). Also, Cronbach's Alpha is 0.87 for firm process innovation and 0.76 for knowledge-based diversity, implying no reliability concerns (Bollen, 1989). Last, Composite Reliability (CR) that considers an items' errors is another often-used assessment, with a value above 0.7 as the threshold (Hair et al., 2010). We have no reliability concerns with the two constructs.

The validity of constructs measures the degree of the quality of the measurement. The Norwegian database has ensured the validity of constructs by using items that have been theoretically suggested and empirically tested (Sellevoll et al., 2007). We have applied additional statistics to check the validity (see Table 2): Average Variance Extracted (AVE) and Maximally Shared Variance (MSV). A desirable validity is obtained when the AVE is higher than 0.50, and the MSV is lower than the AVE (Hair et al., 2010). The construct of knowledge-based diversity has an AVE at 0.49, which is slightly smaller than 0.50. After a careful evaluation, the less desirable AVE of knowledge-based diversity is acceptable for two reasons. First, this construct has met the reliability measurement threshold, where CR is higher than 0.70, and Cronbach's Alpha is also higher than 0.70. Second, the discriminant validity has met its requirement, where the MSV is lower than the AVE (Hair et al., 2010). We, therefore, conclude that the validity of knowledge-based diversity is not optimal but acceptable.

Analysis and Results

We have tested two structural models in the analysis. One is the basic model, and the other is the model addressing the effect of board chairperson gender (see Table 3). Standardized coefficients are reported.

TABLE 3. STRUCTURAL MODEL RESULTS

	Structural Model 1		Structural Model 2			
	Whole dataset		Male Chair		Female Chair	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Board size	0.04	0.03	0.01	0.04	0.29*	0.11
Board chair tenure	0.03	0.01	0.02	0.01	0.08	0.07
Past technology innovation	0.15***	0.04	0.17***	0.04	0.13	0.17
Past product innovation	0.35***	0.04	0.37***	0.04	0.09	0.18
Knowledge-based diversity	0.14***	0.07	0.15***	0.07	0.13	0.22
Gender-based diversity	-0.05	0.32	-0.04	0.33	0.29*	3.18
R^2	0.22		0.24		0.24	
No. of cases	462		428		34	
Model fit indices	CMIN/DF=2.25 CFI=0.95 RMSEA=0.05 PCLOSE=0.41 SRMR=0.06		CMIN/DF=1.45 CFI=0.97 RMSEA=0.03 PCLOSE=0.92 SRMR=0.06			

Notes: *** p-value < 0.01; ** p-value < 0.05; * p-value < 0.10

In Table 3, all five model fit indices satisfy the threshold requirement for a good model fit. The strong model fit has considered the potential correlations between independent and control variables, as suggested by AMOS. They include the correlation between board size and gender-based diversity, past technology innovation and past product innovation, and board chairperson tenure and board size. These three correlations are consistent with the earlier descriptive statistics and data correlations in Table 1.

Hypothesis 1 states that knowledge-based diversity has a stronger impact on process innovation than gender-based diversity. In Table 3, the standardized coefficient of knowledge-based diversity on process innovation is significant and positive at 0.14, and while the coefficient of gender-based diversity is non-significant at -0.05. Hypothesis 1 receives support.

Hypothesis 2a proposes that the impact of gender diversity on process innovation is more pronounced than that of knowledge-based diversity under a female chairperson. The standardized coefficient of gender-based diversity is positive and significant at 0.29, and while the coefficient of knowledge-based diversity is positive at 0.13 but non-significant. We could conclude that gender-based diversity has a stronger effect on process innovation than knowledge-based diversity under a female chairperson. Hypothesis 2a receives support.

Hypothesis 2b proposes that knowledge-based diversity is more important than gender-based diversity influencing process innovation under a male board chairperson. In Table 3, the coefficient of knowledge-based diversity is significant and positive at 0.15. The coefficient of gender-based diversity is -0.04 and non-significant, suggesting gender diversity has little, if any, impact on process innovation. The result supports Hypothesis 2b.

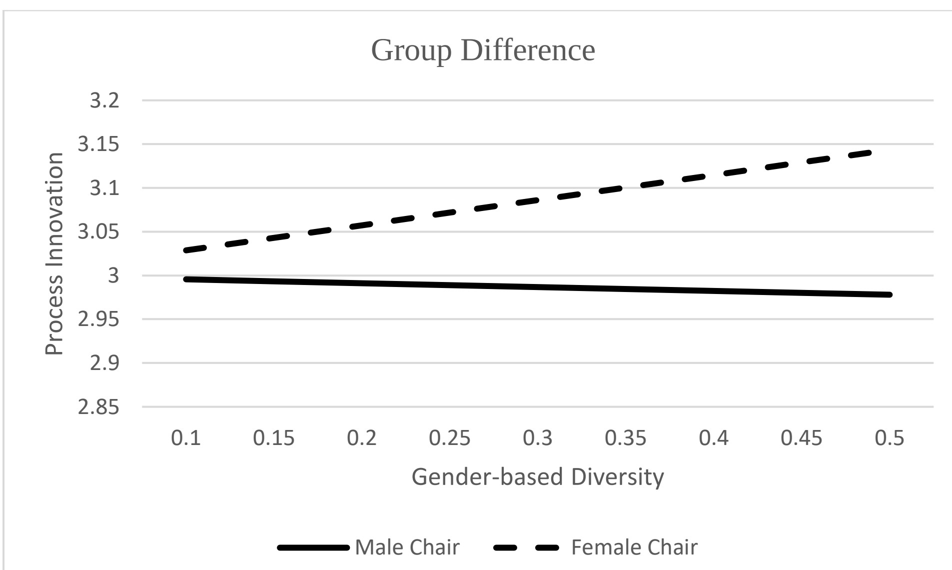

FIGURE 2. GENDER-BASED DIVERSITY ON PROCESS INNOVATION

There is a cautionary note that the subset we used to examine the relationship between board diversity and process innovation under the female board chairperson's impact has limited observations of 34. Albeit a ratio of 10:1 between sample size and the number of variables is generally used to evaluate the adequacy of the sample size for SEM analysis (Bentler & Chou, 1987), there are exceptions for the sample size as small as 30. That is when the model is simple with one or two constructs, and the result is robust with high factor loadings (Wolf et al., 2013). It fits our study, where the model consists of two constructs—board knowledge-diversity and firm process innovation—with a strong model fit (see Table 3). Overall, we may conclude that we have found support for all hypotheses.

DISCUSSION AND IMPLICATIONS

The study applies the categorization-elaboration model (CEM) to examine board diversity's impact on firm process innovation. We hypothesize and have found that knowledge-based diversity has a more significant impact on firm process innovation than gender-based diversity (H1). The conclusion is consistent with the perspective that we should differentiate the dimensions of group diversity (Guillaume et al., 2017; Harrison & Klein, 2007; Homan et al., 2007; Milliken & Martins, 1996). The study results are also in line with board leadership research, where chairpersons can influence the board effectiveness by activating the use of directors' knowledge and skills (e.g., Guerrero et al., 2015; Kakabadse et al., 2015; Zhang, 2010). Specifically, a potential difference exists between female and male chairpersons such that the type of diversity—knowledge-based diversity and gender-based diversity—elicit different responses to process innovation (H2a and H2b). Overall, the study results suggest that it is relevant to investigate gender issues to understand board diversity's effectiveness (Palvia et al., 2015; Nekhili et al., 2016).

Implications

Researchers of group diversity encourage us to investigate barriers in group work such as out-group categorization (Sun et al., 2015), tokenism (e.g., Torchia et al., 2011), perceptions of unequal membership (e.g., Nielsen & Huse, 2010), power asymmetry in decision-making (Haynes & Hillman, 2010; Triana et al., 2013), and cognitive biases in a group (Westphal & Bednar, 2005). CEM equips us with two theoretical approaches to this call: information processing and social categorization, inviting us to explore factors that facilitate and hinder a diverse group. Therefore, the study's theoretical implication suggests that CEM could be seen as a novel approach to shed light on the directors' dynamic relationship.

There are two implications for practitioners. First, the study result highlights the importance of selecting directors based on their diverse knowledge, skills, and experience, as suggested by hypothesis 1. Second, the study result strongly supports that the chairperson's gender matters (Cheng & Groysberg, 2020; Kakabadse et al., 2015). Everything else equal, chairpersons can interact with the social categorization processes due to their gender, potentially changing a board dynamic. In our study, under a female chairperson, gender diversity receives more attention to improve process innovation, and while under a male chairperson, gender diversity receives less attention. As a result, practitioners should be aware of the social barriers due to gender differences.

To limit the hindrance caused by gender, we propose practitioners pay attention to the priority of the type of board diversity, such that they can take coherent actions to meet the need. For example, suppose gender diversity becomes the focus of discussion in a male-dominated board. In that case, selecting a female chairperson is more likely to facilitate practices encouraging gender-diversity.

Limitations and Future Research

There are three limitations to this paper, which also form potential opportunities for future studies. First, the complexity of gender is regarded as a 'double-edged sword' (e.g., Palvia et al., 2015; Nekhili et al., 2016; Triana et al., 2013). It could create both positive and negative influences. In this paper, we have only argued that a female chairperson is likely to create a positive influence on firm process innovation, leaving the potential negative influence to be examined. For example, one study has suggested that female leaders can exacerbate emerging problems between in- and out-groups, reduce the team effectiveness, and ultimately produce an adverse effect on firm performance (Zhang et al., 2015). Therefore, it is interesting to investigate circumstances, such as specific favoritism that a female chairperson shows to women directors, which become detrimental for the firm process innovative.

Second, we have only analyzed social categorization due to a board chairperson's gender. Other demographic observables could also be interesting to explore. For example, researchers have shown the impact of age diversity of boards of directors on the performance of banks (Talavera et al., 2018). Future studies can thus examine how age diversity shapes social categorization between younger and senior board directors, and how the difference further affects firm process innovation.

Third, we have used data collected when hiring female directors started to gain considerable attention. Today, most boards participating in the survey may have more women directors because of the gender quota system (Nielsen & Huse, 2010). The situation could make studying gender issues in boards more relevant when many question the value of increasing female participation (Kanadli, et al., 2018a). If their appointment to the board has less to do with experiences and skills than the law requirement, female directors may face a profound hindrance due to the social-categorization processes (Kakabadse et al., 2015; Matsa & Miller, 2013; Nielsen & Huse, 2010). Investigating contingencies that fully utilize gender diversity amid the impediment may be a fruitful research avenue.

REFERENCES

Åberg, C., & Shen, W. (2020). Can board leadership contribute to board dynamic managerial capabilities? An empirical exploration among Norwegian firms. *Journal of Management and Governance*, *24*(2), 169-197.

Adams, R. B., Haan, J., Terjesen, S., & Van Ees, H. (2015). Board diversity: Moving the field forward. *Corporate Governance: An International Review*, *23*(2), 77-82.

Adner, R., & Levinthal, D. (2001). Demand heterogeneity and technology evolution: Implications for product and process innovation. *Management Science, 47*(5), 611–628.
Balliet, D., Wu, J., & De Dreu, C. K. W. (2010). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin, 140*(6), 1556–1581.
Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural equation modeling. *Sociological Methods and Research, 16,* 78-117.
Blau, P. M. (1977). *Inequality and heterogeneity*. Free Press.
Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods and Research, 17*(3), 303-316.
Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content*. Wiley.
Brodbeck, F.C., Kerschreiter, R., & Schulz-Hardt, A. M. S. (2007). Group decision making under conditions of distributed knowledge: The information asymmetries model. *Academy of Management Review, 32,* 459-479.
Buchanan, B. (1974). Building organizational commitment: The socialization of managers in work organizations. *Administrative Science Quarterly, 19*(4), 533-546.
Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobserved variables: Analysis of covariance structures. In G. W. Bohrnstedt & E.F. Borgatta, (Eds.), *Social measurement: Current issues*. Sage Publications.
Cheng, J. Y., & Groysberg, B. (2020). Gender diversity at the board level can mean innovation success. *MIT Sloan Management Review, 61*(12), 1-8.
Crossan, M. M., & Apaydin, M. (2010). A multi-dimensional framework of organizational innovation: A systematic review of the literature. *Journal of Management Studies, 47*(6), 1154-1191.
De Dreu, C. K. W., Nijstad, B. A., & Van Knippenberg, D. (2008). Motivated information processing in group judgment and decision making. *Personality and Social Psychology Review, 12,* 22-49.
Diamantopoulos, A., & Siguaw, J. A. (2000). *Introducing LISREL*. Sage Publications.
Eagly, A. H. (2007). Female leadership advantage and disadvantage: Resolving the contradictions. *Psychology of Women Quarterly, 31*(1), 1–12.
Eagly, A. H. (2016). When passionate advocates meet research on diversity, does the honest broker stand a chance? *Journal of Social Issues, 72*(1), 199-222.
Forbes, D. P., & Milliken, F. J. (1999). Cognition and corporate governance: Understanding boards of directors as strategic decision-making groups. *Academy of Management Review, 24*, 489-505.
Gabaldon, P., Kanadlı, S. B., & Bankewitz, M. (2018). How does job-related diversity affect boards' strategic participation? An information-processing approach. *Long Range Planning, 51*(6), 937-952.
Gabrielsson, J. (2007). Boards of directors and entrepreneurial posture in mediumsize companies: Putting the board demography approach to a test. *International Small Business Journal, 25*(5), 511-537.
Galaskiewicz, J., & Wasserman, S. (1989). Mimetic processes within an interorganizational field: An empirical test. *Administrative Science Quarterly, 34*(3), 454-479.
Galia, F., & Zenou, E. (2012). Board composition and forms of innovation: Does diversity make a difference? *European Journal of International Management, 6*(6), 630-650.

Guerrero, S., Lapalme, M. È., & Séguin, M. (2015). Board chair authentic leadership and nonexecutives' motivation and commitment. *Journal of Leadership & Organizational Studies*, *22*(1), 88-101.

Guillaume, Y. R. F., Dawson, J. F., Otaye-Ebede, L., Woods, S. A., & West, M. A. (2017) Harnessing demographic differences in organizations: What moderates the effects of workplace diversity? *Journal of Organizational Behavior*, *38*(2), 276-303.

Hair, J. J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: a global perspective* (7th ed.). Pearson Education.

Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, and disparity in the organization. *Academy of Management Review*, *32*, 1199-1228.

Haynes, K. T., & Hillman, A. (2010). The effect of board capital and CEO power on strategic change. *Strategic Management Journal, 31*, 1145-1163.

Hillman, A. J., Cannella, A. A., & Harris, I. C. (2002). Women and minorities in the boardroom: How do directors differ? *Journal of Management*, *28*, 747–763.

Hillman, A. J., & Dalziel, T. (2003). Boards of directors and firm performance: Integrating agency and resource dependence perspectives. *Academy of Management Review, 28*(3), 383-396.

Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, *121*, 43–64

Homan, A. C., Van Knippenberg, D., Van Kleef, G. A., & De Dreu, C. K. W. (2007). Interacting dimensions of diversity: Cross-categorization and functioning diverse work groups. *Group Dynamics: Theory, Research, and Practice, 11*(2), 79-94.

House, M. (2007). *Boards, governance and value creation: The human side of corporate governance*. Cambridge University Press.

Huse, M., (2018). *Value-creating boards: Challenges for future practice and research.* Cambridge University Press.

Huse, M., & Grethe Solberg, A. (2006). Gender-related boardroom dynamics: How Scandinavian women make and can make contributions on corporate boards. *Women in Management Review, 21*(2), 113-130.

Huse, M., Hoskisson, R., Zattoni, A., & Viganò, R. (2011). New perspectives on board research: Changing the research agenda. *Journal of Management & Governance*, *15*(1), 5-28.

Joreskog, K., & Sorbom, D. (2009). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Taylor & Francis.

Kakabadse, A., Kakabadse, N. K., & Barratt, R. (2006). Chairman and chief executive officer (CEO): That sacred and secret relationship. *Journal of Management Development, 25*(2), 134-150.

Kakabadse, N. K., Figueira, C., Nicolopoulou, K., Hong Yang, J., Kakabadse, A. P., & Özbilgin, M. F. (2015). Gender diversity and board performance: Women's experiences and perspectives. *Human Resource Management, 54*(2), 265-281.

Kakabadse, A., Goyal, R., & Kakabadse, N. (2018). Value-creating boards—diversity and evolved processes. *Journal of Creating Value, 4*(1), 22-41.

Kanadlı, S. B., Torchia, M., & Gabaldon, P. (2018a). Increasing women's contribution on board decision making: The importance of chairperson leadership efficacy and board openness. *European Management Journal, 36*(1), 91-104.

Kanadlı, S. B., Bankewitz, M., & Zhang, P. (2018b). Job-related diversity: The comprehensiveness and speed of board decision-making processes—an upper echelons approach. *Journal of Management and Governance*, *22*(2), 427-456.

Kanadlı, S. B., & Kakabadse, N. (2019). Diversity of Boards. In S. Idowu, R. Schmidpeter, N. Capaldi, & L. Zu (Eds.), *Encyclopedia of sustainable management*. Springer.

Labelle, R., Francoeur, C., & Lakhal, F. (2015). To regulate or not to regulate? Early evidence on the means used around the world to promote gender diversity in the boardroom. *Gender, Work & Organization, 22*(4), 339-363.

Martinez-Ros, E. (2000). Explaining the decisions to carry out product and process innovations: The Spanish case. *Journal of High Technology Management Research*, *10*(2), 223-242.

Matsa, D. A., & Miller, A. R. (2013). A female style in corporate leadership? Evidence from quotas. *American Economic Journal: Applied Economics, 5*(3), 136–169.

Milliken, F. J., & Martins, L. L. (1996). Searching for common threads: Understanding the multiple effects of diversity in organizational groups. *Academy of Management Review*, *21*(2), 402-433.

Nekhili, M., Chakroun, H., & Chtioui, T. (2016). Women's leadership and firm performance: Family versus nonfamily firms. *Journal of Business Ethics*, *153*(2), 291-316.

Nielsen, S., & Huse, M. (2010). Women directors' contribution to board decision-making and strategic involvement: The role of equality perception. *European Management Review, 7*, 16-29.

Nunnally, J. C., & Bernstein, I. H. (1994). *Pscyometric theory*. McGraw-Hill.

Oxelheim, L., & Randøy, T. (2003). The impact of foreign board membership on firm value. *Journal of Banking and Finance, 27*(12), 2369-2392.

Page, S.E. (2007). Making the difference: applying a logic of diversity. *Academy of Management Perspectives, 21*(4), 6–21.

Palvia, A., Vähämaa, E., & Vähämaa, S. (2015). Are female CEOs and chairwomen more conservative and risk averse? Evidence from the banking industry during the financial crisis. *Journal of Business Ethics, 131*(3), 577–594.

Pugliese, A., Nicholson, G., & Bezemer, P. J. (2015). An observational analysis of the impact of board dynamics and directors' participation on perceived board effectiveness. *British Journal of Management*, *26*(1), 1–25.

Reichstein, T., & Salter, A. (2006). Investigating the sources of process innovation among UK manufacturing firms. *Industrial and Corporate Change*, *15*(4), 653-682.

Scholten, L., Van Knippenberg, D., Nijstad, B. A., & De Dreu, C. K. W. (2007). Motivated information processing and group decision-making: Effects of process accountability on information processing and decision quality. *Journal of Experimental Social Psychology*, *43*, 539-552.

Sellevoll, T., Huse, M., & Hansen, C. (2007). *The value creating board: Results from the 'follow-up surveys' 2005–2006 in Norwegian firms (*Research Report no. 2-2007). Norwegian School of Management.

Shropshire, C. (2010). The role of the interlocking director and board receptivity in the diffusion of practices. *Academy of Management Review*, *35*(2), 246-264.

Singh, V., Terjesen, S., & Vinnicombe, S. (2008). Newly appointed directors in the boardroom: How do women and men differ? *European Management Journal*, *26*(1), 48-58.

Sun, S. L., Zhu, J., & Ye, K. (2015). Board openness during an economic crisis. *Journal of Business Ethics, 129*, 363–377.

Sweigart, A. (2012) Women on board for change: The Norway model of boardroom quotas as a tool for progress in the United States and Canada, *Northwestern Journal of International Law and Business*, *32*(4):81A-105A.

Talavera, O., Yin, S., & Zhang, M. (2018). Age diversity, directors' personal values, and bank performance. *International Review of Financial Analysis, 55*(C), 60-79.

Tanford, S., & Penrod, S. (1984). Social influence model: A formal integration of research on majority and minority influence processes. *Psychological Bulletin, 95*, 189-225.

Tate, G., & Yang, L. (2015). Female leadership and gender equity: Evidence from plant closure. *Journal of Financial Economics, 117*(1), 77-97.

Terjesen, S., Sealy, R., & Singh, V. (2009). Women directors on corporate boards: A review and research agenda. *Corporate Governance: An International Review, 17*(3), 320-337.

Torchia, M., Calabrò, A., Gabaldon, P., & Kanadlı, S. B. (2018). Women directors contribution to organizational innovation: A behavioral approach. *Scandinavian Journal of Management, 34*(2), 215-224.

Torchia, M., Calabrò, A., & Huse, M. (2011). Women directors on corporate boards: From tokenism to critical mass. *Journal of Business Ethics, 102*, 299–317.

Triana, M. D. C., Miller, T. L., & Trzebiatowski, T. M. (2013). The double-edged nature of board gender diversity: Diversity, firm performance, and the power of women directors as predictors of strategic change. *Organization Science, 25*, 609-632.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell Publishing.

Van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology, 89*, 1008-1022.

Van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology, 58*, 515-541.

Westphal, J. D., & Bednar, M. K. (2005). Pluralistic ignorance in corporate boards and firms' strategic persistence in response to low firm performance. *Administrative Science Quarterly, 50*(2), 262–298.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 76*(6), 913-934.

Zahra, S. A., Neubaum, D. O., & Huse, M. (2000). Entrepreneurship in medium-size companies: Exploring the effects of ownership and governance systems. *Journal of Management, 26*(5), 947-976.

Zhang, P. (2010). Board information and strategic tasks performance. *Corporate Governance: An International Review, 18*, 473-487.

Zhang, X. A., Li, N., Ullrich, J., & Van Dick, R. (2015). Getting everyone on board: The effect of differentiated transformational leadership by CEOs on top management team effectiveness and leader-rated firm performance. *Journal of Management, 41*(7), 1898-1933.

Zhu, D. H., Shen, W., & Hillman, A. J. (2014). Recategorization into the in-group the appointment of demographically different new directors and their subsequent positions on corporate boards. *Administrative Science Quarterly, 59*(2), 240-270.

APPENDIX

The goodness fit statistics of measurement model are presented in the following table

1A. MEASUREMENT MODEL FIT INDEXES

	Observed Value	Ideal Threshold
CMIN/df	2.02	Between 1 and 3
CFI	0.99	>0.95
RMSEA	0.04	<0.06
PCLOSE	0.50	>0.05
SRMR	0.03	<0.09

THE EFFECT OF ENTERPRISE RESOURCE PLANNING SYSTEM (ERP) IMPLEMENTATIONS ON THE PROPERTIES OF ANALYST FORECASTS

Debjeet Pradhan, Tarleton State University
pradhan@tarleton.edu

Joseph F. Brazel, North Carolina State University
jfbrazel@ncsu.edu

ABSTRACT

We know from earlier studies that the implementation of Enterprise Resource Planning (ERP) systems has allowed companies to process, prepare, and disseminate accounting information more quickly and accurately than the legacy accounting systems that they replaced. However, ERP systems also provide a greater opportunity for companies to manage earnings. In this study, we examine if ERP implementations have improved the information environment for equity analysts. Using implementation data from a major ERP software provider, we find that the magnitude of analyst forecast accuracy decreased and its standard deviation increased soon after ERP implementations. These results hold even when we control for discretionary accruals.

INTRODUCTION

Enterprise Resource Planning (ERP) systems are "information system packages that integrate information and information-based processes within and across functional areas in an organization" (Kumar & Hillegersberg, 2002). Frequently, ERP systems replaced ageing legacy systems, which were patchworks of disparate information technology (IT) systems with different operating systems and database architectures. Sharing of data between the IT systems was difficult, often requiring costly custom programming or middleware implementations. Functional departments acted like information "silos," resulting in excessive politicking and operational inefficiencies. The implementation of ERP systems helped alleviate some of these problems by integrating information across an entire organization through the use of a single authoritative database and automated workflows.

ERP systems gather information in a timely manner, help process accounting information efficiently (Davenport, 1998; Hitt et al., 2002) and provide a unified enterprise view of a firm's financial condition (Dillon, 1999). They also help eliminate information silos and allow managers unprecedented access to accounting information (O'Leary, 2000). Brazel and Dang (2008) show that ERP system implementations shorten reporting lags, between quarter-end and earnings release dates, for "good news" firms. Hayes et al. (2001) compare ERP adopters with ERP non-adopters and illustrate that ERP adopters exhibit improved operational performance. ERP systems also enable managers to provide more accurate management forecasts.

The above evidence points to an improved internal information environment for firms with ERP systems which likely creates expectations among external users of accounting information of improved operational and financial performance. Such expectations could raise the optimism among equity analysts regarding earnings forecasts. Consistent with this expectation, the market reacts positively to ERP implementation announcements (Hunton et al., 2003).

Analyst forecasts

Equity analysts play an important role in capital markets by providing earnings forecasts and buy or sell recommendations about the firms that they cover. They gather, analyze and release information, obtained from various sources, in the form of research reports or notes. Such reports or notes are frequently used by investors to make investment decisions. Prior studies have shown that the stock market does react to the information contained in analyst reports (e.g., Park & Stice, 2000).

A primary source of information for equity analyst is the firms themselves. Publicly traded U.S. firms are required to make certain disclosures about their business to the Securities and Exchange Commission, but they also voluntarily provide other information to capital market participants (e.g., analysts). A company with a better internal information environment (ERP adopters) can provide more accurate and useful information, relative to ERP non-adopters, thereby enabling analysts to make better earnings forecasts and stock recommendations.

However, ERP systems also provide managers with powerful tools to test out the effect of different discretionary accrual scenarios and thereby manage earnings (Brazel & Dang, 2008). In addition, there is evidence to suggest that there is a reduction in internal control effectiveness and audit quality subsequent to ERP implementations (Wright & Wright, 2002; Brazel & Agoglia, 2007), which may lead managers to engage in opportunistic behavior to, for example, issue bad news management earnings forecasts around stock option award periods to temporarily depress stock prices (Aboody & Kasznik, 2000), or issue overly optimistic forecasts around secondary equity offerings to ensure higher prices for new shares (Rogers & Stocken, 2005). Such opportunistic behavior by managers may not be entirely transparent to external observers. To the extent that analysts cannot entirely unravel manipulations in reported earnings, earnings management can make it more difficult for analysts to issue accurate forecasts.

This study examines whether the implementation of ERP systems by firms helps them improve the information environment for equity analysts, such that earnings forecasts for firms became more accurate after ERP implementation, relative to prior periods. The information environment consists of publicly available information, including quarterly and annual financial statements, regulatory filings with the Securities and Exchange Commission, macroeconomic reports, and others, which generally tend to be based on historical data. It also includes privately generated information from interviews, surveys, macroeconomic forecasts, and other research conducted by the analysts themselves. Firms can greatly enhance this environment by voluntarily releasing firm- and industry-specific information that is more up to date. To the extent that ERP tools help firms gather and process information efficiently, firms should be able to provide better guidance to financial analysts after ERP implementation, relative to prior periods.

RESEARCH DESIGN

Our research design relies on a time-series regression analysis using a dummy variable to distinguish time periods before and after ERP implementations. We employ this design in preference over cross-sectional matched sample tests to reduce the effect of measurement errors arising from a lack of ERP implementation data for firms outside our sample. Text based keyword searches are not a reliable method of identifying firms that do not have ERP systems in a given year. As more and more firms implement ERP systems over time, as the majority of firms did for competitive reasons, measurement error was expected to grow rapidly during the 1990s. To illustrate, by 1999, 70% of Fortune 1000 firms (most firms in our sample are large firms) had either adopted or were in the process of implementing ERP systems. As such, there would be a low likelihood of finding appropriate matched firms without ERP systems in the latter part of our sample period. For this reason, we avoid cross-sectional designs for our tests and instead rely on ordinary least square regression with a time-based dummy variable. The dummy variable equals 0 in time periods preceding ERP implementation, and 1 otherwise. The main advantage of this design is that firms in the sample serve as their own control groups.

Our main dependent variables are forecast accuracy and the standard deviation of analyst forecasts. Forecast accuracy is defined as the negative of the absolute value of analyst forecast error, deflated by beginning of period stock price.

$$\text{Forecast accuracy} = -\left(|EPS_t - AF_t|\right) / P_t$$

Where, EPS_t = Earnings per share in quarter t
AF_t = median analyst forecast of EPS_t for quarter t
P_t = Price per share in the beginning of quarter t

We define standard deviation of analyst forecasts as the inter-analyst standard deviation of EPS forecasts, deflated by stock price at the beginning of the quarter.

Our control variables are consistent with Lang and Lundholm (1996). Prior studies noted below have shown that these variables are associated with analyst following and firm disclosure policy, which in turn are associated with forecast accuracy and standard deviation of analyst forecasts.

Market Value	= the market value, in $ billions, of the firm's equity at the beginning of the quarter.
Std. Dev. Of ROE	= the historical standard deviation of return on equity (ROE) computed over the preceding ten quarters.
Return-Earnings Correlation	= the historical correlation between quarterly returns and earnings computed over the preceding ten quarters.

Earnings Surprise = the absolute value of the difference between the current quarter's earnings per share and last year's same quarter earnings per share, divided by the price at the beginning of the fiscal quarter.

Market value is a proxy for firm size. Larger firms have been shown to have larger analyst following (Bhushan, 1989; Brennan & Hughes, 1991), more extensive media coverage, and better disclosure policies (Waymire, 1986; Lang & Lundholm, 1993). Analyst forecast accuracy will be positively correlated with market value (Lang & Lundholm, 1996).

The standard deviation of ROE is a proxy for performance variability. There is weak evidence that analysts are more likely to follow firms with low performance variability (Lang & Lundholm, 1996), suggesting that firms with high performance variability have inferior disclosure environments. Analyst forecast accuracy will be weakly associated with performance variability (Lang & Lundholm, 1996).

The incentives for private information acquisition will be greater when the returns-earnings correlation is high because it is easier to forecast future stock price based on earnings forecasts (King, et al., 1990). However, counter to this intuition, Lang and Lundholm (1996) find a weak, negative association between analyst forecast accuracy and returns-earnings correlations.

Earnings surprise is included as a control variable because analyst forecast accuracy is likely to be affected by the magnitude of earnings information (Lang & Lundholm, 1996). For example, if a firm introduces a new product, then realized earnings are likely to deviate substantially from expected earnings. Earnings surprise will be negatively associated with analyst forecast accuracy.

SAMPLE SELECTION

We obtain our sample of ERP system implementations from a proprietary database supplied by a leading international provider of ERP systems. The database contains the names of firms that implemented the ERP system and the dates on which the system went live. We start with 315 unique firms with ticker symbols, CUSIP, and PERMNO in the database that implemented ERP systems between 1994 and 1999.

From Compustat we get 377,341 firm-quarter observations for the period 1991 through 2004. We chose this period because we needed at least 12 quarters of data before the first ERP implementation date and 12 quarters of data beyond the last ERP implementation date. We merged this data with returns data from CRSP and dropped observations with missing values. We were left with 252,866 firm-quarter observations. Then we combined this data with our ERP implementation data and dropped observations with missing values. That left us with 21,229 firm-quarter observations. Finally, we merged analyst forecast data from IBES and limited the observations to plus or minus 3 years from the ERP implementation dates. That left us with 1885 firm-quarter observations to estimate our regressions.

RESULTS

Table 1 shows descriptive statistics for the variables used in our regressions for the periods both before and after the completion of ERP implementation. Variable definitions are provided in Appendix 1. Forecast accuracy has been defined as the negative of the absolute forecast error, deflated by stock price, in order to ensure that a higher value indicates a more accurate forecast. A forecast accuracy of zero is the most accurate. A univariate comparison of mean forecast accuracy (Table 1, Panel A versus Panel B) demonstrates that analyst forecasts become less accurate in the period immediately following ERP implementations. However, a univariate analysis does not reliably explain the change in forecast accuracy because it may be driven by other variables, creating a correlated, omitted variable problem. Therefore, we carry out multivariate regressions to study the relation between our dependent and independent variables.

TABLE 1. DESCRIPTIVE STATISTICS

Panel A. After completion of ERP Implementation:

Variable	N	Mean	Std. Dev.	Median	Min.	Max.
Forecast accuracy	932	-0.3746	0.8333	-0.1023	-9.0275	0
Std. dev. fcast acc	883	0.0021	0.0039	0.0008	0	0.06161
Market value	929	18.911	55.901	16.160	-33.152	600.627
Std. dev. ROE	932	-0.4817	1.6680	-0.0786	-10.1727	0.5123
Ret-earn corr	932	-0.0313	0.2824	0.0058	-0.8828	0.7649
Earn surprise	932	-0.0183	0.0600	-0.0175	-0.5172	0.8921
Disc accrual	743	0.0021	0.0388	0.0012	-0.1713	0.3616

Panel B: Before completion of ERP Implementation:

Variable	N	Mean	Std. Dev.	Median	Min.	Max.
Forecast Accuracy	966	-0.2770	0.6171	-0.0821	-7.4189	0
Std. dev. fcast acc	915	0.0019	0.0071	0.0008	0	0.1959
Market Value	958	9.4897	22.5030	1.6288	-17.8490	166.0252
Std. dev. ROE	966	-0.1317	0.2640	-0.0887	-2.4199	0.3983
Ret-earn corr	966	-0.0349	0.2806	-0.0073	-0.8359	0.7280
Earn surprise	964	-0.0151	0.0372	-0.0122	-0.2837	0.4447
Disc accrual	691	-0.0028	0.0441	-0.0016	-0.5473	0.2413

Table 2 shows correlations between the dependent and independent variables used in this study. In our sample, forecast accuracy is positively correlated with firm size and the standard deviation of return on equity, but negatively correlated with the return-earnings correlation and earnings surprise in the period following ERP implementation. These results are generally consistent with those of Lang & Lundholm, (1996).

TABLE 2. PEARSON (ABOVE)/SPEARMAN (BELOW) CORRELATIONS

Panel A. After completion of ERP Implementation:

Variable	Forecast Accuracy	Std. dev. fcast acc	Market Value	Std. dev. ROE	Ret-earn corr	Earn surprise	Disc accrual
Forecast Accuracy	1	**-0.6229**	**0.1353**	**0.1498**	**-0.0685**	**-0.2752**	0.0551
Std. dev. fcast acc	**-0.5904**	1	**-0.1460**	-0.0498	**0.0935**	**0.1403**	-0.0553
Market Value	**0.3871**	**-0.3811**	1	**-0.2349**	-0.0089	**-0.0799**	0.0008
Std. dev. ROE	**-0.0929**	**0.1009**	**-0.1062**	1	**0.0976**	-0.0126	-0.0120
Ret-earn corr	-0.0531	**0.1045**	-0.0357	0.0248	1	**0.0605**	0.0219
Earn surprise	**-0.2077**	**0.1660**	**-0.2071**	**0.1263**	**0.1008**	1	-0.0097
Disc accrual	**0.0915**	**-0.0751**	0.0387	0.0182	0.0564	-0.0521	1

Panel B: Before completion of ERP Implementation:

Variable	Forecast Accuracy	Std. dev. fcast acc	Market Value	Std. dev. ROE	Ret-earn corr	Earn surprise	Disc accrual
Forecast Accuracy	1	**-0.3134**	**0.1145**	-0.0358	**-0.0995**	**-0.2371**	**0.1102**
Std. dev. fcast acc	**-0.5726**	1	**-0.0625**	0.0475	**0.0299**	**0.3230**	-0.0797
Market Value	**0.3136**	**-0.2760**	1	-0.0112	**-0.0884**	-0.0438	0.0108
Std. dev. ROE	-0.0353	0.0309	**-0.0913**	1	0.0497	0.0364	-0.0291
Ret-earn corr	**-0.1026**	**0.1589**	**-0.1475**	-0.0082	1	0.0506	-0.0192
Earn surprise	**-0.0996**	**0.1551**	**-0.1174**	**0.0865**	0.0523	1	-0.0423
Disc accrual	**0.0681**	-0.0517	0.0226	-0.0052	-0.0287	**-0.0893**	1

Bold: Significant at the 10% level or better

Our main results are shown in Table 3, where we present the results of a regression of analyst forecast accuracy on the ERP dummy and control variables.

$$\text{Forecast accuracy} = b_0 + b_1 \text{(ERP dummy)} + b_2 \text{ Lagged market value} + b_3 \text{ Standard deviation of ROE} + b_4 \text{ Return-earnings correlation} + b_5 \text{ Earnings surprise} + e \qquad (1)$$

A negative sign on the ERP dummy coefficient indicates that analyst forecast accuracy decreased after ERP implementation compared to periods immediately preceding the implementation. These results hold even when we control for discretionary accruals. Discretionary accruals are estimated as the absolute value of the residual from the modified Jones model (Dechow et al., 1995). We report t-statistics and p-values based on robust standard errors to control for firm clustering effects (Petersen, 2009).

TABLE 3. REGRESSION OF ANALYST FORECAST ACCURACY

	Panel 1			Panel 2		
Variable	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value
Intercept	**-0.3476**	-6.78	<.0001	**-0.3518**	-6.48	<.0001
ERP	**-0.1008**	-2.26	0.0266	**-0.0953**	-1.86	0.0665
Market Value	**0.0024**	3.27	0.0016	**0.0024**	3.24	0.0018
Std. dev. ROE	**0.0914**	1.91	0.0592	**0.0982**	1.77	0.0805
Ret-earn corr	**-0.1928**	-2.26	0.0265	**-0.2656**	-2.58	0.0117
Earn surprise	**-3.6434**	-3.16	0.0022	**-3.0660**	-3.6	0.0006
Disc. Accrual				**1.3762**	3.2	0.002
No. of observations	1885			1425		
R-sqr	0.1098			0.1034		

Standard errors are clustered by firm

Bold: Significant at the 10% level or better

Forecast accuracy is positively correlated with market value. This was to be expected because larger companies have better information environments, larger analyst following, and more accurate analyst forecasts. Analyst forecast accuracy is positively correlated with the standard deviation of ROE, contrary to the findings of Lang & Lundholm (1996) who found only a weak, negative relation between forecast accuracy performance variability.

Forecast accuracy is negatively correlated with the return-earnings correlation, consistent with Lang and Lundholm (1996), and suggesting that earnings do a relatively poor job of capturing value relevant information about the stock. Similarly, forecast accuracy is negatively correlated with earnings surprise, indicating that analysts find it harder to make accurate forecasts when the level of earnings is significantly different from the previous period. These results are consistent with Lang and Lundholm (1996).

In Panel 2 of Table 3, we control for the ability of companies to manage earnings by including a variable for discretionary accruals. Earnings management can help make analyst forecasts more accurate if it is used as a mechanism to smooth earnings and provide better management guidance. On the other hand, opportunistic behavior by management can make it more difficult for analysts to make accurate forecasts. In our sample, analyst forecast accuracy is positively correlated with discretionary accruals.

We also test the effect of ERP implementation on the standard deviation of analyst forecasts. In Table 4, we present the results of a regression of the standard deviation of analyst forecasts on the ERP dummy and control variables. The ERP dummy variable is positive and significant only when we control for discretionary accruals. The results suggest that the standard deviation of analyst forecasts increases in the period following ERP implementations. This result is consistent with the results in Table 3, as the standard deviation of analyst forecasts is expected to be negatively correlated with analyst forecast accuracy (see Table 2). As such, all coefficients in Table 4, Panel 2 have the opposite sign when compared to Table 3.

TABLE 4. REGRESSION OF STANDARD DEVIATION OF ANALYST FORECASTS

	Panel 1			Panel 2		
Variable	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value
Intercept	**0.0025**	4.72	<.0001	**0.0019**	7.84	<.0001
ERP	0.0004	1.58	0.1180	**0.0007**	2.25	0.0275
Market Value	**-0.00001**	-2.62	0.0104	**-0.00001**	-3.13	0.0025
Std. dev. ROE	-0.0002	-1.1	0.2746	-0.0003	-1.49	0.1407
Ret-earn corr	0.0008	1.13	0.2612	**0.0014**	2.26	0.0265
Earn surprise	**0.0335**	1.86	0.0671	**0.0132**	2.19	0.0315
Disc. Accrual				**-0.0051**	-3.23	0.0018
No. of observations	1785			1359		
R-sqr	0.06516			0.0770		

Standard errors are clusted by firm

Bold: Significant at the 10% level

ROBUSTNESS TESTS

We test the robustness of our results by arbitrarily shifting the pre-/post- ERP implementation (pseudo-event) window both backward by one year, and forward by one and two years. We expect to see no effect of ERP adoption across the pseudo-events. The results are shown in Tables 5 and 6. When the pre-/post- event window is shifted backward by one year, when ERP was yet to be implemented, the dummy variable is not statistically significant, as expected. However, when we shift the pre-/post- event window forward by one year, the dummy variable is negative and significant, suggesting that the effects of ERP adoption persist for a period of time after ERP implementation as firms transition to a new accounting environment. When we shift the pre-/post-event window forward by two years, the effects of ERP implementation go away. Collectively, these results suggest that the observed changes in analyst forecast accuracy is related to ERP implementation rather than some other uncontrolled events.

TABLE 5. ROBUSTNESS TEST: REGRESSION OF ANALYST FORECAST ACCURACY

	Pseudo event = 1 year earlier than actual ERP implementation			Pseudo event = 1 year later than actual ERP implementation			Pseudo event = 2 years later than actual ERP implementation		
Variable	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value
Intercept	**-0.405**	-7.2	<.0001	**-0.350**	-5.92	<.0001	**-0.4052**	-6.14	<.0001
ERP	0.003	0.07	0.9478	**-0.126**	-1.88	0.0644	-0.0401	-0.45	0.6554
Market Value	**0.002**	2.63	0.0102	**0.003**	2.89	0.005	**0.0023**	2.69	0.0087
Std. dev. ROE	0.142	1.43	0.158	0.080	1.49	0.1392	0.0676	1.4	0.1643
Ret-earn corr	-0.278	-3.37	0.0012	**-0.217**	-1.76	0.0824	-0.1843	-1.46	0.1486
Earn surprise	-4.304	-3.96	0.0002	**-2.525**	-2.85	0.0056	**-3.1193**	-2.26	0.0267
Disc. Accrual	0.221	0.59	0.5589	0.632	1.24	0.2203	-0.1161	-0.25	0.8028
No. of observations	1499			1398			1493		
R-sqr	0.09975			0.0741			0.08006		

Standard errors are clustered by firm

Bold: Significant at the 10% level or better

TABLE 6. REGRESSION OF STANDARD DEVIATION OF ANALYST FORECASTS

	Pseudo event = 1 year earlier than actual ERP implementation			Pseudo event = 1 year later than actual ERP implementation			Pseudo event = 2 years later than actual ERP implementation		
Variable	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value
Intercept	**0.0024**	6.52	<.0001	**0.0021**	6.97	<.0001	**0.0024**	6.52	<.0001
ERP	0.0008	1.42	0.1584	**0.0008**	1.81	0.0737	0.0008	1.42	0.1584
Market Value	**-0.00001**	-2.47	0.0154	**-0.00001**	-2.7	0.0084	**-0.00001**	-2.47	0.0154
Std. dev. ROE	-0.0002	-0.81	0.4221	-0.0002	-1.14	0.2591	-0.0002	-0.81	0.4221
Ret-earn corr	**0.0017**	1.69	0.0952	**0.0014**	1.75	0.0843	**0.0017**	1.69	0.0952
Earn surprise	**0.0182**	1.79	0.0773	0.0102	1.61	0.1121	**0.0182**	1.79	0.0773
Disc. Accrual	0.0004	0.11	0.9095	-0.0018	-0.65	0.5157	0.0004	0.11	0.9095
No. of observations	1408			1328			1408		
R-sqr	0.0572			0.04844			0.0572		

Standard errors are clustered by firm

Bold: Significant at the 10% level or better

CONCLUSION

Previous studies have shown that ERP systems gather and process information more efficiently and accurately, resulting in an improved internal information environment for firms that implement ERP systems. This enables management at these firms to provide guidance that is more accurate to security analysts. However, there is evidence to suggest that ERP systems enable firms to engage in greater earnings management (Brazel & Dang, 2008). Hence, whether or not ERP systems result in a better external information environment is an empirical question.

The results of this study show that analyst forecast accuracy decreased and the standard deviation of analyst forecasts increased in the years following ERP implementation. These results hold even after we control for discretionary accruals. It appears that the information environment for equity analysts has worsened after ERP implementations, even as management forecasts have become more accurate over the same period.

REFERENCES

Aboody, D. & Kasznik, R. (2000). CEO stock option awards and the timing of corporate voluntary disclosures. *Journal of Accounting and Economics*, *29*(1), 73 – 100.

Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting and Economics*, *11*(July), 255–274.

Brazel, J. F., & Agoglia, C. P. (2007). An examination of auditor planning judgements in a complex accounting information system environment. *Contemporary Accounting Research 24*(4), 1059–1083.

Brazel, J. F., & Dang, L. (2008). The effect of ERP System implementations on the management of earnings and earnings release dates. *Journal of Information Systems*, *22*(2), 1–21.

Brennan, M., & Hughes, P. (1991). Stock prices and the supply of information. *Journal of Finance 46*(December), 1665–1691.

Davenport, T. (1998). Putting the enterprise into the enterprise system. *Harvard Business Review 76*, 121–131.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *The Accounting Review 70*(April): 193–225.

Dillon, C. 1999. Stretching towards enterprise flexibility with ERP. *APICS – The Performance Advantage* (October), 38–43.

Hayes, D. C., Hunton, J. E., & Reck, J. L. (2001). Market reaction to ERP implementation announcements. *Journal of Information Systems*, *15*(1), 3–18.

Hitt, L. M., Wu, D. J., & Zhou, X. (2002). Investment in enterprise resource planning: Business impact and productivity measures. *Journal of Management Information Systems, 19*(1), 71–98.

Hunton, J. E., Lippincott, B., & Reck, J. L. (2003). Enterprise resource planning systems: Comparing firm performance of adopters and nonadopters. *International Journal of Accounting Information Systems, 4*(3): 165–184.

King, R., Pownall, G. & Waymire, G. (1990). Expectations adjustments via timely management forecasts: Review, synthesis and suggestions for future research. *Journal of Accounting Literature 9*, 113–144.

Kumar, K., & Hillegersberg, J. V. (2000). ERP experiences and evolution. *Communications of the ACM*, *43*(4), 22-26.

Lang, M. H. & Lundholm, R. J. (1993). Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research*, *31*(Autumn), 246–271.

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

O'Leary, D. E. (2000). *Enterprise resource planning systems: Systems, life cycle, electronic commerce, and risk.* Cambridge University Press.

Park, C. W. & Stice, E. K. (2000). Analyst forecasting ability and the stock price reaction to forecast revisions. *Review of Accounting Studies, 5*(2000), 259–272.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies, 22*(1), 435–480.

Rogers, J. L., & Stocken, P. C. (2005). Credibility of management forecasts. *The Accounting Review*, *80*(4), 1233–1260.

Waymire, G. (1996). Additional evidence on the accuracy of analyst forecasts before and after voluntary management earnings forecasts. *The Accounting Review*, 59(January), 129–142.

Wright, S. & Wright, A. M. (2002). Information system assurance for Enterprise Resource Planning systems: Unique risk considerations. *Journal of Information Systems, 16*(s-1), 99–113.

APPENDIX 1

Variable definitions:

ERP:	An indicator variable that is equal to one for years following ERP implementation, and zero otherwise.
Forecast Accuracy:	The negative of the absolute value of the analyst forecast error deflated by stock price, multiplied by 100 (to convert to %).
Market Value:	Market value of outstanding equity at the beginning of the quarter in billions of dollars.
Std. dev. ROE:	Standard deviation of return on equity (ROE) over prior 10 quarters
Ret-earn corr:	Market adjusted return-earnings per share correlation over prior 10 quarters.
Earn surprise:	The absolute value of the difference between the current quarter's earnings per share and last quarter's earnings for share, divided by the price at the beginning of the quarter.
Disc. Accrual:	Discretionary accruals in the prior quarter, estimated as the residual of the Modified Jones model.
Std. dev. fcast acc:	Standard deviation of accuracy of analyst forecasts.

A SHIFT-SHARE ANALYSIS OF SERVICE EXPORTS OF THE COUNTRIES OF LATIN AMERICA & CARIBBEAN

Philemon Oyewole, Howard University
Poyewole@howard.edu

ABSTRACT

Increased market globalization on the one hand, and rapid advances in telecommunications and information technology on the other, have given a boost to the volume of services marketed across national borders in recent years. This paper analyzes the relative progress made by the countries of Latin America & Caribbean for a share in this expanding sector of the international market over a 15-year period, from 2002 to 2017. A shift-share analysis was carried out to identify the winners and the losers in the market during the study period. Results show that major winners of market share in terms of total service exports were Brazil, Panama, Argentina, Peru, and Colombia in order of magnitude. The major losers of market shares, in order of magnitude (of losses), were Mexico, Dominican Republic, Chile, The Bahamas, and Jamaica. Policy implications of these results for the countries of Latin America & Caribbean are discussed.

INTRODUCTION

The volume of services marketed across national borders in the last decades have dramatically expanded, thanks in part to increased liberalization, market globalization, and rapid advances in telecommunications and information (Bhattacharya & Bhattacharya, 2013; Hongchindaket et al., 2013). Back in 1995, the World Bank predicted that the internationalization of services was expected to continue apace (World Bank, 1995). It has - services have now become a major component of world trade. While service exports stood at US$1,371 billion in 1997, ten years later they more than double growing to US$3,486 billion in 2007! And by the year 2017, the figures rose further to US$5,593 billion, accounting for more than 24% of total world exports for the year (World Bank, 2021). Understandably, countries and regions of the world are competing for a share of this growing revenue. Who are the winners and the losers in the global market for services? What are the public policy implications for the countries and regions involved? This paper analyzes the competition among the developing countries of Latin America & Caribbean for a share in this expanding international market for services over a 15-year period, from 2002 to 2017. A shift-share analysis was carried out to identify the winners and the losers in the market during the study period.

In the past services were thought to be unmarketable across national boundaries. It was even the Marxist prescription and Communist practice to omit them entirely from national income accounts (Bhagwati, 1984). This treatment of services was borne out of the so-called “haircut view.” Due to the simultaneous production and consumption characteristic of services, it was argued that one could not have a haircut long distance. The client and the barber must physically be present at the same place to render and consume the haircut service (Bhagwati, 1991). Thus, international marketing of services was considered not as important as international marketing of physical goods. With advances in technology however, services formerly considered to be unmarketable across national boundaries are now actively traded. This is especially true of, but not limited to, financial services, communications services, and data transmission (Dávila-Vargas-Machuca et al., 2014; Pilat, 1998). It has been said that; "Any activity that can be conducted through a screen and a telephone, from writing software to running a secretarial service, can be carried out anywhere in the world" (The Economist, 1996, p. S33).

Increased service components of manufacturing also gave a boost to international trade in services (Lodefalk, 2013, 2014; Hassan & Nassar, 2018). To be more competitive, manufacturers are shifting more and more to service differentiation, as against tangible product differentiation (Bas, 2014). Most of these service inputs into production are outsourced internationally to cut down costs (World Bank, 1995). International marketing of services was given a further boost by including services in the GATT's Uruguay Round of talks launched in 1986 (I.T.F., 1996). The Uruguay Round extended multilateral disciplines and negotiations to international trade in services for the first time. Services have also been included in recent initiatives of several regional integrations. These include USMCA, the Australia-New Zealand Closer Economic Relations Agreement, the Gulf Cooperation Council, Mercosur, and the European Union. Thus, it was speculated: "the internationalization of services will likely lead the next stage of economic globalization" (World Bank, 1995, p. 3). In view of this development, it is important to examine how developing countries of Latin America & Caribbean are competing in the world market for services. Which countries are benefiting the most from this growing phenomenon? Are some countries being left behind in this era of economic globalization? Who are they (if any), and in which service sector(s) are they lagging behind? What are the managerial and public policy implications of this? These questions of great importance to companies and national governments of these countries are addressed in this paper.

THE LITERATURE

Although rapidly growing in importance in world economy, relatively little research has been done on service exports compared to exports of physical goods (Bhattacharya et al., 2012; Chan & Coulthard, 2005; Jalali, 2013). As far back as 1988, Green et al. (1988, p. 208) called on marketing academics to "begin to study the international dimension of services marketing in greater depth." In 1997, Coviello et al. (1998, p. 8) also bemoaned the fact that "most studies" on export performance issues, "focus on manufacturing firms, with little attention given to service organizations." A few authors have responded to this call for more work on international marketing of services. For example, in their survey of Australian firms, Patterson and Cicic (1995) developed new criteria for classifying services that would be better for strategic international marketing of services. Their criteria classified services into four cells defined according to organizational demographics, and international practices. Likewise, empirical research of Edvardsson et al. (1993) examined foreign market entry strategies of service companies. Useful as these studies are, they do not deal with service export performance, or with developing countries like those of Latin

America & Caribbean. Winsted and Patterson (1998) studied service export performance and prospects, but again, with a focus on developed countries alone. They concluded that most developed countries were not exporting services to their full potential. Through their study of engineering consulting firms in the U.S., they identified barriers to service exports by developed countries. Malhotra et al. (1994) compared the international services marketing of developed and developing countries. However, the basis of their comparison was service quality, not service export performance of the two groups of countries. In a series of essays, Bhagwati (1984, 1987, & 1991) wrote on developing countries and international marketing of services. He argued that developing countries have good prospects for international marketing of services due to what he christened "splintering" and "disembodiment" of services. He thus strongly encouraged developing countries to participate actively in the Uruguay Round of multilateral negotiations in services rather than boycotting them as some developing countries wanted to do. While his essays are enlightening, they are more conceptual than empirical, thus leaving room for validation.

Some other authors have also tried to respond to the call for more research on service exports, however their efforts are often limited to just one service category, and/or concentrated on just one country. For example, Jalali (2013) studied engineering services unveiling six barriers to the international market faced by exporters of that category of services. Lu et al. (2012). studied professional business services discovering factors that contribute to success of exporters, which include management attitude, resource commitment and international experience, among others. In their own study, Tadesse and White (2012) examined tourism services and reported that immigrants significantly enhance exports of such services in the US. Bhattacharya and Bhattacharya (2013) studied software services in India and found that exports of that service category had unidirectional causality to economic growth during the study period. A few other authors have studied exports of a broad range of services, but again, they often concentrate on a single country. These include Minondo (2014) who focused on Spain; Hongchindaket et al. (2013) on Thailand; Lodefalk (2013, 2014) on Sweden; Fryges et al. (2015) on Germany; Castellacci (2014) on Norway; Sahoo and Dash (2014) on India; Sudarsan and Karmali (2011) also on India; and White et al. (2013) on the US. Of particular relevance to the present research is the study of Oyewole (2003). The study covered all the major sectors of service exports and covered several developing countries. However, the study was limited to the region of Sub-Saharan Africa. A further attempt by Oyewole (2016) studied all the major sectors of service exports, covering all the 7 regions of the world. However, the study was an aggregate at regional level, with no direct information on individual countries in each region. Such direct information is needed for effective formulation of public policies at national government level. This is the type of information provided by the current study on the developing countries of the region of Latin America & Caribbean.

METHODOLOGY: THE TECHNIQUE OF SHIFT-SHARE ANALYSIS

This paper used the robust technique of shift-share analysis. First developed in the field of regional economics in the 1940s (Markusen et al., 1991), several applications of shift-share analysis have been reported in the marketing literature. As far back as 1967, Huff and Sherr (1967) used it to determine regional growth rates of markets. Shift-share analysis was used by Yandle (1978) to assess brand performance. Green and Allaway (1985) used it to identify export opportunities for companies trading with the OECD countries. Likewise, Green and Larsen (1986) used the technique to assess the impact of sudden marketing environmental changes (economic, political, and/or social) on export markets. Oyewole (2003, 2016) used it to study service exports of the countries of Sub-Saharan Africa, and of the seven regions of the world, respectively.

The technique of shift-share analysis consists of disaggregating the growth (or change) in an indicator of interest (say employment, or exports) from one period to the other into three main components. These components are referred to as the (i) the national growth component, (ii) the industrial mix component, and (iii) the competitive share component (Sirakaya et al., 1995). This disaggregating is done for the region(s) being studied within a country of interest. The "national growth component" is the expected growth in a region had that region grown at the same rate as the nation over the study period. The "industrial mix component" shows the portion of the growth in a region that is due to its sectoral make-up compared to that of the nation. Thus, if a region had relatively more of the fast-growing industrial sectors than is the average for the nation, this component would be positive, indicating relative structural strength. The inverse would be the case if the region had less of the fast-growing sector(s) (Hustedde et al., 1993). The third component, "competitive share component," shows what portion of the growth in a region is attributable to the strengths or weaknesses of firms located within that region relative to firms within the same industry located elsewhere in the country. The sum of the second and third components is termed the "net shift." It shows what growth (or change) has taken place in the indicator of interest within a given region independent of the general growth (change) in the nation. The proportion of the net shift accounted for by an individual sector within a region is termed the "% net shift."

As adapted from Sirakaya et al. (1995), the shift-share model could be summarized by the equation:

$$e_{ij}^{t} - e_{ij}^{t-1} = \Delta e_{ij} = NG + IM + CS$$

Where:
i = the i^{th} sector in the benchmark economy
j = the j^{th} region of the benchmark economy
Δe_{ij} = total change in sector i, in the jth region
e^{t}_{ij} = value of indicator of interest for sector i in region j at time t
NG = national growth component
IM = industrial mix component
CS = competitive share component.

For a region j, each of these components could be computed as follows:

$$NG = \Sigma\, e_i^o E^t/E^o - \Sigma\, e_i^o$$
$$IM = \Sigma(e_i^o E_i^t/E_i^o - e_i^o\, E^t/E^o)$$
$$CS = \Sigma(e_i^t - e_i^o E_i^t/E_i^o)$$

Where:
$e = \Sigma\, e_i$ = sum of the indictor of interest across sectors i's in the region
$E = \Sigma E_i$ = sum of the indicator of interest across sectors i's in the benchmark economy

t = final year of the study period
o = beginning year of the study period

Over the years, various adjustments have been proposed to the shift-share analysis (e.g., Nazara & Hewings, 2004; Marquez et al., 2009; Zaccomer & Mason, 2011), but the basic original technique remains adequate for the purpose of the present paper, hence used. Although shift-share analysis was developed for examining regional development within a country, it has been shown that it could also be applied to study regional development within a much larger area such as a continent, sub-continent, or even the world as a whole (Dinc & Haynes, 1998a, 1998b; Markusen et al., 1991; Noponen et al., 1998; Sihag & Mcdonough, 1989). Following this reasoning, the shift-share analysis is here applied to studying relative development of the countries within the region of Latin America & Caribbean with regards to their service export performance. When applied to the world as the benchmark economy, the national growth component of the shift-share analysis could be termed the world growth effect (Sihag & Mcdonough, 1989). By the same token, that component will be termed "*regional growth effect*" in this study.

DATA SOURCE

The data for this study were obtained from the World Bank's World Development Indicators (World Bank, 2021). The source gives values of service exports for the Latin America & Caribbean region, and for individual countries in the region, in current U.S. dollars, from 2002 to 2017. Due to serious data gaps for several countries during this period, some 9 countries were removed from the total of 42 countries of Latin America & Caribbean. Thirty-three (33) countries in the region have complete data for the study years. These are the only countries retained for analysis in this this paper. All monetary figures in the paper are given in millions of current US dollars ($m). The World Development Indicators data provide information that break up services exports into four major categories, namely: (i) Transport services; (ii) Travel services; (iii) Insurance and financial services; and (iv) Communications, computer, information, and other services.

The four service categories are defined by World Bank (2021) as follows:

Communications, computer, information, and other services cover international telecommunications; computer data; news-related service transactions between residents and nonresidents; construction services; royalties and license fees; miscellaneous business, professional, and technical services; personal, cultural, and recreational services; manufacturing services on physical inputs owned by others; and maintenance and repair services and government services not included elsewhere.

Insurance and financial services cover various types of insurance provided to nonresidents by resident insurance enterprises and vice versa, and financial intermediary and auxiliary services (except those of insurance enterprises and pension funds) exchanged between residents and nonresidents.

Transport covers all transport services (sea, air, land, internal waterway, pipeline, space, and electricity transmission) performed by residents of one economy for those of another and involving the carriage of passengers, the movement of goods (freight), rental of carriers with crew, and related support and auxiliary services. Also included are postal and courier services. Excluded are freight insurance (included in insurance services); goods procured in ports by nonresident carriers (included in goods); maintenance and repairs on transport equipment (included in maintenance and repair services n.i.e.); and repairs of railway facilities, harbors, and airfield facilities (included in construction).

Travel covers goods and services acquired from an economy by travelers for their own use during visits of less than one year in that economy for either business or personal purposes. Travel includes local transport (i.e., transport within the economy being visited and provided by a resident of that economy) but excludes international transport (which is included in passenger transport. Travel also excludes goods for resale, which are included in general merchandise.

In this paper, "Insurance and financial services" will be referred to as "Financial services" henceforth for brevity. Likewise, "Communications, computer, information, and other services" will be referred to as "Communications services" for the same reason.

RESULTS

Figure 1 shows that in 2002, the total service export of the entire region of Latin America & Caribbean was $63,329m. Of this figure, 48% ($30,645m) were exports of Travel services, followed by exports of Communications services with 28% ($17,910m), Transport services with 18% ($11,099m) and lastly Financial services with 6% ($3,676). A decade and a half later, in 2017, the volume of total service exports for the region rose nearly 300% to $189,500m. Interestingly, there was no marked change in structure from 2002. Travel services export maintained its lead with 48% ($90,550m), followed by Communications services with 29% ($55,597m), Transport services with 16% ($30,243m), and lastly Financial services with 7% ($13,109). Nonetheless, albeit tiny, Communications, and Financial services gained 1% in relative market shares from 2002 to 2017; while Transport services lost 2% in market share over the same period. Travel services had no change from the 48% market share of 2002.

Figure 2 presents the relative proportion of each country's total service exports in the total exports for Latin America & Caribbean in 2002 and 2017. In 2002, Mexico led all other countries with 19.81% of the total service exports, followed by Brazil with 13.57%, Dominican Republic with 7.56%, Chile with 6.93%, and Argentina with 5.40%, to name the top five. The country with the least proportion was Suriname with 0.06%. By the year 2017, Brazil took the lead with 18.19%, pushing Mexico to the second place now with 14.59%, followed by Argentina, that shot up from the 5th place to the 3rd place, with 7.78%, then Panama with 7.03%, and Chile with 5.38%. The country with the least proportion in 2017 remained Suriname with 0.07% of the regional total service exports.

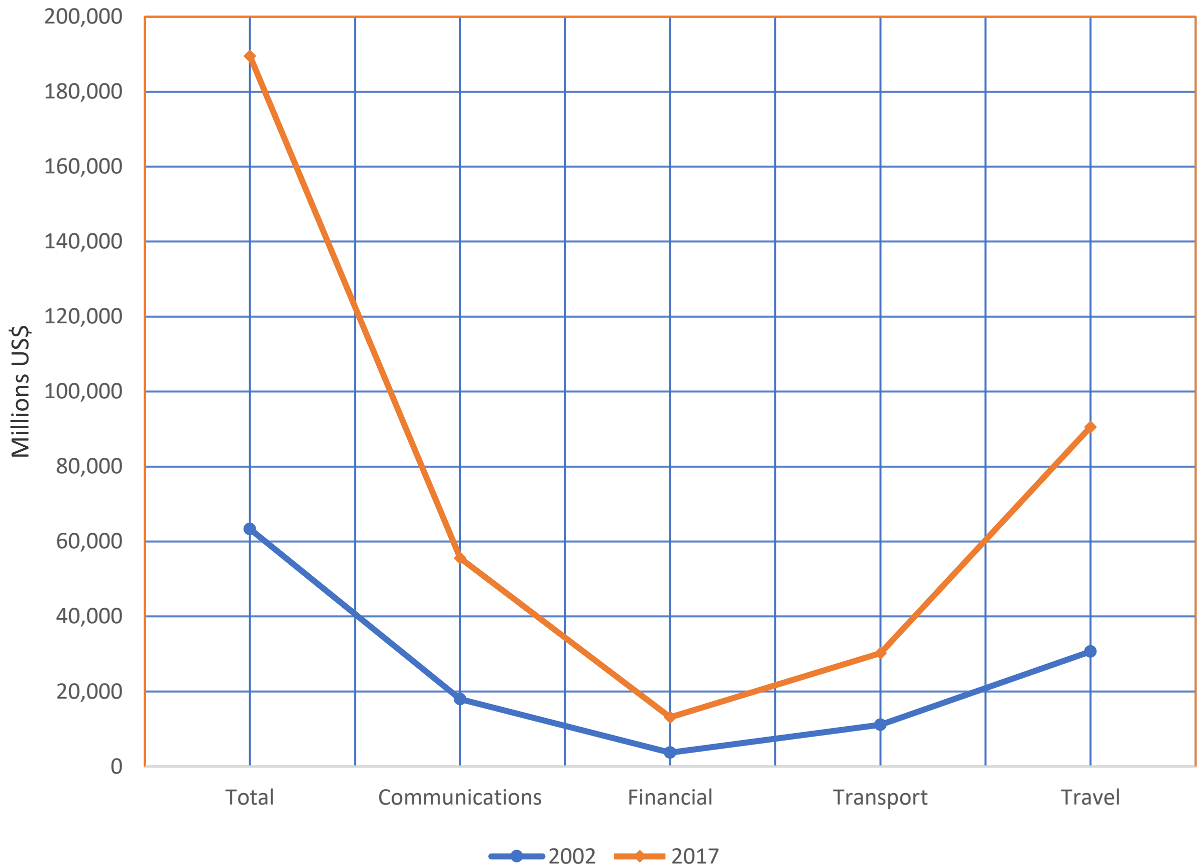

FIGURE 1: LATIN AMERICA & CARIBBEAN SERVICE EXPORTS: 2002-2017 (MIILIONS OF US$)

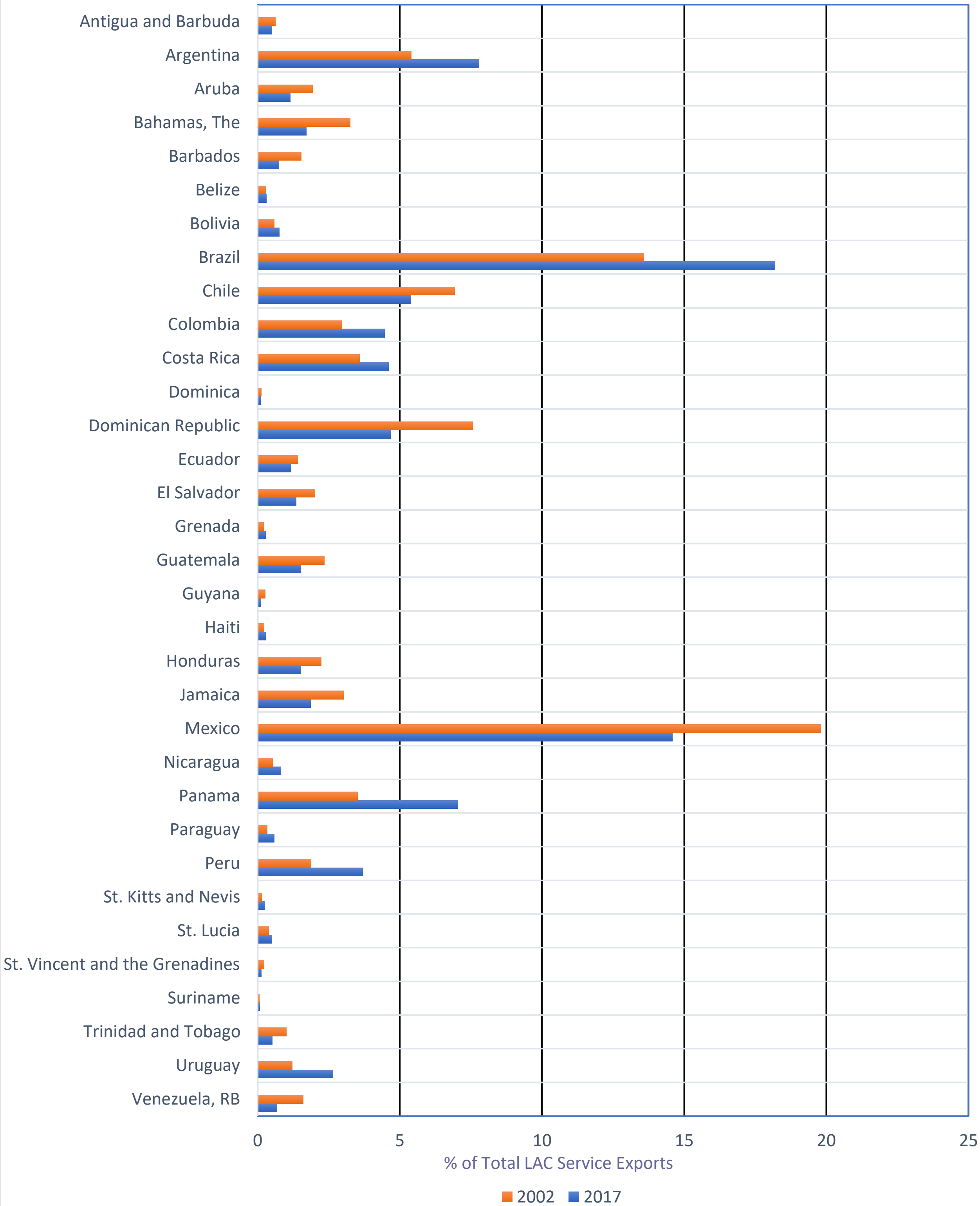

FIGURE 2: STRUCTURE OF TOTAL LATIN AMERICA & CARIBBEAN (LAC) SERVICE EXPORTS BY COUNTRY (%): 2002-2017

While the foregoing analysis is informative it does not reveal the full breakdown of the relative progress (or otherwise) made by individual countries in the region over the 15-year period. This is important for effective policy actions to be recommended, and hopefully taken, for the future of these countries in the international market for services going forward. To fully reveal these details, a shift-share analysis was carried out, with the results discussed below.

Table 1 shows the regional growth effect of the shift-share analysis. The entries indicate the expected change in the volume of service exports from 2002 to 2017 if a country was growing at the same rate as the growth of the regional's total service exports. Table 2 shows the industrial mix component, while Table 3 shows the competitive share component of the shift-share analysis. The sum of the corresponding entries in the two tables gives the net shift presented in Table 4. The net shift is the difference between the actual change in the volume of service exports from 2002 to 2017, and the expected change presented in Table 1. The net shift thus effectively indicates the gainers and the losers of market share among the countries in the region during the study period.

TABLE 1. REGIONAL GROWTH COMPONENT OF THE SHIFT-SHARE ANALYSIS, 2002-2017 (MILLIONS OF $US)

		Total	Communi-cations	Financial	Transport	Travel
1	Antigua and Barbuda	785.41	70.50	17.71	151.76	545.44
2	Argentina	6,814.05	2,180.89	1.79	1,572.98	3,058.39
3	Aruba	2,440.53	703.18	3.94	71.12	1,662.29
4	Bahamas, The	4,107.70	487.04	0.00	114.61	3,506.05
5	Barbados	1,937.95	377.19	208.25	41.83	1,310.69
6	Belize	367.07	88.49	0.14	36.43	242.00
7	Bolivia	739.08	324.85	87.16	126.86	200.21
8	Brazil	17,120.19	8,884.32	1,187.06	3,068.27	3,980.54
9	Chile	8,740.62	2,234.67	323.72	4,393.88	1,788.34
10	Colombia	3,744.80	646.39	81.97	1,089.94	1,926.50
11	Costa Rica	4,515.79	1,415.02	9.26	506.18	2,585.33
12	Dominica	158.51	52.88	4.15	10.77	90.71
13	Dominican Republic	9,538.51	3,912.87	0.00	185.88	5,439.76
14	Ecuador	1,773.98	397.35	0.85	484.85	890.94
15	El Salvador	2,549.94	1,346.59	66.94	647.70	488.71
16	Grenada	261.41	49.44	9.76	19.94	182.27
17	Guatemala	2,970.80	1,444.14	110.84	181.41	1,234.42
18	Guyana	343.27	203.21	23.31	18.53	98.22
19	Haiti	292.21	77.02	0.00	0.00	215.19
20	Honduras	2,825.15	2,064.17	37.06	124.32	599.60
21	Jamaica	3,809.79	591.91	75.23	734.56	2,408.09
22	Mexico	24,994.99	3,117.59	2,415.38	1,814.30	17,647.72
23	Nicaragua	671.40	337.69	4.58	60.96	268.16
24	Panama	4,442.82	418.18	592.91	2,409.68	1,022.05

25	Paraguay	429.76	118.54	48.01	139.29	123.92
26	Peru	2,377.90	65.76	213.79	529.59	1,568.77
27	St. Kitts and Nevis	179.85	41.56	4.47	19.99	113.83
28	St. Lucia	497.83	47.33	9.11	22.99	418.39
29	St. Vincent and the Grenadines	273.12	69.94	3.24	18.63	181.32
30	Suriname	76.70	17.87	0.87	50.55	7.41
31	Trinidad and Tobago	1,270.49	186.08	198.83	403.44	482.14
32	Uruguay	1,536.72	152.59	141.94	543.08	699.11
33	Venezuela, RB	2,026.16	529.95	1.99	629.57	864.66

The industrial mix component presented in Table 2 indicates the relative strength and weakness of the sectoral structure in each country. The table shows that overall (see the "Total" column), 21 of the 33 countries have weaker sectoral structure than the region as a whole. These are Chile, Panama, Argentina. Colombia, Jamaica, Venezuela, Ecuador, The Bahamas, Uruguay, Peru, Costa Rica, Antigua and Barbuda, Suriname, St. Lucia, Belize, St. Kitts and Nevis, El Salvador, St. Vincent and the Grenadines, Paraguay, Grenada, and Aruba. All these countries have negative entries under Total service exports. Their slowest growing sectors were Transport services and Travel services, where Table 2 also shows negative entries for them over the 15-year study period. The Implication is that these countries have relatively more of the weaker (such as Transport), but less of the stronger (such as Communications) service sectors, resulting in overall weaker (negative) sectoral structure. Although all other countries (except Haiti) also show negative entries under these two sectors, these were more than made up with their having relatively more of the stronger sectors in the region (Communications and Financial), resulting in positive entries for their Total services.

TABLE 2. INDUSTRIAL MIX COMPONENT OF THE SHIFT-SHARE ANALYSIS, 2002-2017 (MILLIONS OF $US)

		Total	Communi-cations	Financial	Transport	Travel
1	Antigua and Barbuda	-21.55	3.96	5.10	-20.36	-10.25
2	Argentina	-145.47	122.57	0.52	-211.07	-57.49
3	Aruba	-0.13	39.52	1.13	-9.54	-31.25
4	Bahamas, The	-53.91	27.37	0.00	-15.38	-65.91
5	Barbados	50.93	21.20	59.98	-5.61	-24.64
6	Belize	-4.42	4.97	0.04	-4.89	-4.55
7	Bolivia	22.58	18.26	25.10	-17.02	-3.76
8	Brazil	354.68	499.33	341.89	-411.71	-74.82
9	Chile	-404.37	125.60	93.24	-589.59	-33.62
10	Colombia	-122.53	36.33	23.61	-146.25	-36.21
11	Costa Rica	-34.32	79.53	2.67	-67.92	-48.60
12	Dominica	1.02	2.97	1.19	-1.45	-1.71
13	Dominican Republic	92.72	219.92	0.00	-24.94	-102.25
14	Ecuador	-59.23	22.33	0.24	-65.06	-16.75

15	El Salvador	-1.13	75.68	19.28	-86.91	-9.19
16	Grenada	-0.51	2.78	2.81	-2.68	-3.43
17	Guatemala	65.54	81.17	31.92	-24.34	-23.20
18	Guyana	13.80	11.42	6.71	-2.49	-1.85
19	Haiti	0.28	4.33	0.00	0.00	-4.05
20	Honduras	98.73	116.01	10.67	-16.68	-11.27
21	Jamaica	-88.90	33.27	21.67	-98.57	-45.27
22	Mexico	295.69	175.22	695.66	-243.45	-331.73
23	Nicaragua	7.08	18.98	1.32	-8.18	-5.04
24	Panama	-148.29	23.50	170.76	-323.34	-19.21
25	Paraguay	-0.53	6.66	13.83	-18.69	-2.33
26	Peru	-35.28	3.70	61.57	-71.06	-29.49
27	St. Kitts and Nevis	-1.20	2.34	1.29	-2.68	-2.14
28	St. Lucia	-5.66	2.66	2.62	-3.08	-7.86
29	St. Vincent and the Grenadines	-1.05	3.93	0.93	-2.50	-3.41
30	Suriname	-5.67	1.00	0.25	-6.78	-0.14
31	Trinidad and Tobago	4.53	10.46	57.27	-54.14	-9.06
32	Uruguay	-36.56	8.58	40.88	-72.87	-13.14
33	Venezuela, RB	-70.37	29.78	0.57	-84.48	-16.25

The competitive share component presented in Table 3 indicates the relative *efficiency* of each service sector in each country. It thus shows where a country has high competitive advantage (positive entries in the Table), and where it has low competitive advantage (negative entries in the Table). Table 3 shows that two countries, namely: Peru and Panama have high competitive advantage in all 4 service sectors as indicated by their positive entries in the Table. By contrast, five countries have low competitive advantage in all four sectors as indicated by their negative entries. These countries are Jamaica, Honduras, Trinidad and Tobago, Guyana, and surprisingly Mexico! A closer look at the Table shows that Brazil has the highest competitive advantage (with the largest entry of $7,661.21m) in the Communications sector, followed by Argentina ($3,742.68m), and Costa Rica ($2,257.46m), to name the top three. In the Financial services sector, the country with the highest competitive advantage is Peru (with the largest entry of $877.27m), followed by Panama ($421.96m), and then Argentina ($198.00m). The country with the highest competitive advantage in the Transport services sector is Panama (with $3,153.98m entry), followed by Brazil ($1,593.47m), and then Peru ($758.29m). Finally, in the Travel services sector, Panama again has the highest competitive advantage (with $2,897.17m), but now followed by Colombia ($2,063.45m), and then Uruguay ($1,674.70m). The two strongest sectors for the Latin America & Caribbean region are Travel, and Communications. Panama took the competitive lead in the former, while Brazil took the competitive lead in the latter.

TABLE 3. COMPETITIVE SHARE COMPONENT OF THE SHIFT-SHARE ANALYSIS, 2002-2017 (MILLIONS OF $US)

		Total	Communi-cations	Financial	Transport	Travel
1	Antigua and Barbuda	-226.52	-42.64	0.60	-85.23	-77.69
2	Argentina	4,663.06	3,742.68	198.00	-204.13	926.50
3	Aruba	-1,494.40	-1,060.65	10.33	42.00	-486.08
4	Bahamas, The	-2,857.95	-576.72	0.00	-77.03	-2,204.20
5	Barbados	-1,547.40	-542.18	-345.87	22.99	-682.34
6	Belize	34.59	-18.10	5.33	-20.43	67.79
7	Bolivia	306.69	-277.44	-151.51	245.69	489.95
8	Brazil	8,410.32	7,669.21	-757.89	1,593.47	-94.46
9	Chile	-2,528.25	-232.47	-47.85	-3,029.63	781.70
10	Colombia	2,959.53	729.01	-50.12	217.19	2,063.45
11	Costa Rica	1,970.52	2,257.46	75.04	-239.08	-122.90
12	Dominica	-47.71	-56.07	-3.24	-13.90	25.49
13	Dominican Republic	-5,562.13	-5,187.78	115.27	358.40	-848.02
14	Ecuador	-408.05	-421.62	36.43	-249.55	226.68
15	El Salvador	-1,271.39	-991.70	-22.02	-404.08	146.40
16	Grenada	144.79	-49.33	-10.85	-12.53	217.50
17	Guatemala	-1,673.59	-1,438.27	-123.59	153.37	-265.11
18	Guyana	-323.09	-280.58	-16.03	-7.77	-18.71
19	Haiti	95.49	-48.20	3.39	0.00	140.30
20	Honduras	-1,474.45	-1,195.75	-62.53	-41.66	-174.52
21	Jamaica	-2,114.30	-435.71	-122.40	-798.89	-757.31
22	Mexico	-10,193.58	-4,549.34	-229.16	-577.31	-4,837.77
23	Nicaragua	542.12	124.03	-2.00	-22.68	442.78
24	Panama	6,798.68	325.57	421.96	3,153.98	2,897.17
25	Paraguay	472.40	193.06	-57.44	151.50	185.28
26	Peru	3,471.61	452.75	877.27	758.29	1,383.30
27	St. Kitts and Nevis	209.91	17.64	17.42	-10.17	185.02
28	St. Lucia	199.48	-35.05	-8.78	-18.26	261.56
29	St. Vincent and the Grenadines	-159.18	-93.19	1.93	-13.29	-54.64
30	Suriname	29.93	20.58	2.63	-28.63	35.35
31	Trinidad and Tobago	-936.71	-181.39	-296.37	-197.11	-261.83
32	Uruguay	2,747.29	1,482.34	-103.72	-306.03	1,674.70
33	Venezuela, RB	-1,687.79	-619.74	14.43	-273.09	-809.40

Table 4 shows the net shift of the shift-share analysis. The Table reveals that overall (see the Total column), 17 of the 33 countries lost market share in total service exports between 2002 and 2017 as indicated by their negative entries. In order of magnitude, these were Mexico, Dominican Republic, Chile, The Bahamas, Jamaica, Venezuela, Guatemala, Barbados, Aruba, Honduras, El Salvador, Trinidad and Tobago, Ecuador, Guyana, Antigua and Barbuda, St. Vincent and the

Grenadines, and lastly, Dominica. Total service exports of these countries grew slower than those of the Latin America & Caribbean region as a whole. In order words, they did not meet expectations. The other 16 countries gained market shares as indicated by their positive entries in Table 4 (see the Total column). In order of magnitude of gains, these countries were Brazil, Panama, Argentina, Peru, Colombia, Uruguay, Costa Rica, Nicaragua, Paraguay, Bolivia, St. Kitts and Nevis, St. Lucia, Grenada, Haiti, Belize, and lastly, Suriname. Total service exports of these countries grew at a faster rate than those of the Latin America & Caribbean region as a whole during the study period. In order words, they beat expectations.

As further shown in Table 4, the greatest winners in Communications services were Brazil, Argentina, and Costa Rica with positive net shifts of 8,168.53, 3,865.25, and 2,336.99 respectively (see Column 4). In Financial services, the greatest winners, in order of magnitude of net shifts, were Peru, Panama, and Mexico, with net shifts of 938.84, 592.73, and 466.50 respectively (see Column 5). The top winners in Transport services were Panama, Brazil, and Peru, with net shifts of 2,830.64, 1,181.76, and 687.22 respectively (see Column 6). In Travel services, the greatest were Panama, Colombia, and Uruguay, with net shifts of 2,877.96, 2,027.24, and 1,661.56 respectively (see Column 7).

Dominican Republic, Mexico, and Guatemala were the greatest losers in Communications services, with negative net shifts of -4,967.86, -4,374.12, and -1,357.10 respectively (see Column 4). Likewise, the greatest losers in Financial services were Brazil, Barbados, and Trinidad and Tobago, with negative net shifts of -416.00, -285.89, and -239.11 respectively (see Column 5). For Transport services, the greatest losers were Chile, Jamaica, and Mexico, with net shifts of -3,619.21, -897.45, and -820.76 respectively (see Column 6). And finally, in Travel services, the top three losers were Mexico, The Bahamas, and Dominican Republic, with net shifts of -5,169.50, -2,270.10, and -950.28 respectively (see Column 7). Overall, Panama, and Peru performed outstandingly well, in that they were the only two countries with positive net shift entries in *all* four sectors of the service exports of the Latin America &Caribbean region during the 2002-2017 study period.

TABLE 4. NET SHIFT (ACTUAL-EXPECTED CHANGE) 2002-2017 (MILLIONS OF $US)

#	Country	Total	Communi-cations	Financial	Transport	Travel
1	Antigua and Barbuda	-226.52	-38.68	5.70	-105.59	-87.94
2	Argentina	4,517.59	3,865.25	198.52	-415.19	869.01
3	Aruba	-1,494.54	-1,021.13	11.46	32.45	-517.33
4	Bahamas, The	-2,911.86	-549.35	0.00	-92.41	-2270.10
5	Barbados	-1,496.47	-520.98	-285.89	17.38	-706.98
6	Belize	30.17	-13.13	5.38	-25.31	63.24
7	Bolivia	329.26	-259.19	-126.41	228.67	486.19
8	Brazil	8,765.00	8,168.53	-416.00	1,181.76	-169.29
9	Chile	-2,932.62	-106.88	45.39	-3,619.21	748.08
10	Colombia	2,837.00	765.34	-26.51	70.94	2,027.24

11	Costa Rica	1,936.20	2,336.99	77.71	-307.00	-171.50
12	Dominica	-46.70	-53.09	-2.05	-15.34	23.79
13	Dominican Republic	-5,469.41	-4,967.86	115.27	333.46	-950.28
14	Ecuador	-467.28	-399.29	36.68	-314.61	209.93
15	El Salvador	-1,272.52	-916.01	-2.74	-490.99	137.22
16	Grenada	144.28	-46.55	-8.03	-15.21	214.07
17	Guatemala	-1,608.05	-1,357.10	-91.67	129.03	-288.31
18	Guyana	-309.29	-269.16	-9.31	-10.25	-20.56
19	Haiti	95.77	-43.87	3.39	0.00	136.25
20	Honduras	-1,375.72	-1,079.74	-51.86	-58.34	-185.79
21	Jamaica	-2,203.19	-402.44	-100.73	-897.45	-802.57
22	Mexico	-9,897.88	-4,374.12	466.50	-820.76	-5169.50
23	Nicaragua	549.20	143.01	-0.68	-30.86	437.74
24	Panama	6,650.40	349.07	592.73	2,830.64	2,877.96
25	Paraguay	471.87	199.72	-43.61	132.81	182.95
26	Peru	3,436.33	456.45	938.84	687.22	1,353.81
27	St. Kitts and Nevis	208.71	19.98	18.70	-12.85	182.88
28	St. Lucia	193.82	-32.39	-6.16	-21.34	253.70
29	St. Vincent and the Grenadines	-160.23	-89.26	2.87	-15.79	-58.05
30	Suriname	24.26	21.58	2.88	-35.41	35.21
31	Trinidad and Tobago	-932.18	-170.93	-239.11	-251.25	-270.90
32	Uruguay	2,710.73	1,490.91	-62.84	-378.91	1,661.56
33	Venezuela, RB	-1,758.16	-589.95	15.01	-357.57	-825.66

TABLE 5. % NET SHIFT (ACTUAL-EXPECTED CHANGE) 2002-2017

#	Country	Communi-cations	Financial	Transport	Travel
1	Antigua and Barbuda	-16.66	100.00	-45.47	-37.87
2	Argentina	78.36	4.02	-100.00	17.62
3	Aruba	-66.37	26.10	73.90	-33.63
4	Bahamas, The	-18.87	0.00	-3.17	-77.96
5	Barbados	-34.41	-18.89	100.00	-46.70
6	Belize	-34.15	7.83	-65.85	92.17
7	Bolivia	-67.22	-32.78	31.99	68.01
8	Brazil	87.36	-71.08	12.64	-28.92
9	Chile	-2.87	5.72	-97.13	94.28
10	Colombia	26.73	-100.00	2.48	70.80
11	Costa Rica	96.78	3.22	-64.16	-35.84
12	Dominica	-75.33	-2.90	-21.77	100.00
13	Dominican Republic	-83.94	25.69	74.31	-16.06
14	Ecuador	-55.93	14.87	-44.07	85.13
15	El Salvador	-64.98	-0.19	-34.83	100.00

16	Grenada	-66.70	-11.51	-21.79	100.00
17	Guatemala	-78.13	-5.28	100.00	-16.60
18	Guyana	-87.03	-3.01	-3.32	-6.65
19	Haiti	-100.00	2.43	0.00	97.57
20	Honduras	-78.49	-3.77	-4.24	-13.50
21	Jamaica	-18.27	-4.57	-40.73	-36.43
22	Mexico	-42.20	100.00	-7.92	-49.88
23	Nicaragua	24.62	-2.16	-97.84	75.38
24	Panama	5.25	8.91	42.56	43.27
25	Paraguay	38.74	-100.00	25.76	35.49
26	Peru	13.28	27.32	20.00	39.40
27	St. Kitts and Nevis	9.02	8.44	-100.00	82.54
28	St. Lucia	-54.08	-10.28	-35.64	100.00
29	St. Vincent and the Grenadines	-54.73	100.00	-9.68	-35.59
30	Suriname	36.16	4.83	-100.00	59.01
31	Trinidad and Tobago	-18.34	-25.65	-26.95	-29.06
32	Uruguay	47.29	-14.23	-85.77	52.71
33	Venezuela, RB	-33.27	100.00	-20.17	-46.56

Table 5 presents the % net shift of the shift-share analysis. The entries indicate the relative proportion of gain or loss of market share that comes from each service sector in a given country. The table shows that all the losses of shares by Haiti (-100%) were in Communications. Likewise, all the losses of shares by Colombia (-100%), and Paraguay (-100%) were in Financial services; while all the losses of shares by Argentina (-100%), St. Kitts and Nevis (-100%), and Suriname (-100%) were in Transport services. On the other hand, all the gains in shares by Antigua and Barbuda (100%), Mexico (100%), St. Vincent and the Grenadines (100%), as well as Venezuela (100%) were in Financial services. All the gains in shares by Barbados (100%), and Guatemala (100%) were in Transport service, while all the gains in shares by El Salvador (100%), and Grenada (100%) were in Travel services. All the other countries spread their gains and losses of shares over several service sectors as indicated by their % net shift entries (positive or negative) in Table 5. For example, the gains in shares of Brazil were spread over Communications (87.36%) and Transport services (12.64%), while its losses were spread over Financial (-71.08%) and Travel services (-28.92%). Another example is Jamaica that spread its losses (with no gains at all) over all the four service sectors, namely: Communications (-18.27%), Financial services (-4.57%), Transport services (-40.73%), and (-36.43%). Its greatest loss could be seen to be in Transport services with -40.73% net shift, while its smallest loss was in Financial services with -4.57% net shift.

CONCLUSIONS AND POLICY IMPLICATIONS

The shift-share analysis revealed who the winners and losers of market shares in the international market for services were, during the 15-year study period, among the countries of Latin America & Caribbean. Surprisingly, the region was practically divided into two halves: with one half as winners, and the other half as losers! In order of magnitude, results show that the five major winners, in terms of Total service exports, were Brazil, Panama, Argentina, Peru and Colombia. Following these five are 11 other countries who gained some market shares in much smaller amounts, such as Costa Rica, Haiti, and Suriname. On the other hand, the 5 major losers of market shares, in order of magnitude (of losses), were Mexico, Dominican Republic, Chile, The Bahamas, and Jamaica. Following these five are 12 other countries who also lost some market shares but in smaller amounts, such as Venezuela, Honduras, and St. Vincent and the Grenadines. As once suggested for Sub-Saharan African countries by Oyewole (2003), the losing countries identified in this study should learn from the good practices of the winners and embark on policies that will promote service exports in general (Nwachukwu, 2011), but especially Communications sector in particular. This is the fastest growing sector in the world market for services. It will be important for all the countries but especially the losing countries to adopt appropriate national educational policies that favor expansion of computer literacy and of information technology skills in the general labor force. Nyahoho (2010) for example, found that human capital is clearly related to exports of computer and information services. Thus, school curricula should include courses in basic computer usage and data/information processing (Lopez, 2020). Incentives (such as tax deductions) could also be given to companies to train or retrain their employees in these areas. With appropriate information technology skills, the developing countries of Latin America & Caribbean would be able to take advantage of their relatively low wages in providing labor-intensive long-distance services in the Communications sector.

In addition to targeted education policies, adequate infrastructure and institutional development are essential for effective implementation of any service export expansion initiatives in these countries. For example, there should be 24-hour provision of electricity to operate any installed telecommunications networks if they will work to full capacity. Likewise, institutional building will be necessary to give backbone support to government policies. For example, as once suggested by Oyewole (2001) for developing countries in general, there may be a need for creation of a Department of Special Education in Information Technology, as well as a Department of Trade in Services in these countries.

The losing countries of Latin America & Caribbean region could also learn a lot from the example of India in the region of South Asia. As once noted by Oyewole (2016), India of the South Asia region has been surging in recent decades in the international market for service exports. Apart from increased world demand, the key factors for this surge have been found to include: "human capital, tele-density, financial development, physical infrastructure, and institutions, … exchange rate and foreign direct investment" Sahoo and Dash (2014, p. 1082). Another factor reported by Sudarsan and Karmali (2011, p. 73) was "value of service sector in GDP." All these point to several possible policy initiatives for the countries of Latin America & Caribbean in general, and especially, for the losing countries. For example, there is a need for public policies to increase the share of service sector as % of GDP within these countries. As noted by Bas (2014), while the share of services in GDP in developed countries is 70%, it is only 50% in developing countries of which the countries of Latin America & Caribbean are a part. This % needs to go up as is in India, which enjoys a 60% share of GDP (Bas, 2014). Liberalization and reform policies should be put in place to enhance domestic competition among service firms. This in turn will strengthen their competitiveness in the international market for services.

The countries of Latin America & Caribbean should also consider policies of liberalization of FDI (Foreign Direct Investment) regime in services as was done in India in the 1990s. According to Sudarsan and Karmali (2011), India allows automatic 51% foreign equity in FDI in most services, with provision for approval of up to 74% foreign equity in the case of some special service sectors (e.g., nonconventional energy generations and distribution). The essence of this liberalization is to give automatic controlling power to the foreign investor thus enhancing influx of FDI's in services (Hassan & Nassar, 2017). The countries of Latin America & Caribbean will need this type of influx to supplement government resources needed to implement their policies of increasing % share of services in GDP, which in turn could greatly boost their competitiveness in the international market for exports of services.

REFERENCES

Bas, M. (2014). Does services liberalization affect manufacturing firms' export performance? Evidence from India. *Journal of Comparative Economics, 42*(3), 569–589.

Bhagwati, J. (1984). Splintering and disembodiment of services and developing Nations. *World Economy, 7*(2), 133-143.

Bhagwati, J. (1987). Trade in services and the multilateral trade negotiations. *The World Bank Economic Review, 1*(4), 549-569.

Bhagwati, J. (1991). *Political economy, and international economics.* M.I.T.

Bhattacharya, M. & Bhattacharya, S. N. (2013). Software services export and its implications on economic growth in India: An empirical study. *Journal of Economics and Business, 11*(1), 17-26.

Bhattacharya, R., Patnaik, I., & Shah, A. (2012). Export versus FDI in services. *The World Economy, 35*(1), 61-78.

Castellacci, F. (2014). Service firms' heterogeneity, international collaborations, and export participation. *Journal of Industry, Competition and Trade, 14*(4), 259–285.

Chan, D. C. Y., & Coulthard, M. (2005). *The challenges facing service exporters: Lessons from the Victoria based transportation and travel sectors* [Working Paper]. Department of Management. Monash University.

Coviello, N. E., Ghauri, P. N. & Martin, K. A-M. (1998). International competitiveness: Empirical findings from SME service firms. *Journal of International Marketing, 6*(2), 8-27.

Dávila-Vargas-Machuca, M. A., Moral-Pajares, E. & Muñoz-Guarasa, M. (2014). The export of knowledge intensive services in the OECD countries: An empirical analysis. *Revista de Economía Mundial, 38*(1), 147-173.

Dinc, M., & Haynes, K. E. (1998a). International trade and shift-share analysis: A specification note. *Economic Development Quarterly, 12*(4), 337-343.

Dinc, M., & Haynes, K. E. (1998b). International trade and shift-share analysis: A specification note, rejoinder. *Economic Development Quarterly, 12*(4), 351-354.

Edvardsson, B., Edvinsson, L., & Nystrom, H. (1993). Internationalization in service companies. *The Service Industries Journal, 13*(1), 80-97.

Fryges, H., Vogel, A. & Wagner, J. (2015). The impact of R&D activities on exports of German business services enterprises: First evidence from a continuous treatment approach. *World Economy, 38*(4), 716-729.

Green, R. T., & Allaway, A. W. (1985). Identification of export opportunities: A shift-share approach. *Journal of Marketing, 49*(1), 83-88

Green, R. T., Jeffes, D. L., & Chen, A. (1988). *Problems in the international marketing of services*. Les Actes du 15e Seminaire, I.A.E., France.

Green, R. T., & Larsen, T. L. (1986). Sudden wealth/sudden poverty: Implications for export opportunities. *Columbia Journal of World Business, 21*(4), 3-12.

Hassan, M., & Nassar, R. (2017). An empirical study of the relationship between foreign direct investment and key macroeconomic variables in Mexico. *Journal of International Business Disciplines*, *12*(1), 18-30.

Hassan, M., & Nassar, R. (2018). Relationship between proposed measures of technological change and employment in the manufacturing sector. *Journal of International Business Disciplines*, *13*(2), 19-31.

Hongchindaket, A., Kittisarn, A., & Neck, P. A. (2013). Factors associated with successful export performance: A study of Thai international service firms. *Review of International Comparative Management, 14*(1), 54-62.

Huff, D. L., & Sherr, L. A. (1967). Measure for determining differential growth rates of markets. *Journal of Marketing Research*, *4*(4), 391-395.

Hustedde, R. J., Shaffer, R., & Pulver, G. (1993). *Community economic analysis: A how to manual* (Rev. ed.). North Central Regional Center for Rural Development, Iowa State University.

I.T.F. (1996). *Uruguay Round: The general agreement on trade in services* (Geneva, No.2). International Trade Forum.

Jalali, S. H. (2013). Assessment of the engineering service export barriers: A case study. *The Journal of Commerce, 5*(1), 1-6.

Lodefalk, M. (2013). Servicification of manufacturing—Evidence from Sweden. *International Journal of Economics & Business Research, 6*(1), 87–113.

Lodefalk, M. (2014). The role of services for manufacturing firm exports. *Review of World Economy, 150*(1), 59–82

Lopez, R. M. (2020). Perspectives of technology competency in business instruction. *Journal of International Business Disciplines*, *15*(2), 31-45.

Lu, V. N., Quester, P. G., Medlin, C. J., & Scholz, B. (2012). Determinants of export success in professional business services: A qualitative study. *The Service Industries Journal*, *32*(10), 1637–1652.

Malhotra, N. K., Ulgado, F. M., Agarwal, J., & Baalbaki, I. B. (1994). International services marketing: A comparative evaluation. *International Marketing Review, 11*(2), 5-15.

Markusen, A. R., Noponen, H., & Driessen, K. (1991). International trade, productivity, and U.S. regional job growth: A shift-share interpretation. *International Regional Science Review, 14*(1), 15-39.

Marquez, M. A., Ramano, J., & Hewings, G. J. D. (2009). Incorporating sectoral structure into shift-share analysis. *Growth and Change*, *40*(4), 594-618.

Minondo, A. (2014). The relationship between export status and productivity in services: A firm-level analysis for Spain. *Bulletin of Economic Research, 66*(S1), S138-S146.

Nazara, S., & Hewings, G. J. D. (2004). Spatial structure and taxonomy of decomposition in shift-share analysis. *Growth and Change, 35*(4), 476-490.

Noponen H., Markusen, A., & Driessen, K. (1998). International trade and shift-share analysis: A response to Dinc and Haynes. *Economic Development Quarterly, 12*(4), 344-350.

Nwachukwu, S. L. S. (2011). Assessing the effectiveness of export promotion programs: A research note on Louisiana programs. *Journal of International Business Disciplines, 6*(2), 1-8.
Nyahoho, E. (2010). Determinants of comparative advantage in the international trade of services: An empirical study of the Hecksher–Ohlin approach. *Global Economy Journal*, *10*(1), 1-22.
Oyewole, P. (2001). Prospects for developing country exports of services to the year 2010: Projections and public policy implications. *Journal of Macromarketing*, *21*(1), 32-46.
Oyewole, P. (2003). Winners and losers in the international market for services: A shift-share analysis of Sub-Saharan African service exports. In S. Nwankwo (Ed.), *Dimensions of African business and development* (pp. 41-56), Sheffield Hallam University Press.
Oyewole, P. (2016). Regional competition in the international market for services: A shift-share analysis. *Journal of Global Marketing*, *29*(1), 3-14.
Patterson, P. G., & Cicic, M. (1995). A typology of service firms in international markets: An empirical investigation. *Journal of International Marketing*, 3(4), 57-83.
Pilat, D. (1998). The economic impact of technology. *OECD Observer*, 213(Aug/Sep).
Sahoo, P., & Dash, R. K. (2014). India’s surge in modern services exports: Empirics for policy. *Journal of Policy Modeling*, *36*(6), 1082–1100.
Sihag, B. S., & McDonough, C. C. (1989). Shift-share analysis: The international dimension. *Growth and Change, 20*(3), 80-88.
Sirakaya, E., Uysal, M., & Toepper, L. (1995). Measuring tourism performance using shift-share analysis: The case of South Carolina. *Journal of Travel Research, 34*(2), 55-61.
Sudarsan, P. K., & Karmali, D. (2011). Determinants of India’s services exports: A static and dynamic analysis. *Journal of International Economics, 2*(2), 73-83.
Tadesse, B., & White, R. (2012). Do immigrants enhance international trade in services? The case of US tourism services exports. *International Journal of Tourism Research, 14*(6), 567–585.
The Economist. (1996). The China syndrome. *The Economist, 340*(Sep.), S31-S33.
White, D. S., Ariguzo, G. C. & Curran, C. M. (2013). Using time series analysis to predict U.S. service exports. *Services Marketing Quarterly, 34*(2), 175–190.
Winsted, K. F., & Patterson, P. G. (1998). Internationalization of services: The service exporting decision. *The Journal of Services Marketing, 12*(4), 294-311.
World Bank. (1995). *Global economic prospects and the developing countries.* The World Bank.
World Bank. (2021). *World development indicators*. The World Bank.
Yandle, B. (1978). Identifying brand performance by shift-share analysis. *Journal of the Academy of Marketing Science*, *6*(1-2), 126-137.
Zaccomer, G. P., & Mason, P. (2011). A new spatial shift-share decomposition for the regional growth analysis: a local study of the employment based on Italian Business Statistical register. *Statistical Methods & Applications, 20*(3), 329-356.

ORDER CARTONIZATION AND FULFILLMENT CENTER ASSIGNMENT IN THE RETAIL INDUSTRY

Ehsan Ardjmand, Ohio University
ardjmand@ohio.edu

William A. Young II, Ohio University
youngw1@ohio.edu

Shakil Rahman, Frostburg State University
srahman@frostburg.edu

ABSTRACT

Within the retail industry, after customers place orders, their orders are sent to fulfillment centers and their orders are often cartonized on a first-come-first-serve basis. This strategy has the benefit of reducing shipping costs, but it is far from an optimal solution. This paper explores an integrated mathematical model for fulfillment center assignment and simultaneous cartonization. Considering the complexity of the model for larger-scale problems that arise in real-world scenarios, a decoupled model is proposed. Moreover, a fast and effective heuristic, a genetic algorithm with modified two-stage crossover and mutation operators and a simulated annealing with two cooling schedules are also introduced. The results of applying the integrated and decoupled models using the proposed heuristic and metaheuristics are presented for various testing scenarios. The proposed heuristic and metaheuristics are shown to be more efficient when applied to testing instances that are tightly constrained.
Keywords: E-commerce; Retail; Fulfillment Center Assignment; Cartonization; Metaheuristics

INTRODUCTION

The level of competition within the retail industry is tremendously high. Retail giants like Amazon and Walmart have significantly increased their investments into their e-commerce and fulfillment operations. The sheer magnitude of these company's influence has created pressure for other retailers to implement more innovative and cost-saving operations (Gibson et al., 2015; Owyong & Yih, 2006).

Retailers can significantly reduce their costs by improving their fulfillment process (Onal et al., 2017; Thirumalai & Sinha, 2005). Although there are many goals related to the fulfillment process, an operational goal of fulfillment is to assign a product or a collection of products to as few as boxes as possible. This process, otherwise known as cartonization, can help retailers to ship their products on time at a low cost (Lin & Shaw, 1998; Ricker & Kalakota, 1999), which directly affects customer satisfaction (Espino-Rodríguez & Rodríguez-Díaz, 2014).

The fulfillment process is perhaps simple to define but complex to operationalize efficiently. At an operational level, the fulfillment process consists of several transactions that coordinate activities such as picking, cartonizing, and shipping (Croxton, 2003). Although all of these activities play a vital role in streamlining the fulfillment process, perhaps the most critical part of this process is cartonization. Cartonization is the act of assigning one or multiple products that are ordered by a customer in one or multiple cartons. Most often, the overall objective for this process is to minimize the number of cartons used to ship an order. Minimizing the number of cartons used is common due to the fact that shipping costs are inherently dependent on the number of packages that are sent from the fulfillment center (FC) to the customer and the associated distance that the carton must travel. Consequently, the problem of cartonizing and shipping are entangled problems to solve efficiently. Based on this relationship, an optimal assignment should select the most advantageous fulfillment center that minimizes shipping costs as well as the number of cartons used for a particular order while considering on-hand inventory levels.

When a retailer submits an order, the FC as well as the order's cartonization instructions are assigned based on a first-come-first-serve (FCFS) strategy as long as the FC has the necessary on-hand inventory. However, in real-world situations, inventory is often limited at a given FC. If this is the case, then it might be more advantageous to delay the cartonization and FC assignment temporarily. For example, consider the situation that consists of two orders, three possible products, one carton type option, and two FCs. Figure 1 shows each FCs on-hand inventory as well as two possible scenarios for fulfilling the orders. For this example, assume that the cost of shipping a carton from FC1 to customer 1 is $5 and the shipping cost associated with shipping a carton from FC1 to customer 2 is $7. Likewise, assume that the cost of shipping a carton from FC2 to customers 1 and 2 are $7 and $5, respectively.

Scenario 1 assumes that a FCFS strategy is employed without delaying or buffering any orders. In this scenario, order 1 is fulfilled first, since it is assumed that this order was placed prior to order 2. In this scenario, products A and C from order 1 is assigned to FC 1, and those products are placed in carton 1. Given that FC 1's inventory level for product B from order 1 is zero, the order is assigned to FC 2 and placed in carton 2. In terms of order 2, product C is assigned and cartonized by FC 1 given that FC 2 is out of stock of product C. Finally, the remainder of order 2 is cartonized and fulfilled by FC 2 given the on-hand inventory levels and costs associated with filling this order. As a result, a total of four cartons was used in scenario 1 and the total cost associated with fulfilling the orders would be $24.

Scenario 2 assumes that the cartonization and FC assignment is determined after both orders are received. For this case, when FCFS is not considered, all of the products in order 2 can be cartonized and shipped from FC 1 at a cost of $7. Given the inventory levels and associated fulfillment costs, order 1 is split between the two FCs. For example, products A and B is assigned to FC2 and is placed in carton 2, while product C is fulfilled by FC1 and is placed in carton 3. Therefore, as a result of buffering the orders, the total cost of scenario 2 would be $19, which is a $5 savings when compared to the total cost associated with scenario 1.

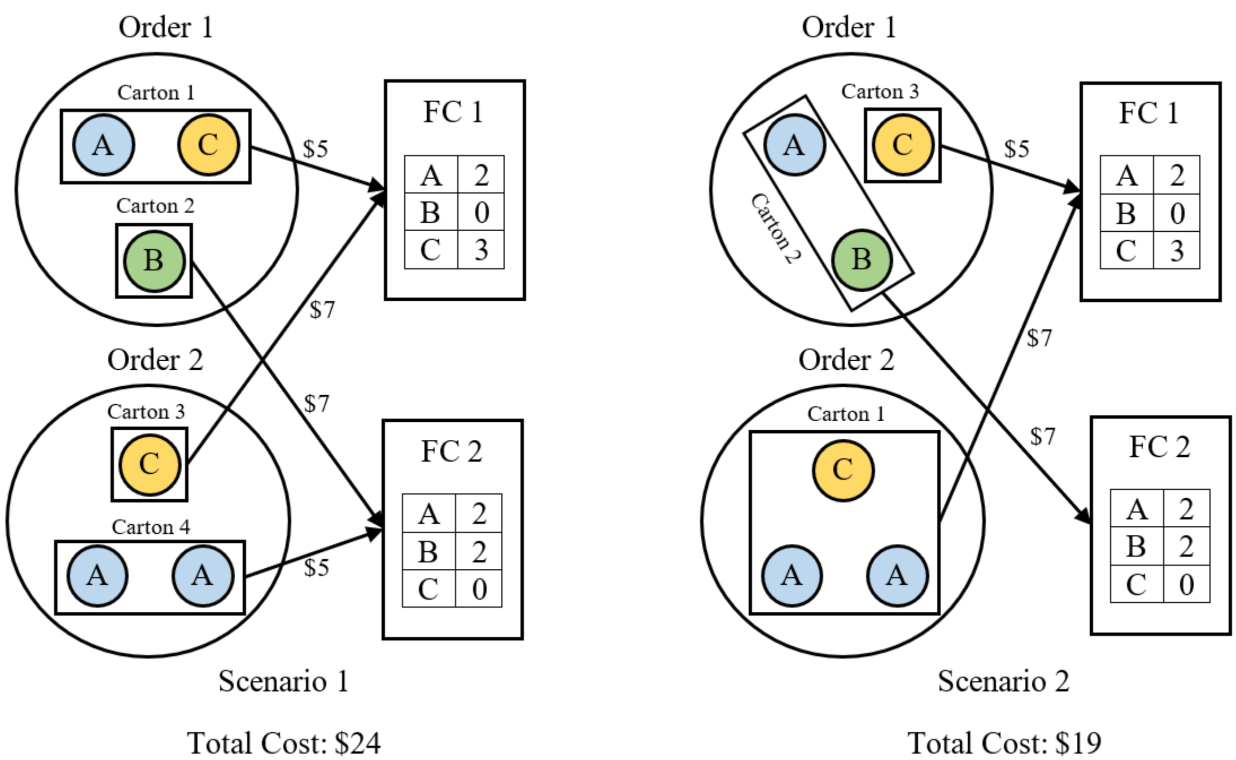

FIGURE 1. TWO CARTONIZATION AND ASSIGNMENT SOLUTIONS BASED ON FIRST-COME-FIRST-SERVE POLICY (SCENRIO 1) AND BUFFERING ORDERS (SCENARIO 2)

Given the example outlined in Figure 1, it is evident that buffering orders before they are assigned to FCs and ultimately cartonized has its benefits from a total cost perspective. However, it is not always ideal to buffer orders long enough to meet or exceed customer expectations in terms of delivery time. The pressure that retailers face of providing customers with their orders as quickly as possible is incredibly high and exacerbated given trends of 'same day delivery' or more ambitious anticipations of 'within the hour delivery.' However, it is conceivable that some retailers might receive thousands of orders to fulfill in a timeframe of less than an hour. Given the high volume, variety, and velocity of these types of situations, the task of processing these orders in a fast manner can become very computationally challenging.

The focus of the research conducted in this paper is to explore the problem of order fulfillment within the retail industry. More specifically, order cartonization and FC assignment will be studied with integrated and decoupled models using the heuristic and metaheuristics. For this particular problem, the results of applying the methods proposed in this paper were applied to various real-world testing scenarios where the utmost importance is placed how quickly orders are assigned to FCs and cartonized. The testing scenarios explored in this research consider a mixture of geographically diverse customers and FCs locations as well as variety of products that can be ordered. In terms order fulfillment, the orders will be partially or completely assigned to FCs, cartonized, and then shipped to the customers. Furthermore, in the event that an item is out of stock across FCs, the product will be outsourced to a third-party wholesaler, which is a customary practice within the retail industry. Ultimately, the objective of the research presented in this paper is to propose an efficient solution to the order cartonization and fulfillment assignment problem, where the total carton-dependent shipping and wholesaler procurement costs are minimized. Therefore, the contributions of this research include:

- The development of an integrated and decoupled mathematical model that provides fast and efficient solutions for order cartonization and FC assignments within the context of the retail industry.

- The introduction and application of a genetic algorithm with modified two-stage crossover and mutation operators as well as a simulated annealing model with two cooling schedules for the cartonization and FC assignment problem.

- The dissemination of the results of applying the proposed mathematical models to small, medium, and larger scaled problems with various testing scenarios using the developed heuristic and metaheuristics.

The remainder of this article is organized as follows: In Section 2, the related literature is reviewed and discussed. In Section 3, architecture of the proposed mathematical model is introduced. In Section 4, the methodology along with the developed and design of the heuristic and metaheuristics are more thoroughly explained. Section 5 is devoted to various numerical experiments under various testing scenarios. Finally, in Section 6, the overall conclusion of this research is presented along with a discussion of future research considerations.

LITERATURE REVIEW

In terms of the fulfillment process, one of the most critical aspects of the process is the assignment of customer orders to FCs. To aid with this important decision, researchers have investigated minimizing the procurement cost when assigning orders to FCs (Xu, 2005). The overall process of assigning orders to FC is very myopic within the retail supply chain. However, with that said, Xu (2005) proposed a method to improve the FCFS order to FC assignment policy based on a periodic review strategy. Similarly, Xu et al. (2009) developed a mathematical model as well as multiple near-optimal heuristics for order to FC assignments. The researchers found that their approach of reevaluating order fulfillment decisions significantly improve total shipping cost.

Research related to minimizing inventory, shipping cost, and customer waiting times have been investigated using a quasi-dynamic allocation policy (Mahar & Wright, 2009). Based on this work, the researchers' concluded that their proposed methodology resulted in a significant improvement upon the existing practices that utilized within a specific retailer. Minimizing outbound shipping costs has also been explored by using dual values of a transportation linear program (Acimovic & Graves, 2014). Researchers have various approaches while attempting to minimize outbound shipping costs associated with assigning customer orders to FCs. For example, this problem has been formulated as a stochastic control problem where researchers have developed solutions using deterministic linear programs (Jasin & Sinha, 2014, 2015). Another approach of minimizing outbound shipping cost can be seen in the work that was conducted by Bhargava et al. (2016). These researchers introduced an agent-based system that operated in a collaborative and geographically distributed network. One differentiating factor of their methodology is that it takes advantage of a best matching protocol. Based on their research findings, they showed promising results that their protocol provided a scalable solution for a sizable supply chain.

Although the literature exploring strategies to assign customer orders to FCs is vast, there is a limited amount of literature dedicated to research exploring assigning customer order to FCs while also simultaneously considering order cartonization. Shipping costs are inherently dependent upon cartons being transported from the FCs to customers. Thus, deriving methodologies that simultaneously consider cartonization along with the assignment of orders to FCs has significant potential in reducing costs as well as maintaining, or even improving, customer satisfaction.

Cartonizing customer orders is a problem that has similar characteristics to the three-dimensional bin packing problem (3D-BPP). The objective of the 3D-BPP involves placing a set of n rectangular-shaped boxes, where each i^{th} box has dimension length l_i, width w_i and height h_i in a minimum number of containers (i.e. cartons) having length, width, and height of L, W, and H respectively (Martello et al., 2000). Given the similarities between the problem associated with cartonization and the 3D-BPP, the related literature involving the 3D-BPP is discussed throughout this section of the article.

The 3D-BPP is a well-studied topic and researchers have proposed many approaches in their attempt to solving this particular problem. One of the first documented attempts to solve the 3D-BPP dates back to the 1960s (Gilmore & Gomory, 1965). The authors proposed an enumeration approach for the 3D-BPP problem that was appropriate for column-generation methods. However, their approach was not considered effective for dealing with larger-scale problems that exist in today's modern retail industry (Crainic et al., 2008). With that said, the methods that attempt to solve the 3D-BPP have certainly evolved tremendously since the 1960s. However, it was not until 2000 that the first exact algorithm for the 3D-BPP was introduced (Martello et al., 2000). The first exact algorithm featured a branch and bound algorithm, where the authors presented several lower bound approaches to solve the 3D-BPP. The researchers showed that the asymptotical worst-case performance of the lower bound obtained by relaxing the integer decision variables to continuous variables was 1/8. In terms of their computational experimentation, the authors showed that their approach was capable of solving instances up to 60 boxes. Since the first exact method was introduced by Martello et al. (2000), other methods have emerged. For example, Fekete et al. (2007) developed a two-level tree search algorithm for the higher dimensional orthogonal packing problem. This specific problem is a general case of 3D-BPP by characterizing the relative position of items (boxes) in a feasible packing using a graph-theoretic schema. Fekete et al. (2007) reported a considerable improvement when compared their method to existing exact methods.

In addition to exact algorithms, several heuristic solutions have been proposed for the 3D-BPP. In general, there are two categories of heuristic methods for the 3D-BPP. These heuristics can be classified as either construction or local search methods (Faroe et al., 2003). Construction heuristics add one item to a carton at a time until all of the items are packed. Local search heuristics attempt to improve a packing solution by systematically searching its neighborhood for better packing solutions (Faroe et al., 2003).

There are a number of strategies involving the process of how to add items to cartons. Wall building is a popular construction heuristic for the 3D-BPP. Wall building involves the process of creating shelves from filled cartons. For this strategy, items are assigned to shelves based on their size and a set of rules (Zhao et al., 2016). The research related to wall building methods is extensive and is the basis of many other 3D-BPP methods (Bortfeldt & Gehring, 2001; Bortfeldt et al., 2003; Chien & Deng, 2004; Moura & Oliveira, 2005; Pisinger, 2002).

Construction methods based on sequencing are also popular within the 3D-BPP literature. In sequencing methods, items are sorted based on a certain criterion before the items are placed into cartons. The assignment of items to cartons is usually determined by a first-fit or best-fit strategy (Crainic et al., 2008). In the first-fit strategy, an item is placed in the first carton that has sufficient space. In contrast, in the best-fit strategy, many cartons are considered, and an item is placed in a carton that optimizes a utility function.

Crainic et al. (2008) presented a successful implementation of construction-based heuristic that takes advantage of a modified corner-point concept that was originally proposed by Martello et al. (2000). According to Martello et al. (2000), corner points are non-dominated locations where an item can be placed into an existing carton. Crainic et al. (2008) showed that the corner points concept potentially ignores some promising placement points in a carton. As a result, the authors introduced the concept of extreme point (EP). The EP concept considers the items already placed in a carton in order to identify the free space available. As a result of using EP, Crainic et al. (2008) proposed several first-fit and best-fit algorithms for 3D-BPP. They were able to show a substantial margin of improvement by using EPs rather than corner-points.

Based on the exceptional results reported by Crainic et al. (2008), their methods will serve as the basis for the cartonization in the proposed study presented in this paper. However, it should be noted that in their original work, items could not be rotated, which is a limitation that is overcome within the presented methodology. Constraining the rotation of an item is considered a very restrictive assumption for real-world scenarios because it can increase shipping cost since more cartons will be generated. By allowing items to rotate, it complicates the 3D-BPP considerably. However, it is a consideration that is necessary for real-world cases given that FC benefit greatly by optimizing the number of cartons that are necessary to fulfil orders.

Many other heuristics exist for the 3D-BPP in the literature. However, many of which are considered beyond the scope of the problem and methods presented in this paper. Readers who are interested in exploring additional literature should refer to a comprehensive review of heuristics published by Zhao et al. (2016).

PROBLEM STATEMENT

The development of two mathematical models (i.e. integrated model and decoupled model) for the problem associated with order cartonization and FC assignment problem are proposed in this section. The first model that is presented is an integrated approach where order cartonization and FC assignments are preformed simultaneously. Although additional details will be provided later in this article, this particular approach is limited to small problems due to the complexities and large dimensionalities associated with the retail industry. To overcome this limitation, a second model is presented. The second model differs in the approach from the first model because order cartonization and FC assignments are preformed separately rather than simultaneously. It will be shown throughout this article, that the decoupled model has the benefit of being able to solve problems that are larger. However, the decoupled approach yields suboptimal results when solutions derived from the integrated approach are compared. The comparison between the integrated approach and the decoupled approach will be described in more detail in Section 5. The remainder of this section will be devoted to the presentation and formulation of the order cartonization and FC assignment problem.

Problem Definition and Assumptions

To define the order cartonization and FC assignment problem, assume that a retailer has F FCs, N different items, and receives orders from C customers. In this scenario, assume that an order is placed by each customer $c \in \{1,2, \dots, C\}$ and these orders can be fulfilled by one or more FCs. Furthermore, once an order is received by a retailer, cartonization is required. After considering the inventory levels of the FC, each carton is assigned to only one FC $f \in \{1,2, \dots, F\}$. Given that cartons are assigned to a single FC based on its ability to fulfill an order, it is possible that a single order may be fulfilled by multiple cartons and FCs. If the cost associated with shipping a carton from a FC to a customer is independent of the number of items in the carton. Thus, it is desirable to minimize the total number of cartons to fulfill orders.

When an item is assigned to a carton, the decision of placement is subject to spatial restrictions of the carton and the physical size of the other items in the carton. To place an item $i \in \{1,2, \dots, N\}$ into a carton, it is assumed that the item's physical shape can be sufficiently described as a three-dimensional cuboid having a length l_i, width w_i, and height h_i. Likewise, the physical shape of each carton can be described as a three-dimensional object having length L, width W, and height H. When items are placed in a carton, it is assumed that the item is either placed parallel or perpendicular to the carton's faces. Finally, it is assumed that that all of the cartons have similar dimensions. To model the spatial constraints of placing items in cartons, a formulation was proposed by Wu et al. (2010), and their approach will be closely implemented into the models presented in this paper.

For the approach explored in this paper, it is assumed that a retailer would prefer to fulfil orders by using the inventory stocked at one of the FCs before considering other options to fulfil the order. As a result, if a customer orders an item and the item is out-of-stock, then it is assumed that the retailer would fulfil the order by outsourcing the out-of-stock item to a wholesaler. The practice of outsourcing out-of-stock items to a wholesaler is common within the retail industry because the cost of outsourcing is dependent on the item itself and is not dependent upon the location of the customer. In other words, based on contracts between the retailer and wholesaler, the wholesaler would charge a retailer a fixed rate to fulfill the item. Thus, given this standard practice, it is assumed that all orders will be first fulfilled by a retailers FCs before out-of-stock items are fulfilled by wholesalers.

Integrated Model (M) Overview

In the following section, the indices, parameters, decision variables, objective function, and constraints for the integrated mathematical model for the order cartonization and fulfillment center assignment are presented. Furthermore, throughout this article, this model will be referred to as model M.

Indices:

$i \in \{1,2,\dots,N\}$ items
$j \in \{1,2,\dots,M\}$ cartons
$c \in \{1,2,\dots,C\}$ customers
$f \in \{1,2,\dots,F\}$ FCs

Parameters:

q^{cf}: Cost of sending a carton from FC f to customer c
I_i^f: Inventory of item i in FC f
E_i^c: 1 if item i is ordered by customer c and 0 otherwise
r_i: cost of fulfilling item i from wholesaler
(l_i, w_i, h_i): Length, width, and height of item i
(L_j, W_j, H_j): Length, width, and height of carton j

Decision Variables:

Q_{ij}^{cf}: 1 if item i ordered by customer c is sent via FC f by carton j
P_j^{cf}: 1 if carton j is used and sent from FC f to customer c
R_i^c: 1 if item i on the order from the customer c is fulfilled by the wholesaler
$(x_{ij}^{cf}, y_{ij}^{cf}, z_{ij}^{cf})$: Continuous variables for coordinates of the item i's left-bottom-behind corner in carton j
$xl_{ij}^{cf}, zl_{ij}^{cf}, yw_{ij}^{cf}, zh_{ij}^{cf}$: Binary variables indicating whether the length direction of item i is parallel to the carton's X and Z axes, the width direction is parallel to the Y axis, or the height direction is parallel to the Z axis, respectively. These variables determine the orientation of item i
$a_{imj}^{cf}, b_{imj}^{cf}, c_{imj}^{cf}$: Binary variables defining the relative placement of item i to item m: variables will be 1 if item i is in front of, to the right of, or on top of item m, respectively; otherwise, 0

Objective Function:

$$(M) \quad Min \quad \sum_{j=1}^{M}\sum_{c=1}^{C}\sum_{f=1}^{F} q^{cf} P_j^{cf} + \sum_{i=1}^{N}\sum_{c=1}^{C} r_i R_i^c$$

Constraints:

$$Q_{ij}^{cf} \leq P_j^{cf} \quad \forall i, j, c, f \tag{1}$$

$$\sum_{j=1}^{M}\sum_{f=1}^{F} Q_{ij}^{cf} + R_i^c = E_i^c \quad \forall i, c \tag{2}$$

$$\sum_{j=1}^{M}\sum_{c=1}^{C} Q_{ij}^{cf} \leq I_i^f \quad \forall i, f \tag{3}$$

$$x_{ij}^{cf} + l_i xl_{ij}^{cf} + w_i\left(zl_{ij}^{cf} - yw_{ij}^{cf} + zh_{ij}^{cf}\right) + h_i(1 - xl_{ij}^{cf} - zl_{ij}^{cf} + yw_{ij}^{cf} - zh_{ij}^{cf}) \leq x_{mj}^{cf} + M\left(1 - a_{imj}^{cf}\right) \forall c, f, j, i \neq m \tag{4a}$$

$$y_{ij}^{cf} + w_i yw_{ij}^{cf} + l_i\left(1 - xl_{ij}^{cf} - zl_{ij}^{cf}\right) + h_i(xl_{ij}^{cf} + zl_{ij}^{cf} - yw_{ij}^{cf}) \leq y_{mj}^{cf} + M\left(1 - b_{imj}^{cf}\right) \forall c, f, j, i \neq m \tag{4b}$$

$$z_{ij}^{cf} + h_i zh_{ij}^{cf} + w_i\left(1 - zl_{ij}^{cf} - zh_{ij}^{cf}\right) + l_i zl_{ij}^{cf} \leq z_{mj}^{cf} + M\left(1 - c_{imj}^{cf}\right) \forall c, f, j, i \neq m \tag{4c}$$

$$x_{ij}^{cf} + l_i xl_{ij}^{cf} + w_i\left(zl_{ij}^{cf} - yw_{ij}^{cf} + zh_{ij}^{cf}\right) + h_i\left(1 - xl_{ij}^{cf} - zl_{ij}^{cf} + yw_{ij}^{cf} - zh_{ij}^{cf}\right) + MQ_{ij}^{cf} \leq L_j + M \;\forall c, f, j, i \tag{5a}$$

$$y_{ij}^{cf} + w_i yw_{ij}^{cf} + l_i\left(1 - xl_{ij}^{cf} - zl_{ij}^{cf}\right) + h_i\left(xl_{ij}^{cf} + zl_{ij}^{cf} - yw_{ij}^{cf}\right) + MQ_{ij}^{cf} \leq W_j + M \;\forall c, f, j, i \tag{5b}$$

$$z_{ij}^{cf} + h_i zh_{ij}^{cf} + w_i\left(1 - zl_{ij}^{cf} - zh_{ij}^{cf}\right) + l_i zl_{ij}^{cf} + MQ_{ij}^{cf} \leq H_j + M \;\forall c, f, j, i \tag{5c}$$

$$-a_{imj}^{cf} - a_{mij}^{cf} - b_{imj}^{cf} - b_{mij}^{cf} - c_{imj}^{cf} - c_{mij}^{cf} + M\left(Q_{ij}^{cf} + Q_{mj}^{cf}\right) \leq 2M - 1 \quad \forall c, f, j, i \neq m \tag{6}$$

$$xl_{ij}^{cf} + zl_{ij}^{cf} \leq 1 \;\forall c, f, j, i \tag{7a}$$
$$zl_{ij}^{cf} + zh_{ij}^{cf} \leq 1 \;\forall c, f, j, i \tag{7b}$$
$$zl_{ij}^{cf} - yw_{ij}^{cf} + zh_{ij}^{cf} \leq 1 \;\forall c, f, j, i \tag{7c}$$
$$zl_{ij}^{cf} - yw_{ij}^{cf} + zh_{ij}^{cf} \geq 0 \;\forall c, f, j, i \tag{7d}$$
$$1 - xl_{ij}^{cf} - zl_{ij}^{cf} + yw_{ij}^{cf} - zh_{ij}^{cf} \leq 1 \;\forall c, f, j, i \tag{7e}$$
$$1 - xl_{ij}^{cf} - zl_{ij}^{cf} + yw_{ij}^{cf} - zh_{ij}^{cf} \geq 0 \;\forall c, f, j, i \tag{7f}$$
$$xl_{ij}^{cf} + zl_{ij}^{cf} - yw_{ij}^{cf} \leq 1 \;\forall c, f, j, i \tag{7g}$$
$$xl_{ij}^{cf} + zl_{ij}^{cf} - yw_{ij}^{cf} \geq 0 \;\forall c, f, j, i \tag{7h}$$

$$xl_{ij}^{cf} + zl_{ij}^{cf} + yw_{ij}^{cf} + zh_{ij}^{cf} + x_{ij}^{cf} + y_{ij}^{cf} + z_{ij}^{cf} \leq MQ_{ij}^{cf} \;\; \forall c, f, j, i \tag{8}$$

$$Q_{i_1 j}^{c_1 f_1} + Q_{i_2 j}^{c_2 f_2} \leq 1 \;\; \forall i_1, i_2, j, c_1 \neq c_2, f_1 \neq f_2 \tag{9}$$

$$Q_{ij}^{cf}, P_j^{cf}, R_i^c, xl_{ij}^{cf}, zl_{ij}^{cf}, yw_{ij}^{cf}, zh_{ij}^{cf}, a_{imj}^{cf}, b_{imj}^{cf}, c_{imj}^{cf} \in \{0,1\}\ \forall c, f, j, i \tag{10a}$$

$$x_{ij}^{cf}, y_{ij}^{cf}, z_{ij}^{cf} \in \mathbb{R} \tag{10b}$$

Integrated Model (M) Summary

In model M, the objective function minimizes the total cost of shipping cartons and fulfilling the items from the wholesaler. Constraint 1 guarantees that an item is placed in a carton if that carton is used. Constraint 2 ensures that all of the orders are fulfilled, whether from the FCs network or the wholesaler. Constraint 3 ensures inventory level does not fall below zero. Constraint 4 prevents items from overlapping each other in a carton. Constraint 5 keeps items inside the cartons' dimensions. Constraint 6 restricts the relative position of items in a carton. Constraint 7 guarantees that items are located in practical positions while controlling the positioning variables. Constraint 8 prevents rotation of an item in a carton if the item is not placed in the carton. Constraint 9 inhibits placement of two items from two different customers or two different FCs in the same carton.

The integrated model represented as model *M* is NP-hard. This relationship can be shown by constructing the order cartonization and FC assignment problem, which will be denoted as P from this point forward in the article, as a 3D-BPP (Martello et al., 2000). Given any instance of 3D-BPP, it is possible to construct an instance of P by assuming one order that includes all items of the 3D-BPP instance and one FC that stores all items of the 3D-BPP instance with infinite inventory. A feasible solution exists for P if and only if there is a solution for the 3D-BPP. When a solution for the 3D-BPP exists, then P has an obvious solution of assigning all items to a FC. In this situation, the solution of the 3D-BPP represents the cartonization solution of problem P. Now, suppose there exists a solution to the instance of problem P. Since there is only one FC with unlimited inventory, only one solution to the assignment part of P exists and the 3D-BPP instance should also have a solution. This shows that the cartonization and FC assignment problem of this study is NP-hard.

Decoupled Model (M1 & M2) Overview

Model M, or the integrated model, was designed to simultaneously generate solutions for the order cartonization and FC assignment problem. As it will be shown later in Section 5, model *M* was capable of finding global optimum solutions for small-scaled problems. However, it will also be shown that model *M* was incapable of finding global optimum solutions for medium or large-scaled problems. To overcome this limitation, model *M* can be separated into two models. Simply stated, the first model can address the problem related to cartonizing orders received by customers by minimizing the number of cartons used, while a second model address the problem of assigning cartons created by the first model to the FCs or the wholesaler by minimizing the overall fulfillment costs. From the perspective of this paper, the first model that addresses the problem of cartonization will be referred to as model *M1*, while the second model addressing the problem of assigning cartons to FCs or wholesaler will be referred to as *M2*.

To further explain the concepts of models *M1* and *M2*, an example is shown in Figure 2. In this example, there are four customers, four items, and three FCs. In order to minimize the number of cartons used, *M1* solves a bin packing problem for each customer order separately. The solutions obtained from *M1* are considered primary cartons, and these results serve as inputs to model *M2*. Model *M2* is designed to minimize the total shipping and wholesaler fulfillment costs by assigning the primary cartons to either the FCs or the wholesaler. In situations where there is not sufficient inventory to fulfill the order from the FCs, *M2* will split a primary carton into two to create two separate secondary cartons. When fulfilling the secondary cartons, *M2* is designed so that it will determine whether the carton is fulfilled by other FCs, or it is fulfilled by outsourcing the order by the wholesaler.

Model M simultaneously considers cartonization and FC assignment problems. As will be shown through numerical examples later, while resulting in the global optimum, model M cannot be used for medium- or large-scale problems. One way to overcome this problem is to break model M into two separate models: $M1$ and $M2$.
FIGURE 2 depicts the order of models $M1$ and $M2$ in an example with four customers, four items, and three FCs. Model $M1$ is only concerned with cartonizing orders received by customers with the objective of minimizing the number of cartons. In other words, model $M1$ solves a bin packing problem for each customer separately. The cartons created by model $M1$ are primary cartons and are used as the input to model $M2$. Model $M2$ is only concerned with assigning cartons created by model $M1$ to the FCs or the wholesaler in order to minimize the overall shipping and wholesaler fulfillment costs. For this purpose, model $M2$ breaks a carton into two (secondary cartons) if there is not enough inventory in any FC to fulfill the whole carton. Moreover, model $M2$ decides if an item in a carton is outsourced to the wholesaler.

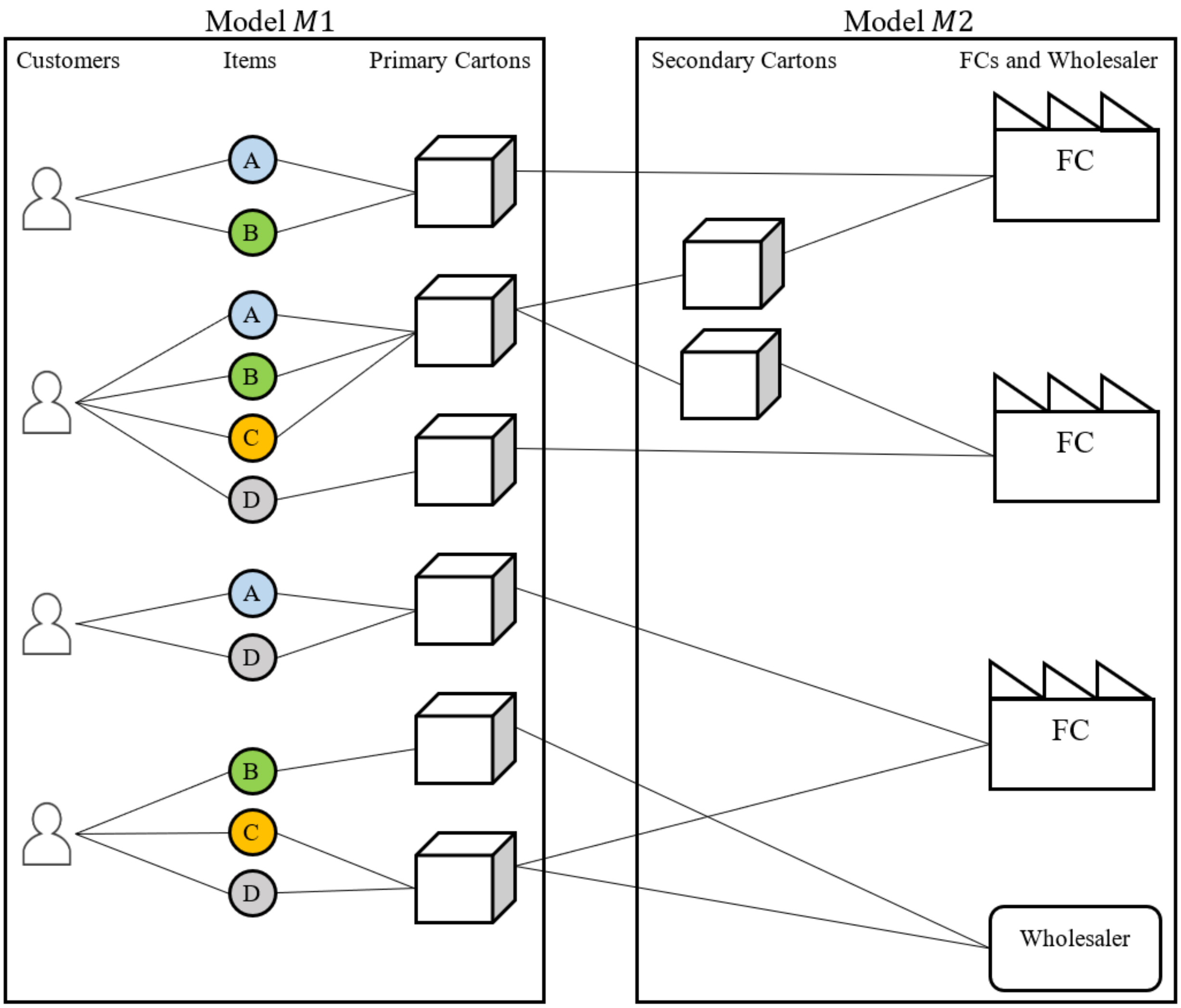

FIGURE 2. MODELS $M1$ AND $M2$ AND THEIR CONFIGURATION FOR ORDER CARTONIZATION AND FC ASSIGNMENT

Decoupled Model (M1) Formulation

Formulating the order cartonization and FC assignment problem as a decoupled model overcomes the complexities of the integrated model. Although the decoupled model has its advantages that will be discussed later, it does have a possibility of a disadvantage of not being able to find global optimal solutions. However, there are two scenarios in which *M1* and *M2* will yield optimum solutions. The first situation where *M2* is guaranteed to produces global optimums is when the output of *M1* places each customer's order in a single carton. The second scenario in which the decoupled model is guaranteed to find global optimums is when a FCFS strategies is implemented and FCs has substantial quantities of on-hand inventory. Although this scenario is capable of finding optimal solutions, it rarely occurs in reality within the retail industry.

In the following sections, the indices, parameters, decision variables, objective function, and constraints for model *M1* are presented.

Indices:
E_c*:* Set of items related to customer c (item i is in E_c if E_i^c is equal to 1*)*
$i \in E_c$
$j \in \{1,2, \dots, M\}$ cartons

Parameters:
(l_i, w_i, h_i): Length, width, and height of item i
(L_j, W_j, H_j): Length, width, and height of carton j
QQ_{ij}^c: 1 if item i ordered by customer c and placed in carton j $(\forall c \backslash \{1\})$

Decision Variables:

Q_{ij}: 1 if item i is placed in carton j

P_j: 1 if carton j is used and 0 otherwise

(x_{ij}, y_{ij}, z_{ij}): Continuous variables for coordinates of item i's left-bottom-behind corner in carton j

$xl_{ij}, zl_{ij}, yw_{ij}, zh_{ij}$: Binary variables indicating whether the length direction of item i is parallel to the carton's X and Z axes, the width direction is parallel to the Y axis, or the height direction is parallel to the Z axis, respectively. These variables determine the orientation of item i, $a_{imj}, b_{imj}, c_{imj}$: Binary variables defining the relative placement of item i to item m: variables will be 1 if item i is in front of, to the right of, or on top of item m, respectively; otherwise, 0

Objective Function:

$$(M1) \quad Min \quad \sum_{j=1}^{M} P_j^c$$

Constraints:

$$Q_{ij} \leq P_j \quad \forall i, j \tag{11}$$

$$\sum_{j=1}^{M} Q_{ij}^c = E_i^c \quad \forall i \tag{12}$$

$$\begin{aligned} x_{ij}^c + \ l_i xl_{ij}^c + \ w_i\left(zl_{ij}^c - \ yw_{ij}^c + zh_{ij}^c\right) + \ h_i(1 - xl_{ij}^c - zl_{ij}^c + \ yw_{ij}^c - \ zh_{ij}^c) \\ \leq \ x_{mj}^c + M\left(1 - a_{imj}^c\right) \forall j, i \neq m \end{aligned} \tag{13a}$$

$$\begin{aligned} y_{ij}^c + \ w_i yw_{ij}^c + \ l_i\left(1 - \ xl_{ij}^c - \ zl_{ij}^c\right) + \ h_i(xl_{ij}^c + zl_{ij}^c - \ yw_{ij}^c) \\ \leq \ y_{mj}^c + M\left(1 - b_{imj}^c\right) \forall j, i \neq m \end{aligned} \tag{13b}$$

$$z_{ij}^c + \ h_i zh_{ij}^c + \ w_i\left(1 - zl_{ij}^c - zh_{ij}^c\right) + \ l_i zl_{ij}^c \leq \ z_{mj}^c + M\left(1 - c_{imj}^c\right) \forall j, i \neq m \tag{13c}$$

$$x_{ij}^c + l_i xl_{ij}^c + w_i\left(zl_{ij}^c - yw_{ij}^c + zh_{ij}^c\right) + h_i\left(1 - xl_{ij}^c - zl_{ij}^c + yw_{ij}^c - zh_{ij}^c\right) + MQ_{ij}^c \leq L_j + M \; \forall j, i \tag{14a}$$

$$y_{ij}^c + w_i yw_{ij}^c + l_i\left(1 - xl_{ij}^c - zl_{ij}^c\right) + h_i\left(xl_{ij}^c + zl_{ij}^c - yw_{ij}^c\right) + MQ_{ij}^c \leq W_j + M \; \forall j, i \tag{14b}$$

$$z_{ij}^c + h_i zh_{ij}^c + w_i\left(1 - zl_{ij}^c - zh_{ij}^c\right) + l_i zl_{ij}^c + MQ_{ij}^c \leq H_j + M \; \forall j, i \tag{14c}$$

$$-a_{imj}^c - a_{mij}^c - b_{imj}^c - b_{mij}^c - c_{imj}^c - c_{mij}^c + M\left(Q_{ij}^c + Q_{mj}^c\right) \leq 2M - 1 \quad \forall j, i \neq m \tag{15}$$

$$xl_{ij}^c + zl_{ij}^c \leq 1 \; \forall j, i \tag{16a}$$

$$zl_{ij}^c + zh_{ij}^c \leq 1 \; \forall j, i \tag{16b}$$

$$zl_{ij}^c - yw_{ij}^c + zh_{ij}^c \leq 1 \; \forall j, i \tag{16c}$$

$$zl_{ij}^c - yw_{ij}^c + zh_{ij}^c \geq 0 \; \forall j, i \tag{16d}$$

$$1 - xl_{ij}^c - zl_{ij}^c + yw_{ij}^c - zh_{ij}^c \leq 1 \; \forall j, i \tag{16e}$$

$$1 - xl_{ij}^c - zl_{ij}^c + yw_{ij}^c - zh_{ij}^c \geq 0 \; \forall j, i \tag{16f}$$

$$xl_{ij}^c + zl_{ij}^c - yw_{ij}^c \leq 1 \; \forall j, i \tag{16g}$$

$$xl_{ij}^c + zl_{ij}^c - yw_{ij}^c \geq 0 \; \forall j, i \tag{16h}$$

$$xl_{ij}^c + zl_{ij}^c + yw_{ij}^c + zh_{ij}^c + x_{ij}^c + y_{ij}^c + z_{ij}^c \leq MQ_{ij}^c \quad \forall j, i \tag{17}$$

$$QQ_{i_1 j}^{c_1} + Q_{i_2 j}^c \leq 1 \quad \forall i_1, i_2, j, c_1 < c \backslash \{1\} \tag{18}$$

$$Q_{ij}, P_j, xl_{ij}, zl_{ij}, yw_{ij}, zh_{ij}, a_{imj}, b_{imj}, c_{imj} \in \{0,1\} \tag{19a}$$

$$x_{ij}^c, y_{ij}^c, z_{ij}^c \in \mathbb{R} \tag{19b}$$

Decoupled Model (M1) Summary

In model $M1$, the objective function minimizes the number of cartons. Constraints 11 to 17 are similar to model M. Note that while solving model $M1$ for each customer, in order to prevent using the wrong cartons for different customers, the models solved for each customer need to be linked. Thus, constraint 18 is added. In this constraint, QQ_{ij}^c is a parameter equal to the value of decision variable Q_{ij}^c related to the previous customers. In the other words, the value of decision variable Q_{ij}^c in model $M1$ for customer c is considered as an input parameter (QQ_{ij}^c) in model $M1$ for the next customer. Obviously, constraint 18 in model $M1$ is not needed for the first customer.

Decoupled Model (M2) Formulation

In the following sections, the indices, parameters, decision variables, objective function, and constraints for model *M2* are presented.

Indices
$i \in \{1,2,\ldots,N\}$ items
$j' \in \{1,2,\ldots,M_1\}$ primary cartons
$j'' \in \{1,2,\ldots,M_2\}$ secondary cartons
$f \in \{1,2,\ldots,F\}$ FCs

Parameters:
${q'}^{f}_{j'}$: Cost of fulfilling carton j' from FC f
r_i: cost of fulfilling item i from wholesaler
I_i^f: Inventory of item i in FC f
$e_i^{j'}$: 1 if carton j' has item i in it

Decision Variables:
$\alpha^f_{ij'j''}$: 1 if item i in carton j' is placed in carton j'' and fulfilled from FC f otherwise 0
$\beta_{ij'}$: 1 if item i in carton j' will be fulfilled by wholesaler, otherwise 0
$\gamma^f_{j'j''}$: 1 if carton j'' includes items from carton j' and is fulfilled via FC f, otherwise 0

Objective Function:

$$(M2) \quad Min \quad \sum_{j'=1}^{M_1}\sum_{j''=1}^{M_2}\sum_{f=1}^{F} {q'}^{f}_{j'}\gamma^f_{j'j''} + \sum_{i=1}^{N}\sum_{j'=1}^{M_1} r_i\beta_{ij'}$$

Constraints:

$$\alpha^f_{ij'j''} \le \gamma^f_{j'j''} \quad \forall i,j',j'',f \tag{20}$$

$$\sum_{j''=1}^{M_2}\sum_{f=1}^{F}\alpha^f_{ij'j''} + \beta_{ij'} = e_i^{j'} \quad \forall i,j' \tag{21}$$

$$\sum_{j'=1}^{M_1}\sum_{j''=1}^{M_2}\alpha^f_{ij'j''} \le I_i^f \quad \forall i,f \tag{22}$$

$$\sum_{f=1}^{F}\gamma^f_{j'j''} \le 1 \quad \forall j',j'' \tag{23}$$

$$\sum_{j'=1}^{M_1}\gamma^f_{j'j''} \le 1 \quad \forall j'',f \tag{24}$$

$$\alpha^f_{ij'j''}, \beta_{ij'}, \gamma^f_{j'j''} \in \{0,1\} \tag{25}$$

Decoupled Model (M2) Summary

In model $M2$, the objective function minimizes the total cost of shipping cartons from FCs to customers and fulfilling items from the wholesaler. Constraint 20 guarantees that if a secondary carton is not used, no item will be placed in it. Constraint 21 ensures that all items are fulfilled whether by the FCs or the wholesalers. Constraint 22 ensures inventory level does not fall below zero. Constraint 23 guarantees that each secondary carton can get items from one primary carton only and can be fulfilled from only one FC.

METHODOLOGY

In this section, a genetic algorithm with modified two-stage crossover and mutation operators and a two-stage simulated annealing are introduced as an order cartonization and FC assignment heuristic (OCFAH).

Order Cartonization and FC Assignment Heuristic (OCFAH)

OCFAH consists of four steps, which include assigning items to FCs cartonization reassigning cartons to FCs and the wholesaler and combining cartons. Each of these steps will be discussed in detail.

Step 1: Assigning Items to FCs

The first step in OCFAH is to determine the FC that each item on a customer's order is supposed to be fulfilled from. In this step, the assignment of items to FCs is done regardless of FCs' inventory. This assignment has a profound effect on the final result of the algorithm. Multiple methods for this assignment can be used. In the first method, each item on a customer's order will be assigned to the closest FC. Hence, for an item $i \in \{1,2, \dots, N\}$ ordered by the customer $c \in \{1,2, \dots, C\}$, the assigned FC is $\underset{f}{\operatorname{argmin}}\, q^{cf}$. In the second method, each item will be randomly assigned to a FC. Approximately in all instances of the problem tried in this study, it was found that closest FC assignment yields superior results compared to random assignment. Thus, only closest FC assignment is used in this paper.

Step 2: Cartonization

After the assignment of items to FCs are determined, it is possible to identify the items that go together in one carton. Obviously, two items ordered by the same customer cannot be placed in the same carton if fulfilled by two different FCs. Similar to the bin packing problem, in the cartonization phase, the goal is to create the least number of cartons. To achieve this goal, after identifying items that can be placed in the same carton, a first-fit heuristic, similar to the one proposed by Crainic et al. (2008), will be used to cartonize the items.

In this method, an item can be placed only on extreme points in a carton. Extreme points (EPs) are the points created after an item is added to a carton, and EPs are the candidates for accommodating the new items. Every time that an item i with length l_i, width w_i, and height h_i is placed on EP (x_k, y_k, z_k), it creates a series of new EPs with coordinates $(x_k + l_i, y_k, z_k)$, $(x_k, y_k + w_i, z_k)$ and $(x_k, y_k, z_k + h_i)$. Note that unlike the algorithm proposed by Crainic et al. (2008), in this study, the item's rotation is allowed.

Every time the placement of an item on an EP is evaluated, the dimensional and geometrical constraints of the carton and the items in it should be considered. For this, it is necessary to measure the distance between the chosen EP and the faces of the items already placed in the carton, as well as the faces of the carton itself. Each existing face g in a carton can be represented by four points (X_{g1}, Y_{g1}, Z_{g1}), (X_{g2}, Y_{g2}, Z_{g2}), (X_{g3}, Y_{g3}, Z_{g3}), and (X_{g4}, Y_{g4}, Z_{g4}), which are located at the extreme ends of it. If an EP cannot be connected to a face F by a perpendicular line, that face does not impose any geometrical constraint on the EP. For a face g to geometrically constrain an EP (x_k, y_k, z_k), only one or none of the following conditions should be satisfied:

$$(x_k + \epsilon - X_{g1} \leq 0) \wedge (x_k + \epsilon - X_{g2} \leq 0) \wedge (x_k + \epsilon - X_{g3} \leq 0) \wedge (x_k + \epsilon - X_{g4} \leq 0) \quad (26)$$

$$(x_k + \epsilon - X_{g1} \geq 0) \wedge (x_k + \epsilon - X_{g2} \geq 0) \wedge (x_k + \epsilon - X_{g3} \geq 0) \wedge (x_k + \epsilon - X_{g4} \geq 0) \quad (27)$$

$$(y_k + \epsilon - Y_{g1} \leq 0) \wedge (y_k + \epsilon - Y_{g2} \leq 0) \wedge (y_k + \epsilon - Y_{g3} \leq 0) \wedge (y_k + \epsilon - Y_{g4} \leq 0) \quad (28)$$

$$(y_k + \epsilon - Y_{g1} \geq 0) \wedge (y_k + \epsilon - Y_{g2} \geq 0) \wedge (y_k + \epsilon - Y_{g3} \geq 0) \wedge (y_k + \epsilon - Y_{g4} \geq 0) \quad (29)$$

$$(z_k + \epsilon - Z_{g1} \leq 0) \wedge (z_k + \epsilon - Z_{g2} \leq 0) \wedge (z_k + \epsilon - Z_{g3} \leq 0) \wedge (z_k + \epsilon - Z_{g4} \leq 0) \quad (30)$$

$$(z_k + \epsilon - Z_{g1} \geq 0) \wedge (z_k + \epsilon - Z_{g2} \geq 0) \wedge (z_k + \epsilon - Z_{g3} \geq 0) \wedge (z_k + \epsilon - Z_{g4} \geq 0) \quad (31)$$

In the inequalities 26-31, ϵ is a small number. If d_{kg}^x represents the distance between an EP k and a face g along the x axis, then $\min_{g \in G} d_{kg}^x$, in which G is the set of all the existing faces in the carton, should be greater than the maximum dimension of an item along the x axis that is placed on EP k. The same argument applies to the y and z axis.

To cartonize items, they are first sorted decreasingly based on their volume. Starting with one empty carton, each sorted item is added to the first carton with an EP that can accommodate the item. Note that EPs that have the lowest z will be evaluated first in a carton. If two EPs have the same z, the one with the lowest y is prior, and finally, if z and y are equal, the EP with lowest x will be evaluated first. If an item cannot be placed on any of the EPs in one of the existing cartons, a new carton will be created and the item will be added to it. Algorithm 1 shows the pseudo-code for extreme point based first-fit heuristic for cartonization.

ALGORITHM 1. EXTREME-POINT-BASED FIRST-FIT HEURISTIC FOR CARTONIZATION

```
Sort the items based on their volume
Create one carton
For each sorted item i:
    For each carton j:
        For each EP k in carton j in order of lowest z,y and then x:
            If item i or any of its rotations can be placed on EP k:
                Place the item i on EP k
                Update list of EPs in carton j
            Else
                Create a new carton
                Place item i in the new carton
                Update the list of EPs in the new carton
```

Step 3: Reassigning Cartons to FCs and Wholesaler

In step 1, the assignment of items to FCs was determined. Hence, after cartonization, it is clear to which FC each carton belongs. However, due to inventory constraints, this assignment may not be feasible. The goal of reassigning cartons to FCs and wholesalers is to turn the solution obtained at the end of step 2 into a feasible, high-quality solution.

Due to inventory constraints, if a FC cannot fulfill the cartons assigned to it, some of the items in cartons need to be placed in another carton(s) and assigned to another FC, or even assigned to the wholesaler. This process can create an extra number of cartons and increase shipping cost. The idea behind dividing the cartons in a way to have a minimum effect on total shipping cost is simple yet effective. It is better to divide a carton that has fewer items, and the difference between cost of fulfilling its items from the wholesaler and FC is minimal. Cartons that contain more items help to save more on shipping cost because the shipping cost is paid per carton. The difference between the cost of FC f and wholesaler fulfillment for item i can be calculated by $r_i - q^{cf}$. The value for all of the items inside a carton j would be $\sum_{i\ in\ carton\ j} r_i - q^{cf}$.

In the carton reassignment step, all cartons are sorted by the number of items in them and then by $\sum_{i\ in\ carton\ j} r_i - q^{cf}$. Then, for a FC f where the inventory constraint for an item i' is violated, the carton with the fewest number of items and $\sum_{i\ in\ carton\ j} r_i - q^{cf}$ that include item i' will be picked and item i' will be placed in a separate carton and assigned to the FC $\underset{f}{\operatorname{argmin}}\, q^{cf}$. If no such FC can be found, item i' will be assigned to the wholesaler. Then, the sorted list of cartons will be updated to include the newly created and modified cartons. This process will be continued for all FCs and items where the inventory constraint is violated until a feasible solution is found. Algorithm 2 shows the pseudo code for reassigning cartons to FCs and the wholesaler.

ALGORITHM 2. PSEUDO CODE FOR REASSIGNING CARTONS TO FCs AND THE WHOLESALER

```
Sort the cartons based on the number of items in them and ∑_{i in carton j} r_i − q^{cf}
For all the FCs f:
    For all the items i:
        If inventory constraint for FC f and item i is not satisfied:
            Find the first carton in the sorted list of cartons that include i and is fulfilled from f
            Separate the item i from the carton and assign it to FC argmin_f q^{cf} that has inventory
            If no FC can be found, assign the item i to the wholesaler
            Update the sorted list of cartons to include the newly created and modified cartons
```

Step 4: Combining Cartons

During step 3, where the cartons are reassigned, it is possible that a carton can be divided into several cartons, and each carton assigned to the same FC. This scenario creates unnecessary additional cartons. To avoid this problem, in step 4, OCFAH searches for such instances where it is possible to combine cartons that are assigned to the same FCs and are ordered by the same customer and, if found, combine them using the first-fit heuristic in step 2.

An Example

To illustrate how OCFAH works, this section provides a numerical example. Consider a problem with three customers, two FCs and five items as shown in Figure 3. Inventory of each item in each FC is shown inside FCs. Cost of shipping one carton from a FC to a customer is shown above dotted lines and the items ordered by each customer is presented above each customer. The cost of fulfilling one unit of items A, B, C, D and E from a wholesaler is assumed to be \$3, \$4, \$4, \$4 and \$3 respectively. Additionally, let us assume that all items ordered by one customer, except A, can fit in on carton.

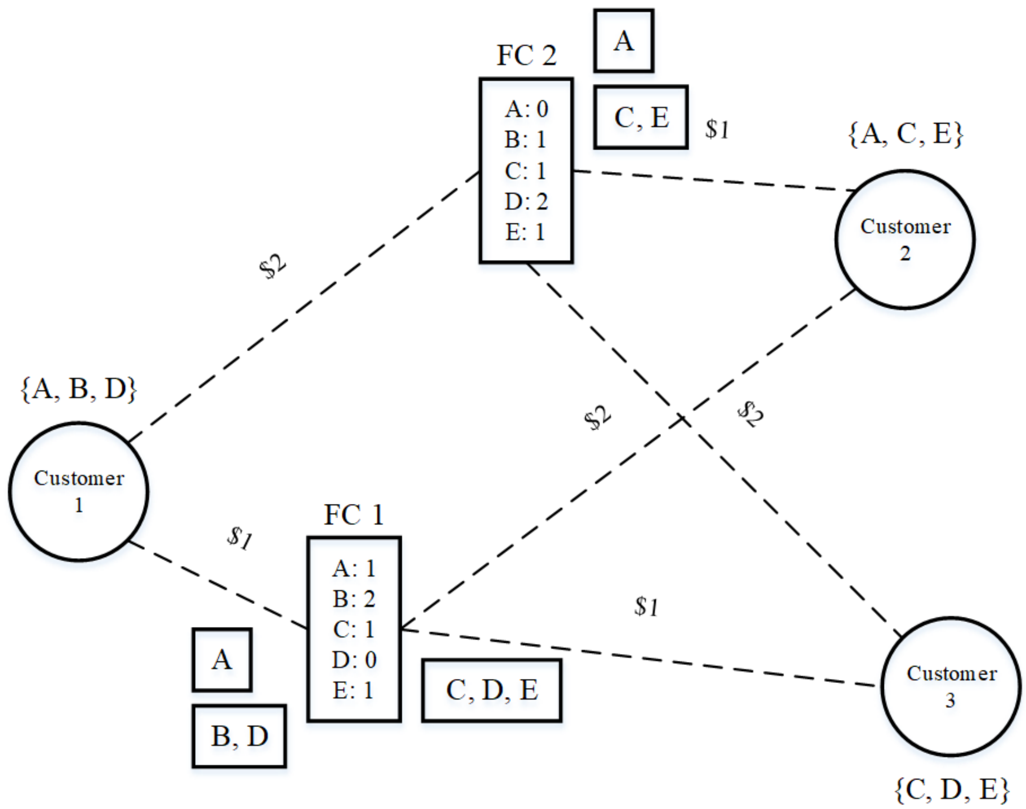

FIGURE 3. AN EXAMPLE WITH THREE CUSTOMERS AND TWO FCs AND FIVE ITEMS

To minimize the total fulfillment cost using OCFAH, the following steps are to be taken:

- *Step 1 - Assigning Items to FCs:* In this step, each item is assigned to the closest FC. Thus, all items ordered by customers 1 and 3 are assigned to FC1 and all items ordered by customer 2 are assigned to FC2.
- *Step 2 - Cartonization:* Since all items ordered by one customer (except A) can fit into one carton, five cartons will be created, as depicted by small rectangular shapes close to each FC. The output of this step is determined using the cartonization algorithm explained previously.
- *Step 3 - Reassigning Cartons to FCs and Wholesaler:* Since there is not enough inventory in the FCs to fulfill all cartons assigned to them, some of the cartons need to be modified. For this purpose, first, cartons are sorted based on the number of items in them and the summation of difference between the cost of FC and wholesaler fulfillment for items in them (i.e. $\sum_{i \ in \ carton \ j} r_i - q^{cf}$ for carton j). For instance, for item C in carton (C, D, E) the difference between the cost of FC and wholesaler fulfillment is 4 – 1 = 3. Thus, for carton (C, D, E) total difference between the cost of FC and wholesaler fulfillment is (4 -

1) + (4 - 1) + (3 - 1) = 8. Sorting cartons based on the number of items and the difference between FC and wholesaler fulfillment costs, carton (C, D, E) will be the last carton that one may want to modify as it is carrying the largest number of items. Similarly, since cartons (A) and (A) contain the smallest number of items and both have the lowest difference in FC and wholesaler fulfillment cost, they will be the first choice for modifying in face of inventory constraints. The list of sorted cartons are (A), (A), (B, D), (C, E) and (C, D, E). Starting from FC1, carton (A) can be fulfilled. However, carton (B, D) cannot be fulfilled as FC1 does not have enough inventory of item D. Thus, item D will be inserted to another carton and assigned to the closest FC where there is sufficient inventory (i.e. FC 2). If there is no such a FC, item D should be assigned to the wholesaler. Now, the cartons created should be sorted again based on the number of items in them and the total difference between FC and wholesaler fulfillment cost. This process continues until all items are properly assigned to a FC or the wholesaler.

- *Step 4 - Combining Cartons:* After cartons are reassigned to the FCs, it is possible that some cartons assigned to the same FC can be combined. To prevent creation of excess cartons, OCFAH checks all cartons that are ordered by the same customer and fulfilled from the same FC to find if they can be combined and hence, reduce the number of cartons created.

Two-Stage Crossover and Mutation Genetic Algorithm (TGA)

In this section, main features of TGA including genetic representation, fitness function, genetic operators and seeding are explained.

Genetic Representation

In TGA, each chromosome represents a list of customers and the items they have ordered (Figure 4). If a customer orders two units of the same item, each unit of the item will be represented separately in the chromosomes. In Figure 4, customer 2 has ordered two units of item 4. Thus, there are two loci representing the same item for customer 2.

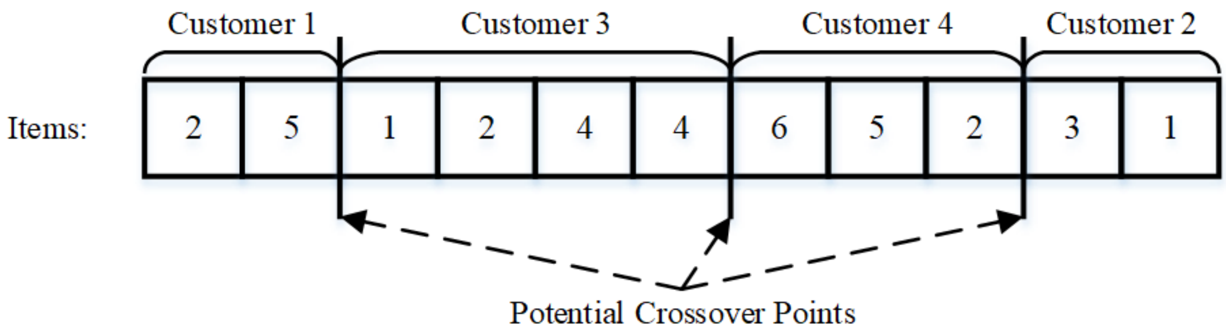

FIGURE 4. A CHROMOSOME REPRESENTING 4 CUSTOMERS AND THE ITEMS THEY HAVE ORDEREXD

Fitness Function

In TGA, the fitness of a chromosome is equivalent to its total fulfillment cost. To evaluate the fitness of a chromosome, it is processed from left to right. For this purpose, each item ordered by a customer is assigned to the closest FC that has sufficient inventory. If there is no such FC, the item will be outsourced to the wholesaler. For instance, in Figure 4, item 2 for customer 1 will be processed first and assigned to the closest FC with sufficient inventory of item 2. Then, item 5 for customer 1 will be assigned to the closest FC where there is sufficient inventory of item 5. Obviously, items ordered by customer 3 will be processed after customer 1.

After all assignments are determined, the items assigned to the same FC for each customer will be cartonized together using the method explained in step 2 of OCFAH. At this point, assignment of items to FCs as well as cartonization decisions are determined and the total fulfillment cost can be calculated and used as the fitness value of a chromosome.

Genetic Operators

The order of customers and items in a chromosome significantly affects the fitness value of the chromosome. Thus, to explore the search space of the problem properly, it is necessary to have genetic operators that can generate diverse solutions while preserving the desirable properties of high-quality chromosomes. During the experimentations, the authors noticed that given a chromosome structure as introduced earlier, conventional GA operators are not capable of finding high quality solutions. To illustrate this, consider the chromosome of Figure 4. A conventional crossover for these types of chromosomes where repetitions should be avoided is the PMX crossover (Goldberg & Lingle, 1985). To preserve the integrity of chromosomes and avoid infeasible solutions, crossover points can only be located between customers. Thus, after crossover, only the order of customers will change. In this case, PMX can explore the feasible space only by reordering the customers while the sequence of items for each customer will be ignored. To overcome this flaw, a two-stage PMX crossover is proposed in this study.

In the two-stage PMX crossover, two chromosomes crossover at two levels before producing an offspring. The first crossover set takes place among parts of two selected chromosomes that correspond to the same customer using a PMX structure. For instance, if a problem has 10 customers, each chromosome will have 10 sections where the ones that correspond to the same customer will crossover. Thus, at the first level there will be 10 PMX crossovers. This will contribute to finding the best sequence of items for each customer in a chromosome. In the second stage of crossover, the two chromosomes undergo a PMX crossover where the order of customers' locations are altered. A two-stage PMX crossover can sufficiently search the feasible space while diversifying the solutions.

Similar to crossover, a two-stage mutation alters a chromosome by first changing the sequence of items for a customer and then the sequence of customers in the chromosome. For items, this is performed by inversing, inserting and swapping the items for each customer and creating three neighbor solutions. Then the solution with the best fitness is selected. For customers, the mutation is performed by inversing, inserting and swapping the customers in a chromosome and then picking the resulting chromosome with the best fitness.

Inverse operation works by inversing the order of units (either items or customers) between positions i and j in a chromosome. Here, depending on the stage where the mutation operator is performing at, a position is considered as the location of items for each customer or customers in the chromosome. If the ith position of chromosome ω is shown by $\omega(i)$, then the inverse operation using positions i and j creates a new chromosome ω' such that $\omega'(i) = \omega(j), \omega'(i+1) = \omega(j-1), \ldots, \omega'(j) = \omega(i)$. Note that $1 \leq i, j \leq n$ and $1 \leq j - i < n - 1$ where n is the chromosome length.

Insert operation moves an order-SKU in position j of a chromosome to position i ($j > i$). Thus, by applying the insert operation on chromosome ω a new chromosome will be generated such that $\omega'(i) = \omega(j), \omega'(i+1) = \omega(i), \ldots, \omega'(j) = \omega(j-1)$. Swap operation swaps two order-SKUs i and j. Thus, by applying the insert operation on chromosome ω a new chromosome will be generated such that $\omega'(i) = \omega(j)$ and $\omega'(j) = \omega(i)$.

The selection process in this study is implemented by choosing the best k chromosomes from a population. A $(\mu + \lambda)$ selection strategy is used in which μ and λ represent the number of parents and children, respectively.

Seeding

In problems with numerous variables and large chromosome size, seeding shortens the convergence time. In this study, the first population generated is seeded by incorporating a chromosome that represents the solution obtained by OCFAH. The authors have observed that this type of seeding significantly contributes to the quality of solutions and convergence time.

Fine-Tuning the TGA's Parameters

In this research, analysis of variance (ANOVA) is utilized to determine the parameters of the GA. Number of parents (μ), number of children (λ), number of generations, crossover and mutation rate are the influencing parameters in the GA. The GA's tuning parameters and their levels considered in ANOVA are listed in Table 1. For each combination of parameters' levels, four replications were generated. The test problem chosen for conducting the ANOVA is a medium tight instance with 15 customers, 15 FCs and 50 items, which will be introduced in section 5. Based on ANOVA's results the TGA's parameters, μ, λ, number of generations, crossover and mutation rate were determined to be 100, 100, 300, 0.5 and 0.05.

TABLE 1. RESULTS OF ANOVA FOR TUNING THE GA's PARAMETERS

Levels	μ	λ	No. Generations	Crossover Prob.	Mutation Prob.
1	50	50	100	0.3	0.05
2	75	75	200	0.4	0.1
3	100	100	300	0.5	0.15

Two-Stage Simulated Annealing (TSA)

In its original form, simulated annealing (SA) starts from a solution x and in each step generates a new neighbor solution y. If y is better in terms of solution quality, it will be accepted as the new solution, otherwise it will be accepted based on the acceptance probability $p = e^{\frac{-(f(y)-f(x))}{t}}$ in which t is the temperature parameter. Solution representation and evaluating the objective function in TSA is similar to TGA. As the SA algorithm moves forward, t will be reduced gradually and it becomes harder to accept a worse solution. The process of decreasing t is called cooling and can be done in various ways. Two commonly used methods in the literature is to decrease t by a fixed number or a specified percentage (geometric cooling schedule) until a low temperature is achieved. SA is very sensitive to cooling schedule and a generic cooling schema might not work for all the problems. Usually, an adaptive cooling schedule is used (Ingber, 2000). It is shown that an adaptive cooling schedule can improve the results of SA (Zhan et al., 2016).

In TSA, the first step is to produce two initial temperature lists L_1 and L_2. L_1 is the items' temperature list while L_2 is the customers' temperature list. Each temperature list includes temperature levels that should be used in the cooling schedule. As mentioned earlier, given a temperature level, it is possible to calculate the probability of accepting a worse neighbor solution. Similarly, by inversing the acceptance probability formula and having the initial acceptance probability p_0, it is possible to calculate the corresponding temperature as $t = \frac{-(f(y)-f(x))}{\ln p_0}$. For L_1 a neighbor solution y is generated by inserting, inversing and swapping items for each customer in solution x and accepting the solution with minimum cost. For L_2 a neighbor solution y is generated by inserting, inversing and swapping customers in solution x and accepting the solution with minimum cost. Thus, if a solution includes n customers, $3n + 3$ neighbor solutions will be evaluated ($3n$ for items and 3 for customers). Here, coefficient 3 corresponds to the three operations of inverse, insert and swap.

The temperature levels obtained using p_0 will be used to create initial temperature lists. Algorithm 3 shows the pseudo-code for producing initial temperature lists L_1 and L_2. If the inverse, insert and swap operations are performed on items, Algorithm 3 generates L_1. Similarly, if the inverse, insert and swap operations are performed on customers, Algorithm 3 generates L_2.

ALGORITHM 3: PSEUDO-CODE FOR PRODUCING INITIAL TEMPERATURE LIST

Create an initial solution x
Define temperature list L = [] # initialize an empty list. L can be either L_1 or L_2
Define the length of temperature lists L_{max}
Define initial acceptance probability p_0
Create an initial solution x
while length(L) $\leq L_{max}$
    create a neighbor solution y from x using inverse, insert and swap operators
        if $f(y) \leq f(x)$:   # f is the objective function
            $x = y$
        else:
            $t = \frac{-(f(y)-f(x))}{\ln p_0}$
            Append t to L

The maximum temperatures t^1_{max} and t^2_{max} in L_1 and L_2 will be used for calculating the probability of accepting a worse solution. In TSA, M is the number of times a neighbor solution is generated in an iteration and is equal for both items and customers. In each iteration of TSA, M new solutions by changing the position of items for each customer using L_1 and M new solutions by changing the position of customers using L_2 are generated. Algorithm 4 depicts the pseudo-code of the TSA. The temperature lists L_1 and L_2 updates in each iteration of TSA. In TSA, the initial solution x is the result of OCFAH.

ALGORITHM 4. TSA PSEUDO-CODE

Define the best found solution $g_{best} = +\infty$
Generate an initial random solution x
Produce initial temperature lists L_1 and L_2
Define number of iterations n
Define number of neighbor solutions generated in each iteration M

for i in (1 to n):
    $c_1 = 0$ # number of worse solutions accepted for items
    $c_2 = 0$ # number of worse solutions accepted for customers
    $t^1_{sum} = 0$ # summation of temperature levels in L_1

```
    $t_{sum}^2 = 0$    # summation of temperature levels in $L_2$

for $m$ in (1 to $M$):
    Generate a neighbor solution $y_1$ from $x$ using $L_1$
    if $f(y_1) \le f(x)$:    # $f$ is the objective function
        $x = y_1$
        if $f(x) \le f(g_{best})$:
            $g_{best} = x$
    else:
        $p = e^{\frac{-(f(y_1)-f(x))}{t}}$    # $p$ is the acceptance probability
        Generate a random number $r \in [0,1)$
        if $r < p$:
            $c_1 = c_1 + 1$
            $t_{sum}^1 = t_{sum}^1 + \frac{-(f(y_1)-f(x))}{\ln r}$
            $x = y_1$
    Generate a neighbor solution $y_2$ from $x$ using $L_2$
    if $f(y_2) \le f(x)$:    # $f$ is the objective function
        $x = y_2$
        if $f(x) \le f(g_{best})$:
            $g_{best} = x$
    else:
        $p = e^{\frac{-(f(y_2)-f(x))}{t}}$    # $p$ is the acceptance probability
        Generate a random number $r \in [0,1)$
        if $r < p$:
            $c_2 = c_2 + 1$
            $t_{sum}^2 = t_{sum}^2 + \frac{-(f(y_2)-f(x))}{\ln r}$
            $x = y_2$
if $c_1$ <> 0:
    delete the highest temperature in $L_1$
    append $\frac{t_{sum}^1}{c_1}$ to $L_1$
if $c_2$ <> 0:
    delete the highest temperature in $L_2$
append $\frac{t_{sum}^2}{c_2}$ to $L_2$
return $g_{best}$
```

Fine-Tuning the TSA's Parameters

In this research, an ANOVA is utilized to determine the parameters of the TSA. Number of iterations, number of generated neighbors in each iteration (M), length of temperature lists L_{max} and initial acceptance probability p_0 are the influencing parameters in the TSA. TSA's tuning parameters and their levels considered in ANOVA are listed in Table 2. Using the results obtained, the TSA's parameters, number of iterations, M, length of temperature list L_{max} and initial acceptance probability p_0 were determined to be *200*, *100*, *20* and *0.9*. The test problem chosen for conducting the ANOVA is a medium tight instance with 15 customers, 15 FCs and 50 items, which will be introduced in section 5.

TABLE 2. LBSA's TUNING PARAMETERS AND THEIR LEVELS CONSIDERED ANOVA

Levels	No. Iterations	M	L_{max}	p_0
1	100	50	10	0.7
2	150	75	15	0.8
3	200	100	20	0.9

NUMERICAL EXPERIMENTS

In this section, the performance of the proposed OCFAH, TGA and TSA are evaluated and compared against the FCFS policy, as well as the exact solution obtained using CPLEX 12.6 for integrated (M) and decoupled ($M1$ and $M2$) models.

The test problems considered in this section are designed to evaluate the strength and weaknesses of the proposed methods while providing a reasonable approximation to real cases. For this purpose, three sets of small, medium, and large problems will be considered. While the small problems are different from real problems in terms of size, they can be solved to optimality and are quite helpful in evaluating the gap among the proposed methods and the optimum solutions. Moreover, the small problems are useful for verifying the mathematical models. In the problem under investigation and in real cases, inventory level of items and carton size are significant factors in determining the total cost and finding the optimum solution. Thus, two types of problems with tight and loose inventory constraints are considered. This will help in evaluating how well the proposed algorithms perform in reassigning cartons when there is not enough inventory or space in the cartons to accommodate the items. In all of the instances, FC fulfillment cost of cartons is a uniform random number in the interval $[2, 10]$ while the wholesaler fulfilment cost of an item is a uniform random number in $[10, 12]$. The cartons' dimensions are considered to be $40 \times 40 \times 40$. In the tight test problems, the items' dimensions are randomly selected from $[10, 35]$ and the inventory of all items is either 0 or 1 in the FCs. In the loose test problems, the items' dimensions are randomly selected from $[5, 25]$ and the inventory level at each FC is a random integer number in $[0, 2]$, $[0, 5]$, and $[0, 10]$ for small, medium, and large instances respectively.

The maximum order size for each customer is set to 10 items. Based on the authors' experience in the retail industry, the average order size in e-commerce businesses is usually less than four items per order. Thus, the maximum order size of 10 covers a significant fraction of all the orders that may be placed in reality. The maximum time limit for OCFAH, TGA and TSA and $M1 - M2$ is five minutes, while the maximum time limit for model M is set on one hour. In real cases, and especially for large retail corporations, the problem of this study needs to be solved in a short time. In the $M1$ model, where each order needs to be cartonized, the time limit for each customer is set to 10 seconds. After 10 seconds, the upper bound found by the CPLEX will be used as the solution. This guarantees that a single customer order cartonization does not consume all of the available five-minute time limit. The computer used for the experiments had a 3.50GHz Xeon CPU and 32.0 GB RAM. Note that all of the times reported in this section are in seconds, except if stated otherwise.

Small Instances

In the small instances (three levels for items), FCs and customers are considered. Table 3 lists the results of applying integrated model M, decoupled model $M1 - M2$, OCFAH, TGA, TSA and the first-come-first-serve (FCFS) methods to the small instances with tight constraints. For TGA and TSA the algorithm has run 10 times for each instance and the best, average, and worst solutions are reported.

As it can be observed, in all instances, the integrated model M, TGA and TSA have successfully obtained the optimum in a short time. However, the optimum or a near-optimum solution (approximately 6% gap) is also obtained by the OCFAH heuristics. A notable observation is the CPU time of the decoupled model $M1 - M2$, which is slightly longer than M in all instances. Although for the small instances, the decoupled model CPU time does not seem faster than the integrated model, but as will be shown, in larger instances, it yields quality solutions while significantly improving the run time.

OCFAH yields about a 6% gap compared to the optimum solution obtained by the $M1$ model and some instances outperform the $M1 - M2$ model while the run time of OCFAH is much shorter. Moreover, OCFAH improves the solutions obtained by FCFS by about 14.5%. TGA and TSA improve upon FCFS by about 19%.

Table 4 lists the results obtained for the small instances with loose constraints. Similar to small instances with tight constraints, the TGA and TSA successfully find the optimum solution in all small instances with loose constraints. One notable observation is the increased gap between OCFAH and the optimum obtained by model M (about 10%), while the gap between FCFS and OCFAH stays approximately the same. With this in mind, one may hypothesize that OCFAH performs better in instances with tight constraints. This hypothesis will be investigated more in the medium size instances.

TABLE 3: SMALL INSTANCES WITH TIGHT CONSTRAINTS

									TGA				TSA					
items	FCs	Customers	M	time	M1-M2	time	OCFAH	time	Best	Average	Worst	avg. time	Best	Average	Worst	avg. time	FCFS	time
3	2	4	63.00	0.23	63.00	0.91	64.73	<0.01	63.00	63.00	63.00	<0.01	63.00	63.00	63.00	<0.01	70.68	<0.01
		5	89.79	0.19	89.79	1.11	89.79	<0.01	89.79	89.79	89.79	<0.01	89.79	89.79	89.79	<0.01	89.79	<0.01
		6	174.91	0.40	174.91	1.60	178.18	<0.01	174.91	174.91	174.91	0.14	174.91	174.91	174.91	0.14	196.34	<0.01
	3	4	39.12	0.32	39.12	0.87	49.66	<0.01	39.12	39.12	39.12	<0.01	39.12	39.12	39.12	<0.01	59.33	<0.01
		5	73.40	0.44	73.40	1.14	79.73	<0.01	73.40	73.40	73.40	<0.01	73.40	73.40	73.40	<0.01	90.79	<0.01
		6	124.79	0.32	124.79	1.50	129.42	<0.01	124.79	124.79	124.79	0.15	124.79	124.79	124.79	0.10	143.12	<0.01
	4	4	28.89	0.49	28.89	0.94	29.50	<0.01	28.89	28.89	28.89	<0.01	28.89	28.89	28.89	<0.01	41.60	<0.01
		5	43.52	0.38	43.52	1.03	47.04	<0.01	43.52	43.52	43.52	<0.01	43.52	43.52	43.52	<0.01	54.14	<0.01
		6	40.28	0.90	40.28	1.42	59.31	<0.01	40.28	40.28	40.28	0.13	40.28	40.28	40.28	0.14	71.79	<0.01
4	2	4	50.12	0.22	50.12	0.86	50.12	<0.01	50.12	50.12	50.12	<0.01	50.12	50.12	50.12	<0.01	59.45	<0.01
		5	127.76	0.26	127.76	1.32	130.98	<0.01	127.76	127.76	127.76	<0.01	127.76	127.76	127.76	<0.01	141.90	<0.01
		6	136.68	0.31	136.68	1.45	138.15	<0.01	136.68	136.68	136.68	0.13	136.68	136.68	136.68	0.19	151.86	<0.01
	3	4	31.23	0.37	31.23	0.96	36.25	<0.01	31.23	31.23	31.23	<0.01	31.23	31.23	31.23	<0.01	48.69	<0.01
		5	31.57	0.36	31.57	1.01	35.11	<0.01	31.57	31.57	31.57	<0.01	31.57	31.57	31.57	<0.01	48.25	<0.01
		6	96.51	0.75	96.51	1.52	96.51	<0.01	96.51	96.51	96.51	0.13	96.51	96.51	96.51	0.12	114.70	<0.01
	4	4	35.00	0.73	35.00	1.29	35.00	<0.01	35.00	35.00	35.00	<0.01	35.00	35.00	35.00	<0.01	41.84	<0.01
		5	30.63	0.82	34.11	1.36	33.84	<0.01	30.63	30.63	30.63	<0.01	30.63	30.63	30.63	<0.01	52.50	<0.01
		6	93.85	0.51	93.85	1.60	94.40	<0.01	93.85	93.85	93.85	0.17	93.85	93.85	93.85	0.17	105.45	<0.01
5	2	4	59.03	0.34	60.56	1.06	59.03	<0.01	59.03	59.03	59.03	<0.01	59.03	59.03	59.03	<0.01	74.43	<0.01
		5	153.52	0.53	153.52	1.40	153.52	<0.01	153.52	153.52	153.52	<0.01	153.52	153.52	153.52	<0.01	175.37	<0.01
		6	196.16	1.18	196.16	1.44	199.12	<0.01	196.16	196.16	196.16	0.17	196.16	196.16	196.16	0.13	215.78	<0.01
	3	4	85.25	0.63	85.25	1.27	97.76	<0.01	85.25	85.25	85.25	<0.01	85.25	85.25	85.25	<0.01	110.08	<0.01
		5	50.06	0.72	55.22	1.18	51.29	<0.01	50.06	50.06	50.06	<0.01	50.06	50.06	50.06	<0.01	64.33	<0.01
		6	156.37	1.75	161.18	1.77	160.99	<0.01	156.37	156.37	156.37	0.10	156.37	156.37	156.37	0.20	183.45	<0.01
	4	4	71.30	5.38	76.30	1.54	82.46	<0.01	71.30	71.30	71.30	<0.01	71.30	71.30	71.30	<0.01	94.70	<0.01
		5	113.27	0.85	113.27	1.70	124.48	<0.01	113.27	113.27	113.27	<0.01	113.27	113.27	113.27	<0.01	131.67	<0.01
		6	200.48	0.79	203.42	2.10	204.41	<0.01	200.48	200.48	200.48	0.16	200.48	200.48	200.48	0.20	227.69	<0.01

TABLE 4. SMALL INSTANCES WITH LOOSE CONSTRAINTS

									TGA				TSA					
items	FCs	Customers	M	time	M1-M2	time	OCFAH-LA	time	Best	Average	Worst	avg. time	Best	Average	Worst	avg. time	FCFS	time
3	2	4	64.50	0.25	64.50	0.93	67.43	<0.01	64.50	64.50	64.50	<0.01	64.50	64.50	64.50	<0.01	82.91	<0.01
		5	43.08	0.30	43.08	0.99	47.46	<0.01	43.08	43.08	43.08	<0.01	43.08	43.08	43.08	<0.01	59.07	<0.01
		6	69.23	0.32	69.23	1.21	73.99	<0.01	69.23	69.23	69.23	0.14	69.23	69.23	69.23	0.18	77.22	<0.01
	3	4	75.80	0.30	75.80	1.02	77.58	<0.01	75.80	75.80	75.80	<0.01	75.80	75.80	75.80	<0.01	80.31	<0.01
		5	69.24	0.43	69.24	1.06	76.68	<0.01	69.24	69.24	69.24	<0.01	69.24	69.24	69.24	<0.01	80.96	<0.01
		6	71.16	0.51	71.16	1.26	71.70	<0.01	71.16	71.16	71.16	0.10	71.16	71.16	71.16	0.19	93.70	<0.01
	4	4	19.76	0.56	19.76	1.06	27.47	<0.01	19.76	19.76	19.76	<0.01	19.76	19.76	19.76	<0.01	33.96	<0.01
		5	42.09	0.39	42.09	1.03	60.57	<0.01	42.09	42.09	42.09	<0.01	42.09	42.09	42.09	<0.01	61.75	<0.01
		6	66.52	0.59	66.52	1.18	78.18	<0.01	66.52	66.52	66.52	0.20	66.52	66.52	66.52	0.15	90.63	<0.01
4	2	4	68.08	0.28	68.08	0.99	68.08	<0.01	68.08	68.08	68.08	<0.01	68.08	68.08	68.08	<0.01	91.27	<0.01
		5	79.94	0.37	79.94	1.05	85.11	<0.01	79.94	79.94	79.94	<0.01	79.94	79.94	79.94	<0.01	99.09	<0.01
		6	116.61	0.57	116.61	1.48	124.65	<0.01	116.61	116.61	116.61	0.20	116.61	116.61	116.61	0.17	148.63	<0.01
	3	4	86.46	0.44	86.46	1.00	89.02	<0.01	86.46	86.46	86.46	<0.01	86.46	86.46	86.46	<0.01	98.47	<0.01
		5	30.44	0.71	30.44	1.04	45.53	<0.01	30.44	30.44	30.44	<0.01	30.44	30.44	30.44	<0.01	57.65	<0.01
		6	82.00	0.78	82.00	1.26	82.85	<0.01	82.00	82.00	82.00	0.11	82.00	82.00	82.00	0.17	96.28	<0.01
	4	4	50.09	0.58	50.09	1.14	62.63	<0.01	50.09	50.09	50.09	<0.01	50.09	50.09	50.09	<0.01	76.18	<0.01
		5	39.68	0.98	39.68	1.25	45.65	<0.01	39.68	39.68	39.68	<0.01	39.68	39.68	39.68	<0.01	65.93	<0.01
		6	37.11	0.85	37.11	1.32	49.47	<0.01	37.11	37.11	37.11	0.19	37.11	37.11	37.11	0.19	46.24	<0.01
5	2	4	80.93	0.48	80.93	1.10	85.25	<0.01	80.93	80.93	80.93	<0.01	80.93	80.93	80.93	<0.01	111.27	<0.01
		5	127.31	0.53	127.31	1.21	130.80	<0.01	127.31	127.31	127.31	<0.01	127.31	127.31	127.31	<0.01	154.39	<0.01
		6	109.93	0.24	109.93	1.32	109.93	<0.01	109.93	109.93	109.93	0.10	109.93	109.93	109.93	0.12	109.93	<0.01
	3	4	59.66	0.62	59.66	1.05	66.92	<0.01	59.66	59.66	59.66	<0.01	59.66	59.66	59.66	<0.01	78.02	<0.01
		5	133.19	0.87	133.19	1.18	139.06	<0.01	133.19	133.19	133.19	<0.01	133.19	133.19	133.19	<0.01	173.76	<0.01
		6	125.80	0.55	125.80	1.29	128.27	<0.01	125.80	125.80	125.80	0.18	125.80	125.80	125.80	0.12	134.09	<0.01
	4	4	47.80	1.10	47.80	1.19	49.95	<0.01	47.80	47.80	47.80	<0.01	47.80	47.80	47.80	<0.01	59.50	<0.01
		5	70.65	1.13	70.65	1.33	92.36	<0.01	70.65	70.65	70.65	<0.01	70.65	70.65	70.65	<0.01	108.74	<0.01
		6	90.04	1.35	92.82	1.35	100.81	<0.01	90.04	90.04	90.04	0.18	90.04	90.04	90.04	0.15	103.23	<0.01

Medium Instances

In the medium instances, three levels for items, FCs, and customers are considered. Table 5 lists results obtained for the medium instances with tight constraints. For the integrated model M, two numbers are reported: lower bound (LB) and upper bound (UB). These numbers are bounds obtained by CPLEX after one hour. As the results show, CPLEX obtained a result or bound in one hour for only a few instances, which shows the integrated model M can be a difficult problem to solve in medium size instances.

According to Table 5, the decoupled model $M1 - M2$ yields high quality solutions and is able to reduce the computation time of M significantly. Note that for each customer, model $M1$ can be solved in parallel and reduce the computation time for the larger problem instances even more. The average gap between OCFAH and the solutions obtained by models M and $M1 - M2$ is about 12% and 9% respectively. Note that the run time of OCFAH is much shorter than M and $M1 - M2$ and is approximately 0.08 second on average. Moreover, the gap between OCFAH and FCFS is increased to 20%, which is higher than what was observed in small instances and can point to the effectiveness of OCFAH.

As was hypothesized earlier, the results of medium instances show that OCFAH works better when applied to tightly constrained instances. Table 5 provides evidence for this hypothesis. If one considers the instances where inventory is scarce compared to the number of customers, such as the cases where there are 15 customers and five FCs, the average gap between OCFAH and the $M1 - M2$ model is about 2%. This shows that OCFAH can handle constrained instances successfully. Moreover, the performance of OCFAH in loosely constrained medium size instances in Table 6 confirms the ability of OCFAH to handle more constrained instances. Similar to tightly constrained instances, Table 6 shows a significant CPU time improvement in $M1 - M2$ as compared to M. The results of Table 6 confirm the same patterns observed in tightly constrained instances.

The results in Table 5 and Table 6 show that TGA and TSA are both successful in finding optimum or near optimum solutions in a reasonable time. For medium instances, no significant evidence regarding the difference in performance of $M1 - M2$, TGA and TSA was observed. However, $M1 - M1$, TGA and TSA both outperform OCFAH on average.

TABLE 5. MEDIUM INSTANCES WITH TIGHT CONSTRAINTS

			M							TGA				TSA					
items	FCs	Customers	LB	UB	time	M1-M2	time	OCFAH	time	Best	Average	Worst	time	Best	Average	Worst	time	FCFS	time
10	5	10	214.89	214.89	457.22	236.73	4.67	234.46	0.05	214.89	214.89	214.89	16.19	252.59	264.87	273.91	15.39	289.35	<0.01
		15	459.20	459.20	617.21	473.91	6.37	479.70	0.10	420.24	429.62	438.33	22.35	483.58	496.26	510.90	19.01	521.77	<0.01
		20	761.46	761.46	1198.30	761.46	19.05	769.77	0.09	761.46	761.46	761.46	25.63	761.46	761.46	761.46	26.00	841.84	<0.01
	10	10	38.55	38.55	1388.81	43.44	2.70	53.39	0.02	38.55	38.55	38.55	12.14	38.55	38.55	38.55	14.47	62.38	<0.01
		15	266.39	266.39	3526.50	280.43	48.74	338.37	0.11	266.39	266.39	266.39	19.15	266.39	266.39	266.39	17.21	374.25	<0.01
		20	601.71	625.31	3600.00	661.91	50.52	700.33	0.10	612.36	629.26	633.01	35.76	612.36	621.94	636.50	27.95	768.75	<0.01
	15	10	20.98	20.98	2378.29	22.46	11.35	27.09	0.02	20.98	21.68	25.20	13.95	20.98	22.82	24.91	11.19	40.91	<0.01
		15	NA	NA	3600.00	74.77	27.83	116.46	0.10	72.93	73.38	75.25	26.42	69.50	69.85	71.25	23.48	152.21	<0.01
		20	NA	NA	3600.00	500.23	29.83	527.89	0.14	486.64	497.24	502.37	32.53	480.10	488.97	497.70	29.51	611.76	<0.01
20	5	10	48.62	48.62	2779.28	61.94	11.40	55.13	0.03	48.62	48.62	48.62	12.31	48.62	48.62	48.62	13.06	76.26	<0.01
		15	223.78	238.12	3600.00	272.94	9.75	302.57	0.05	265.62	276.19	282.25	26.40	276.65	284.06	297.97	15.77	330.17	<0.01
		20	569.58	576.25	3600.00	603.22	39.24	614.56	0.14	561.29	585.16	597.13	35.61	616.09	634.40	663.39	35.73	682.26	<0.01
	10	10	NA	NA	3600.00	72.77	22.69	82.97	0.06	78.27	81.62	83.29	16.69	76.54	80.14	83.97	13.81	113.84	<0.01
		15	NA	NA	3600.00	183.18	44.25	216.31	0.16	156.86	160.54	165.94	19.10	164.60	169.08	173.51	23.22	280.54	<0.01
		20	NA	NA	3600.00	229.84	40.02	253.60	0.10	236.92	246.17	251.66	33.70	223.05	230.89	238.03	29.79	288.85	<0.01
	15	10	NA	NA	3600.00	27.93	2.84	30.18	0.01	24.34	24.90	25.74	12.12	28.77	30.07	30.79	12.58	39.82	<0.01
		15	NA	NA	3600.00	75.34	10.55	92.38	0.05	78.90	81.74	83.71	24.28	66.34	68.33	71.58	17.66	120.37	<0.01
		20	NA	NA	3600.00	93.06	23.95	118.74	0.06	88.84	92.17	94.76	28.76	90.79	94.26	96.80	32.83	149.66	<0.01
50	5	10	NA	NA	3600.00	117.03	15.47	106.72	0.07	105.12	109.54	113.49	14.95	123.85	126.53	129.37	13.94	155.88	<0.01
		15	NA	NA	3600.00	176.26	38.98	186.58	0.09	190.88	195.56	203.93	18.54	182.30	190.59	199.91	18.91	262.45	<0.01
		20	NA	NA	3600.00	404.54	80.33	418.41	0.20	379.45	396.87	410.51	30.46	346.61	354.65	367.81	25.27	504.25	<0.01
	10	10	NA	NA	3600.00	46.36	21.04	43.10	0.03	45.31	47.47	48.83	14.71	43.48	45.51	47.71	13.59	57.59	<0.01
		15	NA	NA	3600.00	106.03	46.24	133.51	0.08	100.00	102.11	105.81	15.77	115.70	119.73	124.35	24.54	169.37	<0.01
		20	NA	NA	3600.00	159.24	73.52	190.53	0.11	164.13	171.03	177.10	38.01	138.35	145.18	149.57	31.81	256.07	<0.01
	15	10	NA	NA	3600.00	41.01	22.53	46.61	0.05	40.06	41.66	42.61	16.42	42.11	43.71	44.78	14.98	56.28	<0.01
		15	NA	NA	3600.00	92.85	47.85	97.34	0.09	83.35	86.89	88.98	21.92	87.54	91.58	95.22	17.81	146.25	<0.01
		20	NA	NA	3600.00	153.56	101.11	174.56	0.16	151.26	155.29	160.17	32.55	147.74	153.44	157.44	31.96	226.87	<0.01

TABLE 6. MEDIUM INSTANCES WITH LOOSE CONSTRAINTS

			M							TGA				TSA					
items	FCs	Customers	LB	UB	time	M1-M2	time	OCFAH	time	Best	Average	Worst	time	Best	Average	Worst	time	FCFS	time
10	5	10	14.67	14.67	12.41	14.67	1.68	18.96	0.02	14.67	14.67	14.67	15.79	14.67	14.67	14.67	13.13	18.96	<0.01
		15	61.18	61.18	780.41	78.78	6.01	111.94	0.14	61.18	61.18	61.18	23.07	61.18	61.18	61.18	15.91	163.39	<0.01
		20	155.93	155.93	1096.14	171.35	9.24	230.56	0.08	155.93	163.63	167.74	24.97	155.93	161.64	167.31	25.95	332.27	<0.01
	10	10	30.00	30.00	695.21	30.00	2.93	31.92	0.02	30.00	30.00	30.00	14.70	30.00	31.48	32.84	13.83	42.48	<0.01
		15	45.76	45.76	847.57	46.86	35.52	62.72	0.11	45.76	47.59	49.80	21.92	45.76	47.62	49.75	17.77	89.34	<0.01
		20	73.58	73.58	1410.03	76.73	58.88	101.81	0.33	73.58	77.13	79.02	31.51	73.58	76.16	79.57	26.69	136.61	<0.01
	15	10	17.16	17.16	1121.97	17.16	2.92	28.43	0.05	17.16	17.77	18.21	15.89	17.16	17.88	18.65	14.71	32.47	<0.01
		15	41.88	41.88	2115.47	42.93	6.41	51.47	0.10	41.88	41.88	41.88	25.69	41.88	41.88	41.88	15.39	62.59	<0.01
		20	50.43	53.92	3600.00	54.64	47.87	78.85	0.15	53.92	55.58	57.29	25.56	53.92	55.63	57.78	24.00	90.76	<0.01
20	5	10	35.10	35.10	156.70	35.10	3.66	43.89	0.11	35.10	36.47	37.57	14.03	35.10	36.09	37.28	12.20	62.89	<0.01
		15	72.64	76.55	3600.00	78.43	8.20	108.29	0.09	74.15	76.93	80.28	21.92	75.98	79.04	80.65	15.50	142.79	<0.01
		20	83.12	84.61	3600.00	84.59	14.27	132.65	0.23	84.61	86.44	89.31	27.59	84.61	88.50	91.13	23.40	159.94	<0.01
	10	10	18.02	18.02	2286.26	18.02	4.78	25.87	0.13	18.02	18.50	18.94	14.25	18.02	18.61	19.30	13.91	32.77	<0.01
		15	38.69	39.47	3600.00	39.82	6.59	47.83	0.06	39.47	40.79	41.60	15.33	39.47	40.30	41.11	14.50	50.92	<0.01
		20	76.19	77.83	3600.00	77.96	17.83	104.35	0.24	77.83	81.62	85.02	28.18	78.01	79.72	83.30	22.68	125.10	<0.01
	15	10	N/A	N/A	3600.00	20.56	3.32	30.76	0.04	22.08	23.05	23.66	11.23	21.37	22.00	22.89	10.20	30.81	<0.01
		15	N/A	N/A	3600.00	30.22	7.12	39.70	0.07	32.69	34.23	35.32	20.81	26.73	27.71	28.61	19.13	46.48	<0.01
		20	N/A	N/A	3600.00	54.41	18.03	82.41	0.18	52.20	53.28	54.75	26.07	50.26	52.40	54.96	22.84	106.37	<0.01
50	5	10	N/A	N/A	3600.00	32.84	4.84	48.41	0.04	35.00	36.58	37.52	16.99	33.71	35.33	36.35	13.91	61.55	<0.01
		15	N/A	N/A	3600.00	47.42	12.12	73.40	0.07	42.85	43.75	45.67	21.42	41.09	42.87	44.22	16.40	77.05	<0.01
		20	N/A	N/A	3600.00	94.90	26.05	141.95	0.10	98.29	102.69	105.94	26.02	84.12	88.09	90.03	25.87	172.51	<0.01
	10	10	N/A	N/A	3600.00	21.39	6.05	33.74	0.09	20.41	21.42	22.15	15.74	19.03	19.49	20.00	10.25	36.81	<0.01
		15	N/A	N/A	3600.00	38.83	11.82	48.40	0.06	35.43	36.35	37.90	17.21	40.41	42.20	43.08	15.27	59.39	<0.01
		20	N/A	N/A	3600.00	55.33	24.92	87.61	0.10	51.37	53.12	55.48	23.05	50.90	52.09	54.40	28.68	105.57	<0.01
	15	10	N/A	N/A	3600.00	21.13	4.62	35.36	0.03	18.70	19.17	19.96	13.90	21.55	22.38	22.94	14.20	38.60	<0.01
		15	N/A	N/A	3600.00	35.01	17.77	44.51	0.15	34.42	35.76	37.28	16.21	35.20	36.75	37.65	16.67	64.94	<0.01
		20	N/A	N/A	3600.00	50.39	29.83	77.01	0.21	55.27	58.01	60.63	25.36	43.18	44.36	45.29	26.59	88.32	<0.01

Large Instances

In the large instances, three levels for items, FCs, and customers are considered. Note that models M and $M1 - M2$ cannot be applied to large instances due to time and memory limitations. Table 7 lists the results obtained for the medium instances with tight constraints where the average improvement obtained using OCFAH compared to FCFS is about 22%. This improvement for loosely constrained instances is about 8%, as tabulated in Table 8, which again, points to the superior performance of OCFAH when applied to tightly constrained instances.

As expected, both TGA and TSA outperform OCFAH while consuming all five minutes allowable CPU time. TGA and TSA outperform OCFAH by approximately 3.6% and 7.9% for tight instances and 4.9% and 7.6% for loose instances. TSA outperforms TGA by approximately 5% for tight instances and 3.3% for loose instances. One possible explanation for superior performance of TSA over TGA can be found in its neighborhood search procedure. While generating neighbor solutions, TSA performs inverse, insert and swap operation on each item set ordered by the customers. However, TGA does not possess such a rigorous neighborhood search procedure and solely relies on information transferred through generations for improving the solutions. In this problem, it seems a rigorous neighborhood search such as the one in TSA is a more reasonable search strategy.

TABLE 7. LARGE INSTANCES WITH TIGHT CONSTRAINTS

					TGA				TSA					
items	FCs	Customers	OCFAH	time	Best	Average	Worst	time	Best	Average	Worst	time	FCFS	time
100	20	100	1143.01	0.6	1080.52	1103.77	1132.57	300.00	1047.49	1058.37	1071.80	300.00	1426.84	<0.01
		250	7239.78	2.1	6637.24	6720.89	6902.06	300.00	6617.86	6672.53	6697.48	300.00	7915.58	<0.01
		500	23715.82	4.04	22643.72	23327.78	23702.04	300.00	20689.48	21745.20	22252.74	300.00	24963.95	<0.01
	50	100	727.86	0.8	686.78	701.16	711.43	300.00	662.32	683.84	694.10	300.00	1003.03	<0.01
		250	1870.23	1.69	1758.93	1811.74	1840.31	300.00	1735.32	1766.66	1797.87	300.00	2401.21	<0.01
		500	9713.62	5.46	9323.27	9435.89	9563.47	300.00	8449.90	8763.77	9012.93	300.00	10948.42	<0.01
	100	100	588.06	0.78	568.21	572.21	577.99	300.00	524.09	547.36	572.47	300.00	824.00	<0.01
		250	1652.34	2.40	1596.16	1603.17	1606.26	300.00	1475.51	1521.98	1587.85	300.00	2206.80	<0.01
		500	3729.76	5.34	3509.31	3626.09	3685.92	300.00	3379.42	3418.33	3496.66	300.00	4887.35	<0.01
250	20	100	863.51	0.76	786.58	807.89	838.37	300.00	766.40	778.77	793.89	300.00	1158.28	<0.01
		250	3295.00	2.55	3155.71	3199.92	3253.78	300.00	2903.00	3075.86	3234.03	300.00	4116.82	<0.01
		500	11905.85	5.67	11173.58	11438.42	11694.45	300.00	11403.99	11539.04	11610.71	300.00	13468.90	<0.01
	50	100	676.03	0.78	662.23	665.78	670.91	300.00	600.69	625.91	654.40	300.00	945.58	<0.01
		250	1775.30	2.42	1708.21	1736.54	1759.88	300.00	1606.52	1654.35	1675.71	300.00	2349.59	<0.01
		500	3937.50	5.98	3709.60	3821.67	3875.45	300.00	3636.54	3666.44	3731.40	300.00	5236.63	<0.01
	100	100	541.00	0.75	530.75	531.44	532.72	300.00	479.84	496.34	511.25	300.00	736.93	<0.01
		250	1505.00	2.30	1443.51	1463.85	1480.94	300.00	1310.67	1336.54	1381.72	300.00	2019.28	<0.01
		500	3253.78	5.75	3034.49	3075.08	3099.72	300.00	2658.25	2785.74	2912.63	300.00	4386.29	<0.01
500	20	100	943.54	0.90	939.21	939.61	940.33	300.00	837.72	847.06	864.98	300.00	1213.74	<0.01
		250	2477.26	2.26	2383.83	2422.13	2442.06	300.00	2191.67	2293.04	2356.60	300.00	3219.38	<0.01
		500	6626.10	6.68	6313.64	6485.70	6608.95	300.00	6221.05	6405.00	6558.99	300.00	8323.60	<0.01
	50	100	618.39	0.93	570.26	584.99	595.99	300.00	549.97	566.16	594.70	300.00	888.26	<0.01
		250	1756.55	2.50	1634.32	1673.19	1731.18	300.00	1586.82	1659.17	1703.32	300.00	2304.41	<0.01
		500	3527.86	5.45	3233.59	3370.40	3449.92	300.00	2922.21	3110.05	3233.10	300.00	4754.51	<0.01
	100	100	584.88	0.70	536.18	557.27	568.24	300.00	516.43	530.72	551.96	300.00	768.47	<0.01
		250	1553.68	2.65	1512.88	1521.05	1532.64	300.00	1397.21	1435.47	1484.85	300.00	2090.46	<0.01
		500	3169.88	5.65	2925.42	2970.48	3011.86	300.00	2869.03	2907.05	2989.04	300.00	4241.22	<0.01

TABLE 8. LARGE INSTANCES WITH LOOSE CONSTRAINTS

					TGA				TSA					
items	FCs	Customers	OCFAH	time	Best	Average	Worst	time	Best	Average	Worst	time	FCFS	time
100	20	100	387.78	1.25	369.75	373.75	382.60	300.00	347.30	353.78	363.68	300.00	424.94	<0.01
		250	982.7	2.2	930.23	953.51	982.46	300.00	892.34	904.90	913.71	300.00	1100.81	<0.01
		500	2172.29	4.77	1972.03	2064.81	2144.41	300.00	2020.78	2059.58	2086.82	300.00	2491.08	<0.01
	50	100	326.72	1.12	315.91	316.48	318.62	300.00	277.51	295.05	306.24	300.00	340.99	<0.01
		250	828.53	2.75	758.26	797.99	818.42	300.00	737.11	785.93	815.39	300.00	897.76	<0.01
		500	1639.32	5.19	1500.75	1552.51	1587.63	300.00	1451.09	1482.18	1498.29	300.00	1854.18	<0.01
	100	100	278.39	0.78	256.22	263.84	268.88	300.00	258.58	261.83	265.41	300.00	291.59	<0.01
		250	789.90	2.01	703.06	725.86	753.40	300.00	667.68	703.34	731.23	300.00	863.86	<0.01
		500	1524.71	5.25	1353.94	1418.55	1462.07	300.00	1356.75	1380.98	1406.14	300.00	1643.37	<0.01
250	20	100	338.57	1.08	311.02	318.03	331.10	300.00	291.55	311.40	320.96	300.00	368.82	<0.01
		250	916.54	2.49	862.91	883.17	901.67	300.00	813.49	844.12	860.34	300.00	1001.05	<0.01
		500	1897.26	4.60	1740.00	1797.42	1848.04	300.00	1665.36	1762.39	1819.06	300.00	2121.05	<0.01
	50	100	326.36	1.01	307.85	312.18	316.90	300.00	285.87	304.39	313.19	300.00	349.32	<0.01
		250	815.86	2.82	780.00	784.44	786.92	300.00	755.82	765.65	781.64	300.00	902.03	<0.01
		500	1550.39	4.91	1423.17	1477.93	1522.44	300.00	1355.51	1429.62	1471.66	300.00	1691.95	<0.01
	100	100	305.89	1.25	282.36	287.17	295.88	300.00	275.93	283.57	287.40	300.00	340.25	<0.01
		250	755.12	2.55	681.82	693.54	717.50	300.00	663.68	665.73	667.91	300.00	839.58	<0.01
		500	1525.05	5.43	1423.93	1463.86	1505.74	300.00	1400.46	1460.41	1486.82	300.00	1652.86	<0.01
500	20	100	342.26	1.11	325.38	333.32	339.60	300.00	303.59	316.89	330.91	300.00	371.66	<0.01
		250	938.06	2.37	877.60	890.96	898.32	300.00	798.64	846.93	889.57	300.00	1014.91	<0.01
		500	1884.31	5.15	1803.94	1845.75	1869.18	300.00	1687.50	1751.17	1804.97	300.00	2065.96	<0.01
	50	100	288.73	0.83	273.15	275.49	280.97	300.00	264.49	267.36	269.38	300.00	314.46	<0.01
		250	762.07	2.69	701.68	712.37	734.70	300.00	685.73	712.74	732.73	300.00	850.55	<0.01
		500	1577.50	5.25	1534.09	1548.01	1574.98	300.00	1420.75	1494.93	1535.01	300.00	1686.13	<0.01
	100	100	318.85	1.21	296.91	308.09	314.84	300.00	286.93	292.70	298.02	300.00	334.13	<0.01
		250	777.36	2.96	736.60	755.11	766.48	300.00	694.85	733.08	752.64	300.00	844.15	<0.01
		500	1515.97	5.26	1380.03	1412.79	1450.49	300.00	1354.59	1368.50	1400.01	300.00	1608.57	<0.01

CONCLUSION AND FUTURE RESEARCH DIRECTION

Recently, retailers are experiencing fierce competition and declining profit margins. In this situation, retailers naturally incline toward increasing the efficiency of their supply chain by improving their fulfillment processes. This study investigates a problem that arises in the retail industry where a retailer is interested in the cartonization of orders it receives from customers and assigning them to fulfillment centers (FCs) in order to minimize the cost of fulfilling demand when there is inventory constraint. If the retailer cannot fulfill a customer's order due to inventory stockout, it will fulfill it by outsourcing the order to a wholesaler or a third party. This is a common practice by the retail industry to retain customers.

For this purpose, two mathematical models were proposed. The first model (M), referred to as the integrated model, solves the problem in one stage and yields the optimum solution. However, for medium and large instances, it was shown that model M is inefficient in terms of time. To overcome this inefficiency, a decoupled model ($M1 - M2$) was proposed that could solve the integrated model M in two stages, and it was shown that this model can improve the run time of M for medium size instances of the problem significantly—at the cost of losing a small degree of optimality. While being efficient for the medium size instances, the decoupled model $M1 - M2$ could not be applied to the large instances due to time and memory constraints.

Since in real settings the investigated problem requires a quick resolution, a fast heuristic, namely the order cartonization and FC assignment heuristic (OCFAH), a two-stage genetic algorithm (TGA) and a two-stage simulated annealing (TSA) were proposed. The idea of implementing a two-stage process in GA and SA was based on the observation that solution representations of the problem usually lead to large data structures which cannot be sufficiently explored by conventional crossover and mutation operators. Thus, it was necessary to apply a method that diversifies the search. It was found that while TGA and TSA both result in optimal or near optimal solutions in small and medium instances of the problem, in large instances, TSA outperforms TGA. It was conjectured that superiority of TSA over TGA in this problem is due to its rigorous neighborhood search procedure.

Several questions and directions remain to be investigated in the future. First, minimizing time to fulfill demand received by customers and incorporating it in the model can be a very attractive extension to the problem. While minimizing fulfillment cost is important, minimizing delivery time can also contribute significantly to the success of a retailer in the market. Cost and time may be conflicting objectives and the resulting extension would be a multi-objective problem.

Sometimes retailers offer a discount or credit if the customer postpones the delivery time. While this practice can help reduce fulfillment cost, a second interesting extension to the problem would be to incorporate the discount options or the pricing of items in the model. The third future research direction can be focused on the application of surrogate models in order to solve $M1 - M2$ more efficiently. Currently, both large-scale $M1$ and $M2$ models are difficult to solve. Replacing one of these models with a surrogate model that efficiently approximates the output of either $M1$ or $M2$ can significantly enhance the CPU time of $M1 - M2$ and makes it applicable for larger instances of the problem.

REFERENCES

Acimovic, J., & Graves, S. C. (2014). Making better fulfillment decisions on the fly in an online retail environment. *Manufacturing & Service Operations Management, 17*(1), 34-51.

Bhargava, R., Reyes Levalle, R., & Nof, S. Y. (2016). A best-matching protocol for order fulfillment in re-configurable supply networks. *Computers in Industry, 82*, 160-169.

Bortfeldt, A., & Gehring, H. (2001). A hybrid genetic algorithm for the container loading problem. *European Journal of Operational Research, 131*(1), 143-161.

Bortfeldt, A., Gehring, H., & Mack, D. (2003). A parallel tabu search algorithm for solving the container loading problem. *Parallel Computing, 29*(5), 641-662.

Chien, C.-F., & Deng, J.-F. (2004). A container packing support system for determining and visualizing container packing patterns. *Decision Support Systems, 37*(1), 23-34.

Crainic, T. G., Perboli, G., & Tadei, R. (2008). Extreme point-based heuristics for three-dimensional bin packing. *Informs Journal on Computing, 20*(3), 368-384.

Croxton, K. L. (2003). The order fulfillment process. *The International Journal of Logistics Management, 14*(1), 19-32.

Espino-Rodríguez, T. F., & Rodríguez-Díaz, M. (2014). Determining the core activities in the order fulfillment process: An empirical application. *Business Process Management Journal, 20*(1), 2-24.

Faroe, O., Pisinger, D., & Zachariasen, M. (2003). Guided local search for the three-dimensional bin-packing problem. *Informs Journal on Computing, 15*(3), 267-283.

Fekete, S. P., Schepers, J., & Van der Veen, J. C. (2007). An exact algorithm for higher-dimensional orthogonal packing. *Operations Research, 55*(3), 569-587.

Gibson, B. J., Defee, C. C., & Randall, W. (2015). *The state of retail supply chain.*

Gilmore, P., & Gomory, R. E. (1965). Multistage cutting stock problems of two and more dimensions. *Operations Research, 13*(1), 94-120.

Goldberg, D. E., & Lingle, R. (1985). Alleles, loci, and the traveling salesman problem. *Proceedings of an international conference on genetic algorithms and their applications.*

Ingber, L. (2000). *Adaptive simulated annealing (ASA): Lessons learned.* arXiv preprint cs/0001018.

Jasin, S., & Sinha, A. (2014). *LP-Based artificial dependency for probabilistic etail order fulfillment* (Paper No. 1250). Ross School of Business.

Jasin, S., & Sinha, A. (2015). An LP-based correlated rounding scheme for multi-item ecommerce order fulfillment. *Operations Research, 63*(6), 1336-1351.

Lin, F.-R., & Shaw, M. J. (1998). Reengineering the order fulfillment process in supply chain networks. *International Journal of Flexible Manufacturing Systems, 10*(3), 197-229.

Mahar, S., & Wright, P. D. (2009). The value of postponing online fulfillment decisions in multi-channel retail/e-tail organizations. *Computers & Operations Research, 36*(11), 3061-3072.

Martello, S., Pisinger, D., & Vigo, D. (2000). The three-dimensional bin packing problem. *Operations Research, 48*(2), 256-267.

Moura, A., & Oliveira, J. F. (2005). A GRASP approach to the container-loading problem. *IEEE Intelligent Systems, 20*(4), 50-57.

Onal, S., Zhang, J., & Das, S. (2017). Modelling and performance evaluation of explosive storage policies in internet fulfilment warehouses. *International Journal of Production Research,* 1-14.

Owyong, M., & Yih, Y. (2006). Picklist generation algorithm with order-consolidation consideration for split-case module-based fulfilment centres. *International Journal of Production Research, 44*(21), 4529-4550.

Pisinger, D. (2002). Heuristics for the container loading problem. *European Journal of Operational Research, 141*(2), 382-392.

Ricker, F., & Kalakota, R. (1999). Order fulfillment: The hidden key to e-commerce success. *Supply Chain Management Review, 11*(3), 60-70.

Thirumalai, S., & Sinha, K. K. (2005). Customer satisfaction with order fulfillment in retail supply chains: Implications of product type in electronic B2C transactions. *Journal of Operations Management*, *23*(3), 291-303.

Wu, Y., Li, W., Goh, M., & de Souza, R. (2010). Three-dimensional bin packing problem with variable bin height. *European Journal of operational research, 202*(2), 347-355.

Xu, P. J. (2005). *Order fulfillment in online retailing: What goes where.* Massachusetts Institute of Technology.

Xu, P. J., Allgor, R., & Graves, S. C. (2009). Benefits of reevaluating real-time order fulfillment decisions. *Manufacturing & Service Operations Management, 11*(2), 340-355.

Zhan, S.-H., Lin, J., Zhang, Z.-J., & Zhong, Y.-W. (2016). List-based simulated annealing algorithm for traveling salesman problem. *Computational Intelligence and Neuroscience, 2016*, 8.

Zhao, X., Bennell, J. A., Bektaş, T., & Dowsland, K. (2016). A comparative review of 3D container loading algorithms. *International Transactions in Operational Research*, *23*(1-2), 287-320.

JOURNAL OF INTERNATIONAL BUSINESS DISCIPLINES

Volume 16, Number 2 November 2021

Published By:
University of Tennessee at Martin and the International Academy of Business Disciplines

ISSN 1934-1822 WWW.JIBD.ORG

Printed in the USA by the University of Tennessee at Martin

Printed in the USA by the University of Tennessee at Martin