



ניתוח החלטות רכישה של לקוחות באמצעות נתוני מיקום



שחר רייכמן

טל שושני

פיטר זובצ'יק

ד"ר פיטר פאל זובצ'יק הוא חבר סגל בפקולטה לניהול ע"ש קולר באוניברסיטת תל אביב. את הדוקטורט סיים ב-INSEAD וכיום הוא חוקר ומרצה בתחום השיווק. במחקרו הוא מתמקד בתפקידן של רשתות חברתיות באינטראקציות עם הצרכן, באפליקציות לניהול קשרי לקוחות, בהפצה של מוצרים חדשים ובחקר קהילות. תחומי העניין שלו כוללים פרסום במכשירים ניידים ואסטרטגיה תחרותית. עבודותיו הופיעו ב-Journal of Marketing, Journal of Marketing Research, Journal of Personality and Social Psychology, Quantitative Marketing and Economics, Journal of Interactive Marketing, and Social Networks.

טל שושני הוא בוגר תואר שני (M.Sc., בהצטיינות) מהפקולטה לניהול ע"ש קולר באוניברסיטת תל אביב. בעל תואר ראשון בכלכלה וניהול (בהצטיינות יתרה) מהמרכז האקדמי רופין, חבר סגל הוראה במרכז האקדמי רופין ועוזר מחקר באוניברסיטת תל אביב. מחקרו מתמקד בהסבר וחיזוי התנהגות צרכנים באמצעות שימוש בנתוני עתק המגיעים מהמובייל.

ד"ר שחר רייכמן הוא חבר סגל בפקולטה לניהול ע"ש קולר באוניברסיטת תל אביב. הוא עוסק בתחום ניהול טכנולוגיה ומידע ומחקרו עוסקים בניית ויעבוד מתקדם של נתוני עתק (Big Data) בארגונים, בסביבות מקוונות וברשתות חברתיות, לצורך תמיכה בקבלת החלטות עסקיות ושיפור תהליכים צרכניים. מחקריו כוללים שימוש בכלי כריית מידע, למידת מכונה (Machine Learning), אקונומטריקה ושיטות ניתוח רשתות מורכבות למציאת תבניות ומאפיינים ייחודיים המאפשרים בניית מודלים לחיזוי ולהסבר של תהליכי מסחר מקוון. בשנת 2011 סיים את לימודי הדוקטורט באוניברסיטת תל אביב, ובשנת 2014 הצטרף לפקולטה לניהול לאחר פוסט-דוקטורט במכון הטכנולוגי של מסצ'וסטס (MIT).

תקציר

האימוץ הנרחב של טלפונים חכמים בשנים האחרונות הביא לשימוש מוגבר בנתוני מיקום הלקוחות הנאספים ממכשירים אלו. השימושים העסקיים בנתוני מיקום מבוססים בעיקר על "גיידור גיאוגרפי" – התאמה אישית של המידע המוצג ללקוחות על סמך מיקומם. מאמר זה מציע גישה חדשה לפילוח קבוצות צרכנים על בסיס דפוסי הניידות שלהם (שינוי במיקום לאורך זמן), ומדגים כיצד חילוף מאפייני ניידות פשוטים מאפשר לאמוד כיצד צרכנים משלבים גורמים מבוססי מיקום בהחלטותיהם. המחקר כלל ניתוח נתוני ניידות לחישוב מסלולי נסיעה יומיומיים של הצרכנים ולניתוח החלטותיהם בבחירת תחנות דלק לתדלוק. תוצאות המחקר מראות כי שילוב של דפוסי ניידות משפר בצורה משמעותית את דיוק המודלים לניתוח החלטות הרכישה, לעומת המודלים הקיימים המסתמכים רק על בחירות היסטוריות, מידע על המותג והמחיר הנוכחי. יותר מזה, ניתוח התוצאות מעיד על היכולת לנתח באמצעות נתוני הניידות בלבד, את הטרונגניות העדפות הלקוחות ורגישותם למחיר ולמרחק הנסיעה. למחקר תרומה אקדמית ופרקטית בהצגת דרך חדשה לשימוש בנתוני ניידות לקוחות להבנת החלטות רכישה.

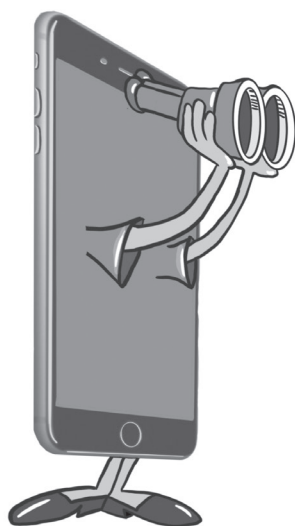
Andrews et al., 2016; Ghose et al., 2019; Molitor) (et al., 2020b; Zubcsek et al., 2017

במחקר זה אנו מציעים גישה חדשה לפילוח קבוצות צרכנים על פי דפוסי הניידות שלהם – דפוס מיקום המכשיר הנייד שלהם לאורך זמן. לשם כך אנו משלבים נתוני ניידות עירונית עם נתוני מחירי דלק כדי ללמוד: (1) אודות מסלולי הנסיעה היומיומיים של צרכנים; (2) לזהות "ביקורי תדלוק" שנצפו באחת מתחנות הדלק הנכללות בבסיס הנתונים (איור 1). לאחר חילוץ מסלולי הנסיעה אנו מחשבים את המרחק של כל תחנת דלק ממסלול הנסיעה השגרתי של כל צרכן באמצעות ניתוח תצפיות מיקום היסטוריות, מה שמאפשר דינמיות בחישוב המסלולים (והמרחקים הנגזרים) העשויים להשתנות מדי יום. במקביל, הניתוח כולל נתונים אודות מחירי הדלק היומייים של תחנות דלק באזורים הכלולים במחקר¹. נתונים אלו אפשרו לנו לבנות מודל סטטיסטי ולאמוד את שקלול התמורות (Trade-off) בין מחיר הדלק לבין מרחק הנסיעה לתחנת הדלק שמבצע הלקוח בעת תהליך קבלת ההחלטה (כלומר, ההחלטה היכן לתדלק).

במחקר זה אנו מרחיבים גישה דומה העושה שימוש במאפייני מסלולי הלקוחות בתוך קניון לשיפור התאמה

השימוש בטכנולוגיות ניידות, ובייחוד בטלפונים חכמים, הולך וגדל במהירות במהלך העשור האחרון. בשנת 2018 הערך הכלכלי של טכנולוגיות ושירותי המובייל היה 3.9 טריליון דולר, או 4.6% מהתמ"ג העולמי (GSMA Intelligence 2019). האימוץ הנרחב של טלפונים חכמים – כ-80% מאוכלוסיית ארה"ב החזיקו במכשיר כזה בשנת 2018 (Statista, 2019) – מספק לחברות ולגופים עסקיים מקור נתונים עשיר אודות צרכנים וסביבתם. ואכן, רוב חברות ה-B2C המובילות אוספות מידע מהטלפונים החכמים של לקוחותיהם אודות מיקומם במרחב הפיזי (Carto, 2018). עם זאת, מאמצייהן של חברות להפיק רווחים מנתוני מיקום של לקוחות התבססו במידה רבה על פרטוקולים פשוטים כמו "נידור גיאוגרפי" – התאמה אישית של תוכן האפליקציה ו"דחיפת" מידע אך ורק על סמך מיקומם הנוכחי (Danaher et al., 2015; Fong et al., 2015; Grewal et al., 2016; Luo et al., 2014; Molitor et al., 2020a). גישה זו מתעלמת מההטרונגיות של הלקוחות בכל מקום נתון, ולכן מחמיצה פוטנציאל הכנסות או "משאירה כסף על השולחן" (Zubcsek et al., 2017). לאור זאת, לאחרונה גבר העניין האקדמי והניהולי בהפקת מידע נוסף מהיסטוריית המיקומים של צרכנים, ומחקרים חדשים בתחום הראו עד כמה חשוב להתייחס להטרונגיות של לקוחות בתגובותיהם למודעות ולתמריצים שיווקיים במקומות שונים

1 הנתונים נאספו מ-Oil Price Information Services - <https://www.opisnet.com/>



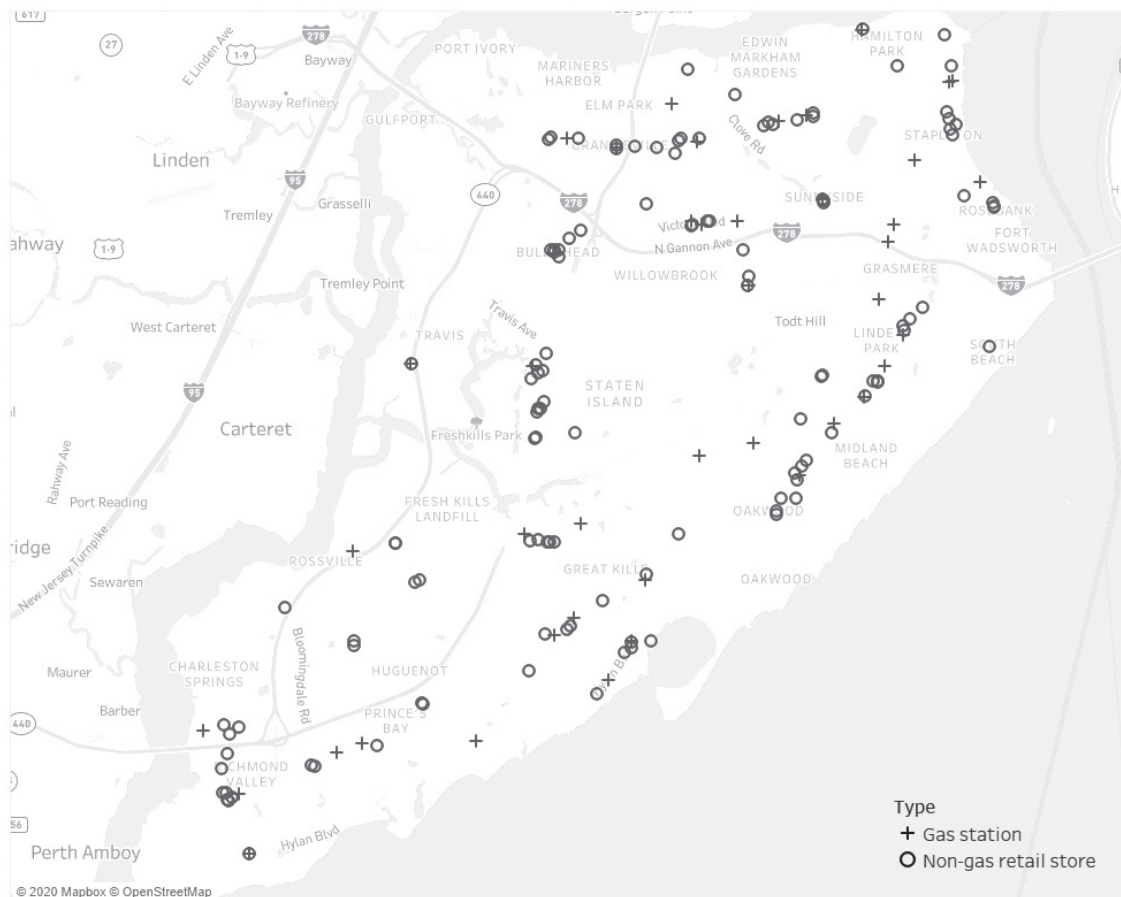
קיימים המללים רק בחירות היסטוריות (נאמנות), מידע על המותג והמחיר הנוכחי. עוד מלמדות התוצאות כי עבור הנחה של סנט אחד במחיר הדלק, הצרכן הממוצע מוכן לסטות ממסלול הנסיעה השגרתי ולנסוע 13 מטרים נוספים. כמו כן, התוצאות מגלות הטרוגניות של לקוחות ברגישותם למחיר ולמרחק, כך שפלח שוק אחד רגיש יותר למחיר ולכן מוכן לנסוע יותר על מנת ליהנות ממחיר תחרותי, ומנגד פלח שוק אחר רגיש יותר למרחק ולכן נוטה פחות להתפשר על נוחות ומעדיף לתדלק בתחנות דלק בקרבת מסלול הנסיעה. הטרוגניות זו מתוארת על ידי משתנים דמוגרפיים (הכנסה חציונית וגודל אוכלוסייה) מבוססי מסלול שאנו משלבים בנינוח.

מעבר לתרומה האקדמית, למחקר יש תרומה פרקטית וניהולית חשובה בהצגת דרך חדשה ליצירת ערך מנתוני ניידות של לקוחות.

אישית של הצעות שיווקיות (Ghose, Li and Liu, 2019), על ידי יישום אופן חישוב המסלול של הלקוח באזור גיאוגרפי רחב שאינו בתוך מבנה וכולל מידה מופחתת של נתוני מיקום. באופן זה אנו מדגימים כיצד במרחב עירוני, חילוף מאפיינים פשוטים מנתוני ניידות דלילים יחסית יכול לאפשר למשווקים לאמוד כיצד צרכנים משלבים גורמים מבוססי מיקום בהחלטותיהם. בפרט, אנו עונים על שאלות המחקר הבאות: (1) מהו שקלול התמורות בין מחיר הדלק לבין מרחק הנסיעה בתהליך קבלת החלטת רכישה בתחנות דלק? (2) כיצד משווקים יכולים להסביר את ההטרוגניות של הלקוחות ברגישותם הן למחיר והן למרחק?

תוצאות המחקר מראות כי שילוב גורמים מבוססי מיקום (כלומר מרחק הנסיעה לתחנה מהמסלול האופייני) במודלים לנינוח החלטות רכישה, משפר את הדיוק בכ-31% לעומת מודלים

איור 1: חנויות קמעונאיות שאינן תחנות דלק (בעיגול) ותחנות דלק (בפלוס) הנכללות בנתוני הניידות; Staten Island, NY



נתונים ומתודולוגיה

נתוני המחקר נאספו משני מקורות: (1) נתוני ניידות עירונית (היסטוריית נתוני מיקום של מכשירים ניידים) שסופקו על ידי Ubimo², חברת שיווק מדיה מבוססת מיקום המספקת פלטפורמה לניתוח התנהגויות צרכנים וקמפיינים פרסומיים; (2) נתוני מחירי דלק שנרכשו מ-Oil Price Information Services. שני המקורות כוללים נתונים אורכיים שנאספו במשך 6 חודשים (נובמבר 2017 – אפריל 2018) במדינת ניו יורק. נתוני הניידות מכילים תצפיות על משתמשי אפליקציות אנונימיים המזהים אך ורק על ידי מזהה המכשיר שלהם. כל תצפית כוללת את המיקום הגיאוגרפי (נקודת ציון GPS) של החנות הקרובה ביותר למכשיר (עד 100 מטרים), המותג של אותה חנות, מרחק המכשיר מהחנות, חותמת זמן ומשך שהייה במיקום. לדוגמה, כפי שמוצג בטבלה 1, המכשיר abc-123 נצפה 13 מטרים מחנות Starbucks הממוקמת בקו הרוחב 40.751317 וקו אורך -73.990844, ב-18 בנובמבר 2017, בשעה 14:45, למשך 1,196 שניות (כ-20 דקות). מידע זה מאפשר ניתוח של ההקשר של כל תצפית. בדוגמה זו אנו יכולים להסיק שצרכן זה ביקר באחד מסניפי Starbucks ובילה בחנות מספר דקות (ייתכן שקנה קפה).

חשוב לציין כי המכשירים הניידים אינם נצפים ברציפות במשך כל התקופה, וכי נתוני המחקר כוללים נתוני תצפיות של מכשירים ניידים רק כאשר היו בקרבת חנויות קמעונאיות מסוימות ותחנות דלק ספציפיות. המותגים הספציפיים נבחרו על פי פריסת רשת החנויות של המותגים (במדינת ניו יורק), כדי להבטיח מספיק תצפיות מכל מכשיר לצורך יישום המתודולוגיה המוצעת.

נתוני מחירי הדלק כוללים תצפיות על כל תחנות הדלק במדינת ניו יורק. כל תצפית מורכבת ממוזהה תחנת הדלק, מותג התחנה, המיקום שבו נמצאת התחנה (נקודת ציון GPS), מחיר הדלק היומי הממוצע, והתאריך שבו נקבע

טבלה 1: דוגמה לתצפית בודדת בסט נתוני הניידות

המחיר על ידי המפעיל. לכל תחנת דלק בוצעה התאמה בין שני מקורות הנתונים על פי המרחק ושם המותג. לאחר ההתאמה בוצע איחוד של מחירי הדלק ממקור נתונים (2) לתחנות במקור נתונים (1).

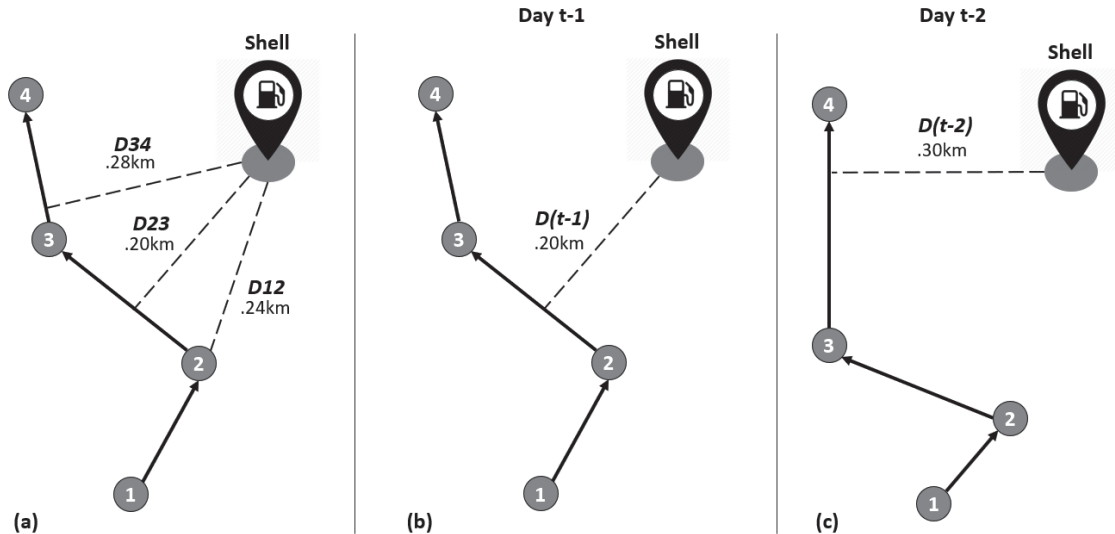
גישת ניתוח מסלולי תנועת לקוח וחישוב המרחקים מתחנות התדלוק

מסלול הנסיעה של צרכן מוגדר כנתיב המחבר את סדרת המיקומים הנצפים בנתונים (מסודרים על פי זמן) באותו היום. כפי שניתן לראות בדוגמה באיור 2, הצרכן שהה ביום מסוים בקרבת חנויות 1 עד 4; המסלול מוגדר כנתיב המחבר את רצף התנועה בין מיקומים אלו. המרחק של תחנת דלק ממסלול הנסיעה השגרתי של הצרכן מחושב על ידי מיצוע המרחק של כל תחנה מכל המסלולים היומיים שנצפו במהלך השבוע שקדם לכך. ראשית, מחשבים את מרחק תחנת הדלק מכל מקטע של המסלול היומי, וכוללים בחישוב רק את הערך המינימלי (באיור 2(a) הערך המינימלי הוא 0.2 ק"מ). שנית, מכיוון שהמסלולים עשויים להשתנות מדי יום, אנו לוקחים בחשבון נתונים ממספר ימים וממצעים את המרחק על פני מספר ימים שקודמים ליום הנתון (t); באיור 2, b ו-c, הערכים המינימליים הם 0.2 ו-0.3 ק"מ. בימים t-1 ו-t-2 בהתאמה, ולכן הממוצע הוא 0.25 ק"מ. לבסוף, כדי ללמוד על הפעילות השבועית של הצרכן אנו ממצעים את הערכים של שבעת הימים הקודמים (t-1 - t-7). החישוב מבוצע עבור כל צרכן וכל תחנת דלק הכוללים בנתוני המחקר.

חשוב לציין כי הנתונים אינם כוללים באופן ישיר אירועי תדלוק של צרכנים, ולכן אנו מסווגים את התצפיות כאירועי תדלוק על פי ההקשר של התצפית. ההגדרה לאירוע תדלוק מבוססת על: (1) מיקום – המכשיר היה בסמוך לאחת מתחנות הדלק הנכללות בנתונים; (2) מרחק מהתחנה – עד 50 מטר; (3) משך שהייה – נע בין 120 ל-1,500

מזהה	תאריך	שעה	מותג	קו רוחב	קו אורך	מרחק (מטרים)	משך שהייה (שניות)
abc-123	18/11/17	14:45:09	Starbucks	40.751317	-73.990844	13	1,196

איור 2: תיאור שלבי מתודולוגיית משתנה המרחק מבוסס המסלול שלנו



אוכלוסייה המהווה תת-מדגם ספציפי הנגזר מהמדגם הכולל במדינה, וכן פריסה מקובלת של תחנות דלק. המיקוד באזור מצומצם נועד למנוע מצב שבו הניתוח כולל תחנות דלק נוספות כחלק מהבחירה של הצרן, בעוד שבפועל תחנות אלו אינן אפשרויות ריאליות לתדלוק על ידי הצרן בעת נסיעתו. לכן בחירת אזור מבודד מבחינה גיאוגרפית הופכת את סט הבחירה של הצרן למציאותי יותר. הניתוח במחקר זה מבוסס על נתוני אזור Staten Island שבמדינת ניו יורק.

נתוני האזור שנבחרו מכילים 11.7 מיליון תצפיות (נתוני מיקום) שנאספו מ-273 אלף מכשירים, שנמצאו בסביבת 123 חנויות קמעונאיות שאינן תחנות דלק ו-43 תחנות דלק (איור 1). לאחר סינון התצפיות לפי הקריטריונים שתוארו לעיל, נותרו עם 935 צרכנים ו-81,168 תצפיות (מתוכן 1,625 אירועי תדלוק). לאור המתאם החזק בין מחירי שלושת סוגי הדלק⁴ שנמכרו בתחנות (איור 3), אנו מתמקדים בדלק מסוג נטול עופרת (Unleaded gas) בלבד. במהלך תקופת הניתוח, מחירי הדלק ב-43 תחנות הדלק נעו בין 2.219 ל-3.519 דולרים לגלון ($M = 2.82$, $SD = .14$; איור 3 מציג את מגמות המחיר לאורך תקופת הניתוח).

שניות. אנו מניחים שתי הנחות: הראשונה – צרכנים שנצפו במרחק של יותר מ-50 מטרים מתחנת הדלק נמצאים מחוץ לתחנה. השנייה – צרכנים ששהו פחות מ-120 שניות הם עוברי אורח, וצרכנים השהים יותר מ-1,500 שניות נמצאים בתחנה שלא לצורך תדלוק (למשל עובדי תחנת הדלק או מועסקים בסביבתה).

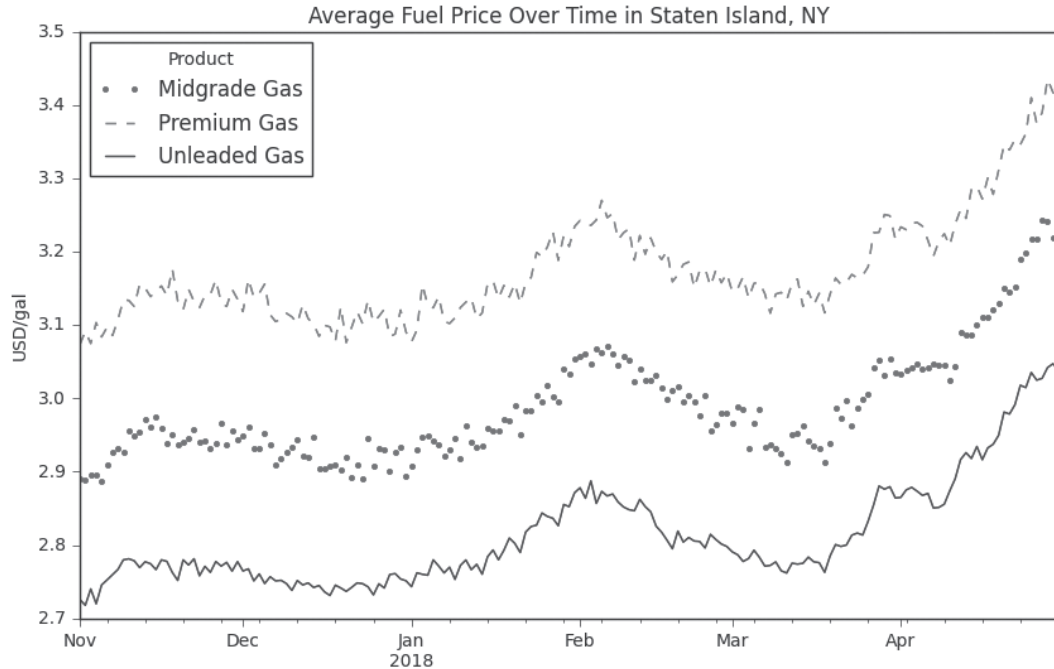
על מנת לקבל תוצאות מהימנות יותר, הניתוח כולל רק צרכנים שיש להם לפחות שני אירועי תדלוק במהלך התקופה (6 חודשים), או שעל פי הקריטריונים שלעיל תדלקו את מכוניתם בימים שאינם רצופים. התנאי הראשון מסנן צרכנים שנסעו באזור באופן חד-פעמי (ולכן לא ניתן לנתח את החלטותיהם לאורך זמן), ואילו התנאי השני מסנן ביקורים שאינם סטנדרטיים להחלטות תדלוק, כגון ספקים של תחנת הדלק.

אזור גיאוגרפי

הניתוח בוצע על אזור גיאוגרפי קטן ומבודד יחסית המהווה את תחום החלטות התדלוק עבור הלקוחות³. אזור זה כולל

4 מקדם המתאם של פירסון בין מחירי סוגי הדלק (unleaded, midgrade), (premium, midgrade), (premium, unleaded) הינו 0.975, 0.978, ו-0.973, בהתאמה (כולם מובהקים ברמת $p < .001$).

3 המחקר מניח שצרכני דלק עושים השוואת מחירים ומוזעים למחירי הדלק השונים, בייחוד המחירים באזור הגיאוגרפי שבו הם נמצאים.



ניתוח תיאורי וכוח ניבוי

במטרה לנתח את שקלול התמורות של צרכנים בין מחיר הדלק למרחק הנסיעה בעת קבלת החלטת רכישה בתחנות דלק באותו אזור גיאוגרפי, אנו בוחנים מודל רגרסיה הקרוי Latent class conditional logit (LCL) המבוסס על עיקרון כלכלי – הצרכנים ממקסמים את התועלת הנובעת מבחירתם (במקרה הזה בתחנת דלק). יתרונותיו העיקריים של מודל זה הם האפשרות לפרש את התוצאות בפשטות על ידי בחינת מקדמי הרגרסיה, וכן האפשרות לנתח את ההטרונגניות בהעדפותיהם של פלחים שונים בשוק.

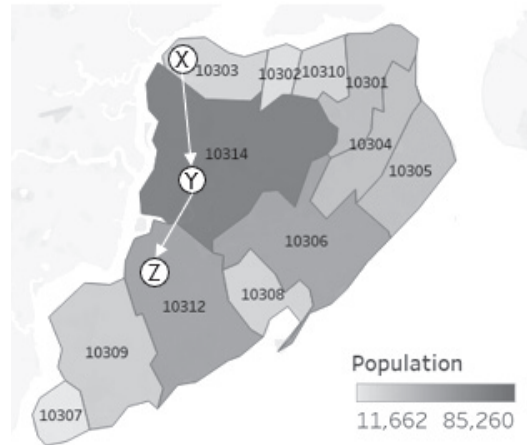
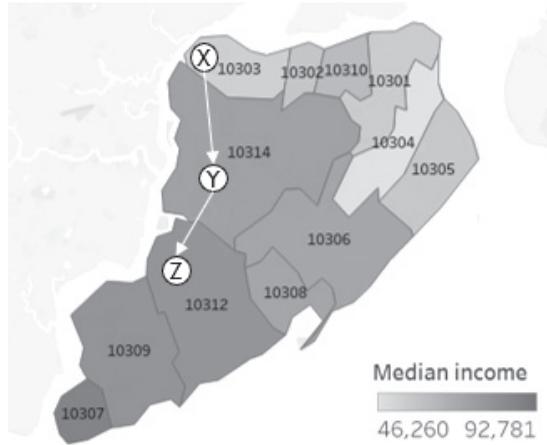
משתני המודל

עבור כל יום בתקופת ניתוח הנתונים, המודל כולל את מחיר הדלק היומי הממוצע בכל תחנה ואת משתנה המרחק שתואר לעיל. המודל מניח שלצרכנים יש העדפות שאינן משתנות על פני זמן. עם זאת, אנו כוללים גם משתני נאמנות לאורך זמן (הן למוטג והן לתחנה) המתוארים במאמרם של

(Guadagni and Little, 1983). באופן בסיסי, ערכיהם של משתני הנאמנות גדלים עבור המוטג או התחנה שבהם הצרכן תדלק בפעם הקודמת וקטנים עבור שאר המותגים/תחנות. כמו כן, המודל כולל כמשתני דמה את מותגי תחנות הדלק המופיעים בסט הנתונים, כדי להתחשב בפופולריות של כל מותג באזור הגיאוגרפי הנחקר. כל אחת מ-81,168 התצפיות מייצגת בחירה שנעשתה על ידי צרכנים ביום נתון – הם בחרו לתדלק באחת מ-43 תחנות הדלק או לחילופין החליטו שלא לתדלק באף אחת מהן (כלומר לא לתדלק כלל).

נוסף על כך, כדי לערוך ניתוח אודות פילוח קבוצות צרכנים בשוק הנתון במודל ה-LCL, אנו משלבים משתנים דמוגרפיים ממפקד אוכלוסין בארה"ב (הכנסה חצינית וגודל אוכלוסייה) ומשייכים אותם לצרכנים על בסיס דפוסי הניידות שלהם. כדי לקשר בין משתני המפקד לבין הצרכנים, אנו מחשבים עבור כל צרכן את הממוצע המשוקלל של ההכנסה החצינית וגודל האוכלוסייה במיקודים (ZIP codes) המכילים כל אחד מהמיקומים הנצפים של המכשיר הנייד. לדוגמה, כפי שמוצג באיור 4, רמת ההכנסה החצינית במיקודים 10304, 10303, ו-10312 (Staten Island) ניו

איור 4: משתני מפקד לפי מיקוד; Staten Island, NY



בתחנה. שנית, נמישות הביקוש לדלק ביחס למחיר (פירוש החישוב אינו נכלל במאמר) תואם לממצאי מחקרים קודמים; Hughes et al., 2008, Levin et al., 2017, Small and Van Dender 2007, Park and Zhao, 2010. שלישית, שילוב גורמים מבוססי מיקום (כלומר משתנה המרחק מבוסס המסלול) בהחלטות רכישה בתחנות דלק, משפר בצורה משמעותית את טיב התאמת המודל ($\chi^2(1)=1,038.56, p<.001$). לבסוף, התוצאות של מודל 2 מלמדות כי בהפחתה של סנט אחד (לגלון) במחיר הדלק, הצרכן הממוצע מוכן לסטות ממסלול הנסיעה השגרתי ולנסוע 13 מטרים נוספים.

תוצאות מודל 3, המחלק את הצרכנים לשתי קבוצות, מראות כי קיימת הטרוגניות ניכרת ברגישות של צרכנים הן למחיר הדלק והן למרחק הנסיעה: כ-60% מהצרכנים נאמנים מאוד לתחנות שבהן הם נוטים לבקר, ורגישים למחיר כמעט פי שלושה מהרגישות הממוצעת בשוק (בממוצע, עבור הנחה של סנט אחד לגלון, הם מוכנים לסטות מהמסלול ולנסוע 35 מטרים נוספים), בעוד שאר הצרכנים מעדיפים לבחור בחלופות נוחות יותר שאינן דורשות סטייה ארוכה מהמסלולים השגרתיים שלהם (בממוצע, עבור הנחה של סנט אחד לגלון, הם מוכנים לסטות מהמסלול ולנסוע רק שני מטרים נוספים) ונותנים פחות משקל למותג התחנה. בנוסף, תוצאות מודל 4 מראות כי צרכנים המשתייכים לקבוצה 1 (רגישה יותר למחיר) נטו יותר לבקר במקומות עם הכנסה חצינית גבוהה וגודל אוכלוסייה קטן יותר. איור 5 מתאר כיצד הימצאות במקומות שונים משפיעה על ההסתברות להשתייך לקבוצה 1 – ככל שהאזור כהה יותר, כך גדלה ההסתברות להשתייך לקבוצה 1.

יורק) היא \$55, \$78 ו-\$85 (באלפים) בהתאמה. כלומר, אם צרכן ביקר במיקומים X, Y, Z המשתייכים למיקודים לעיל (באותו הסדר) ושהיה בהם 1,000, 2,000 ו-1,000 שניות, בהתאמה, אנו נקשר צרכן זה עם הכנסה חצינית של \$74,000 (ערך הממוצע המשוקלל). שיטה זו מיושמת גם על משתנה האוכלוסייה.

תוצאות

ניתוח מחירי הדלק ב-Staten Island מראה כי ההבדל היומי בין המחיר הגבוה ביותר למחיר הנמוך ביותר נע בין 0.38 ל-1.1 דולרים לגלון ($M = 0.615, SD = 0.113$) במהלך תקופת המחקר.

על מנת להעריך את ההשפעה של שילוב משתנה המרחק מבוסס המסלול, וכן את ההתייחסות להטרוגניות בהעדפותיהם של צרכנים על ההסבר לבחירה בתחנת דלק מסוימת, אנו משווים ארבעה מודלים. מודל 1 מקבץ את כלל הצרכנים תחת קבוצה אחת (כלומר אינו לוקח בחשבון את ההטרוגניות שלהם) ואינו כולל את משתנה המרחק. מודל 2 כולל את משתני מודל 1 וגם את משתנה המרחק. מודל 3 מחלק את הצרכנים לשתי קבוצות וכולל את משתנה המרחק. מודל 4 כולל את משתני מודל 3 וגם את משתני המפקד (הכנסה חצינית וגודל אוכלוסייה) שתוארו לעיל.

תוצאות הניתוח מדווחות בטבלה 2. ראשית, למשתנה המחיר אפקט שלילי מובהק ברמת $p<.05$ במודלים 1 ו-2, כלומר עלייה במחיר מקושרת בירידה בסיכוי לתדלוק

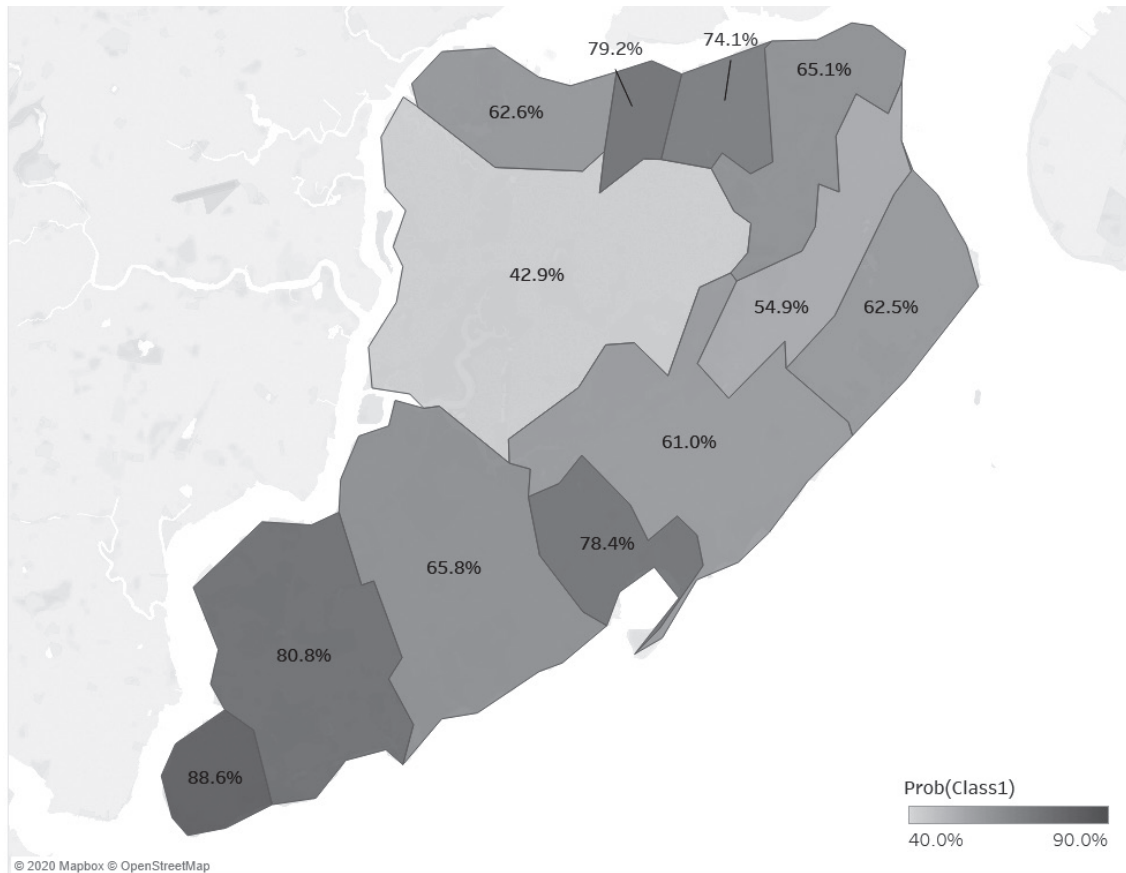
טבלה 2: תוצאות ניתוח הרגרסיה

DV: Station Choice	Model 1 No-Trajectory	Model 2 Full Model	Model 3 Two Latent Classes		Model 4 Two Latent Classes (CP)	
(Staten Island, NY)			Class 1	Class 2	Class 1	Class 2
Price (USD/gal)	-.502** (-2.32)	-.440** (-2.04)	-.638* (-1.86)	-.198 (-.41)	-.580* (-1.65)	-.336 (-.74)
Distance (km)		-.339*** (-27.51)	-.181*** (-10.81)	-.832*** (-9.40)	-.175*** (-7.95)	-.810*** (-9.53)
Station Loyalty	3.765*** (34.54)	2.782*** (24.63)	3.376*** (17.03)	1.555*** (6.11)	3.419*** (15.54)	1.595*** (6.52)
Brand Loyalty	-.130** (-1.19)	-.100 (-.92)	-.137 (-.82)	-.033 (.43)	-.181 (-1.00)	.024 (.12)
7-Eleven ^a	.017 (.12)	-.179 (-1.25)	-.110 (-.43)	-.325 (-1.38)	-0.103 (-.38)	-.313 (-1.38)
ExxonMobil ^a	-.121 (-1.28)	-.151 (-1.60)	-.0078 (-.05)	-.299 (-1.91)	.024 (.90)	-.310* (-1.90)
Shell ^a	-.404*** (-3.64)	-.301*** (-2.70)	.031 (.16)	-.811*** (-3.36)	.061 (.76)	-.799*** (-3.34)
BP ^a	-.253*** (-2.92)	-.253*** (-2.92)	.147 (.85)	-.918*** (-4.27)	.147 (.79)	-.801*** (-3.68)
CITGO ^a	-.566*** (-6.00)	-.564*** (-5.96)	-.410** (-2.37)	-.732*** (-3.97)	-.398** (-2.17)	-.706*** (-3.97)
Est. Class Share			59.36%	40.64%	57.59%	42.41%
Constant					.210 (.84)	0
Wt. Median Income					.255* (1.83)	0
Wt. Population					-.268** (-2.17)	0
LL	-12,802.81	-12,283.53	-12,200.91		-12,197.24	
BIC	25,801.76	24,778.29	24,629.19		24,606.54	

z-scores in parentheses. Observations: 81,168. Groups: 935.

^aBaseline = Speedway.

p<0.1, **p<0.05, ***p<0.01*



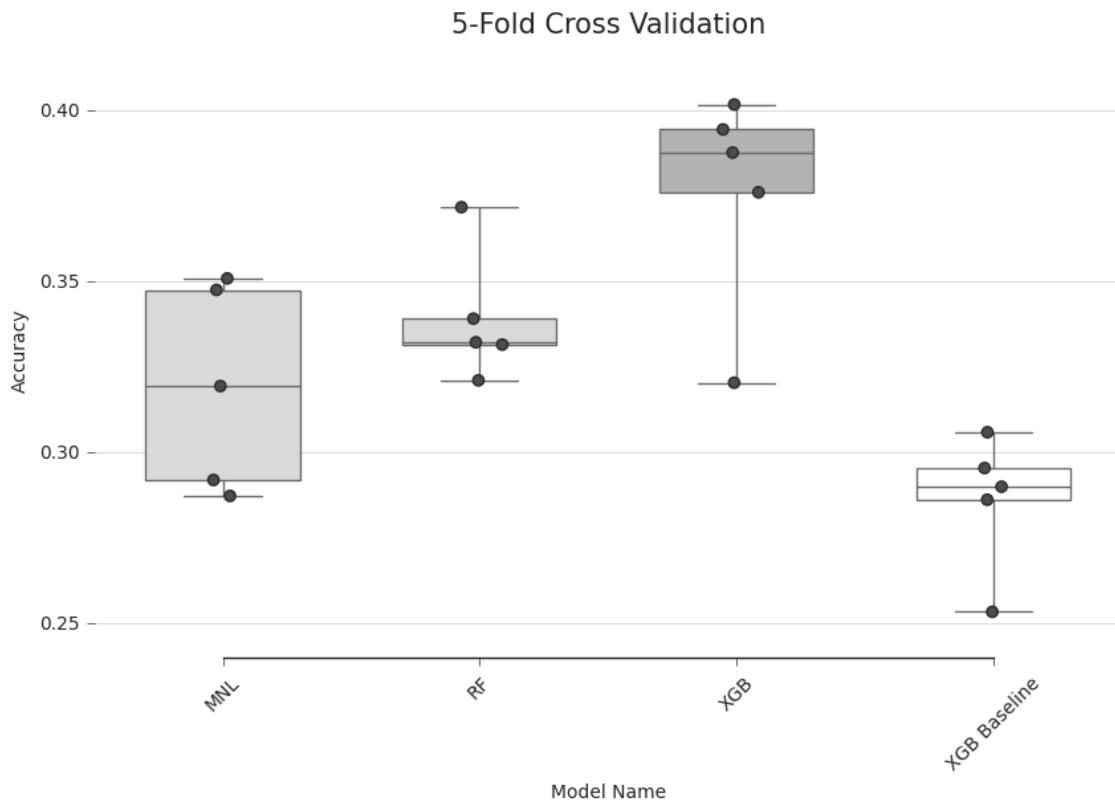
הערכת כוח הניבוי

הבאים: Multinomial Logistic Regression (MNL), Extreme Gradient Boost-*i* RandomForest (RF) (XGB). עבור כל מודל אנו בוחנים את התוצאות על ידי 5-Fold Cross Validation. כדי לכמת את התרומה של שילוב משתנה המרחק מבוסס המסלול לכוח הניבוי, אנו משווים את המודל שזכה לתוצאה הטובה ביותר בשלב השוואת המודלים המסווגים למודל בסיסי שאינו כולל את משתנה המרחק. כפי שמוצג בטבלה 3 ובאיור 6, האלגוריתם XGB נותן את התוצאה הטובה ביותר מבין שלושת האלגוריתמים (דיוק בסיווג ההחלטה של 37.6% בממוצע). בנוסף, שילוב משתנה המרחק מבוסס המסלול מביא לעלייה של 31.38% לעומת המודל הבסיסי.

אחד היישומים האפשריים למודל המתואר הוא ביצוע סימולציות מחירים, וכך קמעונאים בתעשיית תחנות הדלק יוכלו לקבוע מחירים אופטימליים כדי להגדיל את רווחיהם. ניתן לצפות כי אותם מפעילים יאמצו את המודל אם הוא מסוגל לחזות החלטות רכישה בתחנות דלק על ידי צרכנים. בהתאם, ביצענו חיזוי החלטות רכישה בתחנות דלק על נתוני מבחן שלא נכללו בחישוב המודל (Out of sample). מכיוון שהשאלה היא "היכן הצרכן יתדלק", אנו מתמקדים בתצפיות שסווגו כאירועי תדלוק בלבד (1,625 אירועים המהווים 2% מכלל התצפיות). בחרנו להשוות את מודלי הסיווג (Classification)

DV: Station choice (43 alternatives)				
Model	MNL	RF	XGB	XGB baseline
Accuracy	31.94%	33.91%	37.60%	28.62%

איור 6: השוואת תוצאות מודלים שונים לסיווג החלטות תדלוק



ניידים, בוצע חישוב של מסלולי הנסיעה של צרכנים וכן חישוב מרחק של תחנות הדלק באזור מהמסלול השגרתי של כל צרכן. המחקר מציג שיטה להגדרת המסלול של צרכנים באמצעות שימוש נתוני ניידות דלילים יחסית, ומנתח את העדפות הצרכנים באמצעות מודל בחירה פרמטרי. התוצאות מראות כי מודלים הכוללים את נתוני מסלולי הנסיעה של הצרכנים מדויקים יותר ממודלים קיימים. נוסף על כך, התוצאות מראות כי נתוני המיקום

סיכום ומסקנות

מחקר זה מדגים כיצד ניתוח נתוני מיקום צרכנים בסביבה עירונית (ליד חנויות קמעונאיות, מסעדות, תחנות דלק וכדומה) מאפשר ליצור תובנות חדשות אודות העדפותיהם של צרכנים. בעבודה זו אנו מציגים גישה חדשה לפילוח לקוחות על סמך דפוסי הניידות שלהם בלבד. בהתבסס על מאגר נתונים רחב, הכולל נתוני מיקום של אלפי מכשירים

מאפשרים גם לנתח באופן מדויק יותר את הרגישות של צרכנים למחיר ולמרחק ומעידות על הטרוגניות ניכרת בהעדפות של קבוצות צרכנים.

למחקר מספר מגבלות שיש ללמוד ולבחון במחקרי המשך בנושא: (א) המחקר מניח שצרכני דלק משווים מחירים ומוזעים למחירי הדלק השונים, אך נתוני המחקר אינם מאפשרים למדוד את המידע שהיה ברשות כל אחד מהלקוחות בעת החלטת התדלוק. השימוש הנרחב באפליקציות ניווט הכוללות מחירי דלק וכן השימוש באתרים המשווים מחירי דלק, מחזקים את ההנחה לגבי השוואת המחירים, בייחוד המחירים בתחנות הדלק באזור הגיאוגרפי שבו הלקוחות נמצאים. עם זאת, ייתכן שחלק מהצרכנים אינם מבצעים השוואה זו ואינם מודעים לאפשרויות המחיר השונות. (ב) המחקר מתמקד במוצר אחד – דלק לרכב. מדובר במוצר הכרחי (לבעלי רכבים המעוניינים לנסוע ברכב) שנרכש באופן חזרתי. מחקר המשך יוכל לבחון אם ניתן להשליך את תוצאות המאמר על סוגי מוצרים אחרים והאם האפקט המדווח יהיה דומה.

המחקר תורם לספרות בתחומי ניהול וכלכלה, ובייחוד לתחומי השיווק וטכנולוגיית המידע. בעוד שמחקרים

קודמים בחנו כיצד משפיע המרחק על בחירה בחנויות על סמך מיקום סטטי (למשל מקום המגורים של הצרכן; Athey et al. 2018; Bell, Ho, and Tang, 1998; Chan et al. 2007; Black, Ostlund, and Westbrook, 1985; Hortman et al., 1990; Luo et al., 2020; Nishida and Remer, 2018), מחקר זה מציג גישה דינמית שמתחשבת בדפוסי ניידות (מסלול הנסיעה השגרתי של הצרכן), המייצגת תמונה מציאותית יותר של תהליך קבלת ההחלטות בבחירת חנויות.

מלבד התרומה האקדמית, למחקר יש השלכות ניהוליות מעשיות. קמעונאים בתעשיית תחנות הדלק יוכלו לבצע סימולציות של שינוי במחירי המוצרים במטרה להגדיל את הרווחים על ידי קביעת מחירים אופטימליים או בהתאמה אישית של מחירים עבור פלחים שונים בשוק הנתון. כמו כן, חילוץ והפקת דפוסי ניידות מנתוני מיקום של מכשירים ניידים, יכולים לסייע לקמעונאים בתעשיות אחרות (מלבד תחנות דלק), למשל בתמחור הצעות שיווקיות ליום הבא (Luo et al., 2014) או בזמן אמת (Ghose et al., 2019), ולייעד אותן לצרכנים בהתבסס על דפוסי הניידות שלהם.

sr@tauex.tau.ac.il

ד"ר שחר רייכמן

- Andrews, Michelle, Xueming Luo, Zheng Fang, and Anindya Ghose (2016), "Mobile Ad Effectiveness: Hyper-Contextual Targeting with Crowdedness," *Marketing Science*, 35 (2), 33–218.
- Athey, Susan, David Blei, Robert Donnelly, Francisco Ruiz, and Tobias Schmidt (2018), "Estimating Heterogeneous Consumer Preferences for Restaurants and Travel Time Using Mobile Location Data," *AEA Papers and Proceedings*, 108, 64–67.
- Bell, David R., Teck-Hua Ho, and Christopher S. Tang (1998), "Determining Where to Shop: Fixed and Variable Costs of Shopping," *Journal of Marketing Research*, 35 (3), 352.
- Black, William C., Lyman E. Ostlund, and Robert A. Westbrook (1985), "Spatial Demand Models in an Intrabrand Context," *Journal of Marketing*, 49 (3), 106.
- Carto (2018), "The State of Location Intelligence in 2018," CARTO, (accessed 12/11/2019), [available at <https://carto.com/state-of-location-intelligence-2018/>].
- Chan, Tat Y., V. Padmanabhan, and P.B. Seetharaman (2007), "An Econometric Model of Location and Pricing in the Gasoline Market," *Journal of Marketing Research*, 44 (4), 35–622.
- Danaher, Peter J., Michael S. Smith, Kulan Ranasinghe, and Tracey S. Danaher (2015), "Where, When, and how Long: Factors that Influence the Redemption of Mobile Phone Coupons," *Journal of Marketing Research*, 52 (5), 25–710.
- Fong, Nathan M., Zheng Fang, and Xueming Luo (2015), "GeoConquering: Competitive Locational Mobile Promotions," *Journal of Marketing Research*, 52 (10), 35–726.
- Ghose, Anindya, Beibei Li, and Siyuan Liu (2019), "Mobile Targeting Using Customer Trajectory Patterns," *Management Science*, 65 (11), 49–5027.
- Grewal, Dhruv, Yakov Bart, Martin Spann, and Peter Pal Zubcsek (2016), "Mobile Advertising: A Framework and Research Agenda," *Journal of Interactive Marketing*, 34, 3–14.
- GSMA Intelligence (2019), "GSMA Intelligence - Research - The Mobile Economy 2019," (accessed 12/11/2019), [available at <https://www.gsmainelligence.com/research/?file=b9a6e6202ee1d5f787cfebb95d3639c5>].
- Guadagni, Peter M. and John D. C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (3), 38–203.
- Hortman, Sandra Mccurley, Arthur W. Allaway, J. Barry Mason, and John Rasp (1990), "Multisegment analysis of supermarket patronage," *Journal of Business Research*, 21 (3), 23–209.
- Hughes, Jonathan, Christopher R. Knittel, and Daniel Sperling, (2008) "Evidence of a shift in the short-run price elasticity of gasoline demand," *The Energy Journal* 29 (1).
- Levin, Laurence, Matthew S. Lewis, and Frank A. Wolak (2017), "High frequency evidence on the demand for gasoline," *American Economic Journal: Economic Policy* 9 (3), 47–314.

Luo, Xueming, Michelle Andrews, Zheng Fang, and Chee Wei Phang (2014), "Mobile Targeting," *Management Science*, 60 (7), 56–1738.

Luo, Xueming, Yuchi Zhang, Fue Zeng, and Zhe Qu (2020), "Complementarity and Cannibalization of Offline-to-Online Targeting: A Field Experiment on Omnichannel Commerce," *MIS Quarterly* 44 (2).

Molitor, Dominik, Martin Spann, Anindya Ghose, and Philipp Reichhart (2020a), "Measuring the effectiveness of location-based advertising: A randomized field experiment." *Available at SSRN* 2645281.

Molitor, Dominik, Peter Pal Zubcsek, Martin Spann, and Philipp Reichhart (2020b), "Location Retargeting," working paper, Fordham University.

Nishida, Mitsukuni, and Marc Remer (2018), "The determinants and consequences of search cost heterogeneity: Evidence from local gasoline markets," *Journal of Marketing Research*, 55(3), 305-320.

Park, Sung Y., and Guochang Zhao (2010), "An estimation of US gasoline demand: A smooth time-varying cointegration approach," *Energy Economics* 32 (1), 110-120.

Small, Kenneth A., and Kurt Van Dender (2007), "Fuel efficiency and motor vehicle travel: the declining rebound effect," *The Energy Journal* 28 (1).

Statista (2019), "Number of smartphone users in the U.S. 2010-2023," (accessed 12/11/2019), [available at <https://www.statista.com/statistics/201182/forecast-of-smartphone-users-in-the-us>].

Zubcsek, Peter Pal, Zsolt Katona, and Miklos Sarvary (2017), "Predicting Mobile Advertising Response Using Consumer Colocation Networks," *Journal of Marketing*, 81 (4), 26–109.