

עדכונים משמעותיים באומדן להתחלות הבנייה

- הנתונים על התחלות הבנייה במשק משמשים תשומה חשובה לגיבוש המדיניות בשוק הדיור כמו גם לגיבוש המדיניות המוניטרית. ניתוח זה מראה כי בין 2004 ל-2014 האומדן הרבעוני הראשון להתחלות הבנייה מוטה כלפי מטה בכ-10% בממוצע. הנתון מתעדכן כלפי מעלה בעיקר בשנה הראשונה לאחר פרסומו, וכעבור כשנתיים (8 פרסומים רבעוניים) הוא מתייצב.
 - קובעי המדיניות בתחום הדיור חשופים לאומדן הראשון – וגם הציבור הרחב חשוף אליו משום שבשנים האחרונות הוא זוכה לסיקור תקשורתי נרחב – ולכן הם עלולים לגבש תפיסה מוטעית לגבי הפעילות במגזר הבנייה. על כן יש לנקוט משנה זהירות כאשר משתמשים באומדן הראשון, ובד בבד להשקיע משאבים בשיפורו.
 - איסוף הנתונים תלוי בגורמים רבים שעוסקים בייזום הבנייה, באישורה ובהתחלתה בפועל, וחלקם מעבירים ללמ"ס נתונים חלקיים. תהליך מורכב זה מקשה על פרסום נתונים איכותיים בזמן אמת.
 - מודל סטטיסטי לחיזוי הרוויזיות מאפשר לצמצם את גודל העדכון הממוצע לכל הפחות בשליש, ולגבש תחזית שאינה מוטה לכיוון מסוים – כלומר תחזית שמתעדכנת לעתים כלפי מטה ולעתים כלפי מעלה. מודל זה מסתמך על הנתונים הידועים בזמן אמת (Ex Ante), ולכן ניתן לשפר באמצעותו את האומדן בעת הפרסום.
- הלשכה המרכזית לסטטיסטיקה (למ"ס) מפרסמת מדי רבעון נתונים על מספר התחלות הבנייה. בדיקה שערכנו העלתה כי הנתון שמתפרסם לראשונה מוטה כלפי מטה באופן שיטתי ומשמעותי כמותית. נתון זה מתעדכן כלפי מעלה בעיקר בשנה הראשונה לאחר פרסומו, וכעבור כשנתיים (8 פרסומים רבעוניים) הוא מתייצב כשערכו גבוה יותר באופן מובהק סטטיסטית. הערכת חסר זו משפיעה גם על שיעורי השינוי, הן בין רבעונים סמוכים והן בין רבעונים מקבילים, והיא עלולה לעוות את התמונה המצטיירת בנוגע לפעילות במגזר הבנייה בישראל. עבודה זו נועדה להפנות את תשומת לבם של צרכני הנתון על התחלות הבנייה לבעייתיות הקיימת בו, ולהמליץ לעשות בו שימוש זהיר וביקורתי בעת קבלת החלטות. נוסף לכך העבודה מראה כי ניתן להציע מודלים שמאפשרים לתקן את האומדן הראשוני שהתפרסם עד כה.

התחלות הבנייה: מקורות הנתונים ועדכונם

התחלת הבנייה – היינו התחלת החפירה של היסודות לבניין – מתחילה גם את החוליה האחרונה בשרשרת הייצור של יחידת דיור. האטה והאצה במספר ההתחלות אינן תלויות רק בהתאמה בין כוחות הביקוש וההיצע בשוק, אלא גם בחוליות הקודמות (הממשלתיות או המקומיות) בשרשרת הייצור: החל מייזום הפרויקט, עבור דרך אישורו במוסדות התכנון, וכלה בשיווק הקרקע, בפיתוחה וברישוי הבנייה.

גופים רבים מעורבים בתהליך והדבר מקשה על איסוף נתונים איכותיים לגבי התחלות הבנייה ועל אמידת מספרן. הלמ"ס אוספת נתונים מארבעה מקורות: (א) את הנתונים על ההתחלות ביוזמת משרד הבינוי והשיכון היא מקבלת מהמשרד בקובץ מנהלי חודשי. אף שמדובר בקובץ מנהלי המשרד מעדכן ומטייב אותו, כפי שהדגישה הלמ"ס בפרסום האחרון על התחלות הבנייה¹; (ב) את הנתונים על שאר סוגי הבנייה (מרביתם ביוזמה פרטית וחלקם ביוזמה ציבורית, אם של הממשלה ואם של הרשויות המקומיות) הלמ"ס מקבלת מדיווח חודשי על היתרי הבנייה שמוסרות ועדות התכנון המקומיות, ובהתאם לכך היא עורכת סקר טלפוני חודשי בקרב חברות הבנייה; (ג) את הנתונים על ההיתרים שלא נסקרו בסקר הטלפוני הלמ"ס מייחסת לפי מודל סטטיסטי; (ד) את הנתונים על הבנייה

1 בהודעה שהלמ"ס הפיצה לעיתונות ב-15 בדצמבר 2016 היא מבחירה כי משרד הבינוי והשיכון טייב בחודשים האחרונים את נתוני לגבי מספר ההתחלות שהוא יזם בשנים 2014–2016. הדבר השפיע במיוחד על הנתון לגבי ההתחלות ביוזמה ציבורית, אך מאחר שהפילוח לפי יזום אינו מתפרסם עוד, קשה לדעת מהם סדרי הגודל של העדכונים שנבעו מהטיוב.

ללא היתר היא מקבלת מדיווח שנתי שמשדר הפנים מוסר על גזרי דין רלוונטיים. כדי לסבר את האוזן נתבונן בנתון מדצמבר 2016 על התחלות הבנייה בשנת 2015. רובו הגיע מהסקר הטלפוני (55.2% מהתחלות), כשתי חמישיות הגיעו מקובץ משרד הבינוי (22.5%) ומייחוס נתונים (19.3%), והשאר הגיע מנתוני משרד הפנים על בנייה בלתי חוקית (2.6%) וממודל זקיפות (0.4%)².

הלמ"ס משקיעה מאמצים רבים בטיוב הנתונים שהיא מקבלת, אך כרגע עדיין קיימות בהם הרבה מגבלות שמקשות על שיפור משמעותי באומדן הראשון. הנתונים המקוריים על בנייה ביוזמה פרטית חשופים לעדכוני רטרואקטיביים מכמה סיבות: פיגור בדיווח שהוועדות המקומיות לתכנון ובנייה מוסרות על היתרי הבנייה, פיגור בדיווח שהיזמים מוסרים על מועדי התחלת הבנייה וגמר הבנייה, התחלת הבנייה לפני הנפקת היתר³, ופיגור בגילוי היתרים כפולים במאגר המידע⁴. כדי להתמודד עם הפיגור בדיווח על היתרים יצרה הלמ"ס מודל סטטיסטי שזוקף נתונים בהתאם להיתרים שניתנו ביישוב נתון בתקופות קודמות ובהתאם למגמות הקיימות ביישובים דומים. כדי להתמודד עם הפיגור בדיווח על התחלות הבנייה הלמ"ס מתמקדת בייזמים שקיבלו היתר בתקופה מוקדמת יותר, בוחנת כמה זמן חלף עד שהם החלו לבנות, ובהתאם לכך זוקפת את מועד ההתחלה החזוי. שיטות אלו נועדו לצמצם ככל האפשר הטיות אפשריות של נתוני הסקר. בהמשך הלמ"ס מחליפה את הזקיפות בנתונים שמגיעים באיחור⁵. בהקשר זה יש להעיר כי מודל הזקיפות אינו כולל את הנתונים המנהליים המתקבלים ממשרד הבינוי והשיכון על התחלות הבנייה ביוזמה ציבורית. המשרד עדכן את הנתונים ל-2009—2014 בכ-28%+ בממוצע בשעה שהנתונים על הבנייה ביוזמה פרטית התעדכנו בכ-10%+ בממוצע, והדבר מעיד כי הנתונים המתקבלים ממשרד השיכון והבינוי מתאפיינים באיכות פחותה⁶. אולם ההתחלות ביוזמה ציבורית היוו 20% מסך ההתחלות בשנים הנידונות, ולכן לרוויזיות בהן יש השפעה משוקללת קטנה יותר על הנתון הכולל.

הפערים בין הפרסום הראשון לפרסום העדכני

איור 1 משווה בין האומדן הראשוני למספר התחלות הבנייה לבין האומדן העדכני ביותר⁷, והוא מתמקד בתקופה המשתרעת בין הרבעון הראשון של 2004 לרבעון השלישי של 2016. ניתן לראות כי האומדן הראשוני תמיד נמוך מהעדכני. כאשר מבטאים את הפערים במונחי שיעורם מתוך האומדן הראשוני (הקו הסגול), מוצאים כי החל מ-2011 הם גדלו; הפער הממוצע בתקופה כולה עומד על כ-11% מהתחלות הבנייה. עוד ניתן להבחין כי בשנה האחרונה שיעור העדכון נמוך, אך כפי שנראה בהמשך, רוב העדכון נערך בשנה שלאחר הפרסום הראשוני, ולכן ניתן לצפות ששיעורי העדכון ב-2015—2016 ימשיכו לעלות.

איורים 2א ו-2ב בוחנים, בהתאמה, את שיעור השינוי במספר ההתחלות בהשוואה לתקופה המקבילה בשנה הקודמת ובהשוואה לרבעון הקודם. ניתן לראות בבירור כי הנתון סובל באופן קבוע מהערכת חסר; לעתים אף נוצר רושם שמספר ההתחלות ירד ובדיעבד מתברר כי הוא עלה (למשל בנתוני המחצית השנייה של 2013 והרבעון הראשון של 2014). הערכת החסר גבוהה יותר כאשר משווים את מספר ההתחלות למספרן ברבעון המקביל בשנה הקודמת, והדבר נובע בעיקר מכך שהנתון של הרבעון המקביל עובר ארבע רוויזיות מפרסומו הראשוני ועד שמשתמשים בו בהשוואה. אמנם גם הנתון לרבעון המקביל ממשך להתעדכן ולהשפיע על שיעור השינוי, אך במידה קטנה יותר.

2 הלמ"ס, מצגת (2016), "שיטת האמידה של התחלות הבנייה של דירות: מקורות המידע ותהליך העבודה".

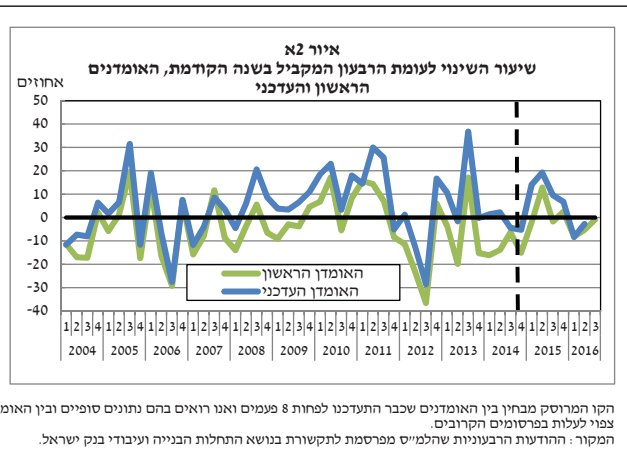
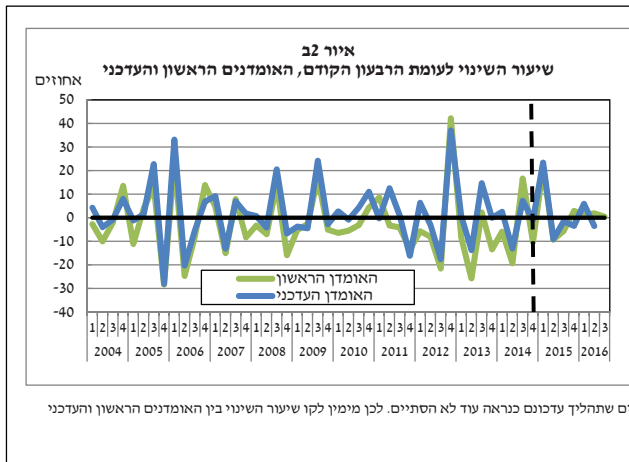
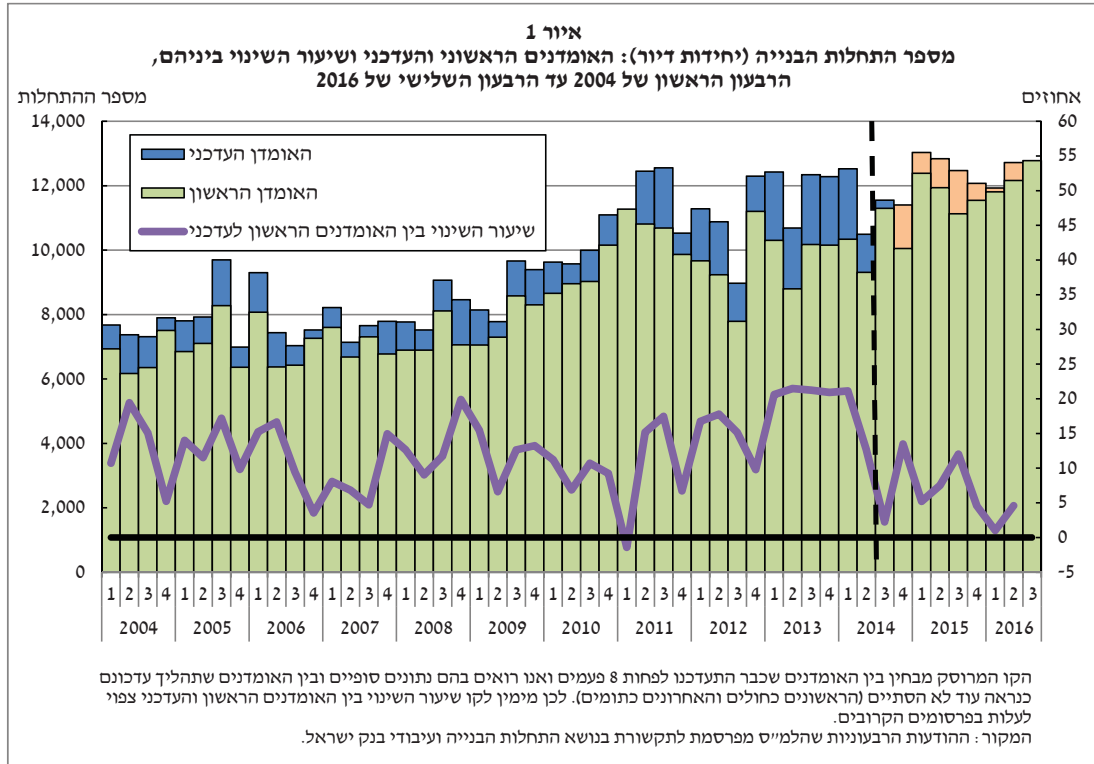
3 התחלות אלו מהוות כ-30% מההתחלות ביוזמה פרטית.

4 מחיקת הכפילויות צפויה לעדכן את הנתון כלפי מטה.

5 הלמ"ס, הודעה לתקשורת (10/02/2015), "התחלת הבנייה וגמר הבנייה בשנת 2014".

6 כך עולה מבדיקה שערך תחום הבינוי בלמ"ס.

7 כלומר הפרסום האחרון שבו הופיע הנתון הרבעוני, או הפרסום בדצמבר 2016 אם מדובר ברבעונים שעדיין מופיעים בלוחות הלמ"ס (2008—2016).



לוח 1 מסכם את סדרי הגודל של העדכונים בנתונים המתקבלים לאורך שנתיים מיום הפרסום של נתון רבעוני חדש⁸. החלק העליון מציג את תקופת החקירה בכללותה, והחלקים שמתחתיו מציגים שתי תת-תקופות: (א) 2004 עד אמצע 2009 – בתקופה זו לא חל שינוי משמעותי בהתחלות הבנייה (ראו איור 1); (ב) אמצע 2009 עד אמצע 2014 – בתקופה זו מסתמנת מגמת עלייה במספר ההתחלות⁹. בחינה ראשונית של ההבדלים בין התקופות מגלה כי מספר ההתחלות הממוצע בתקופה א' נמוך ממספרן הממוצע בתקופה ב'. זאת ועוד, בתקופה א' הפער בין הנתון הראשוני

8 אנו מעוניינים לבדוק את הנתון "הסופי", כלומר לאחר סיום העדכונים. בבדיקה ראשונית זיהינו כי השינויים אחרי שנתיים כבר אינם משמעותיים, ולכן אנו רואים במספר המתפרסם לאחר שמונה רבעונים נתון סופי.

9 קטענו את תקופת החקירה באמצע 2014 מפני שאין עדיין נתונים "סופיים" לתקופה שמתחילה בנקודה זו.

הסקירה הפיסקלית התקופתית ולקט ניתוחים מחקריים

לוח 1						
העדכוניים הממוצעים שנערכו בנתונים במהלך השנתיים שלאחר פרסום האומדן הרבעוני הראשוני, 2004 עד 2014						
מס' הפרסום	מספר התחלות הבנייה ברבעון	שיעור העדכון המצטבר בנתון הרבעוני	שיעור השינוי ביחס לרבעון הקודם	העדכון המצטבר (בנקודות אחוז) של שיעור השינוי ביחס לרבעון המקביל בשנה הקודמת	שיעור השינוי ביחס לרבעון המקביל בשנה הקודמת	העדכון המצטבר (בנקודות אחוז) של שיעור השינוי ביחס לרבעון המקביל בשנה הקודמת
2014Q3-2004						
1	8,371	-2.5	-4.4			
2	8,778	-0.1	4.0	2.4	-0.3	4.0
3	8,951	0.2	5.7	2.6	1.3	5.7
4	9,103	0.6	6.9	3.1	2.5	6.9
5	9,221	1.4	7.7	3.8	3.4	7.7
6	9,266	1.5	7.7	4.0	3.4	7.7
7	9,295	1.6	7.8	4.1	3.5	7.8
8	9,316	1.6	8.0	4.1	3.7	8.0
9	9,341	1.3	8.3	3.8	3.9	8.3
תקופה א': 2009Q2-2004						
1	7,063	-2.2	-5.5			
2	7,301	-1.0	3.0	1.2	-2.5	3.0
3	7,443	-0.5	4.6	1.7	-0.9	4.6
4	7,571	0.2	5.6	2.4	0.1	5.6
5	7,637	0.6	5.9	2.8	0.4	5.9
6	7,673	0.8	5.8	3.0	0.2	5.8
7	7,688	0.6	5.8	2.9	0.3	5.8
8	7,725	0.6	6.1	2.9	0.6	6.1
9	7,762	0.6	6.5	2.9	0.9	6.5
תקופה ב': 2014Q3-2009Q3						
1	9,742	-2.7	-3.1			
2	10,325	0.9	5.1	3.6	1.9	5.1
3	10,532	0.9	6.7	3.6	3.6	6.7
4	10,707	1.1	8.2	3.8	5.1	8.2
5	10,880	2.2	9.6	4.9	6.5	9.6
6	10,935	2.3	9.8	5.0	6.7	9.8
7	10,978	2.6	10.0	5.3	6.8	10.0
8	10,983	2.6	10.0	5.3	6.9	10.0
9	10,994	2.1	10.2	4.8	7.0	10.2

לנתון כעבור שמונה עדכונים עמד במוצע על 700 דירות ברבעון (כ-10%), ואילו בתקופה ב' הוא גדל ל-1,250 דירות ברבעון (כ-13%). אפשר לטעון כי הגידול נובע מכך שהלמ"ס מסתמכת על נתוני שנים קודמות כדי לזקוף נתונים חסרים, ומכך שבתקופה ב' מספר ההתחלות נמצא במגמת עלייה. אולם מכיוון שגם בתקופה א' קיימים עדכונים רטרואקטיביים משמעותיים, ומכיוון שמספר ההתחלות בה גדל לאט במוצע, נראה שמגמת העלייה בתקופה ב' אינה האחראית היחידה להטיה הקיימת כיום.

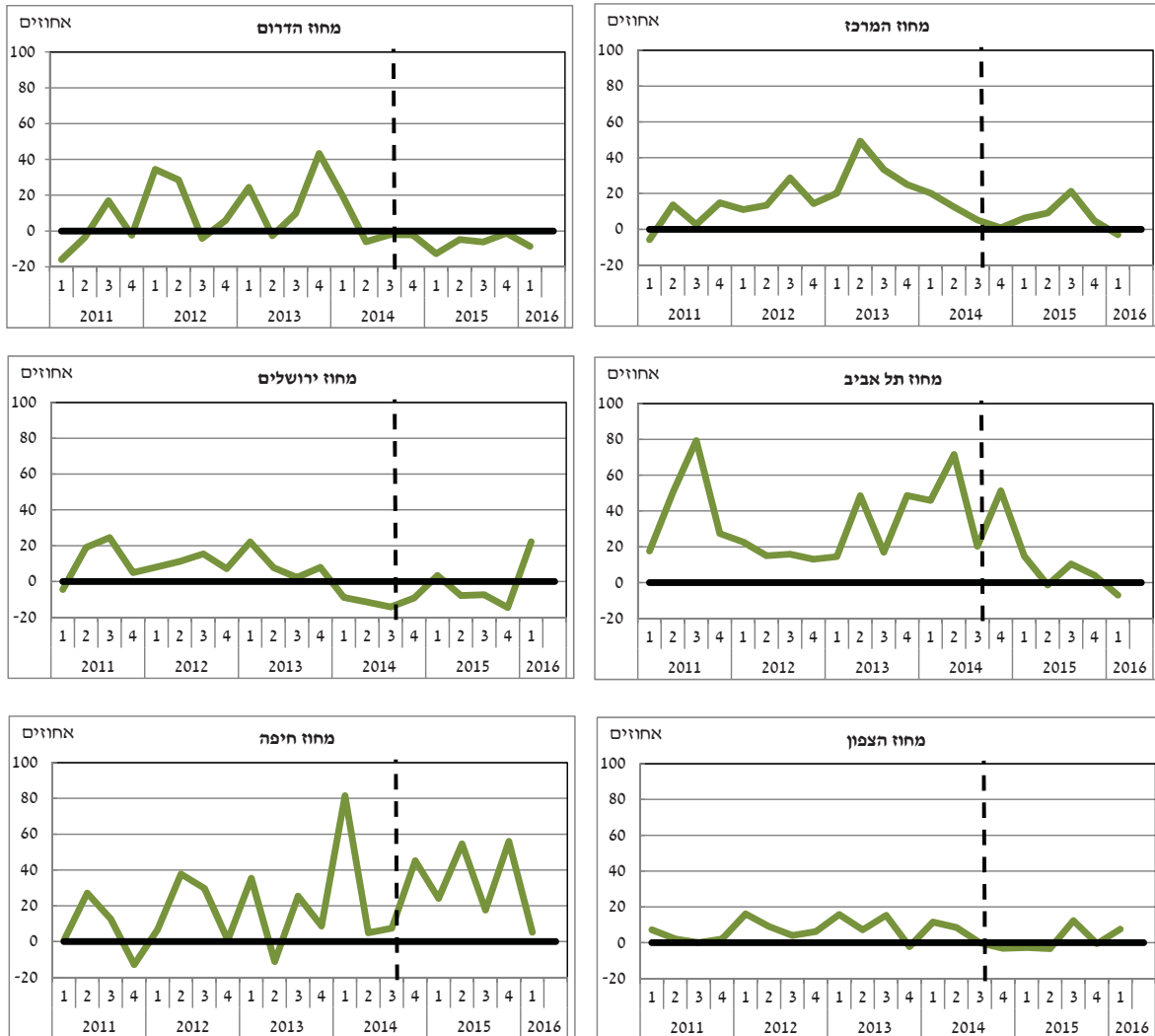
עוד ניתן להבחין כי בשתי תת-התקופות מתפרסמים תחילה נתונים ממוצעים שליליים על שיעור השינוי בהשוואה לרבעון הקודם, ובפרסומים הבאים הם הופכים לחיוביים (עמודה 4). בדיעבד תוקנו הנתונים כלפי מעלה ב-2.9 וב-4.8 נקודות אחוז ברבעונים של תקופות א' ו-ב', בהתאמה. תופעה דומה מתרחשת בנתון על שיעור השינוי ביחס לרבעון המקביל בשנה הקודמת, אך במקרה זה התיקונים משמעותיים יותר ומגיעים, בהתאמה, ל-6.5 ול-10.2 נקודות אחוז. בתקופה הכוללת אנו מבחינים כי לאחר 8 עדכונים שיעור העדכון הממוצע של הנתון הרבעוני עומד על כ-11.4%, והתוספת הממוצעת לשיעור השינוי בין רבעונים מקבילים עומדת על 8.3 נקודות אחוז. לבסוף ניתן לראות כי בשנה הראשונה (פרסומים 2–5) התרומה השולית הממוצעת של העדכונים גדולה מתרומתם בשנה השנייה.

מלבד הנתונים שהתעדכנו לפחות שמונה פעמים בחנו גם את הנתונים הראשוניים שהתעדכנו לפחות ארבע פעמים – כלומר את האומדנים הראשוניים בין הרבעון הרביעי של 2014 לרבעון השלישי של 2015. האומדנים הראשוניים לרבעונים אלו דיווחו על ירידה ממוצעת של 1.4% ביחס לרבעון המקביל בשנה הקודמת, בשעה שארבעת העדכונים מגלים גידול ממוצע של 4% ביחס לרבעונים המקבילים – פער של 5.4 נקודות אחוז. זהו פער משמעותי כמותית, במיוחד כאשר הוא משנה את התפיסה לגבי המגמות בפעילות הענף בשנה האחרונה. האומדנים שהתפרסמו מאז הרבעון הרביעי של 2015 דיווחו במוצע כי נתוני הקצה ירדו ב-2.9% ביחס לרבעונים המקבילים בשנה הקודמת, ועולה השאלה אם ניתן לחזות מראש מה ידווחו הנתונים הסופיים בעוד שנתיים.

אפשר לטעון שהטעות מצטמצמת כאשר משווים תקופות ארוכות מרבעון, למשל המחצית הראשונה של השנה לעומת המחצית הראשונה בשנה הקודמת, תשעת החודשים הראשונים לעומת תשעת החודשים המקבילים, ושנה לעומת שנה. לכן בחנו את שיעור הגידול בין תקופות כאלה. לא זו בלבד שחלוקה כזו אינה מאפשרת לזהות תהליכים בנתוני הקצה, גם בה יש תיקונים כלפי מעלה בדיעבד. הבעיה אינה נפתרת גם כאשר סוכמים את נתוני ארבעת הרבעונים האחרונים, במטרה לבחון את הפעילות במשך שנה שלמה, משום שלרוב הנתונים הטריים יותר מוטים כלפי מטה ומטים את הנתון הכולל. בדיקות אלה מבהירות כי הבעיה אינה טמונה באופן הצגת הנתונים אלא במדידתם.

על מנת לבדוק אם העדכונים משקפים בעיה מקומית במחוז מסוים, בחנו גם את מספר ההתחלות לפי מחוזות (איור 3). ניתן לראות כי מספרי ההתחלות במחוזות תל אביב וחיפה, המהווים יחד כ-27% מסך ההתחלות בתקופה שהאזור מתייחס אליה, נוטים יותר לעדכונים בהשוואה לאומדן הראשון. אולם קיימים עדכונים ניכרים גם במחוזות האחרים.

איור 3
שיעור השינוי באומדן הרבעוני הראשוני של התחלות הבנייה, לפי מחוזות, הרבעון הראשון של 2011 עד הרבעון הראשון של 2016



הקו המרוסק מבחין בין האומדנים שכבר התעדכנו לפחות 8 פעמים ואנו רואים בהם נתונים סופיים ובין האומדנים שתהליך עדכוןם כנראה עוד לא הסתיים. לכן מימין לקו שיעור השינוי בין האומדנים הראשון והעדכני צפוי לעלות בפרסומים הקרובים.
המקור: נתוני הלמ"ס על התחלות הבנייה ועיבודי בנק ישראל.

השפעת הנתון על קובעי המדיניות ועל דעת הקהל

הנתון על מספר התחלות הבנייה משמש גופים ממשלתיים שונים – משרד השיכון והבינוי, משרד האוצר, קבינט הדיור ומנהל התכנון – כדי להעריך את מספר הדירות החדשות שיתווספו למשק בשנים הקרובות. באמצעות הערכות אלו הם אומדים את ההיצע הצפוי בשוק הדיור ומגבשים מדיניות לגבי מידת הצורך בהתערבות ממשלתית בענף הבנייה. הנתון הנידון משמש גם את בנק ישראל, משום שהבנק מבסס את המדיניות המוניטרית שלו על פעילות המשק. אם קיימת בנתון אמידת חסר שיטתית, הדבר עלול ליצור רושם שהפעילות במגזר הבנייה מאטה בשעה שלמעשה מספר הדירות החדשות נמצא במגמת עלייה מתונה לאורך זמן.

העדכונים בנתונים על התחלות הבנייה משפיעים גם על אומדני החשבונאות הלאומית, משום שהנתונים הרבעוניים על שטח התחלות הבנייה והשלמותיהן (גמר הבנייה) משמשים לחישוב אומדני ההשקעה בנכסים קבועים. גם בסעיף זה קיימים עדכונים משמעותיים, ובעבר נמצא כי סדרי הגודל הממוצעים שלהם גבוהים בכ-3.5 נקודות

אחוז מהנתון הראשוני על שיעור הגידול הרבעוני¹⁰. בדיקה שערכנו לאחרונה הראתה כי הערכת החסר הממוצעת בין 2009 ל-2016 מגיעה ל-2.75 נקודות אחוז. המידע המצוי ברשותנו אינו מאפשר לכמת כיצד הנתון על התחלות הבנייה משפיע על אומדן ההשקעה בנכסים קבועים, אך מכיוון שההשקעה בבנייה למגורים מהווה כ-30% מרכיב זה וכ-6% מהתוצר, ניתן לשער כי העדכונים בהתחלות הבנייה משפיעים על נתון ההשקעה בבנייה למגורים ישירות, וכי הם משפיעים במידה מסוימת גם על הרוויזיות הנצפות באומדני התוצר.

הנתון המוטא על מספר ההתחלות מחלחל גם לציבור הרחב, דרך אמצעי התקשורת. לשם המחשה, ב-31 במאי 2009 פרסם העיתון גלובס את הכותרת: "לידיעת רוכשי הדירות: מספר התחלות הבנייה ירד ב-9% ברבעון הראשון". אולם העדכון האחרון של הנתון לרבעון זה (הרבעון הראשון של 2009) מעיד כי מספר ההתחלות עלה ב-3.8%. דוגמה עדכנית יותר נוגעת לנתון שהלמ"ס פרסמה בסוף 2015 וגרר כותרות כגון: "הפער ממשיך לגדול, רק 47 אלף דירות חדשות החלו להיבנות בשנת 2015 – זה לא מספיק לזוגות הצעירים" (גלובס, 28/03/2016). כיום, לאחר שלושה עדכונים בלבד, עלה הנתון מ-47.7 אלף התחלות ל-52.7 אלף – גידול של 10.5%. הדוגמה האחרונה שנוכיר קשורה לנתון שהתפרסם באמצע דצמבר 2016 לגבי ההתחלות ברבעון השלישי של אותה שנה – כ-38.7 אלף בשלושת הרבעונים הראשונים של 2016. במונחים שנתיים מדובר בכ-51.6 אלף התחלות. נתון זה שימש בסיס לכותרת: "נתוני הלמ"ס חושפים: היעד הממשלתי של 60 אלף התחלות בנייה רק מתרחק" (כלכליסט, 15/12/2016). אולם הבדיקה הנאיבית של ממוצע העדכונים עד כה (לוח 1) משמשת יסוד לצפות שהנתון במונחים שנתיים יתעדכן ל-57.2 אלף התחלות – תוצאה קרובה למדי ליעד הממשלתי. אמנם מדובר בתחזית גסה, אך כפי שנראה בהמשך, היא תואמת את העדכונים שהתבררו בדיעבד בשנים האחרונות.

אומדן החסר של התחלות הבנייה טומן בחובו שתי סכנות. ראשית, הנתון עלול להפוך לבלתי אמין וקובעי המדיניות יפסיקו להשתמש בו¹¹. הסכנה השנייה קשורה ביצירת ציפיות בקרב הציבור: אם מדי רבעון מדווחים שהתחלות הבנייה ירדו משמעותית, הדבר יוצר בציבור ציפיות שההיצע יצטמצם בשנים הקרובות והמחירים יעלו.

האם ניתן לחזות את העדכונים מראש?

א. חיזוי עדכונים בנתונים מקרו-כלכליים

מבחן Mincer-Zarnowitz (1969) ליעילות החיזוי בודק אם ניתן לחזות את השינויים הרטרואקטיביים בנתון על סמך המידע הקיים כשהוא מתפרסם לראשונה. השימוש שלנו במבחן זה מבוסס על ההנחה שהפרסומים הראשוניים של הנתונים על התחלות הבנייה מכילים טעות ביחס לנתון הסופי:

$$(1) \quad S_t^{t+1} = S_t^f + \varepsilon_t$$

כאשר S מייצג את מספר התחלות הבנייה; S_t^{t+1} מייצג את מספר ההתחלות ברבעון t שהיה ידוע בזמן הפרסום הראשון – בזמן $t+1$; S_t^f מייצג את הנתון הסופי (עם סיום העדכונים) על מספר ההתחלות ברבעון t ; ו- ε_t מייצג טעות מקרית בעלת תוחלת אפס, כלומר טעות שאי-אפשר לחזות בזמן הפרסום. בהקשר זה Mankiew and Shapiro (1986) מבחינים בין שני גורמים שמשפיעים על העדכון הרטרואקטיבי של נתוני המקרו – "רעש" (noise) מקרי ו"חדשות" (news). גישת ה"רעש" גורסת כי באמידת האומדן הראשון יש רעש שאינו מתואם עם הנתון הסופי,

10 בנק ישראל (2013), "הרוויזיות בנתונים הרבעוניים של החשבונאות הלאומית", ההתפתחויות הכלכליות בחודשים האחרונים, מס' 135.

11 ראו "תגובת משרד הבינוי: קצב התחלות הבנייה 'האמתי' – 47 אלף יח"ד" nrg מעריב, (10/03/2015).

והוא מצטמצם במהלך העדכון. גישת ה"חדשות" גורסת כי הנתון הראשוני משקף תחזית רציונלית לגבי הנתון הסופי בהינתן המידע הקיים בעת הפרסום, וכי העדכונים נובעים ממידע חדש, מתואם עם הנתון הסופי, שאי-אפשר לצפות מראש. נראה כי הלמ"ס מאמצת את גישת ה"חדשות" שכן היא מודעת לעדכונים הצפויים בנתונים ומתריעה על כך בהודעותיה לעיתונות – כלומר נוצר רושם שאי-אפשר לחזות מראש את העדכונים עקב מחסור במידע.

Aruoba (2006) מדגיש כי כאשר תוחלת הרוויזיות שונה מאפס, הרגרסיה הבסיסית של Mincer-Zarnowitz עשויה להשמיט משתנים מסבירים נוספים. לכן הוא מציע לבחון אם בעת הפרסום הראשוני ידוע על עוד משתנים שיכולים לסייע בחיזוי העדכונים – לרבות רוויזיות של נתוני רבעונים קודמים ושל נתוני מקרו אחרים. כפי שראינו, תוחלת הרוויזיות בנתונים על התחלות הבנייה שונה מאפס. לכן נתייחס להלן להשערת ה"חדשות" (ההשערה שהעדכונים הם נתונים מקריים בעלי תוחלת אפס), ונבחן אם היא ההשערה הטובה ביותר במקרה של הנתונים על התחלות הבנייה. בהתאם לגישתו של Aruoba גם אנו נשתמש במשתנים מסבירים נוספים במטרה לבדוק אם הם יכולים להוסיף למודל מידע בזמן אמת. לצורך זה נשתמש רק במשתנים המסבירים העומדים לרשותנו בזמן פרסום הנתון הראשוני. משתנים שמתגלים לנו רק בפרסומים הבאים אינם יכולים לשמש אותנו בניבוי העדכונים העתידיים משום שהם אינם ידועים בעת עריכת האומדן.

בהתאם לגישה זו נבחין בין ארבע גישות לתחזית העדכונים בנתון על התחלות הבנייה:

M0 – השערת ה"חדשות" האפס גורסת כי אין מודל וכי העדכונים מקריים ובעלי תוחלת אפס, כלומר הנתון הראשוני הוא הנתון האפקטיבי ביותר בחיזוי הנתון הסופי.

מודלים M1–M3 מפורטים להלן:

$$(M1) \quad R_t^{t+1+n} = \alpha_0 + \varepsilon_t$$

$$(M2) \quad R_t^{t+1+n} = \alpha_0 + \alpha_1 S_t^{t+1} + \varepsilon_t$$

$$(M3) \quad R_t^{t+1+n} = \alpha_0 + \alpha_1 S_t^{t+1} + \sum_{i=1}^3 \beta_i R_{t-i}^{t+1} + t + \varepsilon_t$$

R_t^{t+1+n} מייצג את העדכון המצטבר של הנתון לרבעון t לאחר n עדכונים, כלומר את ההפרש בין הנתון שמתפרסם לרבעון t אחרי n רבעונים (S_t^{t+1+n}) לבין הנתון שמתפרסם עבורו לראשונה (S_t^{t+1}); מייצג את שיעור השינוי בין התחלות הבנייה ברבעון t להתחלות ברבעון המקביל בשנה הקודמת ($t-4$); R_{t-i}^{t+1} מייצג את העדכון המצטבר של הנתון לרבעון $t-i$ שידוע בעת הפרסום הראשוני של הנתון לרבעון t ; t מייצג מגמת זמן לינארית שגדלה ביחידה אחת בכל רבעון מאז תחילת תקופת החקירה.

מודל M1 גורס כי גודל הרוויזיות אינו תלוי במשתנים נוספים וכדי לחזות את הנתון הסופי ניתן להוסיף לנתון הראשוני שיעור קבוע – את ממוצע הרוויזיות הבלתי מותנה. כאשר עושים שימוש נאיבי בנתונים מלוח 1 משתמשים למעשה במודל M1, שכן מניחים כי ביכולתנו לתקן את הנתונים על סמך ממוצע הרוויזיות הידוע כיום. מודל M2 בוחן אם נוסף לממוצע הבלתי מותנה, גם גודל הנתון הראשוני משפיע על העדכונים הצפויים בו; כלומר אם אומדן ראשוני קטן במיוחד מתעדכן במידה רבה יותר או להיפך. מודל M3 בוחן אם נוסף לכל המשתנים הנזכרים לעיל, גם לזמן ולעדכונים שפורסמו לרבעונים קודמים יש יכולת להסביר את העדכונים הצפויים בנתון של רבעון t .

חשוב להדגיש כי אנו עורכים תרגיל פשוט שאינו מסביר את הגורמים האמתיים לרוויזיות, משום שהוא משתמש רק במשתנה מסביר אחד, היינו מספר התחלות הבנייה. במילים אחרות, מודלים M3—M1 משתמשים אך ורק בידע שנצבר עד כה על הרוויזיות כדי להצביע על תופעה מתמשכת שצפויה להימשך בעתיד. מודל של ממש דורש מערך נתונים רחב יותר, לרבות נתונים על הגורמים האמתיים לגידול או לירידה בהתחלות הבנייה. גם מידע על היקף הזקיפה של נתונים, ועל התיקון הסטטיסטי שכבר נערך בנתונים הגולמיים, יכול לסייע בתיקון הנדרש ובצמצום גודל הטעות. אולם דרך המלך לפתרון הבעיה אינה עוברת באומדנים סטטיסטיים אלא בשיפור האיסוף. לשם המחשה, הלמ"ס צופה כי שימוש בנתונים המגיעים דרך מערכת "רישוי זמין" יאפשר לזרז ולייעל את האופן שבו המידע על היתרי הבנייה עובר אליה מוועדות התכנון המקומיות, והדבר צפוי לסייע בשיפור הסקירה. אם כל הגורמים העוסקים במלאכה ינקטו את הצעדים הדרושים לשיפור העברת המידע בין הגופים השונים, הם יוכלו להשפיע משמעותית על סדרי הגודל של העדכונים ולהביא לצרכני המידע נתונים טובים יותר בזמן אמת.

ב. בדיקה בדיעבד (Ex Post) של השערת ה"חדשות"

על מנת לבחון אם ניתן להציע מודל שחווה את העדכונים ביעילות רבה יותר, נשווה את הדיוק של מודלים M3—M1 לזה של השערת החדשות (M0): נבדוק את ה-Adjusted R² וה-Residual Standard Error (RSE) של כל מודל ביחס לאלה של השערת האפס.

לוח 2 מציג את תוצאות האמידה של מודלים M3—M1 לגבי העדכונים בשיעורי השינוי בין רבעונים מקבילים, מכיוון ששיעור שינוי זה סובל מפערים גדולים יותר ומכיוון שהוא מנחה חלק מהגורמים העונתיים המשפיעים על מספר התחלות הבנייה. הנתונים לקוחים מהלוחות שהלמ"ס פרסמה בהודעותיה לתקשורת מדי רבעון בין 2004 ל-2016 (נתונים מקוריים).

התוצאות בלוח 2 מעידות כי האומדן הראשון לשיעור השינוי בהתחלות הבנייה מוטה כלפי מטה באופן שיטתי, ואומדן ההטיה עומד על כ-8.6%. הטיה זו מתגלה עם התייצבות הנתון כעבור שמונה עדכונים, והיא מייצגת את ממוצע הרוויזיות הבלתי מותנה. תוצאה זו פירושה שמודל האפס נדחה, כלומר העדכונים אינם נובעים רק מחדשות שאי-אפשר לחזות. זאת ועוד, המקדם של האומדן הראשון אינו שונה במובהק מאפס, כלומר לגודלו של הנתון הראשוני אין השפעה על גודל העדכונים הצפויים. כאשר משווים את הטעויות במודלים M3—M1 לטעות במודל האפס, מוצאים כי הן קטנות יותר. דברים אלה תקפים במיוחד לגבי מודל M3: באופן ממוצע הוא מנבא את הנתון הקרוב ביותר לנתון בפועל ומצמצם את הטעות ב-54%. נוסף לכך, העלייה ב-Adjusted R² מעידה כי המשתנים שהוספו למודל משפרים את כושר החיזוי של המודל באמצעות הגדלת שיעור השונות המוסברת.

ערכנו לתוצאות האמידה כמה בדיקות רגישות. ראשית, בדקנו אם ההבדל בין התחזית לגבי עדכון אחד לבין התחזית לגבי ארבעה ושמונה עדכונים אינו נובע מהרבעונים ה"טריים" יותר שעדיין לא עודכנו שמונה פעמים (החל מהרבעון הרביעי של 2015). לשם כך נטלנו את אותן 40 תצפיות שכבר התעדכנו מספר פעמים גדול דיו והשתמשנו בהן גם באמידות של העדכון הראשון והעדכון הרביעי. מצאנו כי התוצאות אינן משתנות בצורה משמעותית. שנית, הוספנו למשתנים המסבירים את הרוויזיות המצטברות של רבעונים מוקדמים יותר, ומצאנו כי לרוב אין להן ערך מוסף בחיזוי העדכונים. לבסוף, הוספנו משתנה דמי לרבעון כדי לבחון אם יש לרוויזיות מאפיינים עונתיים, ומצאנו כי לסוג הרבעון אין השפעה על גודל הרוויזיה.

ג. בדיקה בזמן אמת (Ex Ante) של מודל החיזוי

בחלק ב' השתמשנו במידע שיש לנו כיום כדי לאמוד את ההטיות בדיעבד, אך מעניין לבדוק גם אם המידע שעמד לרשותנו בעת הפרסום הראשוני מספיק על מנת לחזות את העדכונים. לשם כך הרצנו את המודלים השונים על תת-קבוצות של מסד הנתונים באופן שמייצג את המידע שעמד לרשותנו בזמן אמת. גישה זו מאפשרת לדמות את המידע

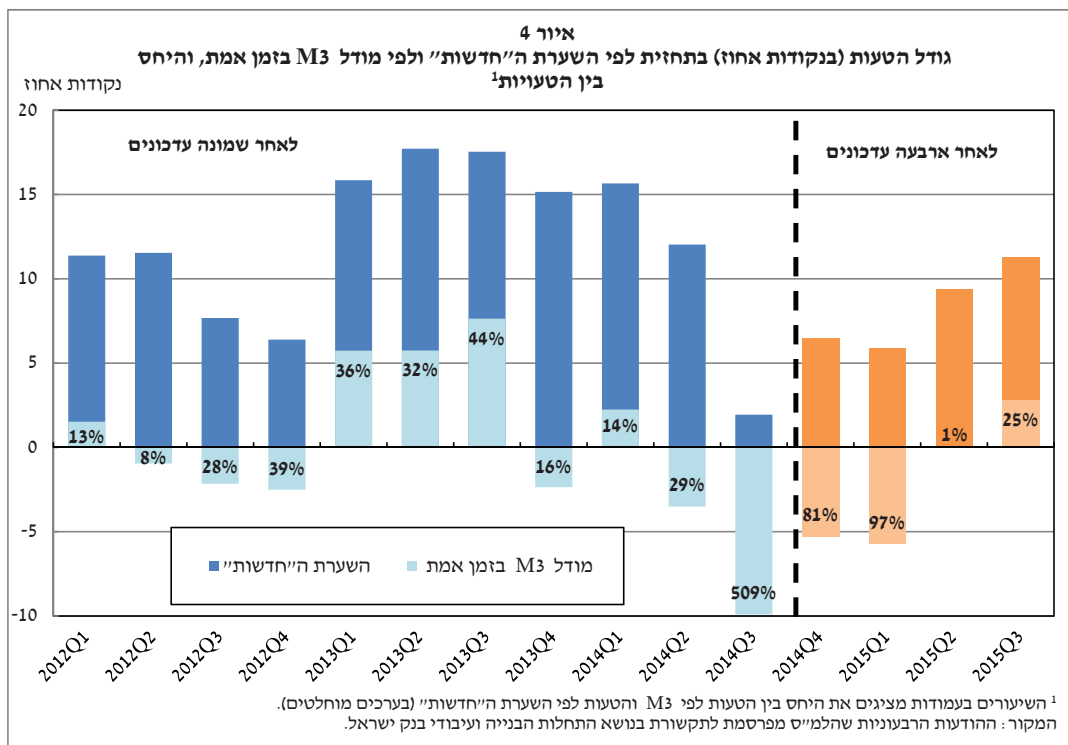
לוח 2									
העדכונים המצטברים (בנקודות אחוז) בנתון על שיעור השינוי בהתחלות הבנייה יחסית לרבעון המקביל בשנה הקודמת, הרבעון הראשון של 2004 עד הרבעון השלישי של 2016									
M3	M2	M1	M3	M2	M1	M3	M2	M1	
אחרי שמונה רויזיות			אחרי ארבע רויזיות			אחרי רויזיה אחת			
3.02*** (1.06)	8.40*** (1.12)	8.56*** (1.13)	3.26** (1.27)	8.01*** (0.90)	8.10*** (0.94)	2.16*** (0.79)	3.92*** (0.70)	3.81*** (0.69)	חותך
0.01 (0.06)	-0.05 (0.04)		0.03 (0.05)	-0.03 (0.04)		0.07 (0.05)	0.03 (0.03)		האומדן הראשון
0.41*** (0.15)			0.41** (0.16)			0.39*** (0.12)			העדכון הראשון של רבעון t-1
-0.23* (0.17)			-0.25 (0.15)			-0.19* (0.1)			העדכון המצטבר של רבעון t-2
0.27* (0.14)			0.25** (0.12)			0.23*** (0.09)			העדכון המצטבר של רבעון t-3
0.14* (0.08)			0.12* (0.07)			-0.01* (0.04)			משתנה דמי למגמת הזמן
40	40	40	44	44	44	47	47	47	מספר התצפיות
0.33	0.01	0	0.27	0.005	0	0.19	0.01	0	R ²
0.23	-0.01	0	0.18	-0.02	0	0.09	-0.01	0	Adjusted R ²
4.62 (df = 4)	5.30 (df = 8)	5.27 (df = 9)	4.56 (df = 8)	5.08 (df = 2)	5.03 (df = 43)	3.64 (df = 41)	3.84 (df = 45)	3.82 (df = 46)	Residual Std. Error
3.32** df = 5;) (34)	0.56 df = 1;) (38)		2.86** df = 5;) (38)	0.20 df = 1;) (42)		1.95 df = 5;) (41)	0.52 df = 1;) (45)		F סטטיסטי
46%	53%	53%	48%	54%	53%	68%	72%	71%	RSEi/RSE0
*p<0.1, **p<0.05, ***p<0.01 הערך בסוגריים מציין את סטיית התקן המתוקנת להטרוסקדסטיות ולמתאם סדרתי לפי Newey-West.									

העומד לרשותנו ברבעון מסוים ולהשתמש בו כדי לחזות את העדכון המצטבר הצפוי בעוד שנתיים. עם ההתקדמות בזמן אנו מתקנים את המודל אך ורק באמצעות המידע החדש שנעשה זמין מדי רבעון (מידע זה כולל גם נתונים על רבעונים חדשים וגם עדכוני נתונים על רבעונים קודמים). איור 4 משווה בין הטעות שנגרמת כאשר משתמשים בהשערת האפס ובין הטעות שנגרמת כאשר משתמשים במודל M3.

עד הרבעון השלישי של 2014 הטעות של מודל M3 שווה להפרש בין גודל העדכון כעבור שמונה רבעונים וגודל העדכון שהמודל חוזה. הטעות של השערת ה"חדשות" שווה לגודל העדכון כעבור שמונה רבעונים, שכן כאשר משתמשים

בנתון של הפרסום המקורי כמות שהוא, מניחים כי העדכונים הצפויים אפסיים בממוצע. מכיוון שהחל מהרבעון הרביעי של 2014 לא ניתן לדעת מהו העדכון בפועל כעבור שנתיים, אנו בוחנים את ההפרש בין הצפי של המודל לגבי העדכון כעבור ארבעה רבעונים לבין העדכון בפועל כעבור ארבעה רבעונים. כפי שממחיש האיור, גם במודל הפשטני שלנו קיימות טעויות, אך הן אינן מוטות לכיוון מסוים (כאשר הטעות שלילית, פירוש הדבר שהמודל הניב הערכת יתר של הנתון). נוסף לכך, הטעויות של מודל M3 קטנות משמעותית מהטעויות של M0: המודל שלנו מתקן בממוצע כ-34% מגודל הטעות הנובעת מהשערת ה"חדשות", והוא מתקן כ-67% ממנה אם משמיטים את התצפית החריגה ברבעון השלישי של 2014.

בדיקת רגישות שערכנו שמרה על מספר תצפיות קבוע: בכל פעם שנוספה תצפית למדגם תצפית אחרת יצאה ממנו. לא נמצאו הבדלים משמעותיים בתחזיות או ביכולת הניבוי בשיטה זו. אולם השיטה הראשונה בכל זאת מתאימה יותר משום שהיא מדמה כיצד המידע הקיים ברשותנו מתרחב בחלוף הזמן.



סיכום

הנתונים על התחלות הבנייה מהווים מרכיב חשוב בתמונה הכלכלית של ישראל, במיוחד בתקופה שבה תחום הדיור ניצב במוקד המדיניות הכלכלית. אם קיימת בנתונים הערכת חסר בזמן אמת, היא משפיעה על קובעי המדיניות ועל הציבור, ודבר זה מצדו יכול להשפיע על שוק הדיור.

כאשר בוחנים את העדכונים ההולכים ומצטברים בסדרת התחלות הבנייה בין 2004 ל-2015, מוצאים כי בנתונים הראשוניים קיימת הערכת חסר שיטתית. עבודה זו מצאה כי ניתן לתקן את האומדנים הראשוניים באמצעות מודלים שבוחנים את העדכונים בנתונים של רבעונים קודמים. תוצאות אלו מזמינות בדיקה מעמיקה יותר של האומדנים הקיימים ובחינה של מודלים סטטיסטיים נוספים לתיקון ההטיה. עד אז, וכפי שמציינת הלמ"ס בפרסומיה על התחלות הבנייה, מומלץ לצרכני הנתונים לנקוט זהירות כשהם מסתמכים על הפרסומים הראשוניים של הנתון הנידון.

ביבליוגרפיה

- בנק ישראל (2013), "הרוויזיות בנתונים הרבעוניים של החשבונאות הלאומית", *ההתפתחויות הכלכליות בחודשים האחרונים* 135.
- גלובס (31/05/2009), "לידיעת רוכשי הדירות: מספר התחלות הבנייה ירד ב-9% ברבעון הראשון".
- גלובס (28/03/2016), "הפער ממשיך לגדול, רק 47 אלף דירות חדשות החלו להיבנות בשנת 2015 – זה לא מספיק לזוגות הצעירים".
- הלשכה המרכזית לסטטיסטיקה, הודעה לתקשורת (10/02/2016): "התחלת הבנייה וגמר הבנייה בשנת 2014".
- כלכליסט (15/12/2016), "נתוני הלמ"ס חושפים: היעד הממשלתי של 60 אלף התחלות בנייה רק מתרחק".
- nrg מעריב (10/03/2015), "תגובת משרד הבינוי: קצב התחלות הבנייה 'האמתי' – 47 אלף יח"ד".
- Aruoba, S.B. (2006): Data Revisions are not Well-Behaved, *Journal of Money, Credit and Banking*, 2008, 40(2-3), 319-340.
- Mankiew, N.G. and M.D. Shapiro (1986): News or Noise: An Analysis of GNP Revisions, *Survey of Current Business*, May 1986, pp.20-25.
- Mincer, J. and V. Zarnowitz (1969): The Evaluation of Economic Forecasts in J. Mincer (ed.), *Economic Forecasts and Expectations*, NBER, New York.