



**אמידה של פעילות כלכלית ריאלית
וחיזוי טווח קצר של צמיחת התוצר בישראל¹**

טים גינקר* וטניה סוחוי**

סדרת מאמרים לדיון 2022.07
פברואר 2022

בנק ישראל – <http://www.boi.org.il>

¹ אנו אסירי תודה לסרקן ארסלן על התדיינותו על המאמר ועצותיו. אנו מודים למשתתפי הסמינר בבנק ישראל, אוניברסיטת בר אילן, הכנס הבינלאומי למודלים כלכליים ומדעי הנתונים (EcoMod2021), וועידת הסטטיסטיקה העולמית של ISI 2021. אנו מודים לאלפונסו רוסאליה על הערותיו המועילות בפאנל ISI לגבי פיתוחים חדשים בסטטיסטיקה של בנק מרכזי. אנו מודים לאריאל מנצורה ולאיל ארגוב על ההצעות והדיונים רבי הערך. כמו כן, אנו מודים ליהודה פורת על הייעוץ בסוגיות הנוגעות לכלכלה הישראלית.

* טים גינקר – בנק ישראל, חטיבת המידע והסטטיסטיקה; אוניברסיטת בר אילן, המחלקה לכלכלה.

דוא"ל: ginker.tim@boi.org.il

** טניה סוחוי – חטיבת המחקר, בנק ישראל. דוא"ל: tanya.suhov@boi.org.il

הדעות המובעות במאמר זה אינן משקפות בהכרח את עמדתו של בנק ישראל

חטיבת המחקר, בנק ישראל, ת"ד 780 ירושלים 91007

Research Department, Bank of Israel, POB 780, 91007 Jerusalem, Israel

אמידה של פעילות כלכלית ריאלית וחיזוי טווח קצר של צמיחת התוצר בישראל

טים גינקר וטניה סוחוי

תקציר

מאמר זה מציג הרחבה של מודל פקטורי דינמי המשלב הורדת ממד, שפותח על ידי Bräuning and Koopman (2014), לצורך חיזוי טווח קצר של צמיחת התוצר בישראל. בשיטה המוצעת, הורדת הממד נעשית בעזרת Partial Least Squares שמשתמשת בפרוקסי לצמיחת התוצר החודשית הבלתי נצפית כמשתנה מוסבר. יישום אמפירי על נתוני ישראל מראה שהשיטה מביאה לשיפור מובהק בדיוק החיזוי של הצמיחה הרבעונית. בנוסף לכך, נבחנו היבטים שונים של שילוב נתוני סקרי המגמות בעסקים במודל. נמצא כי הוספת נתוני הסקר משפרת את דיוק החיזוי, אך רק כאשר אלה משמשים לצורך ההשלמה בקצה של האינדיקטורים המסורתיים. בנוסף לחיזוי הצמיחה הרבעונית, המודל מאפשר יצירת מדד חודשי לפעילות ריאלית שמוצג במונחי צמיחת התוצר החודשית הבלתי נצפית. המדד עקבי עם התחזית לצמיחה הרבעונית במובן שהתחזית הרבעונית לצמיחת התוצר מתקבלת מהסיכום של המדדים לצמיחה החודשית. בהשוואה למדד המשולב למצב המשק שמופק ע"י בנק ישראל, המדד המוצע מבוסס על פאנל נתונים רחב יותר ועל כן עשוי לשקף במידה מדויקת יותר את הדינמיקה של הפעילות הריאלית בתדירות החודשית.

מילות מפתח: חיזוי טווח קצר; מודל פקטורי דינמי; Partial Least Squares.

Nowcasting and monitoring real economic activity in Israel

Tim Ginker and Tanya Suhoy

Abstract

This paper employs an extension of the mixed-frequency collapsed dynamic factor model introduced by Bräuning and Koopman (2014) for nowcasting Israeli GDP growth. In the proposed method, the dimension reduction is performed through a partial least squares regression that uses a proxy for the unobserved monthly GDP growth as the dependent variable. Our results show that this method allows us to improve on the accuracy of quarterly forecasts. In addition, we study the role of business surveys in real-time nowcasting. We find that including business surveys improves the accuracy of nowcasting, but only when they are used for endpoint imputation of the traditional macroeconomic indicators. In contrast, expanding the monthly panel with surveys related to some of the already included variables leads to inferior forecasting performance. Finally, the proposed model allows us to construct a monthly index of real economic activity that is consistent with the nowcast. Compared to the Composite State of the Economy Index currently published by the Bank of Israel, it utilizes a much broader data set and thus is likely to provide a more timely and precise picture of the course of economic activity.

Keywords: Nowcasting; Dynamic factor model; Partial least squares.

1. הקדמה

עבודה זו אנו מציגים מודל חיזוי טווח קצר של צמיחת התוצר בישראל. לחיזוי טווח קצר של התוצר המקומי הגולמי בזמן אמת קיים תפקיד חשוב עבור קובעי המדיניות. הערכות עדכניות ומדויקות של הפעילות הריאלית יכולות לתמוך בקבלת ההחלטות של קובעי המדיניות, והן קריטיות בתקופות של שינויים מהירים במצב המשק ומשברים, כמו משבר הקורונה של 2020.

מאז מחקריהם החלוציים של Stock and Watson (1989, 1991, 1993), שהראו כי ניתן לסכם את התנועה המשותפת של אינדיקטורים מקרו-כלכליים רבים במספר קטן של גורמים חבויים, היכולים לשמש למעקב אחר מסלולה של הפעילות במשק, מודל פקטורי דינמי (Dynamic Factor Model) הפך לכלי פופולרי בקרב בנקים מרכזיים ומוסדות כלכליים אחרים לצורך ניתוח בסיסי נתונים מקרו-כלכליים גדולים (Banbura et al., 2013). בנוסף, מודל פקטורי דינמי יכול לפתור מספר סוגיות פרקטיות חשובות באופן פשוט ואלגנטי. ובפרט, הוא מאפשר לשלב סדרות חודשיות ורבעוניות באותה מסגרת ולטפל בסוגים שונים של תצפיות חסרות שיכולות להיווצר עקב משכי זמן היסטוריים שונים או פרסומי נתונים שאינם מסונכרנים.

למרות קיומן של רוטינות אמידה מתאימות (Banbura and Modugno, 2014), הרי שמידול ישיר משותף של צמיחת תוצר רבעוני יחד עם פאנל גדול של אינדיקטורים חודשיים עדיין מלווה באמידה של מספר רב של היפר-פרמטרים, ועל כן כרוך בשונות חיזוי גבוהה יותר. כדי לטפל בבעיה זו, Bräuning and Koopman (2014) הציגו מודל פקטורי דינמי מכווץ (Collapsed Dynamic Factor Model; מכאן ואילך ייקרא CDFM), המאפשר לשלב מספר גדול של סדרות חודשיות במודל מצומצם יחסית. זהו הליך דו-שלבי, שבו, בשלב הראשון, המידע הכלול במשתנים החודשיים מסוכם במספר קטן של פקטורים על ידי ניתוח רכיבים עיקריים (Principal Component Analysis; מכאן יכונה PCA). בשלב השני, הפקטורים מוכנסים למודל במשותף יחד עם צמיחת התוצר במודל פקטורי דינמי. בנוסף למידול הדינמיקה של הגורמים שחולצו, השלב השני הופך אותם לקשורים יותר למשתנה המטרה.

חיסרון ידוע אחד של PCA יכול להיווצר במקרים שבהם שונות דומיננטית במשתנים החודשיים אינה קשורה באופן משמעותי לשונות במשתנה המטרה, ודבר זה עלול לגרום למתאם חלש עם הפקטורים שחולצו בשלב הראשון. לפיכך, בשעה שהצעד השני של ה-CDFM מאפשר הפחתה נוספת של השונות ברכיבים העיקריים, ביחס לגורמים המשותפים הקשורים באופן קרוב יותר למשתנה המטרה, ייתכן כי מרבית המידע הרלוונטי כבר סוּן החוצה במהלך השלב הראשון.

בעבודה זו אנו משתמשים בהרחבה של CDFM במטרה להגדיל את כמות המידע הרלוונטי המחוץ בשלב הראשון. הגורמים הנגזרים מ-PCA הם למעשה ממוצעים משוקללים של הסדרות החודשיות המתוקננות. הרעיון הוא להתאים את צורת השקלול באופן כזה שהיא תייחס משקל רב יותר לאותם משתנים הקשורים באופן חזק יותר לצמיחה במשק. הדרך להשיג זאת היא על ידי מציאת פרוקסי מתאים עבור צמיחת התוצר החודשית הבלתי נצפית, ולאחר מכן להשתמש בשיטת Partial Least Squares (מכאן יכונה PLS) להורדת ממד, במקום שימוש ברכיבים העיקריים. לפיכך, בשלב הראשון, אנו משתמשים

בסדרה של שינויים חודשיים במדד סך הפדיון בענפי המשק כפרוקסי לתוצר, ואומדים את הפקטורים בעזרת PLS. את הבחירה בסך הפדיון כפרוקסי ניתן להסביר באופן הבא. על פי גישת הערך המוסף, הפרש ענפי בין סך הפדיון לתמ"ג (להלן התוצר) יגדל יחד עם שיעור צריכת הביניים בתוצר. התוצר בישראל בנוי באופן כזה שכ-60% ממנו מיוחס לענף השירותים העסקיים – המתאפיין בשיעור קטן של צריכת ביניים. ההיגיון הזה שולל את השימוש במספר פרוקסים פוטנציאליים אחרים, כמו מדד הייצור התעשייתי, שנחשב באופן מסורתי לאינדיקטור פרו-מחזורי. מנקודת מבט אמפירית, בתדירות רבעונית, עבור המדגם מ-2000:Q1 עד 2020:Q4, להפרכי הלוג של מדד סך הפדיון קיים מתאם גבוה של 0.67 עם צמיחת התוצר, בהשוואה ל-0.17 בלבד עבור מדד הייצור התעשייתי.

הבדיקה שנעשתה מחוץ למדגם מראה שהמודל מסוגל גם לתפוס היטב את תקופת "טרור הקורונה", שבה היו תנודות מתונות יחסית בצמיחת התוצר, ולאחר מכן גם את התנודות הקיצוניות שהתרחשו במהלך המשבר של 2020. בנוסף לחיזוי הטווח הקצר הרבעוני, אנו משתמשים באותו מודל כדי להפיק אומדנים של צמיחת התוצר החודשית הבלתי נצפית, שיכולה לשמש אינדיקטור מוביל למצב המשק (Brave et al., 2019). האינדיקטור המוצע נמצא תואם לדפוס החודשי הצפוי.

יחד עם סדרות מקרו-כלכליות מסורתיות עם היסטוריה ארוכה, אנו משתמשים במספר מקורות נתונים שהושקו רק לאחרונה, כמו סקר הערכת המגמות בעסקים (מכאן ואילך, BTS) או נתונים יומיים של רכישות בכרטיסי אשראי וצריכת חשמל. ביישומי זמן אמת, מדדים מבוססי BTS יכולים לשמש כפרוקסי עדכני עבור משתנים מקרו-כלכליים חשובים המתפרסמים באופן טיפוסי בפיגור יחסית ארוך. ממצאינו מעלים כי הסקרים יכולים לשפר משמעותית את דיוק החיזוי, אך רק כאשר אלה משמשים לצורך ההשלמה בקצה של האינדיקטורים המסורתיים. היבטים שונים של השימוש בנתונים "רכים" לחיזוי זמן אמת נידונים בפירוט בחלק 4.2.

העבודה מאורגנת באופן הבא. חלק 2 מציג את המתודולוגיה האקונומטרית. חלק 3 מתאר את הנתונים. חלק 4 דן בניסיון שלנו בחיזוי מחוץ למדגם ובסוגיות פרקטיות אחרות הקשורות לשימוש בנתונים "רכים" ולשיטות השלמה ביישומי זמן אמת, וחלק 5 מסכם.

2. תיאור המודל

חלק זה מתאר את המבנה של המודל שלנו. אנו מתחילים בסקירה קצרה של מתודולוגיית המודל הפקטורי הדינמי. בהמשך, בדומה למתודולוגיית (Bräuning and Koopman, 2014), אנו מפרטים את הדינמיקה המשותפת של צמיחת התוצר החודשית הבלתי נצפית באמצעות אינדיקטורים חודשיים, וכן את שיטת הורדת הממד (collapsing) המעודכנת. לבסוף, על ידי שימוש בקירוב של Mariano and Murasawa (2003), אנו מרחיבים את ייצוג המצב-מרחב (State-Space) כדי לכלול באותה מסגרת פרקטורית דינמית את צמיחת התוצר הרבעונית הנצפית יחד עם הפאנל של משתנים חודשיים.

נניח כי $x_t = (x_{1,t}, x_{2,t}, \dots, x_{n,t})'$ כאשר $t = 1, 2, \dots, T$ הוא וקטור של n סדרות חודשיות שעברו טרנספורמציה כדי להפוך לסטציונריות ומורכזו. מודל פקטורי דינמי (DFM) מבוסס על ההנחה שניתן לסכם את התלות בין הרכיבים של x_t בעזרת מספר גורמים משותפים בלתי נצפים. המודל מנוסח באופן הבא:

$$(1) \quad x_t = \Lambda F_t + \varepsilon_t, \varepsilon_t \sim N(0, R)$$

כאשר F_t הוא וקטור $(r \times 1)$ של פקטורים בלתי נצפים, Λ הוא מטריצת $(n \times r)$ של המקדמים שלהם, ו- ε_t הוא וקטור $(n \times 1)$ של רעשים. אנו מניחים שהפקטורים מיוצרים על ידי תהליך VAR(p) סטציונרי:

$$(2) \quad F_t = \sum_{s=1}^p \Phi_s F_{t-s} + u_t, u_t \sim (0, Q)$$

כאשר Φ_s הן מטריצות $(r \times r)$ של מקדמים אוטורגרסיביים. בהקשר למודל זה, פרוצדורות ההסקה והחיזוי יכולות להתבצע באמצעות מסנן קלמן (Kalman Filter; ראו, לדוגמה, Hamilton, 1994, פרק 13). כדי לתאר את המודל, אנו מתחילים מספציפיקציה של הדינמיקה של צמיחת התוצר החודשית הבלתי נצפית, המסומנת על ידי y_t . הכללת משתנים רבעוניים וחודשיים באותה מסגרת DFM תידון בהמשך חלק זה. נתחיל ממודל פשוט (benchmark) המבוסס על ההנחה שהלוגריתם של התוצר מתנהל בהתאם למהלך מקרי בתוספת מגמה (drift), שנותן את הדינמיקה הבאה של y_t :

$$(3) \quad y_t = \mu + \varepsilon_{y,t}$$

כאשר $\varepsilon_{y,t} \sim N(0, \sigma_{\varepsilon_y}^2)$, μ מייצג את רכיב המגמה, ו- $\varepsilon_{y,t}$ הוא ההפרעה המקרית. ניתן לשפר את הביצוע של (3) על ידי הרחבת המודל בקבוצת הפקטורים הנגזרים מ- x_t , מה שנותן את הייצוג הבא:

$$y_t = \mu + \Lambda_{yx} F_t + \varepsilon_{y,t}$$

בצורת מצב-מרחב, ניתן לכתוב את משוואת המדידה באופן הבא:

$$(4) \quad \begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{bmatrix} \mu \\ 0 \end{bmatrix} + \begin{bmatrix} \Lambda_{yx} \\ \Lambda \end{bmatrix} F_t + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_t \end{pmatrix}$$

בנוסחה זו, הדרך הפשוטה היא להשתמש במסנן קלמן כדי לטפל בסוגים שונים של תצפיות חסרות שיכולות להיווצר עקב משכי זמן היסטוריים שונים או פרסומי נתונים שאינם מסונכרנים (Durbin and Koopman, 2001).

ביישומים מקרו-כלכליים, מספר האינדיקטורים החודשיים יכול להיות גבוה יחסית למספר התצפיות, מה שיכול להפוך את רוטינת האמידה למורכבת הרבה יותר ולהתבטא בשונות גבוהה של המודל. כדי להתגבר על קושי זה, (Bräuning and Koopman, 2014) הציגו מודל פקטורי דינמי מכווץ המשתמש

בהורדת ממד במשוואת המדידה. הרעיון הוא להשתמש בגרסה מוחלפת (transformed) של משוואת המדידה (4), לאחר שזו הוכפלה מראש במטריצת החילוף:

$$P = \begin{bmatrix} 1 & 0 \\ 0 & A \end{bmatrix}$$

כאשר A הוא מטריצת $(r \times n)$. משוואת המדידה המותאמת הופכת אפוא ל-

$$\begin{pmatrix} y_t \\ Ax_t \end{pmatrix} = \begin{bmatrix} \mu \\ 0 \end{bmatrix} + \begin{bmatrix} \Lambda_{yx} \\ A\Lambda \end{bmatrix} F_t + \begin{pmatrix} \varepsilon_{y,t} \\ A\varepsilon_t \end{pmatrix}$$

ואילו משוואת המצב נותרת ללא שינוי. כדי לצמצם את איבוד המידע שנגרם, Bräuning and Koopman (2014) בונים את A באמצעות משקלי המרכבים העיקריים. נציין על ידי $\hat{F}_t = A_{pc}x_t$ את המרכיבים העיקריים המקושרים ל- r ערכים העצמיים (eigenvalues) הגדולים ביותר של מטריצת הנתונים $(x_1, \dots, x_T)'$. על ידי כתיבת $\hat{F}_t \approx F_t + error$, הכפלה מראש של (4) ב- $P = \begin{bmatrix} 1 & 0 \\ 0 & A_{pc} \end{bmatrix}$ נותנת את המודל הפקטורי הדינמי המכווץ

(5)

$$\begin{pmatrix} y_t \\ \hat{F}_t \end{pmatrix} = \begin{bmatrix} \mu \\ 0 \end{bmatrix} + \begin{bmatrix} \Lambda_{yx} \\ I_r \end{bmatrix} F_t + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_t^{pc} \end{pmatrix}$$

שבו $\varepsilon_t^{pc} = A_{pc}(\varepsilon_t - (A'_{pc} - \Lambda)F_t)$.

מאחר שצמיחת התוצר נצפית על בסיס רבעוני, עלינו להתאים את ייצוג המצב-מרחב בדרך נוספת כדי שנוכל לכלול את צמיחת התוצר הנצפית בתדירות רבעונית והאינדיקטורים החודשיים באותה מערכת של פקטורים דינמיים. לצורך כך, אנו מאמצים את המסגרת שהוצעה על ידי Mariano and Murasawa (2003). הרעיון הוא לבטא את התוצר הרבעוני הנצפה במונחי מקבילו החודשי הנצפה חלקית, שיוביל לווריאנט של בעיית התצפיות החסרות, שניתן לטפל בה בקלות באמצעות מסנן קלמן.

נניח ש- GDP_t^Q מציין את הרמה הרבעונית הנצפית של התוצר ו- GDP_t^M הוא מקבילו החודשי הבלתי נצפה. לפיכך, אם GDP_t^Q נצפה בסוף כל רבעון, אנו מקבלים כי:

$$\begin{cases} GDP_t^Q = GDP_t^M + GDP_{t-1}^M + GDP_{t-2}^M & t = 3, 6, 9, \dots \\ NA & o.w. \end{cases}$$

נגדיר בנוסף כי $Y_t^Q = \log(GDP_t^Q)$ וכן $Y_t^M = \log(GDP_t^M)$. בנוסף, נסמן את שיעור הצמיחה הלוגריתמי החודשי הבלתי נצפה y_t , שווה ל- $-\Delta Y_t^M$. כדי לגשר בין הסדרה הרבעונית הנצפית לנתונים החודשיים, אנו מגדירים גם את צמיחת התוצר כסדרה חודשית שנצפית רק בסוף הרבעון:

$$\begin{cases} y_t^Q = Y_t^Q - Y_{t-3}^Q & t = 3, 6, 9, \dots \\ NA & o.w. \end{cases}$$

ומשתמשים בקירוב של Mariano and Murasawa (2003):

$$y_t^Q = \frac{1}{3}y_t + \frac{2}{3}y_{t-1} + y_{t-2} + \frac{2}{3}y_{t-3} + \frac{1}{3}y_{t-4} \quad t = 3, 6, 9, \dots$$

בהמשך, הסדרות הרבעוניות והחודשיות מבוטאות במודל משותף באמצעות הרחבה מתאימה של משוואת המצב. לנוחות ההצגה, נניח כי $r = p = 1$ (לנוסחה הכללית, ראו נספח ב'). במקרה זה, ניתן לכתוב את צורת מצב-מרחב באופן הבא:

$$\begin{pmatrix} y_t^Q \\ \hat{F}_t \end{pmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ y_{t-4} \\ F_t \end{pmatrix} + \begin{pmatrix} 0 \\ \varepsilon_t^{pc} \end{pmatrix}$$

$$\begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ y_{t-4} \\ F_t \end{pmatrix} = \begin{bmatrix} \mu \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \Lambda_{yx}\Phi_1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Phi_1 \end{bmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ y_{t-4} \\ y_{t-5} \\ F_{t-1} \end{pmatrix} + \begin{bmatrix} 1 & \Lambda_{yx} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \varepsilon_{y,t} \\ u_t \end{pmatrix}.$$

לסיכום, האמידה של CDFM מורכבת משני שלבים. ראשית, המידע הכלול במשתנים החודשיים מסוכם במספר קטן של גורמים באמצעות ניתוח המרכיבים העיקריים. אם חלק מהתצפיות חסרות, מבצעים השלמה באמצעות אלגוריתם ה-EM שהוצע על ידי Stock and Watson (2002). בשלב השני, הגורמים החודשיים מוכנסים למודל במשותף עם צמיחת התוצר הרבעונית במודל פקטורי דינמי עם ההתאמות הנדרשות כדי לשלב נתונים בשתי התדירויות.

עקב המגבלות הנזכרות לעיל של שיטת ה-PCA בחילוץ מידע מאינדיקטורים בעלי תדירות גבוהה, שיכולות להיות חזקות אף יותר בשווקים קטנים כמו ישראל שבה מספר המשתנים מצומצם יחסית, אנו מציעים להשתמש בשיטת ה-PLS בשלב הראשון. PLS היא טכניקה להורדת ממד אשר בשונה מ-PCA, משתמשת במשתנה תגובה נוסף (שיכול להיות רב-ממדי) במבנה שלה ומחלצת פקטורים בלתי תלויים שמטרתם להסביר את מרב השונות בפנל החודשי תוך התחשבות במשתנה המטרה. בנוכחותו של פרוקסי מתאים עבור צמיחת התוצר החודשית הבלתי נצפית שיכול לשמש כמשתנה מטרה ב-PLS, טמון בשיטה פוטנציאל לשפר משמעותית את ביצועי המודל. ראשית, השיטה צפויה לייחס משקל קטן יותר למשתנים עם רעש ייחודי חזק. שנית, צפוי שהיא תוכל לחלץ יותר מידע רלוונטי באמצעות מספר קטן יותר של מרכיבים ובכך להפחית את שונות המודל. מנקודת מבט תיאורטית, (Stock and Watson, 2002) מראים כי PCA יכול להניב אומדנים עקביים של F_t כאשר T ו- n הולכים לאינסוף. לאחרונה, Groen and Kapetanios (2016) הראו כי ל-PLS יש תכונה דומה. יתרה מכך, הוא שומר על מאפיינים של אופטימליות אפילו במקרה של גורמים חלשים, שעבורם ידוע כי רגרסיית PCA הופכת לבלתי עקיבה. בספרות העוסקת בחיזוי טווח קצר נעשה לאחרונה שימוש בממצאים אלה במסגרת דמוית-MIDAS (לדוגמה, Kelly and Pruitt, 2015; Hepenstrick and Marcellino, 2019).

3. תיאור הנתונים

אנו משתמשים ב-31 אינדיקטורים כלכליים, מתוכם רק משתנה המטרה – צמיחת התוצר – הוא רבעוני והשאר חודשיים. סקירה של האינדיקטורים, כולל המקורות, טווחי הזמן הזמינים והטרנספורמציות, מובאת בלוח 4. האינדיקטורים החודשיים נבחרו כדי לייצג מגוון רחב של היבטים של הפעילות הריאלית במשק הישראלי, וככלל הם דומים למה שמקובל בספרות (לדוגמה, Bok et al., 2018). כל האינדיקטורים מותאמים לשניוניים במחיר ולעונתיות, ונעשתה להם טרנספורמציה מתאימה כדי להפוך אותם לסטציונריים. נתוני חו"ל מוצגים במחירים קבועים בדולר ארה"ב.

בנוסף, בחלק 4.2 אנו משתמשים במאזני הנטו של ההערכות (balances of opinions) מסקר המגמות בעסקים. זהו סקר חודשי הנערך משנת 2011 על ידי הלשכה המרכזית לסטטיסטיקה בישראל בקרב 1,600 מנהלי פירמות בתעשיות הייצור, הבנייה, המסחר, המלונאות והשירותים, ובמסגרתו המנהלים מספקים הערכה (על סולם של חמש נקודות, מ"ירידה ניכרת" עד "עלייה ניכרת") של המצב הנוכחי של עסקיהם ואת תחזיותיהם לעתיד הקרוב. השאלונים קשורים לפרמטרים העיקריים של החברות, כמו תפוקה, מכירות וייצוא, העסקת עובדים, ומחירים. סקר המגמות בעסקים מתפרסם במהלך השבוע השני של כל חודש, ולכן יכול להביא מידע מקדים עבור האינדיקטורים המקרו-כלכליים הרלוונטיים המתפרסמים בפיגורים ארוכים יותר.

4. תוצאות אמפיריות

בחלק זה אנו דנים במגוון היבטים של התוצאות האמפיריות שלנו. ראשית, אנו מדגימים את היתרונות של CDFM בחיזוי טווח קצר של צמיחת התוצר. לאחר מכן, אנו מדגימים את השימושים שלו בניטור השינויים החודשיים בפעילות הכלכלית. לסיום, אנו חוקרים סוגיות שונות הקשורות לחיזוי בזמן אמת. חלק 4.1 עוסק בנושא המרכזי שלנו, חיזוי טווח קצר של צמיחת התוצר. אנו משווים את דיוק החיזוי של DFM ו-CDFM באמצעות סימולציה של חיזוי בתנאי זמן אמת, שבה התחזיות עבור כל רבעון מחושבות מתוך סט מתרחב של נתונים זמינים (חודשיים ורבעוניים). סט המידע הקטן ביותר מורכב מנתונים חודשיים עד לחודש הראשון של הרבעון הנחזה, ואינו מכיל מידע על הצמיחה בתקופה הקודמת. לאחר מכן, אנו בוחנים תחזית המבוססת על החודשיים ועל הצמיחה ברבעון הקודם. לבסוף, סט המידע הגדול ביותר מורכב מנתונים חודשיים מלאים. עבור ה-CDFM, הטיפול בנתונים חסרים הוא באמצעות אלגוריתם ה-EM (Stock and Watson, 2002). חלק 4.2 עוסק בנושאים הקשורים לחיזוי זמן אמת, בדגש על השימוש בנתונים "רכים" המתפרסמים קרוב יותר למועד התרחשותם.

4.1 סימולציה של המודל בתנאי זמן אמת

בהמשך לתיאור המודל בפרקים 1 ו-2, אנו אומדים את CDFM (מכאן ואילך, $CDFM^{PLS}$) תוך שימוש בשיטת PLS להורדת הממד, כאשר מדד סך הפדיון משמש כפרוקסי חודשי לצמיחת התוצר הרבעונית. טיב החיזוי של מודל זה משווה ל-CDFM שמבוסס על ה-PCA (מכאן ואילך, $CDFM^{PC}$), מודל פקטורי דינמי (מכאן ואילך, DFM), והמודל הנוכחי של בנק ישראל לחיזוי טווח קצר (מכאן ואילך, Bridge), בו החיזוי מתקבל מצירוף של מודל אקונומטרי (לפרטים טכניים נוספים, ראו: Krief, 2011) עם חיזוי המבוסס על שיפוט. שתי הספציפיקציות של CDFM מבוססות על שני פקטורים. הערכת טיב החיזוי בתת-פרק זה מבוססת על התקופה שקדמה למשבר הקורונה והיא משתמשת בנתונים החודשיים והרבעוניים מינואר 2000 עד דצמבר 2019. כדי ליצור תחזית טווח קצר ב-CDFM ו-DFM, אנו אומדים מחדש את המודל, עבור כל רבעון במדגם הבדיקה, באמצעות הנתונים החודשיים והרבעוניים מינואר 2000 ועד לאחד מבין שלושת הווריאנטים של סט המידע שתואר לעיל. תקופת האימון הראשוני היא מינואר 2000 עד ינואר 2010, והערכת החיזוי מתבצעת עבור התקופה מ-Q1: 2010 עד Q4: 2019. האתגרים הקשורים לחיזוי טווח קצר בזמן אמת ובמהלך תקופות של משבר כגון 2020, יידונו ביתר פירוט בתת-הפרק הבא.

חיזויי הטווח הקצר של מודל Bridge נוצרים בשלושה צעדים. ראשית, מודלים נפרדים נאמדים באמצעות OLS לתתי-הרכיבים של התוצר, על ידי שימוש בסכומים מצרפיים רבעוניים של האינדיקטורים החודשיים. לאחר מכן, המודלים מצורפים יחד כדי לבנות חיזוי טווח קצר לתוצר. לסיום, נעשית התאמה של חיזויי הטווח הקצר על ידי שימוש בקביעה שיפוטית. חיזויי הטווח הקצר מבוצע במהלך המחצית השנייה

של החודש האחרון ברבעון הנבדק. לפיכך, דיוק הניבוי שלו ניתן להשוואה רק לחיזויי טווח קצר של ה-DFM וה-CDFM שמתבצעים בהתבסס על שני חודשי נתונים.

כפי שצוין בפרק 2, הורדת הממד מיועדת להפחית את מספר הפרמטרים וכפועל יוצא את שונות החיזוי. אולם בה בעת, היא גורמת לאובדן מידע (Bräuning and Koopman, 2014). כתוצאה מכך, יהיה מעניין להשוות את טיב החיזוי של ה-CDFM לאלה של ה-DFM, כדי לראות אם היתרונות מהורדת הממד עולים על החיסרון של אובדן המידע. הספציפיקציה שלנו עבור DFM מבוססת על אותו קובץ משתנים כמו ב-CDFM עם מבנה הקיבוץ שהוצע על ידי Bok et al. (2018), והיא נאמדת באמצעות אלגוריתם מקסום הציפיות כמתואר ב-Banbura and Modugno (2014).

טיב החיזוי של המודלים מחוץ למדגם מושווים למודל ייחוס AR שהותאם באמצעות קריטריון המידע של Akaike (AIC). לוח 1 להלן מציג את שורש הטעות הריבועית הממוצעת (RMSE) ואת הסטייה המוחלטת הממוצעת (MAE), לפי היקף המידע הזמין, ביחס למדדי השגיאה התואמים של מודל הייחוס, וכן את תוצאות מבחן Diebold and Mariano עם תיקון למדגמים קטנים (Harvey et al., 1997). בנוסף, איור 1 מציג חיזוי טווח קצר המבוסס על 2 חודשים עבור כל אחד מארבעת המודלים יחד עם האומדן המעודכן ביותר² של צמיחת התוצר.

לוח 1: טעויות ריבועיות ממוצעות ומוחלטות ממוצעות בחיזוי טווח קצר

מודל	לפי 1 חודש		לפי 2 חודשים		לפי 3 חודשים	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
CDFM ^{PLS}	0.898	0.871	*0.872	**0.814	**0.856	**0.809
CDFM ^{PC}	0.980	0.908	0.975	0.911	0.955	0.893
DFM	0.999	0.982	1.007	0.932	1.027	0.934
Bridge			0.925	0.876		

הערה: ה-MAE וה-RMSE עבור המודלים מדווחים כיחסים למדדי הטעות התואמים של מודל ייחוס ה-AR שנאמד באמצעות ה-AIC. ** ו- * מציינים דחייה של השערת האפס של מבחן Diebold and Mariano (1995) (המניחה שאין הבדל בדיוק החיזוי לעומת מודל הייחוס) ברמה של 5% ו-10%, בהתאמה. תקופת ההערכה היא מ-2010: Q1 עד 2019: Q4.

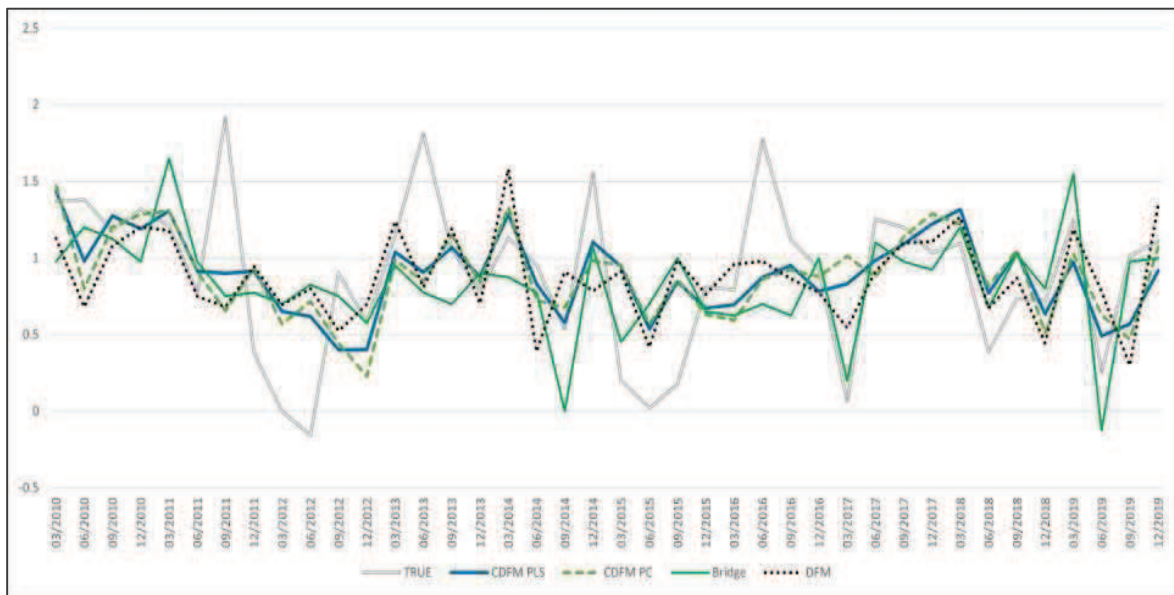
² זמין באפריל 2021.

לוח 2: הדיוק בכיוון החיזוי

מודל	לפי 1 חודש	לפי 2 חודשים	לפי 3 חודשים
CDFM ^{PLS}	74%	82%	77%
CDFM ^{PC}	74%	79%	77%
DFM	77%	82%	77%
Bridge		72%	

הערה: הלוח מדווח על אחוז הפעמים שבהן המודלים ניבאו נכונה אם הצמיחה בתקופה הנוכחית תהיה גבוהה יותר או נמוכה יותר מזו שבתקופה הקודמת. שיעור דיוק הניבוי הנאיבי (הקבוע) הוא 0.56 וזה של מודל לייחוס הוא 0.72.

איור 1: חיזויי טווח קצר לצמיחת התוצר הרבעוני, Q1: 2010 עד Q4: 2019



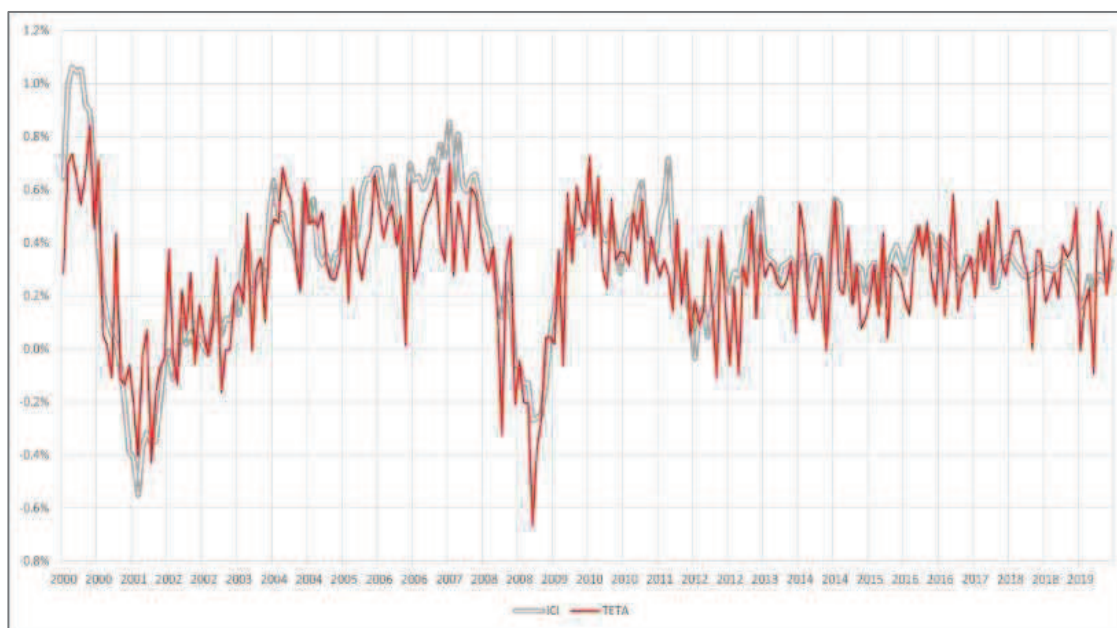
אנו יכולים לראות באופן כללי כי בכל שלושת המצבים, CDFM^{PLS} נותן את טעות הניבוי הנמוכה ביותר והשיפור לעומת מודל הייחוס נע מ-13% (כאשר החיזוי לטווח קצר מבוסס על חודש אחד של נתונים) עד 19% (על הנתונים הרבעוניים המלאים). השיפור מובהק סטטיסטית עבור סט הנתונים המכיל 2 חודשים או יותר. בהקשר לסוגיית הורדת הממד, ניכר כי הן ה-CDFM^{PLS} והן ה-CDFM^{PC} מראים טעות חיזוי קטנה יותר מה-DFM, והשיפור של האחרון לא נמצא כמו בהקשר סטטיסטית. ממצא זה מבליט את היתרון הנובע מצמצום בפרמטרים במודלים של חיזוי.

בנוסף, יהיה מעניין לראות אם המודלים מסוגלים לנבא את השינויים הכיווניים ברמת הצמיחה – ובפרט, האם הצמיחה הנוכחית גבוהה יותר מהקודמת. לוח 2 מציג את התוצאות עבור ארבעת המודלים. מסתבר שהמודלים מגלים רמת דיוק דומה, עם שיפור קל בלבד לעומת הייחוס.

מידול משותף של המשתנים החודשיים והרבעוניים ב-DFM מושג על ידי הרחבת וקטור המצב במונחי צמיחת התוצר החודשית הבלתי נצפית. כפי שהראו (Brave et al. 2019), שיעורי צמיחה חודשיים

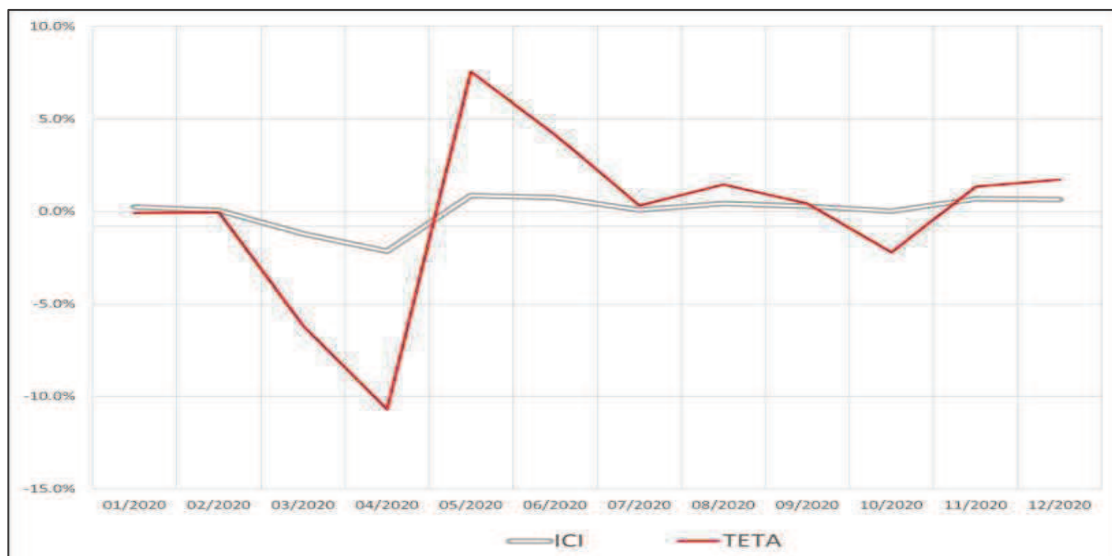
אלה יכולים לספק אינדיקציה רבת-ערך למצב הנוכחי של המשק והם מוצגים במנוחים אינטואיטיביים. מדד הפעילות החדש מחושב מתוך הערכים המוחלקים של וקטור המרחב התואם את צמיחת התוצר החודשית הבלתי נצפית. כפי שניתן לראות מאיורים 2 ו-3, אינדיקטור הפעילות (TETA) שלנו תואם באופן כללי את המדד המשולב למצב המשק של ישראל³ (ICI) במהלך שלבים שונים של המחזור. אולם, יש לו תנודתיות גבוהה בהרבה, הנובעת בחלקה מן השימוש בסט מידע רחב יותר ומהאילוץ הכופה על שיעורי הצמיחה החודשיים להסתכם במדויק לצמיחה הרבעונית המנובאת (או לצמיחה בפועל, כאשר זו זמינה). יתרונותיה של רגישות גבוהה זו בלטו במיוחד במהלך שנת 2020. כפי שניתן לראות מאיור 3, במהלך כלל התקופה, ה-ICI ממשיך לדשדש סביב האפס ואינו מראה כמעט כל אינדיקציה למשבר. במדד החדש, לעומת זאת, אנו רואים ירידה חדה, במקביל לצעדי החירום שהופעלו במאסר, והירידה שממשיכה להעמיק באפריל. השיפור הניכר במאי מקביל להקלות של הגבלות הסגר. ולבסוף, הערכים השליליים באוקטובר מקבילים לסגר השני. כל זה בשונה מ-ICI, שנותר כמעט על אותה רמה מחודש מאי ואילך. יש בכך כדי להמחיש את יכולתו של המדד החדש לתת תמונה מדויקת יותר של המצב הנוכחי של המשק ולהצביע על נקודות תפנית מוקדם יותר.

איור 2: השינוי בפעילות הכלכלית החודשית לפי TETA מול המדד המשולב למצב המשק (ICI) (2000/1-2019/12)



³ <https://www.boi.org.il/en/Research/Pages/ind.aspx>

איור 3: השינוי בפעילות הכלכלית החודשית לפי TETA מול המדד המשולב למצב המשק (ICI) (2020/1-2020/12)



4.2 חיזוי טווח קצר בזמן אמת ושימוש בנתונים "רכים"

בתת-פרק זה אנו דנים בסוגיות הקשורות לשימוש בנתונים רכים בחיזוי טווח קצר בזמן אמת. סדרות חודשיות "מסורתיות" שנכנסות למודל, כמו ייצור תעשייתי, הכנסות, ומדדי התעסוקה, מתפרסמות בפיגורים ארוכים (ראו בלוח 5 פרטים נוספים על מחזור הנתונים). מאידך, קיימים נתונים "רכים" שיכולים לשמש כפרוקסי לאותן סדרות חודשיות ושמתפרסמים בפיגור קצר יותר, כמו סקרי עסקים.

בעבודה הנוכחית, אנו משתמשים בסקר המגמות בעסקים (להלן BTS). זהו סקר חודשי הנערך משנת 2011 על ידי הלשכה המרכזית לסטטיסטיקה בישראל בקרב 1,600 מנהלי פירמות בתעשיות הייצור, הבנייה, המסחר, המלונאות והשירותים. המענה לסקר הוא חובה, ובמסגרתו מנהלי הפירמות מספקים הערכה (על סולם של חמש נקודות, מ"ירידה ניכרת" עד "עלייה ניכרת") של המצב הנוכחי של עסקיהם ואת תחזיותיהם לעתיד הקרוב. השאלונים קשורים לפרמטרים העיקריים של החברות, כמו תפוקה, מכירות פנים וייצוא, העסקת עובדים ומחירים. מאזני הנטו של ההערכות מסקרים אלה נמצאו כשימושיים בחיזוי צמיחת התוצר (ראו, לדוגמה: Chernis and Sekkel, 2017; Pichette and Rennison, 2011) ויכולים לשמש באופן ישיר או להשלמה בקצה של האינדיקטורים ה"מסורתיים".

השימוש בנתוני סקרים לחיזוי תנאים כלכליים הינו פופולרי בקרב בנקים מרכזיים; ראו Anesti et al. (2017), Bok et al. (2018), ו-Chernis and Sekkel (2017) לדוגמאות של יישומים בידי בנקים מרכזיים, ואת Banbura et al. (2013) לסקירה רחבה של בנושא. באופן טיפוסי, משתנים מבוססי סקרים מתווספים ישירות לסט האינדיקטורים החודשיים ה"מסורתיים" או אפילו משמשים להחלפתם. מצד שני, נמצא כי הסקרים נוטים לתגובת-יתר בתקופות של אי-ודאות גבוהה (Anesti et al., 2017). לפיכך, השימוש בהם ישירות או אפילו יחד עם האינדיקטורים ה"מסורתיים" שהם אמורים להוות פרוקסי עבורם, עלול לערער

את יציבות המודל. לדוגמה, בהקשר של רגרסיה לינארית, קל לראות שכאשר המודל מכיל הן את המשתנה המסביר הרלוונטי והן את הפרוקסי שלו (שהוא תערובת של סיגנל ורעש), המקדם של משתנה הפרוקסי ייצג רק את השפעת הרעש. לאור זאת, אנו בוחנים גישה שונה, שבה במקום לאמן את המודל ישירות על הסקרים, אנו משתמשים בהם רק לצורך השלמה בקצה של הסדרות הרלבנטיות. באופן זה אנו מגבילים את השפעת האינדיקטורים ה"רכים", כך שישפיעו רק על אי-הוודאות של החיזוי אך לא על הפרמטרים של המודל.

לצורך כך, שחזרנו את גרסאות הנתונים של זמן אמת (vintages) בתדירות השבועית, ויצרנו רצף תחזיות המבוסס על סט מידע מתרחב (דהיינו, כדי לחקות את מחזור הנתונים האמיתי, החל מהחודש השני של רבעון הייחוס ועד חודש אחד לאחר סוף רבעון, הפקנו חיזויי טווח קצר בתדירות שבועית). במסגרת ה-CDFM, ניתן להוסיף משתני BTS ישירות לפאנל החודשי או להשתמש בהם להשלמה חיצונית של הסדרה שעבורה הם אמורים לשמש כפרוקסי. כדי לבדוק איזו שיטה טובה יותר, אנו מעריכים את הדיוק הממוצע של חיזויי הטווח הקצר על ידי השוואתם לשיטות מבוססות BTS (קרי, עם השלמה והוספה ישירה של משתני BTS) כאשר המודל ללא משתני BTS משמש כמודל הייחוס. עבור השיטה המשתמשת בהשלמת נתונים חיצונית על ידי BTS, בדומה ל-Roash and Suhoy (2019), ההשלמה בקצה נעשית באמצעות מודל של הרכיבים הבלתי נצפים (מכאן ואילך, UCM), שהוצע על ידי Harvey (1990). לרשימת הסדרות עם ההשלמות ומחזור הנתונים, ראו לוח 5.

הניסוי מתבצע עבור התקופה מפברואר 2018 עד ינואר 2021 שכוללת 156 תחזיות טווח קצר; כלומר, עבור כל רבעון אנו יוצרים תחזיות כל שבוע החל מהחודש השני ועד חודש אחד לאחר תום הרבעון. הבחירה בתקופה זו נובעת בעיקר מעומק היסטורי של נתוני BTS. אלה זמינים מאז 2016 והתקופה שלפני 2018 שימשה לאימון ראשוני של ה-UCM.

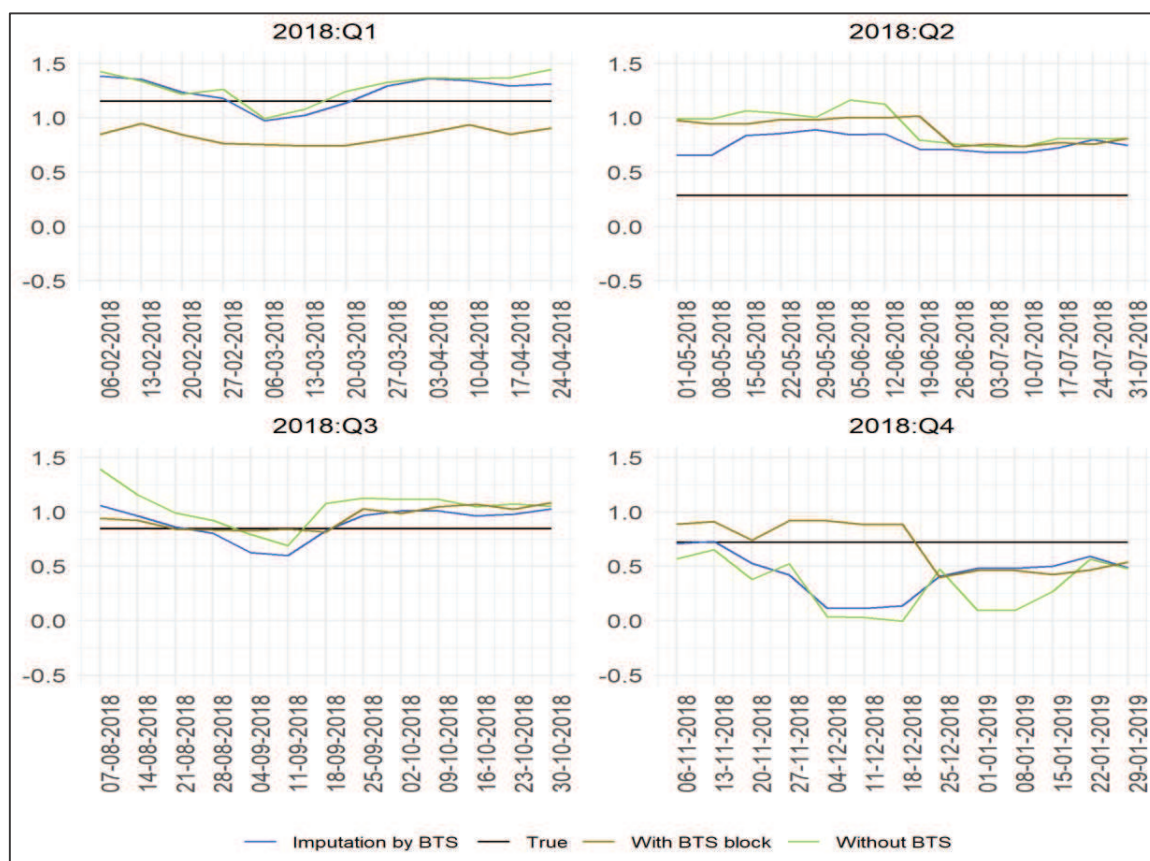
לוח 3 מספק פירוט של הטעות הממוצעת בחיזוי טווח קצר של $CDFM^{PLS}$ לפי שנים ולפי השיטה להכללת BTS. השורה התחתונה מסכמת את ה-RMSE וה-MAE עבור התקופה הכוללת מ-2018 עד 2020 (כולל משבר הקורונה). ככלל, ניתן לראות כי השימוש בהשלמת נתונים באמצעות BTS מפחית את ה-RMSE וה-MAE החזויים בטווח הקצר בהשוואה לשיטה שאינה משתמשת ב-BTS להשלמת נתונים. לעומת זאת, הוספת משתני ה-BTS באופן ישיר מגדילה את הטעות בהשוואה למודל הייחוס. בנוסף, ביצענו את המבחן של Diebold and Mariano כדי לוודא שההפרש לעומת שיטת הייחוס חסרת ה-BTS הוא מובהק. נמצא כי השיטה הכוללת השלמת נתונים על ידי BTS נותנת טעות קטנה יותר ברמת מובהקות של 1%, והשיטה שמשלבת ישירות את נתוני ה-BTS נותנת טעות גדולה יותר ברמת מובהקות הנעה בין 5% ל-1% בהתאם למדדי הטעות. איורים 4-6 מציגים את ההתפתחות של חיזויי הטווח הקצר במהלך התקופה מפברואר 2018 עד ינואר 2021. ניתן לראות כי חיזויי הטווח הקצר המשתמשים בהשלמת נתונים באמצעות BTS הם חלקים הרבה יותר. בנוסף, מודל הייחוס מתכנס באופן כללי לאותה נקודה כמו המודל עם השלמת הנתונים. למרות האמור, טעות חיזוי הנמוכה יותר באופן מובהק במקרה של השלמת נתונים באמצעות BTS מדגישה את השימושיות של נתונים "רכים" כאינדיקטור מהיר.

לוח 3: סיכום טעויות לפי השיטה להכללת BTS

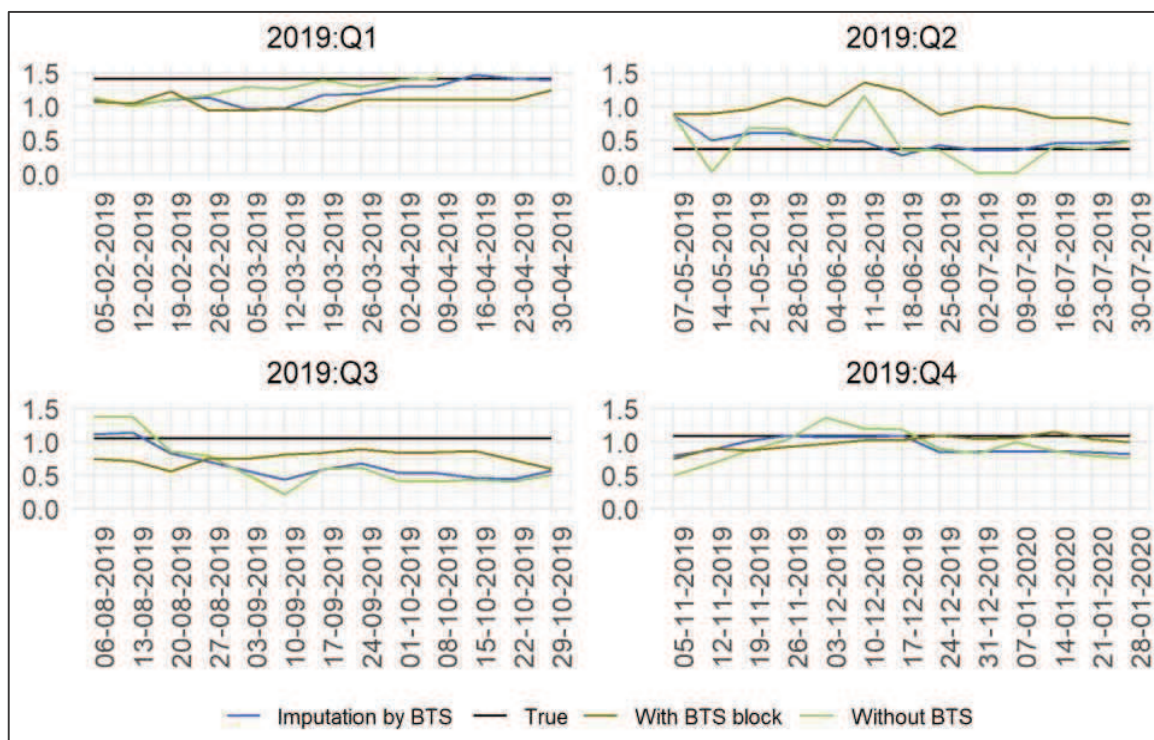
BTS עם גוש		BTS ללא		השלמת נתונים באמצעות BTS		שנה
RMSE	MAE	RMSE	MAE	RMSE	MAE	
0.376	0.313	0.436	0.366	0.320	0.264	2018
0.397	0.336	0.355	0.286	0.294	0.234	2019
4.764	4.001	4.323	3.548	3.882	3.269	2020
*2.768	**1.55	2.517	1.400	***2.225	***1.255	All

הערה: הלוח מדווח את הטעויות הממוצעות לפי שנה ולפי השיטה להכללת נתוני הסקר. תחזיות הטווח הקצר נוצרות בתדירות שבועית מפברואר 2018 עד ינואר 2021. השורה התחתונה מראה את ה-RMSE וה-MAE עבור כלל התקופה מ-2018 עד 2020. *, **, *** מציינים דחייה של השערת האפס של מבחן Diebold and Mariano (1995) (שאינן הבדל בדיוק החיזוי לעומת מודל הייחוס שלא כולל BTS) ברמה של 1%, 5% ו-10%, בהתאמה.

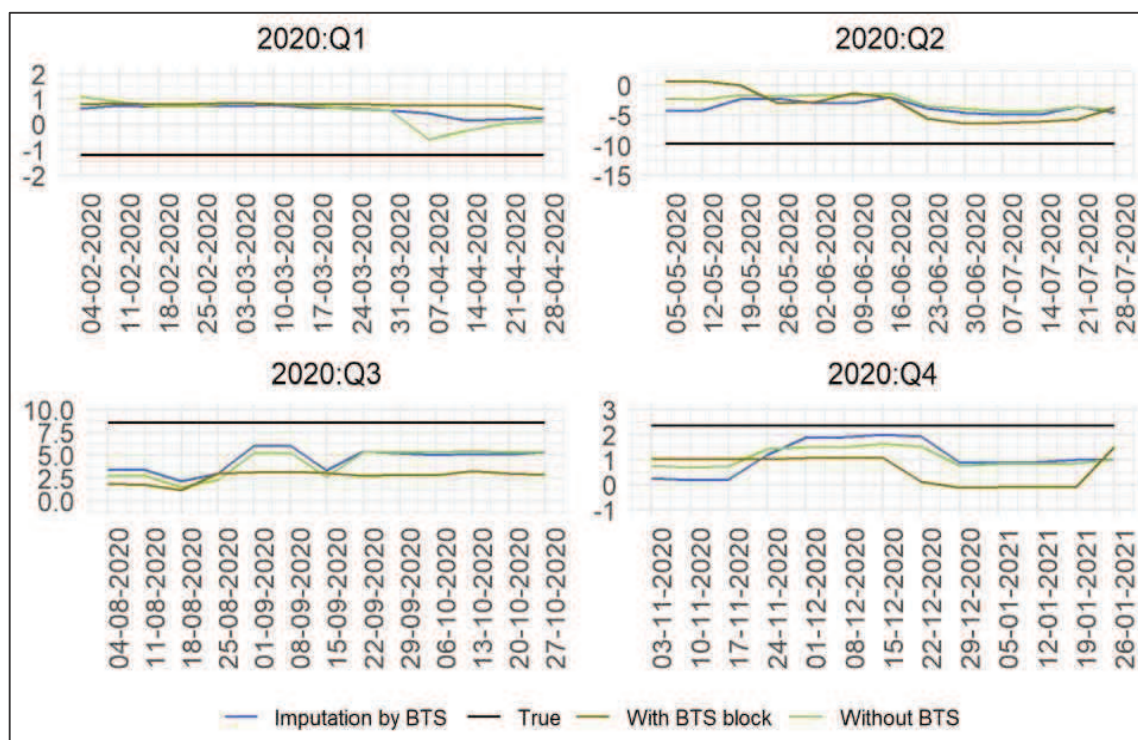
איור 4: חיזויי טווח קצר של התוצר הרבעוני בזמן אמת, 2018



איור 5: חיזויי טווח קצר של התוצר הרבעוני בזמן אמת, 2019



איור 6: חיזויי טווח קצר של התוצר הרבעוני בזמן אמת, 2020



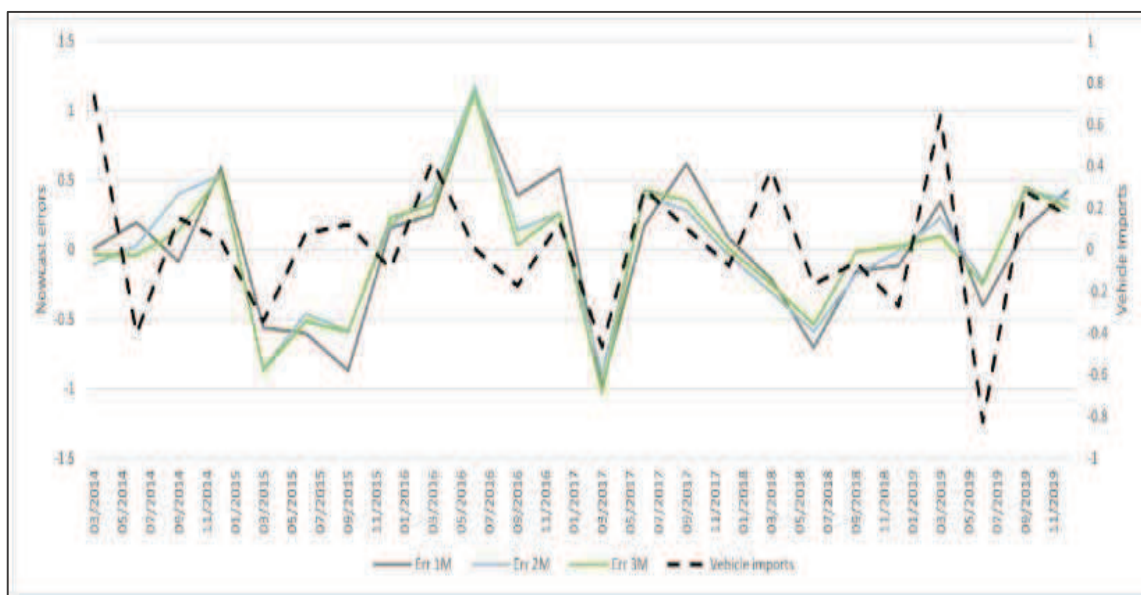
4.3 טיפול בהתערבויות רגולטוריות

בחלק זה, אנו דנים בטיפול בהשפעת מיסוי על יבוא רכבים על צמיחת התוצר. כפי שניכר מאיור 7 להלן, המציג את הטעויות החיזוי על בסיס שלושת הסוגים של סט נתונים ששימשו לניסוי בחלק 4.1, הטעויות נעות יחד עם השינויים בייבוא הרכב באופן ברור. זה קורה בגלל השילוב של כללי החשבונאות הלאומיים עם תקנות המסים הסביבתיים בישראל, שגורם לתנועות פתאומיות בתוצר שאינן קשורות לגורמים כלכליים.

כידוע, התוצר מייצג את סכום ההוצאות המקומיות על מוצרים סופיים בתוספת מאזן הסחורות והשירותים (ייצוא פחות ייבוא). לפיכך, באופן רגיל, גידול בייבוא לא אמור להוביל לשינוי מובהק בתוצר. אך במקרה של ייבוא רכב לישראל, ההשפעה חיובית בגלל שיעורי מס הייבוא על כלי רכב. כדי להבין את המנגנון, נחשוב על מכונת שיובאה זה עתה. מחירה הסופי מורכב מהמחיר המקורי ששולם לסוכן הרכב בחו"ל, מיסי יבוא (כ-90% מהמחיר המקורי), רווח, ומע"מ. המחיר הסופי נכנס לצריכה (או להשקעה עבור חברות ליסינג) ואילו המחיר המקורי ששולם בחו"ל מיוחס לייבוא. מאחר שהמחיר הסופי המתקבל מכפיל ואף יותר את המחיר המקורי, הגידול בייבוא רכבים מוביל לגידול בתוצר דרך השפעתו החיובית על הצריכה ועל ההשקעה. כפי שניתן לראות מאיור 7 להלן, טעויות החיזוי הראו באופן שיטתי ערכים חיוביים גבוהים, שאחריהם ירידה חדה. הסיבה לגלים הללו נעוצה בעדכונים התקופתיים של תקנות המסים הסביבתיים על פליטות כלי רכב, המובילים לעלייה במחירי הרכב. מאחר שרק תאריך העדכון ידוע מראש (בדרך כלל כמה חודשים לפני) אך הפירוט אינו ידוע, אי-הוודאות המתקבלת גורמת לאנשים להקדים את רכישות הרכב לפני עליית המחירים הצפויה, ואנו רואים עליות פתאומיות בתוצר. שינויים אלה אינם קשורים לגורמים הכלכליים במשק, ולכן קשה יותר לבטא אותם במודל רק על ידי הרחבת קבוצת הנתונים החודשיים במשתנה נוסף ואת התנועה המשותפת שלהם ביחס לדינמיקה של צמיחת התוצר החודשית במסגרת של CDFM.

כדי לטפל בסוגיה זו, בדומה ל-Sayag et al. (2021), אנו עושים התאמה לחיזויי הטווח הקצר של ה-CDFM באופן חיצוני בעזרת רגרסיה בין הטעויות הקודמות, נניח $e_{i,t}$, לבין הפרש הלוג של ייבוא הרכבים, נניח $imv_{i,t}$, כאשר $i \in \{1,2,3\}$ מייצג את מספר החודשים הידועים עבור ייבוא בזמן t . באופן ספציפי יותר, בדומה לסימולציה בחלק 4.1, $e_{i,t}$ תואם טעות בחיזוי טווח קצר בתקופה t המבוסס על קובץ מידע של i חודשים. בהתאם לכך, $imv_{i,t}$ מייצג את הצמיחה הלוגריתמית הרבעונית בייבוא הרכב, כאשר הסכום של ייבוא כלי הרכב ברבעון הנוכחי מחושב אך ורק מתוך הנתונים עד לחודש t של הרבעון הנבדק. חיזויי הטווח הקצר המתוקנים נבנים כסכום של חיזויי הטווח הקצר של CDFM והניבוי עבור $e_{i,t}$. לצערנו, בשל מספרם המצומצם מאוד של העדכונים קשה להעריך מהו השיפור בדיוק הניבוי שאנו מקבלים מהתיקון. בנוסף, ראוי לציין כי לאחר התיקון, חיזויי הטווח הקצר הרבעוניים הופכים ללא-עקביים עם האומדנים של צמיחת התוצר החודשית. אולם ניתן לפתור זאת בקלות, על ידי הרצת המחליק של קלמן על הנתונים שכוללים את התחזית המעודכנת במקום הנתון בפועל.

איור 7: טעויות בחיזוי טווח קצר של CDFMPLS על פי קובץ המידע הזמין (2014: Q1-2019: Q4)



5. סיכום

בעבודה זו, יישמנו הרחבה של CDFM כדי להפיק חיזויי טווח קצר עבור צמיחת התוצר הרבעונית של ישראל. תוצאות הניסוי האמפירי שלנו לחיזוי בתנאי זמן אמת מראות כי המודל מסוגל להניב חיזויי טווח קצר באותה רמת הדיוק של החיזוי המשלב שיפוטיות (Bridge). בנוסף, הראינו כי השימוש בפרוקסי עבור צמיחת התוצר החודשית הבלתי נצפית יכול להקטין את אובדן המידע הנובע מהורדת ממד וכך לשפר את הדיוק של התחזית. הניסוי של חיזוי טווח קצר בזמן אמת, מבליט את חשיבות השימוש באינדיקטורים מבוססי סקר המתפרסמים בפיגור קטן יחסית, למטרות השלמת הקצה של נתונים מסורתיים. מנגד, הרחבה ישירה של סט הנתונים באמצעות סקרים מובילה דווקא לירידה בטיב החיזוי. כמו כן, בדומה ל-Brave et al. (2019), יישמנו את המסגרת של CDFM כדי להפיק מדד חודשי לפעילות כלכלית ריאלית. נמצא כי ביחס לכלים הקיימים היום, המדד מסוגל לתת תמונה מדויקת יותר של המצב הנוכחי של המשק ולהצביע על נקודות תפנית מוקדם יותר.

נספח א': תיאור הנתונים

לוח 4: רשימת משתנים חודשיים

משתנה מסביר	מקור נתונים	תחילת הנתונים	ניכוי עונתיות	מחירים קבועים	טרנספורמציה
סך הפדיון	מדד, למ"ס	2000-01	✓	ש"ח	log-dif
ייצור תעשייתי – סה"כ	מדד, למ"ס	2000-01	✓	ש"ח	log-dif
פדיון בענפי המסחר	מדד, למ"ס	2000-01	✓	ש"ח	log-dif
פדיון בענפי השירותים	מדד, למ"ס	2000-01	✓	ש"ח	log-dif
פדיון בענפי שירותים עסקיים ופרטיים (א)	מדד, למ"ס	2000-01	✓	ש"ח	log-dif
פדיון בענפי שירותי מזון ואירוח	מדד, למ"ס	2000-01	✓	ש"ח	log-dif
ערך המכירות של רשתות השיווק	מדד, למ"ס	2000-01	✓	ש"ח	log-dif
רכישות בכרטיסי אשראי (א)	למ"ס/שב"א	2002-01 / 2016-01	✓	ש"ח	log-dif
התחלות בנייה	למ"ס	2000-01	✓		log-dif
צריכת דלק	משרד האנרגיה, מינהל הדלק	2000-01	✓	ש"ח	log-dif
ייבוא מוצרי צריכה	למ"ס	2000-01	✓	\$	log-dif
ייבוא תשומות לייצור	למ"ס	2000-01	✓	\$	log-dif
ייבוא נכסי השקעה	למ"ס	2000-01	✓	\$	log-dif
ייצוא סחורות	למ"ס	2000-01	✓	\$	log-dif
ייצוא שירותים עסקיים ושירותי תיירות	למ"ס	2009-01	✓	\$	log-dif
ייצוא שירותים	מערכת ב"י	2000-01	✓	\$	log-dif
ייבוא שירותים	מערכת ב"י	2000-01	✓	\$	log-dif
תעסוקה, לא כולל נעדרים זמנית	למ"ס	2000-01	✓		log-dif
משרות שכיר במגזר הפרטי	למ"ס	2000-01	✓		log-dif
סה"כ תשלומי שכר (ריאלי) (ג)	למ"ס	2000-01	✓	ש"ח	log-dif
משרות פנויות	למ"ס	2009-06	✓		log-dif
מדד מנהלי הרכש של ישראל	בנק הפועלים	2000-01	✓		
מדד מנהלי הרכש של ארה"ב	בלומברג	2000-01			
צריכת חשמל (06:00-11:00), מותאם לתנאי מזג האוויר (ב)	מודל ב"י	2002-01	✓		log-dif
מסים : מע"מ נטו (ג)	משרד האוצר	2000-01	✓		log-dif
מסים : עקיפים (ג)	משרד האוצר	2000-01	✓		lag(1), log-dif
מסים : בריאות (ג)	ביטוח לאומי	2000-01	✓		log-dif
מדד אמון הצרכנים	בנק הפועלים	2002-02			
מחיר נפט Cushing	בלומברג	2000-01			log-dif
מדד מניות כללי ת"א	בורסת תל אביב	2000-01			log-dif

הערות: א – מנוכה במדד המחירים לצרכן; ב – מבוסס על נתונים שעהיים מקוונים מחברת חשמל לישראל ומהשירות המטאורולוגי הישראלי; ג – מותאם למדד המחירים לצרכן ולשינויי חקיקה.

לוח 5: ציר זמן לפרסום נתונים

משתנה מסביר / יום בחודש	7>	7-14	15 - 22	22<
סך הפדיון	-2	E (2)	E (2)	E(1)
ייצור תעשייתי – סה"כ	-2	E (2)	E (2)	E(1)
פדיון בענפי המסחר	-2	E (2)	E (2)	E(1)
פדיון בענפי השירותים	-2	E (2)	E (2)	E(1)
פדיון בענפי שירותים עסקיים ופרטיים	-2	E (2)	E (2)	E(1)
פדיון בענפי שירותי מזון ואירוח	-2	E (2)	E (2)	E(1)
ערך המכירות של רשתות השיווק	-1	E (1)	E (1)	E(1)
רכישות בכרטיסי אשראי	✓	✓	✓	✓
התחלות בנייה	-5	E (4)	E (4)	E(4)
צריכת דלק	-1	-1	-1	-1
ייבוא מוצרי צריכה	-1	-1	-1	✓
ייבוא תשומות ייצור	-1	-1	✓	✓
ייבוא נכסי השקעה	-1	-1	✓	✓
ייצוא סחורות	-1	-1	✓	✓
ייצוא שירותים עסקיים ושירותי תיירות	-2	-2	-2	E(2)
ייצוא שירותים	-1	-1	-1	✓
ייבוא שירותים	-1	-1	-1	✓
תעסוקה, לא כולל עובדים נעדרים	-1	E (1)	E (1)	E(1)
משרות שכיר במגזר הפרטי	-2	E (2)	E (2)	E(2)
סה"כ תשלומי שכר (ריאלי)	-2	E (2)	E (2)	E(2)
משרות פתוחות	-1	-1	-1	✓
מדד מנהלי הרכש של ישראל	-1	-1	-1	✓
מדד מנהלי הרכש של ארה"ב	-1	-1	✓	✓
צריכת חשמל (06:00-11:00), מותאם לתנאי מזג האוויר	✓	✓	✓	✓
מסים : מע"מ נטו	✓	✓	✓	✓
מסים : עקיפים	-1	✓	✓	✓
מדד אמון הצרכנים	-1	-1	✓	✓
מסים : בריאות	-1	✓	✓	✓
מחיר נפט Cushing	✓	✓	✓	✓
מדד מניות כללי ת"א	✓	✓	✓	✓

הערות: ✓ מציין זמינות מלאה של המנבא עבור חודש היעד. 1-, 2- וכו' – מציין תצפיות חסרות בפיגור של 1 או 2 חודשים, בהתאמה. E (1), E (2) מציין תצפיות חסרות של 1 או 2 חודשים עם השלמת קצה באמצעות BTS כמתואר בחלק 4.2.

נספח ב'

כאן אנו מציגים נוסחה כללית של ייצוג המצב-מרחב בשכיחות מעורבת מחלק 2. נניח כי $\Lambda_{yx} = [\Lambda_1, \Lambda_2, \dots, \Lambda_r]$ הוא מטריצת $(1 \times r)$ של מקדמי הגורמים. אנו מגדירים את המצב-מרחב באופן הבא:

משוואת מדידה

$$\begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ y_{t-4} \\ F_t \\ F_{t-1} \\ \vdots \\ F_{t-p+1} \end{pmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & \mathbf{0}_{(1 \times r)} & \mathbf{0}_{(1 \times r(p-1))} \\ \mathbf{0}_{(r \times 1)} & \mathbf{0}_{(r \times 1)} & \mathbf{0}_{(r \times 1)} & \mathbf{0}_{(r \times 1)} & \mathbf{0}_{(r \times 1)} & I_r & \mathbf{0}_{((r \times (p-1))} \end{bmatrix} \begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ y_{t-4} \\ F_t \\ F_{t-1} \\ \vdots \\ F_{t-p+1} \end{pmatrix} + \begin{pmatrix} 0 \\ \varepsilon_t^{pc} \end{pmatrix}$$

משוואת מצב

$$\begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ y_{t-4} \\ F_t \\ F_{t-1} \\ \vdots \\ F_{t-p+1} \end{pmatrix} = \begin{pmatrix} \mu \\ 0 \\ 0 \\ 0 \\ 0 \\ \mathbf{0}_{(r \times 1)} \\ \mathbf{0}_{(r \times 1)} \\ \vdots \\ \mathbf{0}_{(r \times 1)} \end{pmatrix} + \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ y_{t-4} \\ y_{t-5} \\ F_{t-1} \\ F_{t-2} \\ \vdots \\ F_{t-p} \end{pmatrix} + \begin{bmatrix} 1 & \Lambda_{yx} \\ \mathbf{0}_{(4 \times 1)} & \mathbf{0}_{(4 \times r)} \\ \mathbf{0}_{(r \times 1)} & I_r \end{bmatrix} \begin{pmatrix} \varepsilon_{y,t} \\ u_t \end{pmatrix}$$

לקלות ההצגה, תתי-רכיבים של מטריצת שינוי המרחב מוגדרים בנפרד.

$$T_{11} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad T_{12} = \begin{bmatrix} \Lambda_{yx}\Phi_1 & \Lambda_{yx}\Phi_2 & \cdots & \Lambda_{yx}\Phi_p \\ \mathbf{0}_{(4 \times r)} & \mathbf{0}_{(4 \times r)} & \cdots & \mathbf{0}_{(4 \times r)} \end{bmatrix}$$

$$T_{21} = \mathbf{0}_{(rp \times 5)}, \quad T_{22} = \begin{bmatrix} \Phi \\ I_{p-1} \otimes I_r & \mathbf{0}_{(r(p-1) \times r)} \end{bmatrix},$$

כאשר Φ הוא מטריצת $(r \times pr)$ של המקדמים האוטורגסיביים

$$\Phi = \Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_p$$

- Anesti, N., S. Hayes, A. Moreira, and J. Tasker (2017). Peering into the present: the Bank's approach to GDP nowcasting. *Bank of England Quarterly Bulletin* 57(2), 122-133.
- Banbura, M. and M. Modugno (2014). Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics* 29 (1), 133-160.
- Banbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013). Now-casting and the real-time data flow. In G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 2, pp. 195-237. Elsevier.
- Banbura, M. and M. Modugno (2014). Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics* 29 (1), 133-160.
- Bok, B., D. Caratelli, D. Giannone, A. M. Sbordone, and A. Tambalotti (2018). Macroeconomic nowcasting and forecasting with big data. *Annual Review of Economics* 10 (1), 615-643.
- Brauning, F. and S. J. Koopman (2014). Forecasting macroeconomic variables using collapsed dynamic factor analysis. *International Journal of Forecasting* 30, 572-584.
- Brave, S. A., R. A. Butters, and D. Kelley (2019). A new "big data" index of US economic activity. *Economic Perspectives, Federal Reserve Bank of Chicago* 43, 1-30.
- Chernis, T. and R. Sekkel (2017). A dynamic factor model for nowcasting Canadian GDP growth. *Empirical Economics* 53 (1), 217-234.
- Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13 (3), 253-263.
- Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford Statistical Science Series.
- Groen, J. J. and G. Kapetanios (2016). Revisiting useful approaches to data-rich macroeconomic forecasting. *Computational Statistics & Data Analysis* 100, 221-239.
- Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press.
- Harvey, A. C. (1990). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

- Harvey, D., S. Leybourne, and P. Newbold (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting* 13 (2), 281-291.
- Hepenstrick, C. and M. Marcellino (2019). Forecasting gross domestic product growth with large unbalanced data sets: The mixed frequency three-pass regression filter. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 182 (1), 69-99.
- Kelly, B. and S. Pruitt (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics* 186 (2), 294-316.
- Krief, T. (2011). Nowcasting model for GDP and its components. Bank of Israel Research Department Discussion Paper 2011.01.
- Mariano, R. S. and Y. Murasawa (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* 18, 427-443.
- Pichette, L. and L. Rennison (2011). Extracting information from the business outlook survey: A principal-component approach. *Bank of Canada Review* 2011 (Autumn), 2128.
- Roash, D. and T. Suhoy (2019). Sentiment Indicators Based on a Short Business Tendency Survey. Bank of Israel Discussion Paper 2019.11.
- Sayag, D., D. Ben-hur, and D. Pfeffermann (2021). Reducing revisions in hedonic house price indices by the use of nowcasts. *International Journal of Forecasting*.
- Stock, J. and M. Watson (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual* 4, 351-394.
- Stock, J. and M. Watson (1991). A probability model of the coincident economic indicators. *Leading Economic indicators: new approaches and forecasting records* 66.
- Stock, J. and M. Watson (1993). Business cycles, indicators and forecasting, Chapter A procedure for predicting recessions with leading indicators: econometric issues and recent experience. University of Chicago Press.
- Stock, J. and M. Watson (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20, 147-162.