

בחינת הגידול באבטלה במעבר לסקר כוח אדם חודשי

נועם כהן, לואיזה בורק ויצחק מקובקי

סקר כוח אדם הוא המקור העיקרי למידע על תכונות כוח העבודה בישראל, היקף האבטלה ומגמותיהם, והוא נערך באופן שוטף ע"י הלשכה המרכזית לסטטיסטיקה (למ"ס) מאז 1954. החל מינואר 2012 שונתה מתכונת סקר כוח אדם מסקר רבעוני לסקר חודשי, כך שעומד לרשות מקבלי ההחלטות וחוקרים מידע עדכני ומייצג על שוק העבודה בישראל ברמה חודשית. בסקר החודשי הוכנסו גם שינויים בגודל ובפרמטרים נוספים של המדגם, הועסקו סוקרים חדשים רבים, נעשה שימוש בשיטות אמידה חדשות והגדרות הסקר הותאמו למקובל ב-OECD. בפרט, הסקר מודד כעת את תכונות כלל כוח העבודה במקום את תכונות כוח העבודה האזרחי כפי שהיה נהוג בסקר הרבעוני. בעקבות המעבר נרשמה עלייה של כ-20% במספר הבלתי מועסקים בישראל, ושיעורם שעמד על 5.4% כפי שנמדד ברבע האחרון של 2011 רשם עלייה משמעותית לשיעור של כ-6.7% בתחילת 2012. לאומדן זה חשיבות רבה בקרב מקבלי ההחלטות בממשלה לצורך קביעת מדיניות כלכלית. חשיבות זו באה לידי ביטוי גם בכתבות רבות שפורסמו בעיתונות הכלכלית והעלו תהיות לגבי הסיבות לעלייה במספר הבלתי מועסקים, אך כללו אי דיוקים רבים. מסתבר כי ההבדל התבטא בעיקר בגידול חד במספר הבלתי מועסקים במגזר הערבי לעומת גידול מינורי במגזר היהודי. במאמר נטען כי ממצא זה תומך בהשערה כי אין מדובר בבעיה במערכות הדגימה, האמידה או עיבוד הנתונים, אלא מצביע על בעיה הקשורה לתהליך איסוף הנתונים, תהליך שקיימת בו הפרדה ברורה בין שני המגזרים. במאמר זה אנו מציגים מחקר מדעי על ההבדלים והעדר ההבדלים בין ממצאי הסקר החודשי והרבעוני, ונתמקד בנושא האבטלה.

א. מבוא

אוכלוסיית סקר כוח אדם הנערך ע"י הלשכה המרכזית לסטטיסטיקה (להלן: הלשכה) כוללת את התושבים הקבועים של מדינת ישראל בגיל 15 ומעלה, ביניהם תיירים ותושבים ארעיים השוהים בישראל יותר משנה ברציפות. הסקר מאפשר מעקב אחר ההתפתחות בכוח העבודה, גודלו ותכונותיו, היקף הבלתי מועסקים ועוד.

שיטות הדגימה, האמידה והחקירה של סקרי כוח אדם, שהחלו בשנת 1954, השתנו במשך השנים. החל מתחילת שנת 2012 שונתה מתכונת סקר כוח אדם מסקר רבעוני לסקר חודשי ומרואיינים בכל חודש באופן שוטף כ-21,500 בני 15 ומעלה. הלשכה ערכה בסקר באותה הזדמנות שיפורים נוספים. בין השינויים שהוכנסו ראוי לציין הרחבת הגדרות לכלל כוח העבודה כפי שמקובל בסקרי ה-OECD במקום לכוון העבודה האזרחי בלבד, שימוש בשיטות אמידה יעילות יותר, הגדלת המדגם השנתי ושינויים בעבודת השדה. התוצאה היא מדגם יעיל ומייצג ברמה

החודשית של כלל האוכלוסייה בישראל. לעומת זאת, לא הוכנסו שינויים בשאלון הסקר ובתהליך איסוף הנתונים.

מהשוואת נתוני הסקר החדש לנתונים שנאספו במתכונת הסקר הישן התברר כי בעקבות המעבר גדל שיעור הבלתי מועסקים במשק באופן משמעותי. הפרשים מסוימים בין תוצאות של שני סקרים האומדים אותם פרמטרים באוכלוסייה הם עניין שבשגרה. יש לזכור כי בכל סקר מדגמי קיים גורם של אקראיות המבוסס בעיקר על כך שמשיבי הסקר עשויים להיות שונים קמעא בתכונותיהם מאלה שלא השתתפו בסקר, אך גם על גורמים אזוריים שתלויים כמעט בכל פרמטר אפשרי של הסקר. למעשה הסיכוי ששני סקרים יתנו את אותן התוצאות בדיוק הוא בדרך כלל זניח. לשם כך עושה הסטטיסטיקה שימוש בטעות הדגימה כדי להכריע האם שתי תוצאות, שני אומדנים שונים לאותו פרמטר, הן למעשה זהות או לא. עם זאת, ההשוואה בין נתוני שני הסקרים מראה כי שיטת המדידה החדשה הביאה עבור חלק מהפרמטרים הנחקרים בסקר לתוצאות שונות באופן מהותי. בעיקר, כאמור, בלט השינוי באומדן הבלתי מועסקים במשק שגדל מ-5.4% כפי שנמדד ברבעון האחרון של 2011 לשיעור של כ-6.7% בתחילת 2012. בנוסף, בלט גם גידול בשיעור השייכות לכוח העבודה בקרב יהודים וערבים.

נקדים את המאוחר כדי ליצור שיח ממוקד. בקרב יהודים, גברים ונשים, התמונה המתקבלת מניתוח הנתונים במאמר זה היא כי לא ניתן להסיק בכלים מדעיים שהסקר החדש הביא לשינוי משמעותי באומדן שיעור הבלתי מועסקים לעומת הסקר הישן. בקרב ערבים, לעומת זאת, הייתה קפיצה מובהקת באומדן שיעור הבלתי מועסקים: בערך פי שניים אצל הגברים ובערך פי שלושה אצל הנשים. היות שהערבים מהווים כ-10% מכוח העבודה במשק, עלה עקב כך גם אומדן שיעור הבלתי מועסקים הכללי במשק. בהתאם, גדל גם אומדן השייכים לכוח העבודה, שכן זה מחושב כסכום של המועסקים והבלתי מועסקים. כמו כן, סיבה נוספת לגידול בשיעור השייכות לכוח העבודה שתידון בהמשך היא הכללת משרתי הצבא במועסקים, הכללה המגדילה באופן מלאכותי את שיעור השייכות לכוח העבודה במשק. תחת שני גורמים אלה, החלטנו להתמקד באומדן הבלתי מועסקים בלבד ולא להתייחס לשינויים הנגזרים ממנו על שיעור השייכות לכוח העבודה.

הסקר החדש של הלשכה נחשב למדויק ולאמין יותר מן הסקר הישן, לכן נראה ששיעור הבלתי מועסקים האמיתי בקרב האוכלוסייה הערבית אכן גבוה ממה שנאמד בעבר. עם זאת, נסייג ונאמר כי גם אם הוכחנו כי שני אומדנים שונים באופן מובהק, עדין לא קבענו בוודאות מי מהם הוא הנכון. השינוי, מובן מאליו, הפתיע רבים במערכת הכלכלית והחברתית. רבים הציעו הסברים שחלקם מתקבלים על דעתנו אך חלקם גם נובעים מחוסר ידע או מאי הבנה או שלא בוצעו בכלים מדעיים. בלמ"ס בוצעו בדיקות אנליטיות רבות ומגוונות מכל היבט הגיוני כדי להסביר את ההפרש באומדני שיעור האבטלה בין הסקרים הרבעוני והחודשי, לשם את האצבע על הסיבה הנכונה ולטפל בגורמים להיווצרותה. במאמר נציג את תוצאות הבדיקות שנערכו.

בפרקים הבאים נתאר בקצרה את השינויים המרכזיים שבוצעו בסקר החודשי, ולאחר מכן נציג את תוצאות הבדיקות שנערכו. ההשוואות המגוונות שערכנו לבחינת ההבדלים בין אומדני הסקרים נעשו כל אחת בפני עצמה ולא בצורה של בדיקת השערות מרובות. ראשית מפני שהמאורעות בדרך כלל לא היו על אותו מרחב הסתברות, ושנית מפני שזה היה דורש אמידה של מספר רב של פרמטרים של שונות ושונות משותפת שאין לנו דרך לאומדם בקירוב סביר. פרק ב מוקדש להגדרות. בפרק ג נתאר בתמציתיות את השינויים שבוצעו בדגימה, באמידה ובשיטות החקירה של הסקר, ובפרק ד נבחן אם היו להם השלכות על האומדנים. בפרק ה נציג השוואות שנערכו בין אומדני הסקר החודשי והרבעוני, ונעלה השערות לגבי הסיבות להבדלים שנמצאו. פרק ו יציג את מסקנותינו.

ב. הגדרות

כוח העבודה השבועי: בני 15 ומעלה שהיו מועסקים או בלתי מועסקים בשבוע הקובע, לפי ההגדרות המפורטות להלן: **השבוע הקובע:** השבוע המסתיים ביום שבת שלפני בוא הסוקר למשק הבית. **מועסקים:** אנשים שעבדו, לפחות שעה אחת בשבוע הקובע, בעבודה כלשהי עבור תמורה. הקבוצה מורכבת משלוש תת-קבוצות: עבדו מלא, עבדו חלקית או שנעדרו זמנית מעבודתם. **בלתי מועסקים:** אנשים שלא עבדו כלל וחזפו עבודה באופן פעיל בארבעת השבועות האחרונים, והיו יכולים להתחיל לעבוד בשבוע הקובע לו הוצעה להם עבודה. בעיתונות ובשיח הציבורי מקובל להשתמש במונח אבטלה, אך כדי לשמור על אחידות עם פרסומי הלמ"ס אנו נשתמש במושג "בלתי מועסקים" במאמר זה מעתה ואילך. **אינם בכוח העבודה השבועי:** כל בני 15 ומעלה שלא היו מועסקים או בלתי מועסקים בשבוע הקובע.

ג. תיאור המתודולוגיה של סקר כוח אדם והשינויים המרכזיים בסקר החודשי

ג.1. פנלים ודגם החקירות בסקר

מקובל בלשכות סטטיסטיקה מובילות בעולם, סקר כוח אדם הוא סקר פנלים. מדגם הדירות המוצא מדי שנה מחולק לקבוצות זרות ומשלימות הנקראות פנלים. הפנלים נכנסים לחקירה בתקופות עוקבות, פנל יחיד מדי תקופה, כשכל פנל נחקר במשך מספר תקופות, לעתים עם הפסקה בין החקירות. הדירות שנדגמות משובצות לפנלים השונים באופן כזה שבכל פנל יתקבל בתוחלת אותו מספר של דירות.

השוני בין שיטות רוטציה שונות של הפנלים הוא באחוז החפיפה (שמתורגם לאחוז משקי בית משותפים) בין החודשים הסמוכים, בין חודשים מקבילים מרבעונים סמוכים ובין חודשים מקבילים משנים סמוכות. מבנה החקירה של הפנלים משפיע על טעויות הדגימה של האומדים השונים. ככל שהחפיפה של פנלים בין שתי תקופות זמן גדולה יותר כך טעות הדגימה של אומדי שינוי בין תקופות (למשל שינוי בבלתי מועסקים בין חודש t לחודש $t-1$) קטנה יותר בהשוואה למדגמים זרים באותו הגודל. לגבי אומדי רמה (למשל אומד סה"כ לרבע) ההפך הוא הנכון וחפיפה חודשית רבה תביא לטעויות דגימה גבוהות בהשוואה למדגמים באותו גודל ללא חפיפה.

בסקר הרבעוני המדגם שהוצא מדי שנה חולק לארבעה פנלים שנכנסו לחקירה בארבעה רבעים עוקבים, פנל יחיד מדי רבע. כל פנל נחקר ארבע פעמים לפי הדגם: חקירה ברבע הראשון (חקירה 1), ברבע העוקב (חקירה 2), הפסקה בשני הרבעים העוקבים וחקירות נוספות בשני הרבעים הבאים (חקירות 3 ו-4 בהתאמה), המקבילים בהפרש של שנה לחקירות 1 ו-2 (בהתאמה). מאחר שהסקר החדש הוא חודשי, המדגם השנתי מחולק לשנים עשר פנלים. מבנה הרוטציה החודשי של הפנלים,

1 להגדרה מדויקת של המועסקים והבלתי מועסקים עם התייחסות לשהים במוסדות, בצבא או במילואים, אנשי קיבוץ, מרכזי קליטה, משך התעסוקה השבועית וכו' ראה באתר הלמ"ס: http://www.cbs.gov.il/publications11/1460/pdf/intro04_h.pdf

המוכר בסימון 4-8-4, נבחר לפי צורכי האמידה תחת קביעת סדר העדיפויות שמקטין את טעות הדגימה של אומדני שינוי בין שני חודשים עוקבים ושתי שנים עוקבות. במבנה זה נעשה שימוש גם בסקר כוח אדם בארה"ב. בהתאם למבנה רוטציה זה תתבצענה שמונה החקירות לכל פנל באופן הבא: חקירה בחודש הכניסה ובשלושת החודשים העוקבים לו (חקירות 1-4), הפסקה של שמונה חודשים וארבע חקירות רצופות נוספות בחודשים הבאים (חקירות 5-8). לוחות 6 ו-7 בנספח א מתארים את מבנה החקירות בסקרים החודשי והרבעוני.

2.2. דגימה

רוב אוכלוסיית הסקר מכוסה במדגמים שוטפים המוצאים בדרך כלל אחת לשנה. בשל הקושי לפקוד באופן שוטף קבוצות מסוימות השייכות לאוכלוסיית הסקר, חלק קטן של האוכלוסייה מכוסה במדגמים קבועים שתכונותיהם אינן משתנות לאורך הזמן, ואשר הוצאו ממפקד האוכלוסין האחרון שהתקיים.²

בשלב תכנון המדגם השנתי נקבע שבר הדגימה הארצי לאותה שנה, המבטא את היחס שבין גודל המדגם המתוכנן לבין גודל האוכלוסייה לאותה שנה. רוב יחידות הדגימה הסופיות הן דירות אשר נדגמות בשיטה דו-שלבית: בשלב הראשון דוגמים יישובים. בשלב השני דוגמים דירות ביישובי המדגם, כך שהסתברות הדגימה הסופית לכל דירה תהיה שווה לשבר הדגימה הארצי שנקבע לאותה שנה. מדגם היישובים בסקר מוצא מרשימת יישובים (הנקראת מסגרת לדגימת יישובים), והיישובים משובצים לשתי קבוצות עיקריות בהתאם לגודל היישוב (נקבע על סמך הנתונים של בני ה-15 ומעלה). בקבוצה הראשונה נכללים כל היישובים (הגדולים) הנדגמים בוודאות מדי שנה. בקבוצה השנייה נכללים יישובים עירוניים קטנים וכל היישובים הלא עירוניים. יישובים אלה משובצים לקבוצות הומוגניות הנקראות שכבות דגימה, ובכל שכבת דגימה מוצא מדגם באופן מקרי שיטתי יחסית לגודל היישוב, כך שהסתברות היישוב להידגם היא יחסית לגודלו (לכן הם נקראים יישובים הסתברותיים). נציין כי כל יישוב ודאי מהווה שכבת דגימה בפני עצמו. המדגם בכל יישוב משופץ לפנל מסוים בהתאם לשכבת הדגימה וכך ייווצר איזון בין הפנלים מבחינת מספר היישובים ותכונותיהם.

ככלל, שיטת הדגימה בסקר החודשי נותרה כפי שהייתה בסקר הרבעוני. עם זאת, המדגם גדל, ומספר הדירות החדשות הנדגמות מדי שנה גדול פי 1.5. גידול זה מאפשר פיזור גדול יותר של המדגם על פני האזורים הגיאוגרפיים השונים והגדלת מספר היישובים הנדגמים ברבעון מכ-240 יישובים לרבעון בסקר הרבעוני לכ-310 יישובים בסקר החודשי. כמובן שהגדלת המדגם נעשתה באופן כזה שהמדגם ימשיך לייצג את אוכלוסיית הסקר כפי שהיה בסקר הרבעוני. במילים אחרות, האומדים המבוססים על 310 יישובי הסקר החודשי הם חסרי הטיה בדיוק כפי שהיו האומדים המבוססים על 240 היישובים בסקר הרבעוני.

2 בקטגוריה זו נכללים מוסדות מסוימים שאינם נחקרים באופן שוטף בסקר, שבטי בדווים בדרום והגרים מחוץ ליישובים. ראה סעיף 3.4 בפרסום המיוחד של סקר כוח אדם באתר ה"למ"ס: http://www.cbs.gov.il/publications13/1504/pdf/intro02_h.pdf

ג.3. אמידה ומערכת מקדמי הניפוח

הנתונים של הסקר מתייחסים לנחקרים בסקר בלבד. על מנת לעבור לאומדנים של כלל האוכלוסייה, עלינו לחשב מקדם ניפוח לכל פרט משיב במדגם. מקדם הניפוח הוא למעשה מספר האנשים באוכלוסייה שאותו פרט במדגם מייצג. בעזרת מקדמי ניפוח ניתן לחשב אומדנים עבור כלל האוכלוסייה המבוססים על משיבים במדגם.

מקדמי הניפוח נקבעים בתהליך סדרתי שבסיכומו מתקבלת התאמה מלאה בין התפלגות הנפשות המנופחות לבין האומדנים הדמוגרפיים השוטפים של הלשכה עבור קבוצות אוכלוסייה (קבוצות ניפוח) שנקבעו מראש. החלוקה לקבוצות הניפוח היא בהתאם לחלוקות גיאוגרפיות ולפי תכונות גיל ומין של האוכלוסייה.

עבור כל נפש משיבה במדגם השוטף מחושב מקדם הניפוח הראשוני בהתאם לשייכותה לקבוצת הניפוח, וזאת באמצעות חלוקת האומדן הדמוגרפי השוטף, בסך גודל המדגם שהתקבל באותה קבוצת הניפוח. כל הנפשות באותה קבוצת ניפוח מקבלות אותו מקדם ניפוח ראשוני. לאחר מכן לכל הנפשות במשקי הבית במדגם השוטף מתבצע, כאמור, תהליך חישוב סדרתי לחישוב מקדמי ניפוח סופיים שיקיים את האילוצים הבאים: (א) מקדמי הניפוח הסופיים קרובים ככל האפשר למקדמי הניפוח הראשוניים; (ב) סך הנפשות במדגם המנופח לפי מקדמי הניפוח הסופיים בכל קבוצת ניפוח בנפרד שווה לסך הנפשות באוכלוסייה הידוע על פי האומדן הדמוגרפי השוטף; (ג) כל הנפשות השייכות לאותו משק בית מקבלות את אותו מקדם ניפוח, כלומר מקדם ניפוח סופי אחיד למשק הבית. חישוב כזה יוצר הרמוניזציה בין נתונים הידועים בוודאות (או כמעט בוודאות) באוכלוסייה, לבין אומדנים שיתקבלו עבורם מהמדגם. מרבית הלשכות הסטטיסטיות המובילות בעולם כיום משתמשות בתהליכים דומים. להסבר מפורט ומסכם של שיטות הניפוח ראה (1992) Deville and Särndal.

תהליך ניפוח מסוג זה מכונה לעתים תהליך כיוול.

בסקר כוח אדם חודשי הוחלט לשלב גם מעבר לשיטת אמידה חדשנית. שיטת אמידה זו פורסמה במאמרם של Fuller and Rao (2001) וכונתה בשם Regression Composite Estimation. לאומדן המתקבל באמצעות שיטה זו נקרא בשם אומד משולב. באמידה המשולבת נתוני המשיבים בחודש מסוים מכילים, הן לסך אוכלוסייה הידוע לפי חתכים של אזור גיאוגרפי, מין וקבוצות גיל (כפי שנעשה עד כה בסקר הרבעוני) והן כיוול נוסף המבוסס על האומדנים של נתוני כוח העבודה של החודש הקודם. היינו בעוד שהאומד שהיה בשימוש עד כה מחושב על סמך נתוני ההווה (זמן t) בלבד ואינו מנצל בנקודת הזמן הנוכחית את האינפורמציה שקיימת בעבר לגבי הפרטים הנחקרים בסקר, האומד המשולב מנצל גם אינפורמציה רלוונטית זו מהעבר. לדוגמא: פרט שהיה מועסק בחודש קודם, בסיכוי גבוה מאוד יישאר מועסק גם בחודש הנוכחי. ראה הסבר על כך בנספח ב. השינויים העיקריים בין הסקר החודשי לבין הסקר הרבעוני מוצגים בלוח 1.

לוח 1: שינויים עיקריים בסקר כוח אדם (סקר חודשי בהשוואה לסקר רבעוני)

נושא	סקר רבעוני	סקר חודשי
הגדרות תכונות כוח העבודה	תכונות כוח העבודה האזרחי	תכונות כלל כוח העבודה (כולל אלה שבשירות צבאי חובה או קבע)
דגם החקירה	כל דירה נחקרת 4 פעמים: 2 רבעים סמוכים, הפסקה של 2 רבעים, ועוד 2 רבעים סמוכים	כל דירה נחקרת 8 פעמים: 4 חודשים סמוכים, הפסקה של 8 חודשים, ועוד 4 חודשים סמוכים
גודל מדגם שוטף	כ-22,500 נפשות ברבע באופן שוטף	כ-21,500 נפשות בחודש באופן שוטף
מדגמים קבועים	תת-מדגם של כ-3,900 נפשות מתוך הקובץ המפקדי (נתונים חברתיים-חברתיים-כלכליים) של מפקד האוכלוסין והדירור 1995	כ-26,000 נפשות מתוך הקובץ המפקדי (נתונים חברתיים-כלכליים) של מפקד האוכלוסין 2008
מסגרת דגימה עיקרית	קובצי הארנונה של העיריות והרשויות המקומיות	מרשם מבנים ודירות לאחר עיגון הדירות לאזורים סטטיסטיים
שימוש במידע מרחבי להוצאת מדגם הדירות העיקרי	מיון הקבצים לפי המידע הגיאוגרפי המצוי בקובצי ארנונה	שילוב של האזור הסטטיסטי המתקבל מתהליך עיגון הדירות בנוסף למידע הגיאוגרפי בקובצי הארנונה
מספר יישובים במדגם	כ-240 יישובים ברבע	כ-310 יישובים בחודש
תקופת החקירה ליחידת הדגימה	בשבוע הקובע ועוד 3 שבועות	בשבוע הקובע ועוד שבוע אחד
שיטת האמידה	שכבות ניפוח לפי גיאוגרפיה, גיל ומין ושכבה נפרדת לעולים. אין שימוש במידע על תכונות כוח עבודה מרבעים קודמים בסקר	שימוש באומד משולב המבוסס גם על מידע לגבי תכונות כוח עבודה מחודש קודם בסקר; שינוי בשכבות ניפוח (כמו ביטול שכבת ניפוח נפרדת לעולים)

ד. בחינת ההשלכות של השינויים המרכזיים על אומדני הסקר

המטרה העיקרית של השינויים שבוצעו בסקר היא הקטנת טעות הדגימה של האומדים, אם על ידי הגדלת המדגם ואם על ידי ייעול שיטת האמידה. לשינויים כה משמעותיים בשדרותיו המרכזיות של הסקר – שינוי בהגדרות אוכלוסיית הסקר כגון הכללת הצבא בכוח העבודה, הגדלת המדגם שמאפשרת הגדלת מספר היישובים הנכללים כל חודש במדגם, והשינויים שהוכנסו בשיטת האמידה – אם לא בוצעו בזהירות יכולות להיות השלכות בלתי צפויות. גם התכיפות הגדולה יותר של חקירות בסקר החודשי, במידה שהיא גורמת להגדלת אי ההשבה בסקר, וכל עוד זו אינה מקרית (לכל נדגם באותו הסיכוי – Missing at Random), עלולה להביא להטיה באומדנים. בפרק זה נבחן את ההשפעות של השינויים המרכזיים בסקר.

ד.1. רוטציית הפנלים ותכיפות החקירות

כדי לקבל מדגם מייצג של אוכלוסיית הסקר יש לדאוג לאיזון המדגם, כך שבכל יחידת זמן הפנלים המשתתפים ייצגו יחדיו היטב את כלל אוכלוסיית הסקר כל חודש, כפי שהיו ארבעת הפנלים בכל רבעון של הסקר הרבעוני. כדי לשמור על אותו האיזון בסקר החודשי לאורך כל חודשי האמידה, מספיק לדאוג כי כל שני פנלים הנכנסים לחקירה במרווח של ארבעה חודשים (לדוגמה הפנלים שנחקרו לראשונה ביוני 2012, אוקטובר 2012, פברואר 2013 וכו' יכסו את אותם אזורים גיאוגרפיים, ללא הכרח שיכסו את כלל האזורים במדינת ישראל. בצורה כזו מתקבלות 4 קבוצות פנלים. כדי לעמוד בדרישת האיזון לעיל, מספיק כי כל קבוצה תכסה רבע אחר של האזורים וכל חודש אמידה יהיה מורכב משני פנלים מכל קבוצה, בהתאם לדגם חקירה 4-8-4.

משך השנות כבלתי מועסק בישראל הוא קצר. החקירה של כל פנל בסקר החודשי מתפרשת על פני 16 חודשים ובה כל נדגם נחקר 8 פעמים, בעוד שבסקר הרבעוני נחקר כל נדגם רק 4 פעמים באותה התקופה. תכיפות כה גדולה של חקירות מאפשרת לזהות יותר נדגמים שהיו בלתי מועסקים רק תקופה קצרה, נניח שבוע אחד בלבד, במהלך אותה תקופה. במילים אחרות, גדל הסיכוי לזהות עבור נדגם מסוים לפחות מופע אחד שבו הוא היה בלתי מועסק. הצהרה זו אינה עומדת בסתירה לעובדה שהשפעה זו של תכיפות החקירות רק מקטינה את טעות הדגימה של אומד הבלתי מועסקים ולא מטה את תוחלתו. הסיבה לכך היא שכאשר זיהינו בלתי מועסק בסקר הרבעוני, ייחסנו את היותו בלתי מועסק לכל הרבעון. בסקר החודשי לעומת זאת, ייחסנו את היותו בלתי מועסק רק עבור אותו החודש וייתכן שבחודשים האחרים ברבעון השיב אחרת. נסביר זאת בעזרת החישוב להלן. נניח כי נדגם מסוים נחקר בסקר החודשי ובסקר הרבעוני ונחשב את תרומתו לאומדן הבלתי מועסקים בשני המקרים. נניח, בלי הגבלת הכלליות, כי אותו נדגם היה בלתי מועסק (k_1, k_2, k_3) שבועות בחודשים (m_1, m_2, m_3) שכללו $(4, 4, 5)$ שבועות בהתאמה באותו הרבעון. היות שהסקר הרבעוני מגיע לבית הנדגם בשבוע אקראי במהלך הרבעון, הנדגם ייחקר בשבוע שבו היה בלתי מועסק בסיכוי $13 / (k_1 + k_2 + k_3)$, וזו תהא גם תרומתו לתוחלת אומדן הבלתי מועסקים.

בחישובי תוחלת לתקופה מסוימת מחשבים בלמ"ס את מספר הבלתי מועסקים שנמצאו לאורך אותה התקופה מחולק באורכה. יחידת ההתייחסות הבסיסית במקרה זה היא שבוע ולכן נחשב את תוחלת מספר השבועות שהנדגם היה בלתי מועסק ברבעון ונחלק ב-13, מספר השבועות ברבעון. היות שהסקר החודשי מגיע בשבוע אקראי במהלך החודש, הוא ימצא את הנדגם בלתי מועסק בסיכוי $(\frac{k_1}{4}, \frac{k_2}{4}, \frac{k_3}{5})$ בחודשים (m_1, m_2, m_3) בהתאמה. בכל חודש שבו נחקר הנדגם בשבוע שבו היה בלתי מועסק, הוא ייחשב כבלתי מועסק בכל אותו החודש, אחרת ייחשב כמועסק בכל אותו החודש. לכן, תוחלת מספר השבועות שבו היה בלתי מועסק בחודש הראשון תחושב על פי 4 שבועות בסיכוי $\frac{k_1}{4}$ ועוד אפס שבועות בסיכוי המשלים. תרומתו לתוחלת אומדן הבלתי מועסקים ברבעון תהא לכן $\frac{k_1}{13} = \frac{k_1}{13} \cdot \frac{4-k_1}{4} + 0 \cdot \frac{4-k_1}{13}$ עבור m_1 , באותה צורה עבור יתר החודשים, ובסך הכל $\frac{k_1}{13} + \frac{k_2}{13} + \frac{k_3}{13}$, בזהה לסקר הרבעוני.

ד.2. דגימה ביישובי הפריפריה

אספקט אחר הנוגע לדגימה הוא יישובי הפריפריה. הודות להגדלת המדגם גדל גם מספר היישובים הנדגמים כל חודש. אם בסקר הרבעוני נחקרו כ-240 יישובים ברבעון, אזי בסקר החודשי נחקרים

כ-310 יישובים כל חודש. ההפרש רובו ככולו ביישובים הפריפריאליים, כך שלכאורה ניתן לחשוב כי הם מקבלים ייצוג גדול יותר בסקר החודשי, אך לא כך הדבר. בסקר החודשי, כמו בסקר הרבעוני, מסגרת היישובים נשארה אותה מסגרת. בכל שנה, לכל דירה מכל יישוב יש אותו סיכוי להידגם, בין אם דירה זו שייכת ליישוב פריפריאלי ובין אם לאו, ובכלל זה גם יישובים שלא נדגמו בסקר הרבעוני בעבר עקב אקראיות הדגימה של היישובים ההסתברותיים. עודף היישובים הפריפריאליים בסקר החודשי רק מאפשר פיזור רב יותר בין הדירות הנכללות בהם ואינו מגדיל את החלק היחסי שלהן בסקר, כפי שנראה בהסבר להלן. מספר כתבות שהתפרסמו בעיתונות הכלכלית האלקטרונית והכתובה אף תלו בכך את הגידול בבלתי מועסקים. למשל, "האבטלה עלתה ל-6.9% במאוס; הפריפריה שינתה התוצאות"³, או "הזינוק באבטלה, בשיעור של 1%, שנרשם [...] הפתיע רבים [...] התברר, כי הסיבה היא הרחבת הסקר ל-100 יישובים בפריפריה"⁴.

היישובים הפריפריאליים הם יישובים הסתברותיים (ראה ג.2). כדי להכליל יחדיו במדגם יישובים הסתברותיים וכאלה שאינם הסתברותיים ו/או להגדיל את מספרם בכפוף להבטחה על שמירה של ייצוג נאות של כל הדירות באוכלוסייה, מספר הדירות הנדגמות בכל יישוב צריך להיות תלוי בסיכוי של אותו יישוב להיכלל במדגם. במילים אחרות, בגודל היישוב. תהי i דירה כלשהי במסגרת הדגימה (בכל הארץ), נסמן ב- π_i את ההסתברות של דירה i להיכלל במדגם. יהי j יישוב כלשהו. יהי S המדגם החודשי. מתקיים:

$$\begin{aligned} \pi_i &= P(Dira_i \in S) = P(Dira_i \in S | Yishuv_j \in S) \times P(Yishuv_j \in S) \\ &\equiv \pi_{ij \in S} \times p_j \end{aligned}$$

כאשר $p_j = 1$ עבור יישוב ודאי, ו- $p_j < 1$ עבור יישוב הסתברותי. אם כך מתקבל הקשר $\pi_{ij \in S} = \pi_i / p_j$ שמשמעותו היא שביישוב הסתברותי סיכויו להיכלל במדגם הוא p_j , אנו נדגום דירה בהסתברות π_i / p_j . בדרך זו אנו מבטיחים שלכל דירה בארץ יש אותו סיכוי להיכלל במדגם, ללא תלות במספר היישובים ההסתברותיים שנכלול במדגם.

3.3. אמידה

בסקר כוח אדם חודשי הוחלט, כאמור, לעבור לשיטת אמידה חדשנית. היתרון בשיטה החדשה על פני השיטה ששימשה את סקר כוח האדם הרבעוני נעוץ בניצול אינפורמציה מחודש קודם בחישוב אומדן החודש הנוכחי. בדרך זו מקבלים יעילות גדולה יותר של האומדן, הווה אומר שטעות הדגימה של האומדן המשולב, עבור מדגם באותו הגודל, היא בדרך כלל קטנה יותר מזו של האומדן בסקר הרבעוני עבור הפרמטרים המסוימים שבהם יש לנו עניין מיוחד (ראה גם Cochran, 1977, פרק 12.10). אפיון השיטה בלמ"ס מייעל את אומדי הפרמטרים של תכונות כוח עבודה נבחרות – מועסקים, בלתי מועסקים, לא שייכים לכוח עבודה, עבור גברים ונשים באזורים הגיאוגרפיים השונים הכוללים חלוקה ליישובים יהודיים וערביים. שיטת האמידה החדשה מורכבת יותר מזו שהייתה נהוגה בסקר הרבעוני. היא מנצלת מידע מחודש קודם עבור כל נדגם שקיים לו מידע כזה, ובשל כך בעלת אופי

3 כתבת TheMarker (מקבוצת "הארץ") מיום 30 באפריל 2012.
4 כתבת Ynet (מקבוצת "ידיעות אחרונות") מיום 9 באוקטובר 2012.

רקורסיבי שיכול לעורר חשש שמא תיתכן סיטואציה שבה ייווצר שוני שיטתי (סיסטמטי, קבוע) בין האומדן למציאות. חשש זה בוסס על תיעוד של אומדני תעסוקה בקנדה, ופורסם ב-Gambino, Kennedy and Singh (2001). מאמר שנכתב עי חוקרים מהלשכה הסטטיסטית הקנדית. בכל אופן נציין כי המקרה שתואר בו השוני השיטתי במאמר התרחש תחת תנאים יוצאי דופן אחרים (ראה המאמר לעיל סעיף 3) ועבור אומדן משולב מסוג אחר. בנוסף, רצינו לשלול חשש כי הטיה קטנה באחד האומדנים יכולה להוביל לתגובת שרשרת שתיצור הטיה מצטברת משמעותית לאחר מספר חודשים. גם תיעוד של מצב כזה קיים במאמר של חוקר מאוסטרליה (ראה Bell, 2001, סעיף 6.1). העדויות הנ"ל אף שהן קיימות הן בגדר מקרים חריגים בספרות המקצועית, וכנראה גם היו בכפוף למקרי קצה היוצאים מן הכלל. בכל אופן, לאור החשיבות של פרסומי הלשכה, וכדי לשלול סיבה זו, אמדנו פרמטרים רבים של הסקר גם בשיטת האמידה שהייתה נהוגה בסקר רבעוני (השיטה הישירה) וערכנו השוואה בין האומדנים שהתקבלו בשיטת השיטות.

את ההשוואה בין שתי שיטות האמידה ערכנו במשך שנים עשר חודשים, מאוקטובר 2011 ועד ספטמבר 2012, עבור האומדנים השייכים ללוחות 1-4 בפרסום החודשי הרשמי של הסקר, כלומר, סה"כ 12 תת-לוחות בחתכים שונים ו-144 אומדנים לפרמטרים שונים. ההתייחסות בסעיף זה לפרמטרים אחרים מלבד אחוז הבלתי מועסקים נועדה רק לצורך השוואת שיטות האמידה ולא לצורך הדיון בהן. בצירוף 1 להלן מוצגות סדרות עתיות של אומדן אחד מכל תת לוח, כך שמתקבל ייצוג של אומדנים בחתכים שונים ובסדר גודל שונים. בצירוף, הקו הבהיר מייצג את סדרת האומדנים לפי השיטה הישירה, והכהה לפי השיטה המשולבת.

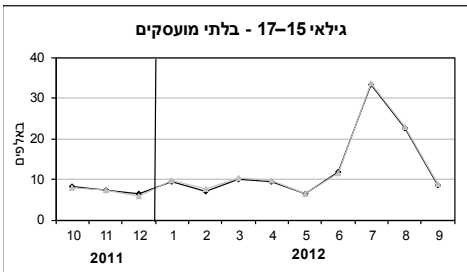
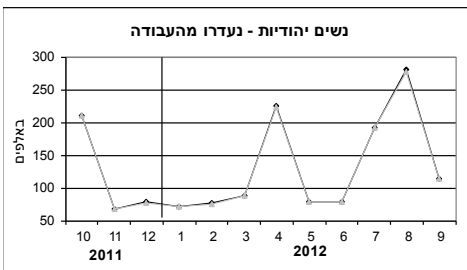
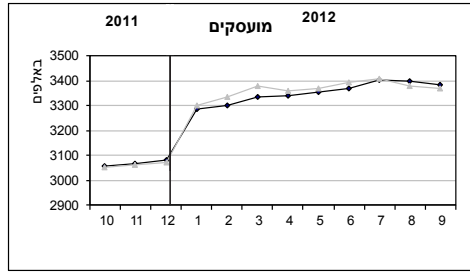
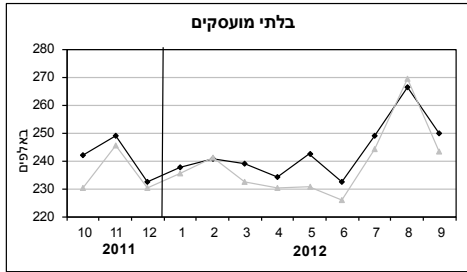
מהתבוננות בצירורים ניתן להתרשם כי בכל האומדנים ללא יוצא מהכלל אין עדות לשוני שיטתי בין השיטות או להטיה מצטברת כלשהי לאורך הזמן, ולמעשה עבור רוב הפרמטרים שתי השיטות נותנות תוצאות דומות מאוד. אמנם ניתן להתרשם כי סדרות אומדני הפרמטרים שבהם יש לנו עניין מיוחד, היינו בלתי מועסקים ומועסקים, אכן שונות במעט בשיטות האמידה. בסדרות אחרות, הגם שנאמדים פרמטרים קטנים יותר, ההבדל בין השיטות אינו בא לידי ביטוי. בכל אופן, מבלי להיכנס לשאלה האם השיטה החדשה אכן יעילה יותר או לא, ניכר כי ההבדל אינו מהותי.

4.4. הגדרת אוכלוסיית הסקר לפי זרישות OECD

סברה נוספת שהועלתה כסיבה לגידול בבלתי מועסקים התבססה על מעבר הלמ"ס למדידה לפי הקריטריונים הנהוגים בארגון לשיתוף פעולה ופיתוח כלכלי – OECD. "הלמ"ס הודתה כי שיטת מדידת האבטלה שלה לא מקובלת על ה-OECD – ושינתה אותה ■ התוצאה: זינוק פתאומי של 20% באבטלה ל-6.5%".⁵

בעקבות הצטרפות ישראל לארגון בשנת 2010 שינתה הלשכה הגדרות של פרמטרים רבים על מנת שתיווצר הרמוניזציה של הסטטיסטיקה הרשמית של ישראל עם זו הנהוגה בארגון, והאומדנים שיתקבלו יהיו בני השוואה לאומדנים המפורסמים על ידי יתר המדינות החברות בו. בנושא הסטטיסטיקה של עבודה נכללו אנשי הצבא הסדיר (חובה וקבע) כמועסקים, ולכן נכללו כשייכים לכוח העבודה. זאת אומרת, שינוי הגדרת השייכים כך שתכלול את כוח העבודה הכללי כולו ולא רק את כוח העבודה האזרחי כפי שהיה נהוג עד כה. שיעור הבלתי מועסקים מחושב כיחס בין מספר

ציור 1: השוואה ארוכת טווח בין שתי שיטות האמידה בסכ"א חודשי



הבלתי מועסקים לבין מספר השייכים לכוה העבודה. לכן, עם הכללת הצבא בסכ"א החודשי וסיווג אנשיו כמועסקים, אחוז המועסקים מתוך כוח העבודה צריך לעלות ואחוז הבלתי מועסקים לרדת, שכן המונה (מספר הבלתי המועסקים) נותר ללא שינוי, והמכנה (מספר השייכים לכוה העבודה) גדל כמספר המשרתים בצבא, וכאמור שניהם נמדדים באופן יחסי מתוך השייכים לכוה העבודה. ואכן, אם גדלים אופייניים למספר המועסקים והבלתי מועסקים בסקר הרבעוני היו 3,000,000 ו-180,000 איש, בהתאמה, ואם נעריך לפי פרסומים זרים את המשרתים בצבא ב-175,000 איש, אזי אחוז הבלתי מועסקים בתחילת 2012 נאמד בכ-180,000/3,355,000, לעומת 180,000/3,180,000 ללא הכללת הצבא. היינו, השינוי בהגדרה מקטין את האומדן לאחוז הבלתי מועסקים בכ-0.3% ואינו מגדיל אותו. אם כך, הגידול בבלתי מועסקים אינו נובע מהכללת הצבא.

ד.5. אי השבה

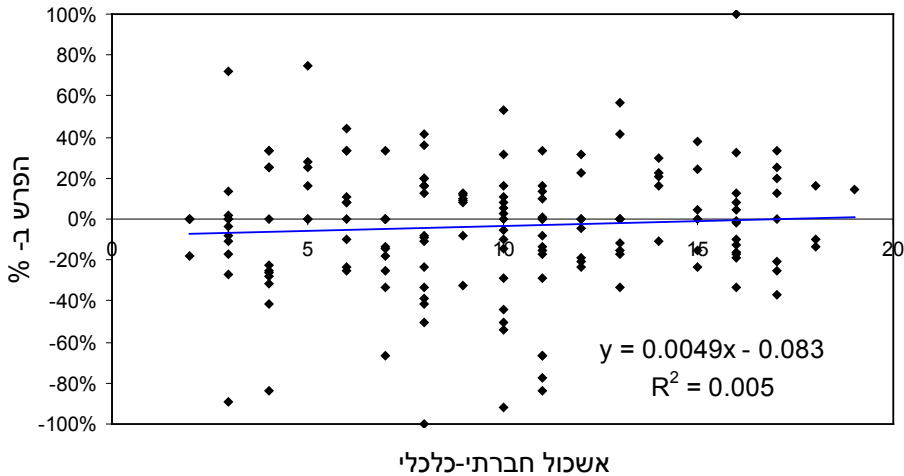
אי-ההשבה בסכ"א חודשי גדולה מזו של סכ"א רבעוני. מאז תחילת הפקידה, קיים פער ממוצע של כ-5%–10% באי-ההשבה בין הסקר החודשי לרבעוני. אחוזי השבה מחושבים מתוך נדגמים פחות מקרי אפס – מקרים שאינם שייכים לאוכלוסיית הסקר. לאור זאת היה קיים חשש שמא אי-השבה עודפת זו היא דיפרנציאלית ביחס למצב החברתי-כלכלי של האזור, ובפרט לאחוז הבלתי מועסקים בו. היינו שבמקומות (המוגדרים ע"י אזורים סטטיסטיים) השייכים לאשכול חברתי-כלכלי נמוך קיימת אי השבה עודפת גבוהה בסקר החודשי ביחס לרבעוני, ולהפך. לכל אזור סטטיסטי הגדרנו משתנה X – האשכול החברתי-כלכלי של האזור על פי מפקד 2008, משתנה המשקלל 16 משתנים כלכליים חברתיים בתחום ההשכלה, ההכנסה, רמת החיים וכו', ושאו מאמינים כי הוא מתואם ברמה זו או אחרת עם אחוז הבלתי המועסקים באזור. חישבנו את המשתנה Y – ההפרש באחוז אי-ההשבה בין הסקר החודשי והרבעוני, וצוירים של Y על X הוצגו לכל אחד מ-14 היישובים הגדולים במדינה. באף יישוב לא נמדד שיפוע משמעותי של קו הרגרסיה של Y על X . המסקנה המתבקשת מכך היא שאי-ההשבה אינה דיפרנציאלית ביחס לאשכול החברתי-כלכלי של האזור. גם אם זו אינה הוכחה לגבי חוסר מתאם עם אחוז הבלתי מועסקים, עדיין יש בה כדי לחזק סברה זו.

עבור ירושלים ניתן לראות בציר 2 כי אין מגמה ברורה של שינוי באחוזי ההשבה בין הסקרים. בניתוח הסטטיסטי, שיפוע קו הרגרסיה הוא 0.005. ז"א, ההפרש המקסימלי באי-השבה בין אזורים השייכים לאשכול הח"כ הכי גבוהה (17) לבין הכי נמוך (2) הוא $0.075\% = (17 - 2) \times 0.005\%$. גם אם נניח הערכה שמרנית למדי, כי אחוז הבלתי מועסקים גדל ב-1% כאשר האשכול החברתי-כלכלי קטן ב-1, והערכה של התפלגות אחידה של מספר האנשים באזורים סטטיסטיים השונים, הדבר יבוא לידי ביטוי בשינוי בבלתי מועסקים של $0.015 \approx \sum_{i=2}^{17} 0.005 \times 0.01 \times (i - 1)$. הווה אומר, גידול מזערי של 0.015% בבלתי מועסקים בירושלים.

את אותם הנתונים בדקנו גם לאור התיאוריה של בדיקת קשרים מונוטוניים בין שני משתנים, יצחקי וגולן (2009) וגם Yitzhaki and Olkin (1988). השיטה משתמשת בעקומת הריכוז ACC ובעקומת LMA כדי לבדוק אם קשר לינארי מסוים שנמדד בין שני משתנים הוא מונוטוני, או שקיימים תחומים זרים על ציר ה- X שבכל אחד מהם הקשר הוא אחר בעוצמתו או אף בסימנו. בדיקה זו אמינה יותר ככל שיש יותר קטגוריות בציר ה- X , כך שניתן להבחין במגמה השונה בתחומים שונים. למשתנה אשכול חברתי-כלכלי יש לכל היותר 20 קטגוריות, כך שהניתוח אינו

מתאים. בכל זאת, מלבד העיר תל אביב-יפו שבה ניתן היה אולי לזהות איזו שהיא מגמה שונה עבור תחומים שונים, לא מצאנו עדות לחוסר מונוטוניות באף אחד מהיישובים הגדולים האחרים שנבדקו.

ציור 2: הפרש בין % השבה רבעוני לחודשי לפי אשכול חברתי-כלכלי, ירושלים



ה. הבדלים באומדני הבלתי מועסקים של משתני מפתח דמוגרפיים

גידול בבלתי מועסקים על פני כל אופנויות של משתנה נחקר יעידו בדרך כלל על שינוי או בעיה מהותית באחד או יותר מהפרמטרים הבסיסיים של הסקר החדש, כמו למשל בעיה בהגדרות האוכלוסייה, בשיטת הדגימה, בשיטת האמידה כדוגמת בעיה באחת הנוסחאות, בעיית מחשב, בעיה בתוכנה או במסדי הנתונים, בעיה בהנחיה של הסוקרים וכדומה. שכן, בעיות מהסוגים הנ"ל חזקה עליהן שישפיעו במידה שווה על הנתונים הנאספים ולא יתבטאו רק באופנויות מסוימות של משתנה כלשהו. לכן, חשוב היה לבחון האם הגידול בבלתי מועסקים היה ספציפי עבור מין, גיל או מגזר מסוים, או אופנות כלשהי של משתנה אחר שניתן להעלות על הדעת בצורה הגיונית.

ה.1. מגזר

נתוני הסקר החודשי, כאמור, הראו שינוי משמעותי באומדן שיעור הבלתי מועסקים לעומת הסקר הרבעוני. שיעורי הבלתי מועסקים במשק כפי שנמדדו ב-2011 היו 5.7, 5.2, 6.1, ו-5.4 עבור

רבעונים I, II, III ו-IV בהתאמה. עם זאת, בחלוקה לפי מגזר, אחוז הבלתי מועסקים היה גבוה יותר דווקא בקרב היהודים בכל הרבעונים הנ"ל. בלוח 2 מוצג אחוז הבלתי מועסקים ברבעון IV/2011 לפי מין ומגזר, וכן ממוצע אחוז הבלתי מועסקים בחודשים ינואר-מרץ (רבעון I) 2012.

לוח 2: אחוז בלתי מועסקים לפי מין ומגזר ברבעונים IV/2011 ו-I/2012 לפי הסקרים הרבעוני והחודשי בהתאמה

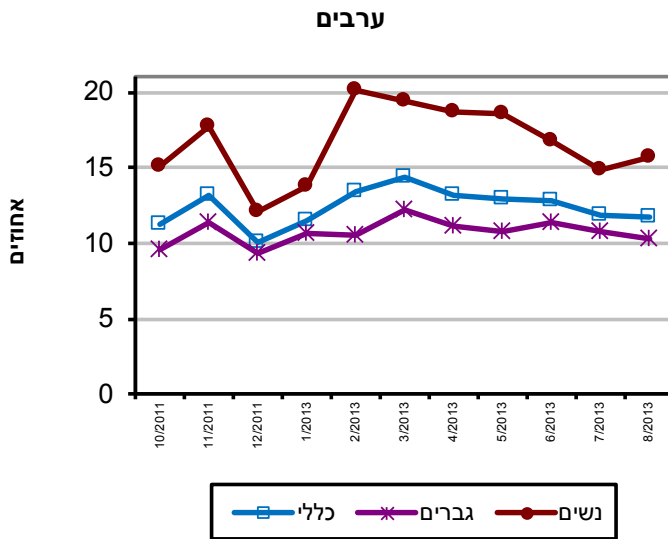
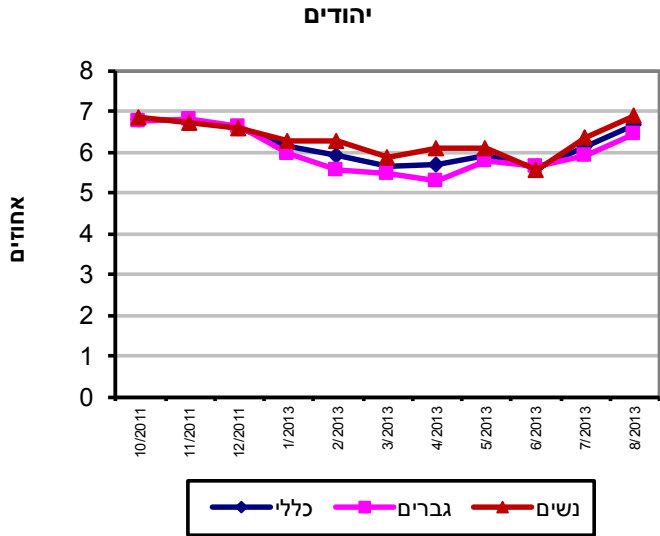
I/2012		IV/2011		
ערבי	יהודי	ערבי	יהודי	נשים
17.80	6.15	6.17	5.75	גברים
11.50	5.65	4.53	5.31	סה"כ לפי מגזר
13.65	5.90	4.94	5.53	סה"כ כללי
	6.70		5.47	

כפי שהקדמנו להזכיר בפרק המבוא וכפי שמשקף בלוח, השינוי באחוז הבלתי מועסקים במעבר לסקר חודשי היה גדול במיוחד בקרב הערבים, וגדול הרבה פחות בקרב היהודים. אחוז הבלתי מועסקים הממוצע ברבעון I של 2012 בקרב היהודים עלה ל-5.9%, היינו גידול של קצת פחות מ-0.4 אחוז לעומת הבלתי מועסקים ברבעון IV של 2011, או בשקלול השפעת הצבא (ראה סעיף 4.ד) עלייה של קצת מעל לחצי אחוז. במגזר הערבי, לעומת זאת, גדל אחוז הבלתי מועסקים ל-13.65% עבור אותה התקופה, ראה גם ציור 3. זאת אומרת, באופן יחסי גדל אחוז הבלתי מועסקים אצל גברים ונשים במגזר היהודי ברבעון I של 2012 פי 1.1 (10%) לעומת שיעורו ברבעון שקדם לו, אך בערך פי שניים אצל הגברים ובערך פי שלושה אצל הנשים במגזר הערבי. ללווחות מפורטים יותר ראה גם נייר מדיניות של מרכז טאוב מאת ישיב ערן (2012).

זאת ועוד, נשים לב כי המשרעת של נתוני הבלתי מועסקים הרבעוניים של 2011 אינה זניחה ועומדת על 0.9% – לפי 5.2% ברבעון II לעומת 6.1% ברבעון III. הנתונים הם נתוני מדגם ולכן חשובים לטעות דגימה, אך בנוסף גם להשפעות של עונתיות בתעסוקה הנובעת מסיבות שונות. כדי להקטין השפעה זו מפרסמת הלמ"ס גם נתונים מנוכי עונתיות המתקבלים לאחר ניכוי השפעת העונתיות והשפעת החגים וימי הפעילות מהנתונים המקוריים. הסבר על תהליך ניכוי עונתיות ואמידת המגמה מתפרסם באתר הלמ"ס באינטרנט.⁶ כפי שציינו בתחילת הסעיף, בשנת 2011 נתון הבלתי מועסקים של רבעון I גדול מזה של רבעון IV. כך היה גם בשנת 2010. לאור זאת ייתכן כי יש בהשפעת העונתיות כדי להסביר חלק מהגידול באחוז הבלתי מועסקים בין רבעון IV/2011 לרבעון I/2012 אצל היהודים, או אף את כולו.

6 הגורמים העונתיים וגורמי ההתאמה מראש ל-2011: מגמות ל-2007–2011. <http://www.cbs.gov.il/publications/tseries/seasonal11/presentationh11.pdf>

ציור 3: אחוז הבלתי מועסקים לפי מין ומגזר בסקר החודשי

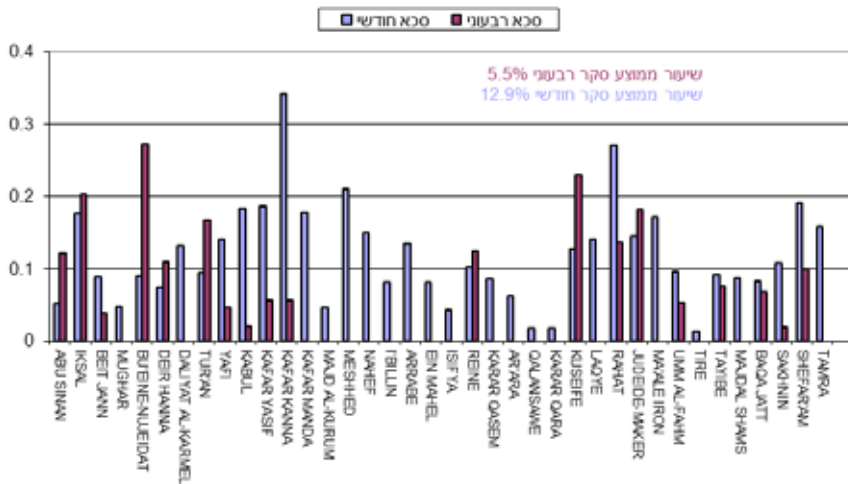


מבחינה סטטיסטית-מתודולוגית יכלו מתכנני הסקר, במידה מסוימת ומסויגת, אף לשאוב עידוד מהממצא הנ"ל. שוני כה מובהק באומדנים של קטגוריה אחת של משתנה מסוים שולל את האפשרות שהשוני בין הסקרים נובע מבעיה במערכת הדגימה, בשיטת האמידה או באחת ממערכות המחשב והתוכנה וכדומה. בעיה באחת ממערכות אלו כאמור הייתה גורמת לשינוי באומדנים על פני כל הקטגוריות של אותו משתנה, שכן שיטות הדגימה, נוסחאות האמידה ותוכנות המחשב אינן תלויות

מגזר, מגדר, גיל, אזור גיאוגרפי וכיוצא בזה. הנתונים שנאספו ועובדו גזרה אחת היא לכולם שאינה תלויה בקטגוריה כלשהי של משתנה דמוגרפי זה או אחר.

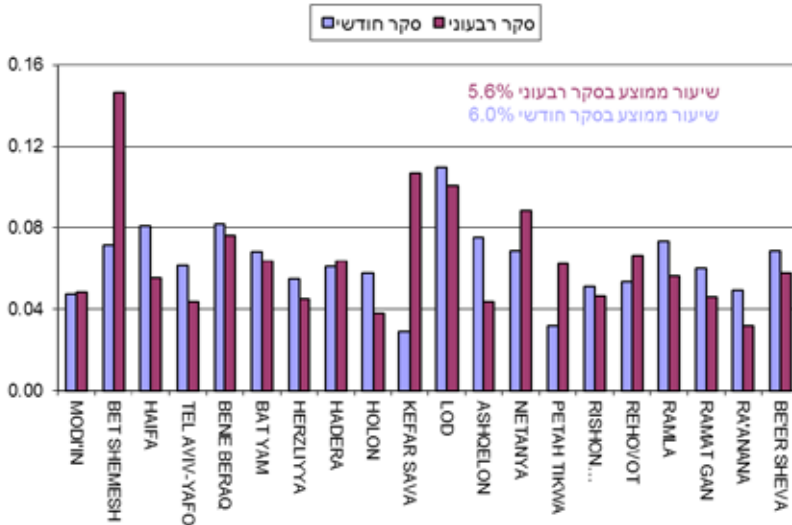
נציין כי על אף שהנתונים הרשמיים של סכ"א במתכונתו החודשית החלו להתפרסם רק בתחילת 2012, הדגימה וחקירת משקי הבית החלה כבר באפריל 2011 (זאת על מנת לאפשר קבלת אומדנים חודשיים לשלושת החודשים האחרונים של 2011). למרות שהשוואת האומדנים מתקופת חפיפה זו בין שני הסקרים יכולה להיות מאוד אינפורמטיבית, השתדלנו במאמר זה להימנע מהסקת מסקנות על פיה שכן אחוז אי-ההשבה בסקר החודשי היה גבוה בתקופה זו באופן משמעותי מהחודשים שלאחר מכן, מאחר שלא כל מערכות הסקר נכנסו בה עדיין לפעולה חלקה (בעיקר בחודש 10/2011 ופחות בסוף הרבעון). בכל זאת, על מנת לסבר את העין, נציג כאן שתי השוואות. ציור 4 מציג יישובים ערביים עירוניים גדולים המונים פחות מ-50 אלף תושבים שהיו משותפים למדגמי שני הסקרים ברבעון IV/2011. אחוז הבלתי מועסקים הממוצע בסקר הרבעוני עומד על 5.5%, בשעה שבאותם יישובים ממש בסקר החודשי נמדד אחוז בלתי מועסקים של 12.9%. קיימת גם שכיחות גבוהה של יישובים שבהם נמדד אחוז אפס בסקר הרבעוני לעומת שכיחות חיובית בסקר החודשי. ב-20 מתוך 38 יישובים אלה לא נחקר בסקר הרבעוני ברבע האחרון של 2011 ולו בלתי מועסק אחד. מציור 5 ניתן לראות כי תופעות אלו אינן קיימות עבור היישובים היהודיים המונים מעל 50,000 תושבים (ראה שאול-מנע ואחרים, 2012). אמנם עדיין רואים נטייה לאחוז בלתי מועסקים גדול יותר בסקר החודשי לעומת הרבעוני, אך היא מינורית, אינה גורפת, וכפי שצוין לעיל הפקידה ברבעון הייתה עדיין בשלב הבחינה, והאומדנים לא נכללו בנתונים הרשמיים שמפרסמת הלשכה.

ציור 4: אחוז הבלתי מועסקים ביישובים ערביים גדולים⁷ שבהם 2–50 אלף תושבים, רבעון רביעי 2011



7 שמות היישובים צוינו בכתב אנגלי כדי להתגבר על בעיות הגיה.

ציור 5: אחוז הבלתי מועסקים ביישובים יהודיים, מעל 50 אלף תושבים, רבעון רביעי 2011



ה.2. איסוף נתונים ושונות בין סוקרים

ההבדל המשמעותי בין המגזרים יכול להצביע על השפעה של הסוקרים עצמם על איסוף הנתונים. בעיקרון היינו רוצים להאמין כי כל הסוקרים שווים, אך בפועל ידוע כי שני סוקרים יכולים להזין שתי תשובות שונות לאותה שאלה מאותו הנדגם באותו הזמן. במילים אחרות, תוצאת הקירה של סוקר היא משתנה מקרי, והשפעת הסוקר על המדידה יכולה להתפרש כאפקט מקרי (Random Effect) בתהליך ניתוח שונות (ANOVA). נסמן ב- y_{ij} את תוצאת הקירה j של סוקר i . המודל הוא $y_{ij} = m_i + \epsilon_{ij}$. אם ניתן לכל סוקר תת-מדגם מייצג של האוכלוסייה היינו רוצים שיתקיים $m_i = m$ לכל i , כלומר אינו תלוי בסוקר i , שכן לכל תת-מדגם מייצג יש אותה התוחלת. אך בפועל אין כך הדבר. מתברר כי $m_i = m + \alpha$ כאשר α משתנה מקרי בעל תוחלת 0 וסטיית תקן σ שנובעת מההבדל האינדיבידואלי בין הסוקרים. מסיבה זו, כאשר מטרתו העיקרית של סקר היא מדידת שינוי בין שתי תקופות זמן, אין מניעה כי הנתונים יאספו על ידי מספר מצומצם של סוקרים, אך כאשר חשוב ערכו האבסולוטי של הפרמטר הנחקר, יש חשיבות לביצוע הסקר על ידי מספר רב של סוקרים.

כדי להבין מניין נובעת סטיית תקן זו נסביר כי שאלת הבלתי מועסקים בסקר כוח אדם מורכבת ממספר רב של שאלות ששקלולן מכריע לגבי היותו של הפרט הנחקר בלתי מועסק אם לאו (ראה שוב נספח א לגבי ההגדרה המדויקת של בלתי מועסק). אפילו אם השאלות מוגדרות היטב, ישנם לעתים מקרים שבהם קיימת השפעה לדרך שבה נשאלת השאלה על ידי הסוקר או לשיקול דעתו

בסיווג התשובה של הפרט הנחקר. לדוגמא, על מנת שפרט ייחשב בלתי מועסק, תנאי הכרחי הוא שחיפש עבודה באופן פעיל בארבעת השבועות האחרונים. למרות שהתשובה לשאלה היא דיכוטומית – כן חיפש או לא חיפש, הגדרת הביטוי "באופן פעיל" היא מעורפלת משהו. למשל, האם דפדוף אקראי במודעות הדרושים בעיתון נחשבת פעילות אקטיבית?

המדגם לאמידת סטיית התקן של אפקט הסוקר בוסס על 22 סוקרי טלפון מהסקר החודשי (לא כולל סוקרים בכירים המטפלים לרוב במקרים מיוחדים וסוקרים ברוסית וערבית) שקיבלו כל אחד מדגם מייצג ככל הניתן של נדגמים לסקר, כך שהצפי היה לקבל מכולם את אותם האומדנים. הבדיקה מצאה סטיית תקן של 0.54% עבור פרמטר הבלתי מועסקים ו-2.9% עבור הפרמטר מועסקים. חלוקה ב- \sqrt{n} נותנת את סטיית התקן של אפקט הסוקרים כאשר מספר הסוקרים הוא n . בלוח 3 מחושבים ערכי סטיית התקן זו עבור $n=1, 4, 40$ סוקרים.

לוח 3: סטיית התקן של אפקט הסוקר

סטיית תקן של סוקר אחד	סטיית תקן של 4 סוקרים	סטיית תקן של 8 סוקרים	סטיית תקן של 40 סוקרים
0.54%	0.27%	0.19%	0.08%
2.9%	1.45%	1.02%	0.45%
בלתי מועסקים			
מועסקים			

בסקר החודשי מוקצים כרגע 4 סוקרים טלפוניים למגזר הערבי ו-40 ליהודי, בעוד שבסקר הרבעוני הוקצה רק סוקר טלפוני יחיד לכל המגזר הערבי לעומת 8 סוקרים במגזר היהודי. נבהיר כי סטיית התקן של אפקט הסוקר אינה מלמדת על ההשתנות של נתונים שנאספו על ידי סוקר מסוים אלא על השוני בין נתונים שנאספו על ידי שני סוקרים שונים שחקרו את אותה האוכלוסייה באותו זמן. סטיית תקן של 0.54% אין בה כדי להסביר שינוי של יותר מ-10% באחוז הבלתי מועסקים במגזר הערבי, ובפרט שגם עבור 8 סוקרים שהיו במגזר היהודי סטיית התקן של 0.19% אינה שונה באופן מהותי. לכן אין בנתון זה לבדו כדי להסביר את ההבדל בין הסוקרים. אמנם ייתכן כי סטיית התקן של אפקט הסוקר עבור הפרמטר "בלתי מועסקים במגזר הערבי" היא גדולה הרבה יותר, אך השערה זו לא נבדקה. עם זאת, כאשר מדובר בסוקר יחיד, כל התרחישים אפשריים ולא מן הנמנע כי לאותו סוקר היה פירוש שונה וחריג במיוחד של הבנת השאלות הקשורות להגדרת הפרמטר "בלתי מועסק". יתכנו גם סיבות נוספות, כגון רשלנות, טעות שיטתית כלשהי בהבנתו כיצד לעבוד עם התוכנה ועוד. לא התייחסנו כאן לסוקרי השטח בעיקר מכיוון שכל סוקר שטח עובד עם אוכלוסייה אחרת באזור גיאוגרפי אחר. כמו כן, מספר סוקרי השטח הוא גדול בשני הסוקרים ולכן ההשפעה הפוטנציאלית של אפקט סוקר השטח לעומת זו של סוקר הטלפון היא קטנה.

ה.3. ראיונות שטח וטלפון

כל פנל בסקר החודשי נחקר כאמור שמונה פעמים. על מנת לחסוך בעלויות, רק החקירה הראשונה היא חקירת שטח פנים-אל-פנים שבה פוקדים את דיירי הדירה שנדגמה, ואילו יתר החקירות

מבוצעות בטלפון בעזרת הפרטים שהתקבלו בחקירת השטח. במגזר הערבי היה כאמור סוקר טלפוני יחיד, אך עשרה סוקרי שטח. בנוסף, ראיונות השטח והטלפון אינם מבוצעים על ידי אותם החוקרים ולכן קיימת ביניהם הסטייה שדגנו בה בסעיף הקודם. בלוח 4 להלן ניתן לראות כי סטייה זו במגזר הערבי אכן באה לידי ביטוי בסקר הרבעוני אך לא בסקר החודשי. בכל אופן אין בכך כדי להסביר את כל ההבדל בין הסקרים במגזר הערבי, שכן אחוז הבלתי מועסקים גם בראיונות השטח היה קטן יותר לעומת הסקר החודשי.

לוח 4: אחוז בלתי מועסקים לפי מגזר וסוג ראיון

ערבים		יהודים		
שטח	טלפון	שטח	טלפון	סקר*
7.3	5.1	5.6	5.6	רבעוני IV/2011
13.9	13.6	5.7	5.7	חודשי ינואר-מרץ 2012

* נתוני הסקר החודשי הותאמו כך שיכללו רק את כוח העבודה האזרחי, כפי שהיה בסקר הרבעוני.

ה.4. בדיקת משתני חקירה נוספים

ניסיונות לאפיין את ההבדל בין הסקרים בעזרת משתני חקירה נוספים לא סיפקו תובנות חדשות. בין היתר נחקרה ההשבה לפי מעמד המשיב (ראש משק בית, בן/בת זוג, וכו'), לפי שלב הפקידה (ראשון עד רביעי בסקר הרבעוני, ראשון עד שמיני בסקר החודשי), גודל משק בית, צורת יישוב המגורים, משך חיפוש העבודה וסוג ראיון (אישי או Proxy⁸). באף אחד לא מצאנו ממצאים חד משמעיים שיכולים לשפוך אור נוסף על ההבדל בין הסקרים. ניתוח ההשבה במגזר הערבי לפי שעת הפקידה העלה כי קיים הבדל באחוז הבלתי מועסקים בין איסוף נתונים שנעשה על ידי הסוקר בשעות הבוקר לבין איסוף שנעשה אחר הצהריים (ראה אטקס וחפץ, 2013). מכל מקום, מאחר שבמגזר הערבי בסקר הרבעוני היה רק סוקר יחיד שבחר לעבוד בשעות הבוקר, לא מן הנמנע כי למעשה ההשפעה היא של הסוקר ולא של השעה, וממצאים אלה יכולים להיות מוסברים לפחות בחלקם על ידי אפקט הסוקר שנדון בסעיף ה-2. בבדיקה לפי סוג ראיון, למשל, הבדל של כ-1.2% באחוז הבלתי מועסקים המדווח בראיון האישי לזה המדווח בראיון ה-Proxy בסקר הרבעוני מצטמצם לכ-0.95% בסקר החודשי. כמובן אין בכך כדי להצביע על הסבר כלשהו. ראה לוח 5.

8 ראיון Proxy הוא ראיון שבו בן משק בית אחד נשאל על בן משק בית אחר.

לוח 5: אחוז בלתי מועסקים לפי סוג ראיון

סקר רבעוני			סקר חודשי				
ממוצע	רבעון IV/2011	רבעון III/2011	ממוצע	מרץ 2012	פברואר 2012	ינואר 2012	
6.42%	6.02%	6.81%	7.03%	6.74%	7.04%	7.23%	ראיון אישי
5.25%	4.92%	5.57%	6.08%	5.84%	6.19%	6.23%	ראיון Proxy
1.17%	1.10%	1.24%	0.95%	0.90%	0.85%	1.00%	הפרש

ו. סיכום ומסקנות

מעבר הלשכה למדידת סקר כוח אדם במתכונת חודשית התאפיין בגידול במספר הבלתי מועסקים במשק. ההבדל התבטא בגידול חד במספר הבלתי מועסקים במגזר הערבי ולעומתו גידול לא מובהק במגזר היהודי. במאמר זה ניסינו לחקור את הסיבות להבדל בין הסקרים. הממצאים במאמר שוללים את האפשרות שהשוני בין הסקרים נובע מבעיה בדגימה או ברוטציית הפנלים, משיטת האמידה החדשה או ממערכות המחשב והתוכנה, בעיות שהיו מביאות כאמור להבדל גורף באומדנים על פני כל הקטגוריות של המשתנים שנחקרו, ובפרט על פני כל המגזרים, בשונה מכפי שקרה בפועל. בנוסף לכך שללנו את האפשרות שהגידול באי-השבה היה תלוי באשכול החברתי-כלכלי של האזור. ההבדל המשמעותי בין המגזרים מצביע לדעתנו על בעיה הקשורה לתהליך איסוף הנתונים. זהו הנדבך היחיד בשדרותיו של הסקר שבו ישנה הפרדה ברורה בין שני המגזרים. בקרה או הנחיה שונה, לחיוב או לשלילה, של סוקרים חדשים מהמגזרים השונים יכולה לגרום לשוני בכלל האומדנים אך גם רק באלו של אחד המגזרים. בפרט, העובדה כי בסקר הרבעוני בוצעו כל החקירות הטלפוניות במגזר הערבי על ידי סוקר טלפוני יחיד מחזקת את ההשערה כי ייתכן שכאן נעוץ ההבדל.

מקורות

- אטקס ה' וחפץ א' (2013), סקרי כוח אדם בישראל: ממצאים ראשוניים אודות הטיות אפשריות עקב זמינות מוגבלת לפקידה. מצגת להרצאה בלמ"ס.
- הלשכה המרכזית לסטטיסטיקה (2011), סקרי כוח אדם 2010, פרסום מיוחד מס' 1460, http://www.cbs.gov.il/publications11/1460/pdf/h_print.pdf
- יצחקי ש' וגולן י' (2009), קשרים מונוטוניים ולא מונוטוניים בתחום העבודה והמגדר, הרבעון לכלכלה, יוני 56(2), 143–168.
- ישיב ע' (2012), שינויים בנתוני סקר כוח אדם ומשמעותם, מרכז טאוב, נייר מדיניות מס' 2012.09.
- שאוּל-מנע נ', פוגל נ', פרידמן י' ורומנוב ד' (2012), השוואות סקר כוח אדם חודשי וסקר כוח אדם רבעוני, מסמך פנימי, למ"ס.
- Bell P. (2001), "Comparison of Alternative Labour Force Survey Estimators", *Survey Methodology* 27(2), 53–63.
- Cochran W.G. (1977), *Sampling Techniques*, 3rd ed., New York.

- Deville J.C. and Särndal C.E. (1992), "Calibration Estimators in Survey Sampling", *Journal of the American Statistical Association* 87, 376–382.
- Fuller W.A. and Rao J.N.K. (2001), "A Regression Composite Estimator with Application to Canadian Labor Force Survey", *Survey Methodology* 27, 45–51.
- Gambino J., Kennedy B. and Singh M.P. (2001), "Regression Composite Estimation for the Canadian Labour Force Survey: Evaluation and Implementation", *Survey Methodology* 27, 65–74.
- Singh A.C. (1996), "Combining Information in Survey Sampling by Modified Regression", *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 120–129.
- Yitzhaki S. and I. Olkin (1988), "Concentration Curves", W.P. no. 179, Dept. of Economics, Hebrew University.

בלוח 7 מוצג דגם החקירה בסקר כ"א רבעוני לשנים 2010 ו-2011.

לוח 7: דגם חקירת הפנלים, לפי רבע החקירה ושנת הדגימה בסקר רבעוני

רבע חקירה בשנת 2011				רבע חקירה בשנת 2010				רבע כניסה לחקירה	שנת דגימה
IV	III	II	I	IV	III	II	I		
							4	4/08	2008
							3	1/09	
					4	3		2/09	
				4	3			3/09	
			4				2	4/09	2009
		4	3				1	1/10	
	4	3			2	1		2/10	
4	3			2	1			3/10	
			2					4/10	2010
			1					1/11	
	2	1						2/11	
2	1							3/11	
								4/11	2011

נספח ב: שיטת האמידה Composite Estimation

אומד משולב הוא מושג בתחום סקרי פנלים לתיאור אומד המחושב על סמך נתוני ההווה ובנוסף גם על סמך נתוני העבר של הנדגמים המשיבים בסקר. בשונה מאומד ישיר המחושב על סמך נתוני ההווה (זמן t) בלבד, האומד המשולב מנצל גם אינפורמציה רלוונטית מהעבר (זמן $t - 1$) של המשיבים בסקר. ישנה אבחנה בין שני סוגים של תצפיות: קבוצת התצפיות הראשונה כוללת את קבוצת הפנלים החופפים, היינו את הפרטים אשר השתתפו בסקר הן בזמן t והן בזמן $t - 1$, ושנייה הכוללת את קבוצת הפנלים החדשים, היינו את אלו אשר השתתפו בסקר בזמן t אך לא בזמן $t - 1$, כשהדבר תלוי במבנה הרוטציה של הסקר.

האומד המשולב שנציג להלן הוא עקיב ואנו מצליחים בעזרתו להקטין, לעתים באופן משמעותי, את טעויות הדגימה של אומדים לפרמטרים מסוימים החשובים לנו במיוחד (כדוגמת האבטלה) לעומת האומד הישיר המבוסס על תצפיות מזמן t בלבד. השיטה גם מאפשרת הרמוניזציה של האומדנים המופקים בעזרתה, במובן שהם מסתכמים לערכים הידועים ממקורות אחרים. בפרט, סכום האומדנים המשולבים למצבי התעסוקה השונים מסתכם לסך האוכלוסייה הידוע (לדוגמא: סכום אומדן לסה"כ מועסקים ואומדן לסה"כ בלתי מועסקים שווה לאומדן לסה"כ שייכים לכוח העבודה). האומד המשולב שנשתמש בו הוא אומד רגרסיה מוכלל (Generalized Regression) על משתנים מהחודש הנוכחי בתוספת משתני עזר המוגדרים על נתוני החודש הקודם (Singh, 1996),

כך שמבחינה טכנית ניתן להפיק את מקדמי הניפוח הסופיים כווריאציה על הרגרסיה הליניארית, כפי שנעשה עבור האומד הישיר עד כה בסקרי כוח אדם.

נסמן:

- S – מדגם המשיבים.
- X – נתונים דמוגרפיים ידועים (דוגמת סך כלל אוכלוסיית המסגרת).
- x – נתוני המדגם המתאימים למשתנים הדמוגרפיים שב- X .
- z – נתוני המדגם המתאימים לפרמטרים המסוימים החשובים לנו במיוחד.
- d_k – מקדם הניפוח הראשוני של פרט k .
- w_k^k – מקדם הניפוח הסופי של פרט k .
- \hat{Y} – אומדן למשתנה הנחקר.

אומד הורוביץ תומפסון (HT) הוא אומד המתקבל ע"י הכפלת נתוני המדגם במקדם הניפוח של כל פרט במדגם. נסמן $\hat{X}^{HT} = \sum_{k \in S} d_k x_k$ עבור המשתנה x . ברוב המקרים האומדן המתקבל אינו זהה לנתונים הדמוגרפיים הידועים X . לכן, בתהליך הניפוח, אנו מחשבים לכל פרט במדגם מקדם ניפוח סופי באופן כזה שאומדן המחושב מנתוני S יתאים לנתונים הדמוגרפיים הידועים. זאת אומרת שבזמן t אנו מחשבים מקדמים $w_{k,t}$ כך שיתקיים $\sum_{k \in S_t} w_{k,t} x_{k,t} = X_t$.

יהי $y_{k,t}$ הנתון מהמדגם של פרט k בזמן t עבור משתנה נחקר החשוב לנו במיוחד. שיטת האמידה המשולבת היא כאמור רקורסיבית. בהינתן האומדן \hat{Y}_{t-1} מזמן $t-1$, נרצה כי מקדמי הניפוח $w_{k,t}$ יהיו כאלה שגם יתאימו את נתוני המדגם מחודש $t-1$ לאומדן \hat{Y}_{t-1} מהחודש שעבר, כאילו היה נתון ידוע, כפי שנעשה עבור הנתונים הדמוגרפיים. זאת אומרת תנאי נוסף לקביעת מקדמי הניפוח הוא $R\hat{Y}_{t-1} = \sum_{k \in S_{t-1} \cap S_t} w_{k,t} z_{k,t-1} = \hat{Y}_{t-1}^{HT}$, כאשר R הוא פרופורציית הפנלים המשותפים עם חודש $t-1$ מכלל הפנלים בחודש t . בצורה כזו ניתן להציג את האומד המשולב כאומד רגרסיה באופן הבא:

$$\hat{Y}_t = \sum_{k \in S_t} w_{k,t} y_{k,t} = \sum_{k \in S_t} d_{k,t} y_{k,t} + \left((X_t, \hat{Y}_{t-1}) - (\hat{X}_t^{HT}, \hat{Y}_{t-1}^{HT}) \right) \hat{\beta}_t$$

ומקדם הרגרסיה נתון ע"י:

$$\hat{\beta}_t = \left[\sum_{k \in S_t} d_{k,t} (x_{k,t}, z_{k,t})' (x_{k,t}, z_{k,t}) \right]^{-1} \sum_{k \in S_t} d_{k,t} (x_{k,t}, z_{k,t})' y_{k,t}$$

כאשר $(x_{k,t}, z_{k,t})$ הוא וקטור המורכב ממשתני המפתח הדמוגרפיים $x_{k,t}$ וממשתני המפתח

המיוחדים $z_{k,t}$. מקדמי הניפוח הסופיים ניתנים בהתאם על ידי:

$$w_{k,t} = d_{k,t} \left\{ 1 + \left((X_t, \hat{Y}_{t-1}) - (\hat{X}_t^{HT}, \hat{Y}_{t-1}^{HT}) \right) \left(\sum_{k \in S_t} d_{k,t} (x_{k,t}, z_{k,t})' (x_{k,t}, z_{k,t}) \right)^{-1} (x_{k,t}, z_{k,t}) \right\}$$

ניתן להוכיח כי האומדים (הרקורסיביים) המתקבלים בצורה זו הם עקיבים. אומד המבוסס על שיטת האמידה המשולבת כולל שונות הנובעת מכיול לאומדן של החודש הקודם, \hat{Y}_{t-1} , נתון שאינו ודאי, ועשוי אם כך לתרום להגדלת שונותו לעומת האומד הישיר. לעומת זאת האומד מאפשר לנו לעשות שימוש בנתונים בעלי רמת ודאות גבוהה שיתרמו להקטנת שונותו לעומת האומד הישיר. תוצאת שקלול שני ההפכים אינה ידועה מראש, אך מחקרים מדעיים בספרות וסימולציות שנערכו בלמ"ס הראו כי במקרים רבים התועלת גדולה מההפסד. ואכן, התוצאות האמפיריות מראות כי סדרת האומדנים החודשיים של תכונות כוח העבודה שהתקבלו היא בעלת השתנות קטנה יותר מסדרת האומדנים שהייתה מתקבלת משימוש בשיטה הקודמת.