***עבודה יישומית***

**אוניברסיטת בר אילן**

**המחלקה לכלכלה**

גל גשקוביץ

**פרק ראשון: בניית מודל כלכלי ואקונומטרי**

המשתנה המסביר המרכזי בו בחרתי הוא **st011q06ta** . משתנה זה הוא קטיגורי ובודק האם לפרט במדגם יש חיבור ביתי לאינטרנט. הערכים שהמשתנה מקבל הם 1 עבור התשובה "כן" (כלומר קיים חיבור אינטרנט ביתי) ו-2 עבור התשובה "לא" (כלומר לא קיים חיבור אינטרנט ביתי). בהתאם לכך שאלת המחקר שאני בוחר לנסח היא – **כיצד חיבור ביתי לאינטרנט משפיע על ציוני המתמטיקה במבחני PISA של התלמידים במולדובה?**

בחרתי במשתנה זה מפני שלדעתי הוא משפיע באופן משמעותי על משתנה התוצאה **pv1math –** ציון המתמטיקה במבחן פיזה. קיימים מחקרים רבים שבדקו את הקשר שבין שימוש אינטרנטי לבין הצלחה בלימודים. בחלק מהם התוצאות שהתקבלו הם מעניינות ביותר. מחקר אחד שמרכז בתוכו את שתי התיאוריות המרכזיות בנושא נערך בקוריאה בשנת [[1]](#footnote-1)2017. הוא בדק תלמידים מרקע סוציו-אקונומי מגוון, שלחלקם הייתה גישה לאינטרנט ולחלקם לא. התוצאות במחקר חשפו שלאלו שהשתמשו באינטרנט למטרותלימודיות חיוביות הייתה הצלחה גדולה יותר בלימודים לעומת אלו שלא גלשו באינטרנט כלל. כלומר התגלתה קורלציה חיובית בין שימוש חיובי באינטרנט לבין הצלחה בלימודים. לעומת זאת, תלמידים שהשתמשו באינטרנט למטרות שליליות כמו משחקים וכדומה, היו חשופים לפיתויים רבים יותר שמנעו מהם מלנצל את הזמן הפנוי ללמידה, ולכן הם הצליחו פחות בלימודים מאשר אלו שהיו ללא אינטרנט כלל. כלומר התגלתה קורלציה שלילית בין שימוש שלילי באינטרנט לבין הצלחה בלימודים.

כמובן שקשר סיבתי לא הוכח במחקר הנ"ל, ולכן ברור שישנם עוד גורמים נוספים שהשפיעו על התוצאות שהתקבלו. לכן לדעתי במדינה ענייה כמו מולדובה איני יודע לומר בוודאות האם אני מצפה לקבל קשר חיובי או שלילי במחקר. יכול להיות שרוב הפרטים במדגם ינצלו את האינטרנט לצרכים לימודיים , ויכול להיות שחלקם יתלהבו מהשפע הגדול שהמרשתת מציעה ולא ינצלו את הזמן הראוי ללמידה.

שלושת משתני הבקרה שבחרתי הם:

**mmins** – זמן הלימוד השבועי (בדקות) במקצוע המתמטיקה. זהו משתנה רציף. זמן הלימוד השבועי מתקשר למשתנה התלוי בכך שסביר להניח שזמן לימוד שבועי גדול יותר יגרום להצלחה במבחן. משתנה זה מתקשר גם למשתנה המסביר שבחרתי בסעיף א' בכך ששימוש באינטרנט לצרכים חיוביים/שליליים יכול להיכלל או לגרוע בהתאמה מתוך זמן הלימוד השבועי ובכך להשפיע על הציון הסופי.

**compete** – תחרותיות של הפרט במדגם. זהו משתנה רציף. הפרטים במדגם נדרשו לדרג שאלות כמו " אני נהנה להתחרות באחרים" או " זה חשוב לי להצליח יותר מאחרים במבחנים" , ובהתאם לתשובותיהם נבנה מדד התחרותיות. מספר גבוה מציין תחרותיות גבוהה. משתנה זה מתקשר למשתנה התלוי בכך שתחרותיות יכולה להיות גורם מדרבן שמשפיע בסופו של דבר על הצלחה במבחן. משתנה זה גם מתקשר למשתנה המסביר שבחרתי בסעיף א' בכך שתלמיד שהוא בעל תחרותיות ינצל את כל המשאבים שברשותו, ובפרט האינטרנט, בשביל שימוש חיובי והצלחה במבחן.

**st013q01ta** – מספר הספרים שיש בבית. זהו משתנה קטיגורי. מספר הספרים מחולק לקטגוריות כמו 0-10 ספרים בבית או 11-25. הקשר בין משתנה זה למשתנה התלוי הוא בכך שספרים הם כלי משמעותי לצורכי למידה, וזמינות גבוהה של ספרים יכולה להשפיע על הלמידה וההצלחה במבחן. הקשר בין משתנה זה למשתנה המסביר מסעיף א' הוא בכך שספרים יכולים להוות תחלופה לאינטרנט.

מלוח 1 ניתן ללמוד שאוכלוסיית המדגם היא מגוונת. מבחינת ציוני המתמטיקה- ניתן לראות שעל פי המינימום והמקסימום טווח הציונים נע מהנמוכים ועד לגבוהים ביותר כאשר הממוצע נע סביב ציון האמצע האפשרי במבחן ואף מעט מעליו (429 לעומת 400). מלוח 2 ניתן לראות שאכן רוב הפיזור הוא סביב הממוצע, כאשר צורת ההתפלגות היא בעלת שיא אחד וללא זנבות. קיימת שונות בין מספר התלמידים שיש להם חיבור אינטרנט ביתי לבין אלו שאין להם- דבר שטוב לצורך בדיקת שאלת המחקר. זמן הלימוד השבועי נע בין 0 ל-1305 , ועל פי לוח 3 ניתן לראות שרוב ההתפלגות נעה באזורים שבין 0-500 ובפרט בסביבות 100-200. מבחינת התחרותיות, ניתן לראות שקיימים ערכים שליליים וחיוביים , ועל פי לוח 4 רוב ההתפלגות נעה בין מינוס אחת ועד לאחת. ניתן לראות גם שמספר הספרים בבית מחולק לשש קטגוריות.

בלוח 5 בחרתי להציג את הקשר בין זמן הלימוד השבועי (בדקות) במקצוע המתמטיקה לבין הציון שהתקבל במתמטיקה במבחן PISA. ניתן ללמוד מהגרף שהקשר הוא חיובי, כלומר ככל שהתלמיד השקיע יותר זמן לימוד שבועי כך הציון שהוא קיבל היה גבוה יותר.

**לוח 1: אפיונים סטטיסטיים שונים של התלמידים במדגם**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **המשתנה** | **מס' תצפיות** | **ממוצע** | **סטיית תקן** | **מינימום** | **מקסימום** |
| ציון במתמטיקה | 4629 | 429.395 | 90.407 | 138.539 | 740.258 |
| האם יש חיבור אינטרנט ביתי | 4629 | 1.059 | .235 | 1 | 2 |
| זמן לימודי שבועי (בדקות) במתמטיקה | 4629 | 193.072 | 50.722 | 0 | 1305 |
| תחרותיות | 4629 | .158 | .853 | -2.345 | 2.005 |
| מס' ספרים בבית | 4629 | 2.286 | 1.201 | 1 | 6 |
|  | | | | | | | |

**לוח 2: תרשים היסטוגרמה של הציון במתמטיקה**



**לוח 3: תרשים היסטוגרמה של זמן הלימוד השבועי (בדקות) במקצוע המתמטיקה**



**לוח 4: תרשים היסטוגרמה של משתנה התחרותיות**



**לוח 5: גרף שמתאר את הקשר בין זמן הלימוד השבועי(בדקות) במקצוע המתמטיקה לבין הציון במתמטיקה במבחן PISA**



המודל האקונומטרי:

זמן לימוד שבועי

האם יש חיבור אינטרנט ביתי

הציון במתמטיקה

תחרותיות

מספר הספרים בבית

**לוח 6: תוצאות הרגרסיות**

|  |  |  |
| --- | --- | --- |
|  | (1) | (2) |
| **משתנים** | **מודל מצומצם** | **מודל מלא** |
|  |  |  |
| יש חיבור לאינטרנט | 68.18\*\*\* | 53.75\*\*\* |
|  | (5.561) | (5.314) |
| זמן לימוד שבועי |  | 0.0552\*\* |
|  |  | (0.0244) |
| תחרותיות |  | 10.11\*\*\* |
|  |  | (1.452) |
| 11-25 ספרים בבית |  | 20.92\*\*\* |
|  |  | (3.219) |
| 26-100 ספרים בבית |  | 49.08\*\*\* |
|  |  | (3.269) |
| 101-200 ספרים בבית |  | 64.31\*\*\* |
|  |  | (4.869) |
| 201-500 ספרים בבית |  | 108.5\*\*\* |
|  |  | (6.778) |
| מעל 500 ספרים בבית |  | 55.63\*\*\* |
|  |  | (9.732) |
| החותך | 365.2\*\*\* | 337.5\*\*\* |
|  | (5.395) | (6.954) |
|  |  |  |
| מספר התצפיות | 4,629 | 4,629 |
| R-squared | 0.031 | 0.134 |

בסוגריים- סטיות תקן

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

באמידה של המודל המצומצם- התקבל שהתלמידים שהיו בעלי אינטרנט ביתי קיבלו 68 נקודות יותר מאשר אלו שהיו ללא אינטרנט. כלומר על פי אמידה זו, השימוש באינטרנט הביא לעלייה בציון במתמטיקה ולכן ניתן להניח ששימוש זה היה חיובי ולצרכי למידה. אומד זה יצא מובהק ברגרסיה ברמת מובהקות של 5%.

באמידה של המודל המלא- האפקט של האינטרנט הביתי פחת, כעת התקבל שבעלי אינטרנט ביתי קיבלו 53 נקודות יותר מאשר אלו שהיו ללא אינטרנט. כלומר הקורלציה היא עדיין חיובית אך מתונה יותר. אומד זה יצא מובהק ברגרסיה ברמת מובהקות של 5%. כמו כן, התקבל שזמן הלימוד השבועי גם כן בעל קורלציה חיובית, כל תוספת של דקת למידה שבועית תורמת לעלייה של חצי נקודה בציון הסופי. אומד זה גם כן יצא מובהק ברגרסיה ברמת מובהקות של 5%. בנוסף, משתנה התחרותיות מובהק גם כן ברמת מובהקות של 5% ובעל קורלציה חיובית- עלייה של יחידה אחת במדד התחרותיות גורמת לעלייה של 10 נקודות בציון במתמטיקה. בנוגע למספר הספרים בבית המצב הוא יותר מעניין. כל הקטגוריות יצאו מובהקות ברמה של 5%, וכולן בעלות קורלציה חיובית. כלומר ככל שלתלמיד היו יותר ספרים בבית , כך הציון שלו במבחן עלה. תלמיד שהיו לו בבית בין 11-25 ספרים קיבל 20 נקודות יותר מאשר תלמיד שהיו לו רק 0-10 ספרים. תלמיד בעל 26-100 ספרים בביתו קיבל כבר 49 נקודות יותר מאשר תלמיד שהיו לו 0-10 ספרים, כלומר קצב הגידול עלה. הקצב עולה וממשיך גם בקטגוריות של 101-200 ספרים ו- 201 עד 500 ספרים, עם תוספת של 64 ו-108 נקודות בהתאמה , אך לתלמיד שהיו לו בבית יותר מ-500 ספרים המגמה מעט משתנה, והציון גדל רק ב -55 נקודות כלומר קצב הגידול פחת. ניתן להניח מכך שלכמות הספרים יש השפעה משמעותית עד רף מסוים, ומעליו האפקט הולך ונחלש. מה שמתאים גם לתפיסה המציאותית, שכן סביר להניח שתלמיד שבביתו יש מעל 500 ספרים, לא באמת יקרא את כולם ויוכל לשאוב מכולם ידע שיוכל לעזור לו לקבל תוצאות גבוהות יותר במבחן.

התוספת של משתני הבקרה אכן מוסיפה כוח להסבר של המודל. זה בעצם עזר לנקות חלק מההשפעות שהמודל המצומצם תפס דרך המשתנה המסביר המרכזי של הגישה לאינטרנט ביתי. כלומר עצם זה שהאפקט של הגישה לאינטרנט הביתי נחלש ב-15 נקודות מוכיח את הטענה הזו , ובעצם ניקה השפעות שוליות אחרות שנכללו והשפיעו על הקורלציה. דבר זה יכול לתרום להוכחה של סיבתיות, כלומר שהגישה לאינטרנט הביתי היא אכן משפיעה על הציון שהתקבל במבחן.

**פרק שני: משתנה תלוי בינארי**

המשתנה החדש שיצרתי הוא **math\_bin** אשר מקבל את הערך 1 אם הציון שהתלמיד קיבל במתמטיקה גבוה מהציון הממוצע (429.395**)** שמופיע בלוח 1 , ומקבל את הערך 0 אם הציון נמוך מהממוצע. על ידי שימוש בפקודת codebook ניתן לגלות שקיימים 2,309 תלמידים שציונם גבוה מהממוצע. לכן ניתן לחשב ששיעור y=1 במדגם הוא .

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

**לוח 7: תוצאות LPM**

|  |  |
| --- | --- |
|  | (1) |
| **משתנים** | **LPM** |
|  |  |
| יש אינטרנט ביתי | 0.238\*\*\* |
|  | (0.0301) |
| זמן לימוד שבועי | 0.000347\*\* |
|  | (0.000138) |
| תחרותיות | 0.050\*\*\* |
|  | (0.00822) |
| 11-25 ספרים בבית | 0.117\*\*\* |
|  | (0.018) |
| 26-100 ספרים בבית | 0.237\*\*\* |
|  | (0.018) |
| 101-200 ספרים בבית | 0.316\*\*\* |
|  | (0.027) |
| 201-500 ספרים בבית | 0.471\*\*\* |
|  | (0.038) |
| מעל 500 ספרים בבית | 0.256\*\*\* |
|  | (0.055) |
| החותך | 0.0568 |
|  | (0.039) |
|  |  |
| מספר התצפיות | 4,629 |
| R-squared | 0.094 |

בסוגריים-סטיות תקן

כל המקדמים ברגרסיה יצאו מובהקים ברמת מובהקות של 5%. ברגרסיית LPM המקדמים מייצגים את הקשר בין המשתנה המסביר להסתברות של . לכן הקשר בין חיבור האינטרנט הביתי לבין המשתנה התלוי הבינארי שיצרתי הוא ערכו של המקדם. כלומר לתלמידים שיש להם חיבור אינטרנט ביתי יש סיכוי של 23.8% יותר לקבל ציון גבוה מהממוצע מאשר לאלו שאין להם חיבור אינטרנט בבית. כלומר אינטרנט ביתי מעלה את ההסתברות לקבל ציון גבוה מהממוצע, ב-23.8%. בהתאם לרגרסיה של סעיף 3 אנו למדים שהקורלציה החיובית בין חיבור האינטרנט הביתי לבין הצלחה במבחן עדיין נשמרת. ההבדל המרכזי הוא שברגרסיה 3 האומדנים כולם מייצגים עלייה בניקוד ואילו ברגרסיית LPM האומדנים מייצגים הסתברויות. למשל לגבי מדד התחרותיות- על פי רגרסיית ה-LPM , כל נקודה נוספת במדד מעלה ב-5% את ההסתברות לקבל ציון גבוה מהממוצע במבחן. הקורלציה החיובית של כל המשתנים נשמרת. לגבי מספר הספרים בבית, קצב הגידול ההולך ופוחת נשמר גם כאן. כלומר עד רמה של בין 201-500 ספרים הגידול בתוספת להסתברות הולך וגדל, מ 11% ועד 47%, ואילו במעבר לקטגוריה של מעל ל-500 ספרים הגידול באחוזים פוחת ל -25% בלבד.

במקרה זה לא סביר להשתמש במודל LPM מפני שניתן לראות בהרצה של גרף כמו בלוח 8 , שקיימת בעיה של הטרוסקדסטיות והערכים מתפזרים מעבר ל- 0 או 1. זוהי בעצם בעיית השונות השונה. כיוון שהמשתנה התלוי הוא בינארי, על פי הגדרת השונות המותנית, השונות תלויה בסיכוי שהמשתנה התלוי שווה ל-1. וסיכוי זה הוא פונקציה של המשתנים המסבירים. לכן פעמים רבות יכול להיווצר מצב שנקבל ערכי הסתברויות שליליים או לחילופין מעל 1 – וזה דבר שהוא לא אפשרי. לכן יש לאמוד את המודל הנ"ל בעזרת רגרסיה שאינה ליניארית.

**לוח 8: גרף פיזור וקו רגרסיה בין זמן הלימוד השבועי לבין המשתנה התלוי הבינארי**



**לוח 9: תוצאות רגרסיית LOGIT**

|  |  |
| --- | --- |
|  | (1) |
| **משתנים** | **LOGIT** |
|  |  |
| יש חיבור אינטרנטי | 1.178\*\*\* |
|  | (0.157) |
| זמן לימוד שבועי | 0.00159\*\* |
|  | (0.000645) |
| תחרותיות | 0.221\*\*\* |
|  | (0.036) |
| 11-25 ספרים בבית | 0.502\*\*\* |
|  | (0.079) |
| 26-100 ספרים בבית | 0.993\*\*\* |
|  | (0.080) |
| 101-200 ספרים בית | 1.341\*\*\* |
|  | (0.125) |
| 201-500 ספרים בבית | 2.225\*\*\* |
|  | (0.215) |
| מעל 500 ספרים בבית | 1.087\*\*\* |
|  | (0.243) |
| החותך | -2.068\*\*\* |
|  | (0.199) |
|  |  |
| מספר התצפיות | 4,629 |

בסוגריים- סטיות תקן

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

על פי הרגרסיה, כל האומדים מובהקים ברמת מובהקות של 5%. ניתן לראות שהחותך שלילי. כמו כן ניתן לראות שההשפעה של כל המשתנים המסבירים היא חיובית, כלומר כל הבטות חיוביות. ניתן לראות שהקשר השולי של חיבור אינטרנטי ביתי הוא גדול יותר מאשר תחרותיות, ושניהם גדולים יותר מאשר זמן הלימוד השבועי. לגבי מספר הספרים בבית, שוב ניתן לראות את מגמת הגידול עד לקטגוריה של פחות מ-500 ספרים. לא ניתן להסיק מעבר לכך מפני שמדובר במודל לא לינארי ועל פי הנלמד בכיתה הבטא עצמה היא רק חלק מתוך ההשפעה של המשתנה המסביר על המשתנה התלוי. על פי הנוסחה:

כלומר ערך התרומה השולית תלוי ומשתנה לפי .

מדובר בחישוב ההשפעה החלקית הממוצעת. יש לחשב את ההשפעה השולית עבור כל פרט ופרט במדגם, ואז לקחת את הממוצע של ההשפעות השוליות הללו. כלומר נריץ את רגרסיית LOGIT מחדש, לאחר מכן נחשב את לכל פרט ופרט בעזרת פונקציית Predict. בהתאם לכך נחשב את , ולאחר מכן ניצור משתנה אשר *ערכיו מכילים את הערך המתקבל מהנוסחה הבאה :*

*לבסוף נסכום את המשתנה החדש ונגלה כי הערך שהתקבל הוא 0.266. כלומר ההשפעה השולית הממוצעת של אינטרנט ביתי זמין היא 26.6%.*

*מצורף צילום מסך של קובץ Stata עם הפעולות הנ"ל.*

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

**\*סעיף e בתחתית ה-Do File.**

**פרק שלישי: אנדוגניות- נתוני פאנל או IV**

סיבות אפשריות לאנדוגניות של המשתנה המסביר שבחרתי הם טעויות מדידה או משתנים מושמטים. כלומר חסרים במודל משתנים שמתואמים עם המשתנה הקיים. למשל, חיבור אינטרנטי ביתי הוא דבר שעולה כסף ומשאבים. סביר להניח שברוב המקרים במדגם שבהם לפרט לא היה אינטרנט ביתי , הדבר נובע מבעיות כספיות בבית. לדוגמא, ישנם מחקרים[[2]](#footnote-2) שטוענים שלאנשים עשירים יש גישה לאינטרנט מהיר ולעומת זאת לאנשים עניים יש גישה לאינטרנט חלש או בכלל לא. כלומר הכנסת הורים , אפילו השכלתם , הם דוגמאות למשתנים שיכולים להשפיע ולגרום למשתנה המסביר שבחרתי להיות מתואם עם השגיאה במודל ובכך לגרום לאנדוגניות. כמו כן, בקרב אלו שיש להם אינטרנט ביתי, משתנים כמו חריצות או משמעת עצמית יכולים גם כן להוות גורמים שמשפיעים על המודל ובכך גורמים לבעיית אנדוגניות.

המשתנים המסבירים הנוספים אכן מסייעים בפתרון בעיית האנדוגניות. ההוכחה לכך היא שכאשר הרצתי את המודל המלא בסעיף 3, המקדם של המשתנה המסביר העיקרי ירד ב-15 מאשר במודל המצומצם. כלומר ההשפעה של חיבור אינטרנט ביתי על ההצלחה במבחן במתמטיקה ירדה ב-15 נקודות. ישנם עוד משתנים שאינם כלולים בבסיס הנתונים, אך אם היו קיימים היו עוזרים לשפר את בעיית האנדוגניות- כמו משמעת עצמית או הכנסת הורים חודשית. אך לדעתי לא ניתן לעולם יהיה לפתור את בעיית האנדוגניות בצורה מושלמת, כי תמיד יהיו עוד משתנים וגורמים שיכולים להשפיע בדרך כזו או אחרת על המודל.

הפתרון החלקי לבעיית האנדוגניות שאבחר להשתמש בו הוא שימוש במשתני עזר ואמידה בשני שלבים 2SLS.

בחרתי בפתרון זה מפני שבהינתן הנתונים הזמינים ניתן להוסיף עוד משתני עזר שקיימים בקובץ הנתונים והם יכולים להיות מתואמים עם המשתנה המסביר האנדוגני אך לא מתואמים עם ההפרעות האקראיות. הדבר עונה על בעיית ההשמטה , ובעצם בצורה זו ניתן יהיה לצמצם את המתאם בין המשתנה המסביר האנדוגני לבין ההפרעה ולהתקרב לאפקט האמיתי שלו על המשתנה התלוי.

ראשית, נבצע מבחן Hausman לבדיקת אנדוגניות. נרצה לבדוק האם המשתנה המסביר הוא אנדוגני או לא. לכן בחרתי 2 משתני עזר מתוך מסד הנתונים – **escs** משתנה רציף שמהווה אינדקס לסטטוס כלכלי חברתי ותרבותי של הפרט, ומשתנה נוסף **gfofail** שמהווה מדד לפחד של הפרט במדגם מכשלון. לאחר שבחרתי את משתני העזר, יש לאמוד את המשוואה המצומצמת:

כלומר את המשתנה המסביר כפונקציה של משתני העזר.

לאחר מכן, יש לחשב את סדרת הטעויות של הרגרסיה בעזרת פקודת Predict , ואז לחשב שוב את המודל המלא המקורי בתוספת השאריות, כלומר יש להריץ את המודל הבא:

כעת יש לבדוק את ההשערה האם המקדם של מובהק או לא. על פי לוח 10 ניתן לראות שהמקדם אכן מובהק , ולכן השארית המחושבת מהמשוואה המצומצמת מסבירה את המשתנה המוסבר במודל המקורי. לכן השמטתה מכניסה את השפעתה לשאריות במודל המקורי וכתוצאה מכך נוצר מתאם בין המשתנה המסביר הנבדק לבין השאריות במודל המקורי- כלומר המשתנה אנדוגני.

כעת לאחר שווידאתי שהמשתנה אכן אנדוגני, נבצע את האמידה בשני שלבים. נאמוד מודל שבו המשתנה האנדוגני הוא המוסבר , והמשתנים המסבירים הם משתני העזר וגם המשתנים מהמודל המקורי. כלומר נריץ את המודל הבא:

למודל זה נחשב את התחזיות, ואז בעזרת פקודת ivreg נריץ את המודל המקורי תוך שימוש בערכים החזויים מהשלב הראשון.

את תוצאות הרגרסיה ניתן לראות בלוח 11.

**לוח 10: תוצאות הרגרסיה לצורך מבחן Hausman**

|  |  |
| --- | --- |
|  | (1) |
| **משתנים** | **רגרסיה** |
|  |  |
| יש חיבור אינטרנט ביתי | 364.9\*\*\* |
|  | (18.38) |
| זמן לימוד שבועי | 0.0416\* |
|  | (0.023) |
| תחרותיות | 9.033\*\*\* |
|  | (1.413) |
| 11-25 ספרים בבית | 10.04\*\*\* |
|  | (3.189) |
| 26-100 ספרים בבית | 27.34\*\*\* |
|  | (3.394) |
| 101-200 ספרים בבית | 35.10\*\*\* |
|  | (5.019) |
| 201-500 ספרים בבית | 68.37\*\*\* |
|  | (6.946) |
| מעל 500 ספרים בבית | 19.08\* |
|  | (9.752) |
| השאריות החזויות | 340.1\*\*\* |
|  | (19.25) |
| החותך | 61.00\*\*\* |
|  | (17.12) |
|  |  |
| מספר התצפיות | 4,592 |
| R-squared | 0.187 |

בסוגריים-סטיות תקן

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

**לוח 11: תוצאות 2SLS**

|  |  |
| --- | --- |
|  | (1) |
| **משתנים** | **2SLS** |
|  |  |
| האם יש חיבור אינטרנטי ביתי | -344.8\*\*\* |
|  | (22.33) |
| זמן לימוד שבועי | 0.020 |
|  | (0.031) |
| תחרותיות | 10.22\*\*\* |
|  | (1.877) |
| מספר ספרים בבית | 15.14\*\*\* |
|  | (1.422) |
| החותך | 754.6\*\*\* |
|  | (26.14) |
|  |  |
| מספר התצפיות | 4,592 |

בסוגריים – סטיות תקן

לפי התוצאות של האמידה בשני שלבים, האפקט של החיבור האינטרנטי השתנה ונהפך לשלילי. לעומת תוצאות ה-OLS שלפיהן לחיבור לאינטרנט ביתי הייתה קורלציה חיובית. משמעות הבדל זה היא שלאחר הניקוי של חלק נוסף מההשפעות האנדוגניות, מתגלה היפוך עצום בהשפעה. אם זאת עדיין יש לקחת תוצאות אלו בערבון מוגבל, שכן לא ניתן היה להריץ את המודל בצורה זו ולקבל את ההבחנה המדויקת בין אלו שנהנים מחיבור ביתי לבין אלו שלא, לעומת מודל ה-OLS שבו ניתן היה לבצע הבחנה זו על שימוש ב- i. במהלך הרצת הרגרסיה. באמידה הנוכחית כל האומדים יצאו מובהקים ברמה של 5% , למעט המקדם של זמן הלימוד השבועי שיצא לא מובהק. זאת בניגוד לתוצאות ה-OLS בהן כל האומדים היו מובהקים ברמה של 5%. ניתן לשייך את זה לעובדה שיכול להיות שמשתני העזר שבחרתי מתואמים עם המשתנה של זמן הלימוד השבועי, ולכן נגרמה הטיה זו. הדבר מתאים גם לתיאוריה מציאותית שלפיה משתנה העזר של הפחד מכשלון יכול להשפיע ולגרום לתלמיד להשקיע יותר או פחות זמן לימוד בשבוע. כמו כן, גם משתנה העזר של המצב הכלכלי חברתי ותרבותי של הפרט הוא גם דבר שיכול להשפיע על זמן הלימוד, שכן תלמיד בעל רקע כלכלי נמוך יכול לעבוד במהלך השבוע בכדי לעזור לפרנסת משפחתו וכתוצאה מכך להשקיע פחות זמן שבועי בלימודים. בנוגע למשתנה התחרותיות, האפקט החיובי שלו נשמר מרגרסיית ה-OLS, וכך גם בנוגע למשתנה של מספר הספרים בבית. אך גם כאן כמו במשתנה המסביר העיקרי של המודל לא ניתן היה להריץ במודל ה- 2SLS את הקטגוריות בנפרד בעזרת השימוש ב- i. ולכן לא ניתן לדעת בוודאות את ההבדל בין כל קטגוריה וקטגוריה.

לסיכום, גיליתי שכמו בשאר המחקרים שקראתי עליהם טרם הכתיבה, אין תשובה חד משמעית בנוגע לשאלה כיצד חיבור אינטרנטי ביתי משפיע על הציונים של התלמידים במולדובה במקצוע המתמטיקה. הדבר תלוי בהמון גורמים נוספים, קיימת בעיה של אנדוגניות , ובכל מודל שנריץ נוכל להגיע בסופו של דבר לתוצאות שונות.

**ביבליוגרפיה:**

* Kim SY, Kim M-S, Park B, Kim J-H, Choi HG (2017), Edited by Stefano Federici, Universita degli Studi di Perugia, ITALY <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174878>

מקור זה מדבר על המחקר שבו השתמשתי בחלק הראשון של העבודה. זמן העיון בו היה כ-40 דקות, והוא נתן לי רקע לגבי הנושא שבחרתי והכווין אותי לגבי שאלת המחקר שניסחתי ואילו תוצאות הייתי מצפה לקבל בסוף העבודה.

* <https://www.oecd-ilibrary.org/sites/0a428b07-en/index.html?itemId=/content/component/0a428b07-en>

מקור זה עוסק בהסבר נוסף ומרחיב על המשתנים שקיימים במסד הנתונים, וכיצד הם נאספו. זמן העיון בו היה כחצי שעה , והוא עזר לי להבין כיצד לבחור את המשתנים המתאימים למודל שלי ואחרי זה כיצד אני יכול לפרש את הערכים שמופיעים בקובץ נתונים בצורה הנכונה.

* [Allan Holmes](https://publicintegrity.org/author/allan-holmes/), [Eleanor Bell Fox](https://publicintegrity.org/author/eleanor-bell-fox/), [Ben Wieder](https://publicintegrity.org/author/ben-wieder/) and [Chris Zubak-Skees](https://publicintegrity.org/author/chris-zubak-skees/), "The Center For Public Integrity, may 12 2016 <https://publicintegrity.org/inequality-poverty-opportunity/rich-people-have-access-to-high-speed-internet-many-poor-people-still-dont/>

מקור זה עוסק במחקר שבו השתמשתי בחלק האחרון של העבודה בנושא האנדוגניות. זמן העיון בו היה כחצי שעה והוא עזר לי להבין טוב יותר אילו משתני עזר אני מעוניין לבחור בחלק האחרון של העבודה.

1. **Kim SY, Kim M-S, Park B, Kim J-H, Choi HG (2017)** [↑](#footnote-ref-1)
2. [**Allan Holmes**](https://publicintegrity.org/author/allan-holmes/)**,**[**Eleanor Bell Fox**](https://publicintegrity.org/author/eleanor-bell-fox/)**,**[**Ben Wieder**](https://publicintegrity.org/author/ben-wieder/)**and**[**Chris Zubak-Skees**](https://publicintegrity.org/author/chris-zubak-skees/)**, "The Center For Public Integrity, may 12 2016**  [↑](#footnote-ref-2)