

Oktatási reformok hatása survey adatokon

Befolyásolják-e a reformok a mintát?

Adamecz-Völgyi Anna*

KRTK KTI és UCL Social Research Institute. E-mail: adamecz.anna@krtk.hu

2022. szeptember

JEL: I2, C21, J13

*A szerző köszöni az Országos Tudományos Kutatási Alapprogramok (OTKA) támogatását (PD 128850). Köszönöm a kitűnő kutatási asszisztensi segítséget Sarkadi-Nagy Lillának. Hálás vagyok a MKE 2021 konferencia a kötet tanulmányai számára szervezett napján kapott kommentekért, és köszönöm Elek Péter, Lindner Attila szerkesztő és két anonim lektor javaslatait, amelyek sokat segítettek a tanulmány fejlesztésében.

Absztrakt

Az iskolázottság hatásait oktatási reformok segítségével mérő irodalom nagy része felmérésekből származó adatokkal dolgozik. A tanulmány azt vizsgálja, okoz-e problémát ezekben a mérésekben, hogy az adatfelvételek mintáit nem azzal a céllal hozták létre, hogy egyes reformok kohorszaira reprezentatívak legyenek. A tanulmány az ELSA adatfelvétel és egy sokszor vizsgált oktatási reform segítségével vizsgálja, hogy a kötelező iskolalátogatási korhatár felemelése befolyásolja-e a felmérésben való részvételt, a lemorzsolódást és a válaszadást. Az eredmények azt mutatják, hogy bár a reform valóban befolyásolta a lemorzsolódást, és a mintát is pont úgy bővítették, hogy a reformnak kitett emberek száma mechanikusan megnőtt, a szokásos eljárással végzett becslési eredmények mindezek ellenére stabilak maradnak az egyes hullámok között. Az első hullámban azonban olyan változóknak is van különbség a szakadási pont két oldalán, amelyeket nem okozhatott a reform, és ezek valószínűleg torzítják a becsült hatás mértékét.

Education reforms as instrumental variables using survey data

Effects on survey samples

Abstract

Most of the literature on the causal effects of education using education reforms as instrumental variables makes use of survey data that would only cover a small sample of the population. However, survey samples are not meant to be representative of covering those born before vs. after the reform equally. This paper investigates whether this phenomenon might introduce a bias. I use the ELSA survey and exploit an education reform in England to test whether the reform affected the probability of survey participation, attrition and non-response. Results show that the reform indeed affected attrition, and a replacement wave increased the number of participants born right after (but not before) the reform. Usual reduced form estimates stay stable across waves. However, I find differences among those born before and after the reform in characteristics that are independent of the reform already in the first wave indicating that the estimated effects of the reform could be biased.

1. Bevezetés

Széles irodalom vizsgálja az iskolázottság hatását egyéni, gazdasági és társadalmi kimenetekre úgy, hogy instrumentális változóként oktatási reformokat használ. A leggyakrabban ilyen célra használt reformok úgy bővítették az iskolázottságot, hogy megemelték az iskolalátogatási korhatárt, azaz azt az életkort, ameddig az iskolakötelezettség tart. Ennek következtében a reform után született kohorszok átlagosan hosszabb ideig maradtak iskolában, mint a reform előttié, így a reform a kohorszok között exogén módon növelte az iskolázottságot. Ha empirikusan igazolható, hogy az iskolázottság a reform következtében valóban nőtt és néhány további identifikációs feltevés is teljesül, a reform használható az iskolázottság instrumentumaként (Angrist és Krueger 1991). Az instrumentum ebben az esetben azt ragadja meg, hogy valaki a reform bevezetése előtt (kontrol) vagy után (kezelt) született-e, azaz a születési idő exogenitását használja ki.

Ezen tanulmány nem célozza a szokásosan szükséges identifikációs feltevések kérdéseit, így ezek tárgyalásától eltekintek. Az identifikációs feltevések teljesülésén túl további problémák forrása lehet a kutatáshoz rendelkezésre álló adatok kérdése. A stratégia megköveteli, hogy a születés ideje véletlen legyen, ne lehessen manipulálni. Ez egyben azt is jelenti, hogy a születés idejének a felhasznált adatbázisban is véletlennek kell lennie, azaz az adatbázisnak a reform előtt és után születetteket ugyanolyan arányban kell lefednie, de legalábbis nem vihet a reform szisztematikus torzítást a mintába. Ha a reform előtt és után születettek összetétele olyan ismérvekben tér el, amelyeket nem befolyásolhatott a reform, és

amelyek nem függetlenek a vizsgált kimeneti mutatótól, akkor a minta torzítást visz a becslésbe (Lee és Lemieux 2010).

Adminisztratív adatbázisok esetén, amelyek egy ország teljes népességét vagy annak egy véletlenszerű mintáját fedik le, három probléma adódhat: ha a kimeneti mutató mérésének idején a reform után születettek nagyobb valószínűséggel vannak még életben, nagyobb valószínűséggel élnek még az adott országban, vagy más valószínűséggel hiányosak az adataik a mérendő kérdésben¹. Ezekben az esetekben a megfigyelések száma és a minta összetétele az adatbázisban megtörik, és már önmagában a mintavétel endogén lesz. Bár ebben a tanulmányban ezt a kérdést nem vizsgáljuk, valós probléma lehet: a magasabban iskolázott emberek tovább élnek (Roy és mtsai. 2020) és nagyobb valószínűséggel váltanak országot (Vakhitova és mtsai. 2013). Minél későbbi életkorban mérjük az adott kimeneteket, a probléma potenciálisan annál nagyobb lehet.

Survey adatok használata esetén ezeken kívül az a kérdés is felmerül, hogy befolyásolta-e a reform az adatfelvételben való részvétel valószínűségét, az abból való lemorzsolódást, vagy a mérendő kimenetre adott válaszadást. Ha igen, az további törést okoz a megfigyelések számában és a minta összetételében a reform előtti és utáni kohorszok között. Ez a probléma elméletileg fennállhat: magasabban iskolázott emberek nagyobb valószínűséggel vesznek részt ezekben a felmérésekben, sőt, általában a társadalmi-gazdasági háttér szerinti eloszlás legalját ezek a felmérések kifejezetten rosszul érik el. Ez a korhatár-reformok esetében különösen problémás lehet, mert az ilyen reformok hatása pont az eloszlás alján a legnagyobb (Brunello, Fort, and Weber 2009). Ha magasabb társadalmi háttérből származó emberek nagyobb valószínűséggel kerülnek a minta kezelt csoportjába, akkor az a különbség, amit a reform hatásának mérünk, részben a mintaösszetétel eltéréseiből adódhat.

A Függelék F.1. táblázata azokat a főbb tanulmányokat tekinti át, amelyek korhatár-reformokat használtak az iskolázottság instrumentálására és survey adatokon dolgoztak. A táblázat azt foglalja össze, hogy tárgyalják-e ezt a problémát, mint az identifikáció egy lehetséges veszélyét, illetve a cikkek vizsgálják-e a reform hatását a mintában való részvételre és/vagy a lemorzsolódásra. A vizsgálat keretében 14 folyóiratban publikált vagy IZA/NBER műhelytanulmány formájában közzétett tanulmány felelt meg a feltételeknek (Függelék F.1. táblázat). Ezek között négy olyan tanulmány van, amelyik megnézi, hogy a megfigyelések

¹ Például, ha egészségügyi kimeneteket akarunk mérni orvoslátogatási adatokból, és a reform előtt született (alacsonyabban iskolázott) emberek kisebb valószínűséggel mennek adott probléma esetén orvoshoz, akkor náluk kevesebb orvoslátogatást fogunk látni. Nem azért, mert egészségesebbek lennének, hanem mert betegség esetén is kisebb valószínűséggel mennek orvoshoz. Ez az iskolázottság hatásának alulbecsléséhez vezetne.

száma folytonos-e a reform által okozott szakadási pontban, egy sem vizsgálja azonban, hogy az általa használt reform befolyásolhatta-e a mintában való részvétel valószínűségét.

Jelen tanulmány elsőként vizsgálja, befolyásolja-e egy oktatási reform egy adatfelvétel mintájában való részvételt, a minta összetételét, a lemorzsolódást és a válaszdást. Egy sokat vizsgált oktatási reformot használok, az iskolalátogatási korhatár 1947-es emelését Angliában, illetve az ELSA (English Longitudinal Study of Ageing, J. Banks és mtsai. (2021)) felmérés első öt hullámának adatait. Nem csak azt vizsgálom, folytonos-e a megfigyelések száma a szakadási pontban, de azt is megnézem, hogy a reform befolyásolta-e a minta összetételét, a lemorzsolódást és a válaszdást. Annak ellenére, hogy a megfigyelések száma nem folytonos a szakadási pontban, és a mintát pont úgy bővítették a negyedik hullámban, hogy a közvetlenül a szakadási pont felett született emberek száma nőtt meg, illetve a reform csökkentette a lemorzsolódást az ötödik hullámban, a szokásos módon végzett becslések (a reform hatása az iskolázottságra és a gyermekek számára) az egyes hullámok között stabilak maradnak. A minta összetétele azonban már az első hullámban eltér a szakadási pont két oldalán olyan ismérvekben, amelyeket nem befolyásolhatott a reform, és ez okozhat torzítást a mért hatásokban.

A tanulmány további struktúrája a következő. A 2. fejezet bemutatja a felhasznált reformot, az adatbázist és az empirikus módszereket. A 3. fejezet részletezi, a 4. fejezet pedig megvitatja az eredményeket.

2. A reform, adatok és empirikus módszerek

Az 1947-es reform (Raising of the school-leaving age, RoSLA) 14 évről 15 évre emelte az iskolakötelezettség korhatárát Angliában. Az 1933. szeptemberében születettek alkotják az első kohorszot, akikre már a megemelt korhatár vonatkozott. Az ELSA adataiban, amit jelen tanulmány használ, csak a születési év elérhető, a születési hónap nem. Mivel az iskolakezdési kohorsz év közben (szeptemberben) vált, az 1933-ban születettek egy része a reform előtt, egy része a reform után született. Ebből adódóan az 1934-es születési évfolyam az első, akik közül már mindenkire a megemelt korhatár vonatkozott. Fort, Schneeweis, és Winter-Ebmer (2016)-ot követve ezt a születési évet tekintem a szakadási pontnak, és az 1933-as vegyes évfolyamot kidobom a mintából. A Függelék F.2. táblázata foglalja össze, amit a reform iskolázottágra és más kimenetekre kifejtett hatásáról az irodalomból tudunk.

Az ELSA mintája az 50 éven felüli lakosság (és partnereik) mintája, szerkezete és mutatói harmonizáltak az európai SHARE-rel (J. Banks és mtsai. 2021). Az első öt hullám adatait

használok, melyeket 2002 és 2011 között vettek fel. Az alapminta egy korábbi longitudinális felmérés (Health Survey for England, HSE) legalább 50 éves válaszadóiból állt, azaz olyan emberekből, akik a HSE-ben korábban már részt vettek. Bár a HSE háztartási szintű alapmintáját területi stratifikáció után véletlenszerűen választották, az ELSA-ba ebből csak olyan háztartások kerültek át, akik korábban már interjút is adtak. Ezek az emberek alkotják az alapmintát, akiket meghívtak ez első interjúra (17 591 fő). Közülük 65 százalék adott végül interjút az első (11 459 fő) hullámban, és 37 százalékuk (6 585 fő) maradt bent az ötödik hullámig (1. táblázat, A blokk).

A minta létszáma két módon bővült az ELSA-ban: ha új partner lépett a háztartásba, illetve, ha új embereket vontak be a mintába. Az első öt hullám alatt, amit vizsgálok, a harmadik és a negyedik hullámban vontak be új kohorszokat. A minta bővítésének célja, hogy pótolják azokat az embereket, akik lemorzsolódtak vagy elhunytak, illetve, hogy megtartsák az 50-52 éves emberek arányát (hiszen az első hullámban 50-52 évesek időközben idősebbek lettek). A lemorzsolódás és a bővülés együttes eredményeként a minta nagysága az első és az ötödik hullám között nagyjából 10 százalékkal csökkent, és az ötödik hullámban az első hullámban részt vettek 57 százaléka (az alapminta 37 százaléka) maradt benne (1. táblázat, A blokk).

1. táblázat: Az ELSA mintája

Adatgyűjtés éve	Hullám	Eredeti résztvevők	Lemorzsolódás az alapminta-hoz képest	Új emberek az alapmintához képest	Összes válaszadó a hullámban	Kumulált mintanagyság
A. Teljes minta						
	Alapminta	17591		-	-	
2002/2003	1. hullám	11459	35%	-	11459	11771
2004/2005	2. hullám	8995	49%	38	9033	11809
2006/2007	3. hullám	7767	56%	1685	9452	13457
2008/2009	4. hullám	6902	61%	3865	10767	15679
20010/2011	5. hullám	6585	63%	3453	10038	15255
B. A szakadási pont előtti 10 évben születettek						
	Alapminta	3988				
2002/2003	1. hullám	2601	35%		2601	2601
2004/2005	2. hullám	1982	50%	3	1985	2604
2006/2007	3. hullám	1658	58%	2	1660	2603
2008/2009	4. hullám	1416	64%	84	1500	2685
20010/2011	5. hullám	1223	69%	61	1284	2662
C. A szakadási pont utáni 10 évben születettek						
	Alapminta	5192				
2002/2003	1. hullám	3492	33%		3492	3492

2004/2005	2. hullám	2822	46%	7	2829	3499
2006/2007	3. hullám	2465	53%	35	2500	3527
2008/2009	4. hullám	2261	56%	794	3055	4286
20010/2011	5. hullám	2214	57%	683	2897	4175

Forrás: ELSA.

A harmadik hullámban olyanokat vontak be, akik az 50-es években születtek. Mivel a szakadási pont, ami körül a megfigyelések számát vizsgáljuk, 1934-ben van, ez a kohorsz csak abban az esetben befolyásolhatta a szakadási pont körül mért eredményeket, ha a bevont emberekkel egy háztartásban idősebb partnerek éltek. A negyedik hullámban azonban pont egy olyan kohorszot vontak be, akik a szakadási pont felett, 1934 és 1958 között születtek. A bővítés 2008/2009-ben történt, ekkor a szakadási ponthoz közel született emberek 74-75 évesek voltak. Elképzelhető tehát, hogy ez a bővítés olyan, erősen szelektált részmintával bővítette az alapminta még le nem morzsolódott halmazát, ami torzítást visz a reform mért hatásába. Az 1. táblázat B és C blokkja azok között mutatja a mintanagyság változását, akik a szakadási pontot megelőző (B) vagy követő (C) 10 évben születtek (azaz, akiknek a mintáján az elemzés történik). Látható, hogy a negyedik hullám bővítése főként a szakadási pont után születetteket érintette.

3. Empirikus módszerek

Az ELSA adatain négy kérdést vizsgálok. Először, hogy a mintában a megfigyelések száma születési évenként folytonos-e a szakadási pont körül. Az irodalomhoz hasonlóan az alapmodellben egy nemparametrikus eljárást, lokális lineáris regressziót használok (Hahn, Todd, és Van der Klaauw 2001; Imbens és Lemieux 2008), azaz a szakadási pont alá és fölé egy-egy lineáris függvényt illeszttek, és a két függvény a szakadási pont két oldalán felvett értékei közötti különbséget a reform hatásaként értelmezem. A becsült modell formálisan a következő:

$$y_t = \alpha + \beta_1 * reform_t + \beta_2 * relatív\ szülev_t + \beta_3 * reform_t * relatív\ szülev_t + u_t, (1)$$

ahol y_t azoknak a száma, akik t évben születtek; $reform_t$ egy bináris változó, ami a szakadási pontban, az 1934-es születési évben és felette 1, alatta pedig nulla; $relatív\ szülev_t$ a születési év 1934-hez képesti távolsága (1934-ben 0, alatta/felette mindig eggyel

csökken/nő), u_t pedig a szokásos robusztus hibatag. Ezeket a modelleket egy olyan aggregált mintán becslem, ahol egy megfigyelés egy születési év adatát tartalmazza, és a minta +/- 10 évet fed le a szakadási ponthoz képest, azaz 20 megfigyelésből áll. A β_1 becsült értéke ragadja meg a törés nagyságát a szakadási pont körül.

Mivel azonban a születési időt többek között például Fort, Schneeweis, és Winter-Ebmer (2016) és Fletcher (2015)-hez hasonlóan években mérem, hiszen az ELSA adataiban a születési évnél finomabb információ (hónap, nap) nem érhető el, a szakadási pontot meghatározó futó változó diszkrét (Lee és Lemieux 2010). Az irodalom nem egyértelmű abban a kérdésben, hogy diszkrét futó változó esetén nemparametrikus eljárások (mint az (1) egyenletben leírt lokális regresszió vagy a hasonló esetekben gyakran használt McCrary (2008) teszt, ami szintén lokális lineáris regressziót használ a szakadási pont két oldalán) használhatók-e (lásd Lee és Card (2008) vs. Dong (2015)). Robusztusság vizsgálatként ezért megismétlem a becsléseket egy olyan globális parametrikus eljárással is, ami korigálja a diszkrét futó változó okozta esetleges specifikációs torzítást (Dong 2015). Ebben a születési év negyedfokú függvényére kontrollálok, az előzőekhez hasonlóan külön a szakadási pont alatt és felett, majd a Dong (2015) által javasolt módon állítom elő a becslésekből a kerekítési torzítással korigált becsült hatást. Mivel azonban ez az eljárás feltételezi, hogy a megfigyelések eloszlása születési éven belül egyenletes, ez azonban nem feltétlenül igaz (ez év egyes hónapjaiban többen születnek), ezeket alternatív eredménynek tekintem. A becsült függvény a következő:

$$y_t = \alpha + \beta_1 * reform_t + \sum_{i=1}^4 \beta_{2i} * relatív\ szülev_t^i + \sum_{i=1}^4 \beta_{3i} * reform_t * relatív\ szülev_t^i + u_t \quad , \quad (2)$$

ahol Dong (2015) módszerét követve a törés nagyságának torzításmentes becslését a $\beta_1 - \frac{1}{2} * \beta_{31} + \frac{1}{6} * \beta_{32} - \frac{1}{30} * \beta_{34}$ paraméter adja meg (ahol a β_3 paraméter második alsó indexe a hatványkitevő).

További problémát okoz, hogy a populációhoz képest az ELSA (vagy bármilyen más adatfelvétel) mintája nagyon kicsi, kb. fél százalékos, ezért esetleges, hogy ki az, aki belekerül, és ki az, aki nem. Ilyen típusú valószínűségi változók eloszlásának leírására a Poisson eloszlás használatos, ezért két további becslésben a függő változó eloszlására ezt teszem fel. Az (1) és a (2) egyenletet megbecslem tehát lineáris regresszió mellett Poisson regresszióval is az alábbiak szerint:

$$\ln E(y_t) = \alpha + \beta_1 * reform_t + \beta_2 * relatív\ szülév_t + \beta_3 * reform_t * relatív\ szülév_t + u_t, \quad (3)$$

és

$$\ln E(y_t) = \alpha + \beta_1 * reform_t + \sum_{i=1}^4 \beta_2 * relatív\ szülév_t^i + \sum_{i=1}^4 \beta_3 * reform_t * relatív\ szülév_t^i + u_t, \quad (4)$$

ahol y_t várható értékének logaritmus ($\ln E(y_t)$) Poisson eloszlást követ. Az eredményeket a 2. táblázat mutatja be, és mind a négy módszer hasonló eredményre vezet.

Másodszor, megnézem, hogy a reform befolyásolta-e a lemorzsolódás és az egyes gyakran vizsgált kérdésekre (oktatás, fertilitás) való válaszadás valószínűségét. Ezekben a modellekben a kimeneti változó bináris, ezért lineáris valószínűségi modelleket használok, továbbra is az elsőként leírt lokális lineáris regressziós keretben, de egyéni szintű adatokon:

$$y_i = \alpha + \beta_1 * reform_i + \beta_2 * relatív\ szülév_i + \beta_3 * reform_i * relatív\ szülév_i + u_i, \quad (5)$$

ahol y_t bináris változó, ami azt ragadja meg, hogy i résztvevő lemorzsolódott-e/válaszolt-e adott kérdésre; $reform_i$ egy bináris változó, ami a szakadási pontban, az 1934-es születési évben és felette 1, alatta pedig nulla; $relatív\ szülév_i$ a születési év 1934-hez képesti távolsága (1934-ben 0, alatta/felette mindig eggyel csökken/nő), u_i pedig a szokásos robusztus hibatermék, Lee és Lemieux (2010) nyomán születési év szerint klaszterezve. A β_1 becsült értéke ragadja meg a reform hatását a lemorzsolódás és az egyes kérdésekre történő válaszadás valószínűségére.

Harmadszor, megnézem, hogy a minta összetétele változik-e a szakadási pont körül olyan ismérvek tekintetében, amelyeket nem befolyásolhatott a reform (de a tanulmányban vizsgált kimenetekkel, az iskolázottsággal vagy a fertilitással kapcsolatba hozhatók). Ha a mintavétel vagy a lemorzsolódás nem különbözik szisztematikusan a szakadási pont két oldalán, ilyen ismérvekben nem lehet különbség. Ha a szakadási pont két oldalán olyan ismérvekben is különböznenének az emberek egymástól, amelyeket a reform nem befolyásolhatott és amelyek befolyásolhatják az iskolázottságot vagy a fertilitást (vagy azt a kimeneti mutatót, amit a kutató mérni szeretne), azt jelentené, hogy a minta a reform hatásainak megbecslésére (legalábbis szakadásos regressziós keretben) nem alkalmas. Ha az eltérés már az alapmintában is meglenne, akkor utalhatna arra, hogy bizonyos típusú emberek

nagyobb valószínűséggel egyeztek bele az adatfelvételbe való bekerülésre (és ez torzíthatja a mért hatást, amennyiben ezek az ismérvek befolyásolják az iskolázottságot vagy a fertilitást). Ha később lenne eltérés, akkor a lemorzsolódás vagy a mintabővítés vihetett volna torzítást a mintába. Három olyan változót tartalmaz az ELSA, amelyek befolyásolhatják az iskolázottságot vagy a gyermekvállalást, de a reformtól függetlenek. Kettő a gyermekkori családi háttér mutatója (szülői társadalmi háttér, könyvek száma gyermekkorban otthon), a harmadik a nők biológiai fertilitásával lehet kapcsolatos (az első menstruáció életkora).²

Negyedszer, megpróbálom megítélni, hogy a mintabővítés és a lemorzsolódás, illetve az alapminta kialakítása okozhat-e torzítást a reform hatásait mérő becslésekben. Attól függetlenül, hogy a minta összetétele nem változik a szakadási pont körül megfigyelhető ismérvek tekintetében az egyes hullámok között, nem megfigyelhető ismérvekben még változhat. Nem megfigyelhető ismérveket nem lehet tesztelni, azt viszont lehet vizsgálni, különbözik-e a reform mért hatása attól függően, hogy melyik hullám adatait használjuk.³ Ezért a következő lépésben újrabecslem a reform az elvégzett iskolaévekre és a gyermekek számára (befejezett fertilitásra) kifejtett hatását, amelyet az irodalomban már sokszor dokumentáltak (Függelék F.2. táblázat). A becsléseket minden egyes hullámban elvégzem. Ha a lemorzsolódás és a mintabővítés nem vitt szisztematikus torzítást a mintába, az eredményeknek nem szabad különbözni az egyes hullámokban. Mivel a reform hatása a lemorzsolódásra az ötödik hullámban volt először szignifikáns, a minta bővítése pedig a negyedik hullámban növelte a reform után születettek számát (és okozhatott szelekciót azzal, hogy a kezelt csoport átlagosan fiatalabb lett), ha ez ebből fakadó torzítás számottevő lenne, akkor a negyedik-ötödik hullámban más eredményeket kapnánk a reform hatására, mint az első háromban.

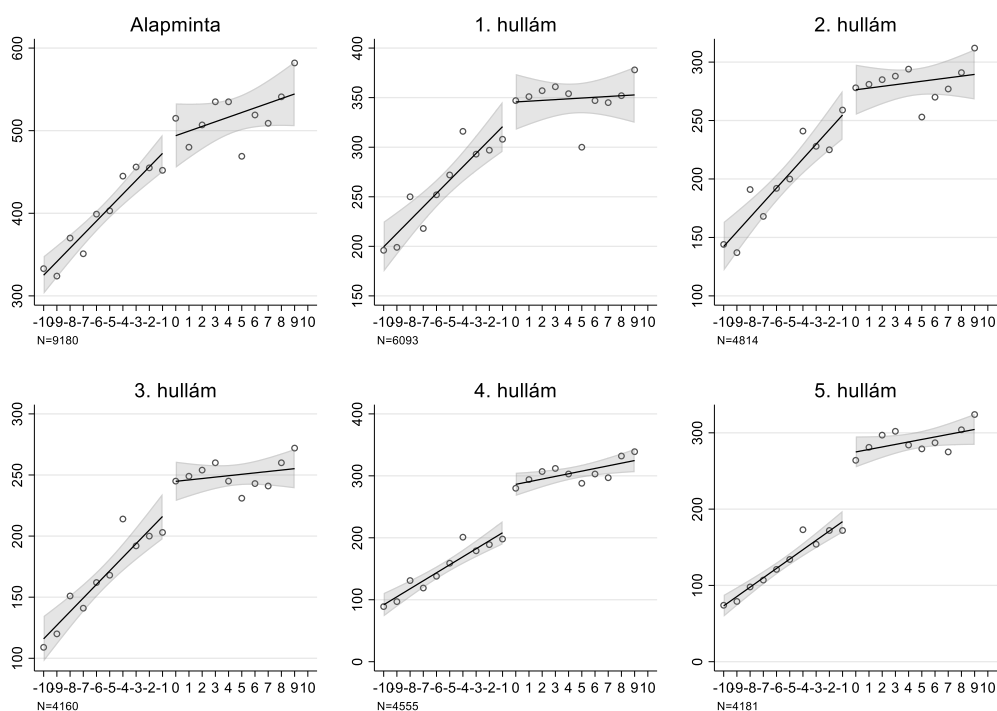
A háttérváltozók szakadási pont körüli folytonosságát és a reform iskolázottságra és a gyermekek számára kifejtett hatását az (5) egyenlethez hasonló lokális lineáris regressziókkal tesztelem, azzal a különbséggel, hogy ezekben az esetekben a függő változó nem bináris, hanem folytonos. Érdekes hozzájárulás az irodalomhoz, hogy ezeket a modelleket nőkre és férfiakra külön is megbecslem. Iskolai reformok hatását férfiak fertilitására tudomásom szerint még nem dokumentálta az irodalom.

²Sajnos ezek a változók nem elérhetők a teljes mintára, ezért ezekben a megfigyelések száma alacsonyabb a fő becslésekhez képest. A szerző tesztelte, hogy ezeket az adathiányokat nem befolyásolta a reform, de ezeket a számításokat a tanulmány nem tartalmazza.

³ Ez a módszer arra nyilvánvalóan nem alkalmas, hogy az alapmintában vizsgálja az esetleges torzítást.

Végül azt is megpróbálom megítélni, hogy bár az egyes hullámok között a becült hatások nagysága stabil, lehet-e már az első hullámban becült hatás torzított. Két olyan ismérven találok különbséget a szakadási pont körül, amit nem okozhatott a reform: a férfiak nagyobb valószínűséggel származnak alacsony családi háttérből a szakadási pont jobb oldalán, mint a bal oldalán, a nők között pedig gyermekkorban a könyvek számában van különbség. Először megnézem a reform heterogenitását a két változó mentén az (5) egyenlet segítségével, majd megvizsgálom a becült hatás változását ezekre való kontrollálás után.

1. ábra: A megfigyelések száma az ELSA egyes hullámaiban a szakadási pont körül



Forrás: saját becslés az ELSA-ból. Az x tengely a szakadási ponthoz viszonyítva jelzi a születési éveket (0=1934). Az adatpontok az adott évben született emberek számát mutatják a mintában, amelyekre a szakadási pont alatt és felett lineáris függvény illesztettem, 95%-os konfidencia-intervallummal. Az illesztett egyesek az (1) egyenlet grafikus reprezentációi.

4. Eredmények

4.1 A megfigyelések száma, lemorzsolódás és válaszadás

Az 1. ábra a megfigyelések számát mutatja a szakadási pont körül az ELSA egyes hullámaiban. Szembetűnő, hogy főként a negyedik és az ötödik hullámban van jelentős törés a megfigyelések számában a szakadási pontban (ami részben a célzott mintabővítésből adódik), míg az előző hullámoknál egyéb olyan pontokat is lehet találni, ahol a megfigyelések száma megugrik.

Ha az előző fejezetben ismertetett regressziós módszerekkel vizsgáljuk a töréspont nagyságát, hasonló képet látunk (2. táblázat): a negyedik és az ötödik hullámban a megfigyelések száma minden módszer szerint szignifikánsan megugrik a szakadási pont körül. Az lineáris és a Poisson regressziós becslések között ((1) és (3) egyenlet) nincs nagy különbség, mindkettő szerint csak a két utolsó hullámban szignifikáns a törés nagysága. Az lineáris modellben 65-79 fővel, a Poisson modellben 0,20-0,27 logponttal nagyobb a megfigyelések száma a szakadási pont jobb oldalán, mint a bal oldalán. A Dong (2015) módszertan szerinti becslések eredménye hasonló a két utolsó hullámot illetően, azonban ezekben a becslésekben az alapmintában és a harmadik hullámban is szignifikáns a törés.

2. táblázat: A megfigyelések száma a reform függvényében az ELSA egyes hullámaiban

	(1)	(2)	(3)	(4)	(5)	(6)
	Alapminta	1. hullám	2. hullám	3. hullám	4. hullám	5. hullám
Lokális lineáris regresszió, OLS ((1) egyenlet)						
A törés nagysága a szakadási pont körül (β_1 , fő)	5.012 (18.381)	11.455 (13.008)	9.061 (10.632)	17.618 (10.611)	65.206*** (10.692)	79.079*** (11.166)
Globális parametrikus modell Dong (2015) alapján, OLS ((2) egyenlet)						
A törés nagysága a szakadási pont körül (β_1 , fő)	87.776* (30.781)	53.996 (38.652)	-7.491 (36.581)	61.580* (27.590)	93.963* (31.714)	97.191** (24.250)
Kerekítéssel korrigált együttható ($\beta_1 - \frac{1}{2} * \beta_{31} + \frac{1}{6} * \beta_{32} - \frac{1}{30} * \beta_{34}$, fő)	75.807* (32.713)	29.577 (25.867)	0.619 (20.153)	38.409* (15.284)	67.030** (17.846)	74.227*** (12.824)
Lokális lineáris regresszió, Poisson regresszió ((3) egyenlet)						
A törés nagysága a szakadási pont körül (β_1 , logpont)	-0.005 (0.038)	0.009 (0.044)	-0.001 (0.043)	0.036 (0.057)	0.199** (0.061)	0.267*** (0.067)
Globális parametrikus modell Dong (2015) alapján, Poisson regresszió ((4) egyenlet)						
A törés nagysága a szakadási pont körül (β_1 , logpont)	0.182*** (0.054)	0.163 (0.100)	-0.051 (0.129)	0.288** (0.105)	0.402** (0.129)	0.434*** (0.107)
Kerekítéssel korrigált együttható ($\beta_1 - \frac{1}{2} * \beta_{31} + \frac{1}{6} * \beta_{32} - \frac{1}{30} * \beta_{34}$, logpont)	0.156*** (0.049)	0.093 (0.060)	-0.006 (0.061)	0.182** (0.054)	0.303*** (0.063)	0.360*** (0.050)
Megfigyelések száma	20	20	20	20	20	20

Forrás: saját becslés az ELSA adataiból. Zárójelben robusztus sztenderd hibák. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Bandwidth: 10-10 születési év a szakadási pont alatt és felett.

A kép akkor is hasonló, ha nem az egyes hullámok mintáit, hanem a hullámok kumulált mintáit vizsgáljuk (Függelék F.5. táblázat). A kumulált hullámokat azért érdemes nézni, mert a mintaelemszám maximalizálása érdekében gyakran úgy használják az ilyen típusú adatokat, hogy az egyes hullámok mintáit egyesítik. Az ilyen mintákban minden olyan ember pontosan egyszer szerepel, aki legalább egy hullámban szerepelt. Természetesen ez csak akkor lehetséges, ha az elemzéshez szükséges változókat minden hullámban felvették; a születési év, az iskolázottság és a gyermekek száma ilyen változók. Fort, Schneeweis, és Winter-Ebmer (2016) mintája például az első öt hullám kumulált mintája, azaz minden olyan nő szerepel benne, aki legalább egy hullámban részt vett.

Ahogy ez korábban elhangzott, két úton jöhet létre törés a megfigyelések számában a szakadási pont körül: ha a reform befolyásolta a lemorzsolódást az alampintából (a reform után született emberek nagyobb valószínűséggel maradtak a mintában), illetve, ha a reform után született emberek nagyobb valószínűséggel kerültek bele a felmérésbe valamelyik későbbi hullámban (amit a negyedik hullám bővítéséről tudunk is). A Függelék F.3. táblázata azt vizsgálja, befolyásolta-e a reform az alampintán belül annak a valószínűségét, hogy valaki az első öt hullámban interjút adott. Itt tehát kizárólag azokra az emberekre szűkítem a mintát, akik az alampintában már benne voltak, és a mintabővítés hatását kizárom. Az első négy hullámban nincs a reform és a mintában maradás között szignifikáns kapcsolatot, bár a becslést együttható nagyságrendje hullámról-hullámra kissé nő. Az ötödik hullámban azonban már szignifikáns a reform és a részvétel közötti kapcsolat mintabővítés nélkül is: ekkorra a reform 5,5 százalékponttal növelte a mintában maradás valószínűségét (Függelék F.3. táblázat).

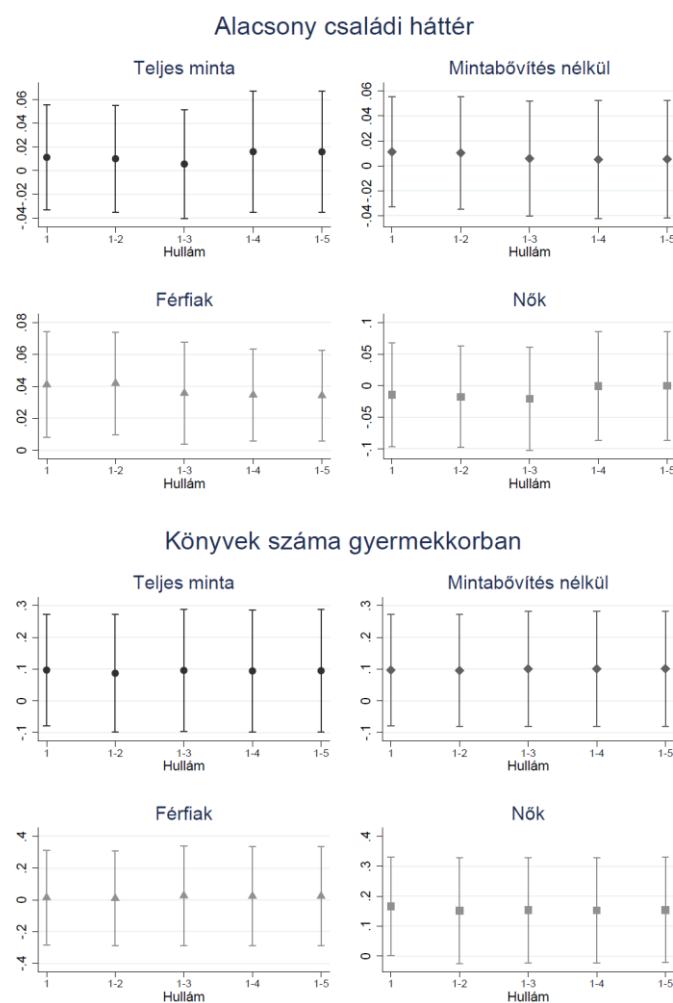
A Függelék F.4. táblázata azt vizsgálja, hogy a 2-5. hullámban interjút adók között nagyobb arányban voltak-e újonnan bekerült emberek azok között, akik a reform után születtek. A második és harmadik hullámban nem látunk ilyen kapcsolatot, a negyedik-ötödik hullámban viszont igen, ami nem meglepő, hiszen a negyedik hullámban közvetlenül a szakadási pont után született emberekkel bővítették a mintát.

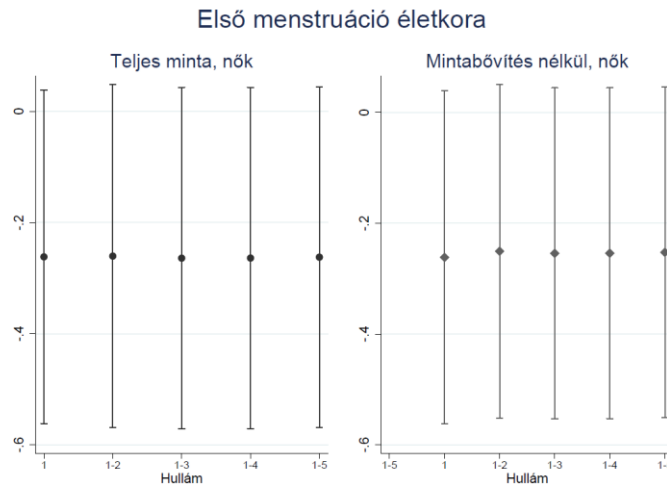
A reform az elvégzett osztályok számára való válaszadást egyik almintában sem befolyásolta, a gyermekek számára vonatkozó kérdésre azonban némileg csökkentette a válaszadás valószínűségét a nők és a férfiak között is, de a teljes mintában ezek együtthatók nem szignifikánsak (Függelék F.1. ábra). Elképzelhető, hogy a válaszmegtagadás növekedése a fertilitásra kifejtett negatív hatásból adódik.

4.2 A minta összetétele a szakadási pont körül

A 2. ábra azt vizsgálja, eltér-e a minta összetétele a szakadási pont két oldalán olyan ismérvekben, amelyeket nem befolyásolhatott a reform, és befolyásolhatják az iskolázottságot vagy a gyermekvállalást. Az alacsony gyermekkori családi háttér tekintetében átlagosan nincs különbség a mintában a szakadási pont két oldalán, kivéve a férfiak almintáját. A férfiak között nagyobb az alacsony családi háttérből származók aránya a szakadási pont felett, mint alatta, és elméletileg ezt nem okozhatta volna a reform (hacsak a gyermekkori emlékek

2. ábra: A minta összetétele a szakadási pont körül az ELSA kumulatív hullámaiban





Forrás: saját becslés az ELSA-ból. Minden ábrázolt együttható külön becslésből származik az (5) egyenlet alapján. A 95 százalékos konfidencia-intervallumok születési év szerint klaszterezett robusztus sztenderd hibákból származnak. Bandwidth: 10-10 születési év a szakadási pont alatt és felett.

visszahívását nem befolyásolta). A könyvek száma gyermekkorban otthon átlagosan nem különbözik a szakadási pont két oldalán, kivéve a nőknél az első hullámban (a becslt együttható értéke a kumulált hullámokban is hasonló nagyságrendű, de nem szignifikáns). Hasonlóképpen, az első menstruáció életkora némileg alacsonyabb a szakadási pont felett, de ez a különbség öt százalékos szignifikancia szinten nem szignifikáns. Ez adódhat abból is, hogy az időben az első menstruáció életkora csökken, de mivel az együtthatók 90 százalékos szinten sem szignifikánsak, ennek vizsgálatától eltekintek.

4.3 A torzítás mértéke

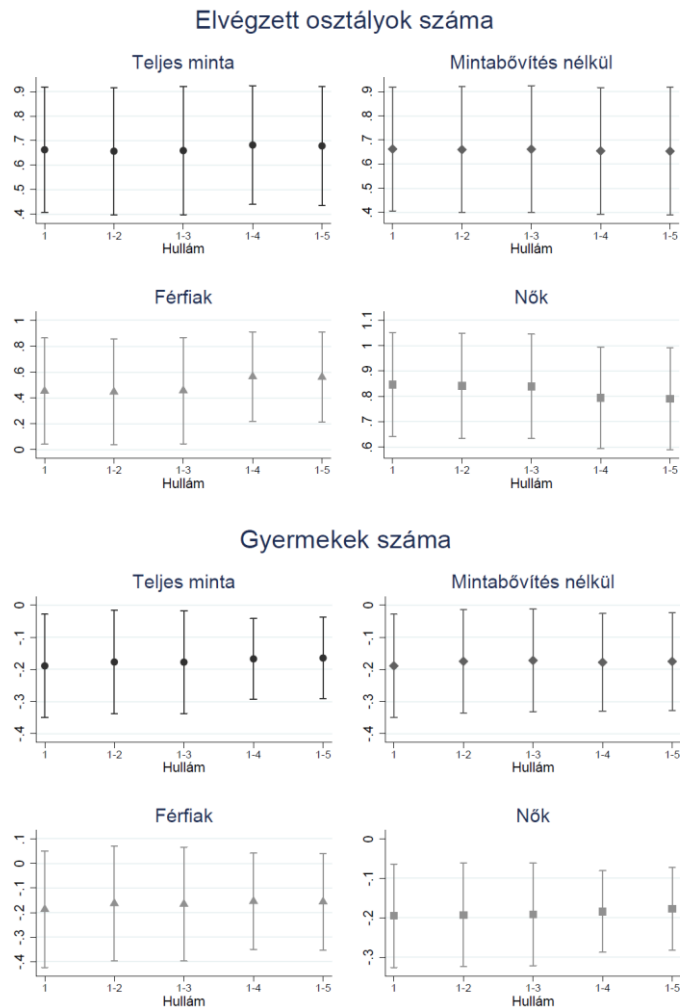
Ez a fejezet azt vizsgálja, hogy az eddig bemutatott eltérések a szakadási pont két oldalán hogyan befolyásolják a torzítás mértékét, ha a reform iskolázottságra és a gyermekek számára kifejtett hatását szeretnénk megbecsülni.

4.3.1 A lemorzsolódásból és a minta bővítéséből fakadó torzítás

A lemorzsolódásból és a minta bővítéséből fakadó torzítás mértékét úgy vizsgálom, hogy kumulált hullámonként becslem meg a reform hatását a két kimenetre. A reform becslt hatása mind az elvégzett osztályok, mind pedig a vállalt gyermekek számára meglehetősen stabil az egyes hullámok között (3. ábra). Hasonló eredményt kapunk, ha a később a felvételbe került embereket kizárjuk a mintából (3. ábra, *Mintabővítés nélkül* blokk). Az elvégzett osztályok számára kifejtett hatásban a férfiak között látunk némi emelkedést a 4. hullám bővítését követően, ami abból adódhat, hogy a férfiak között alacsonyabb várható

élettartam következtében a bővítés életkori korlátja erősebben szűrhetette meg a férfiakat, mint a nőket.

3. ábra: A reform hatása az elvégzett osztályok és a gyermekek számára az ELSA kumulatív hullámaiban



Forrás: saját becslés az ELSA-ból. Minden ábrázolt együttható külön becslésből származik az (5) egyenlet alapján. A 95 százalékos konfidencia-intervallumok születési év szerint klaszterezett robusztus sztenderd hibákból származnak. Bandwidth: 10-10 születési év a szakadási pont alatt és felett.

4.3.2 A mintavételből származó torzítás

Természetesen az a tény, hogy a becsült hatások az egyes hullámokban nem különböznek statisztikai értelemben egymástól, nem ad választ arra a kérdésre, fakadhat-e már az alapminta összeállításából torzítás. Az eddig becsült hatások nagyon hasonlóak azokhoz, amiket eddig az

irodalom a RoSLA hatásaként talált: a reform az elvégzett osztályok számát átlagosan 0,7-tel növelte, míg a befejezett fertilitást 0,2-vel csökkentette. Azonban az összes tanulmány survey adatokat használt (F.2 táblázat), néhány pedig szintén az ELSA-t.

Mivel az első hullámban is már csak az alapminta 65 százalékára sikerült adatot gyűjteni, az alapminta 35 százalékának oktatási és fertilitási kimeneteit nem tudjuk mérni. Azt viszont tudjuk, hogy a reform nem befolyásolta az alapmintán belül az első hullámban történő válaszadás valószínűségét (F.3. táblázat). Az első hullám adataiban korábban láttuk, hogy

- 1) a férfiak között az alacsony társadalmi státuszból származó emberek aránya nagyobb a szakadási pont jobb oldalán (2. ábra);
- 2) a nők között a könyvek száma otthon nagyobb volt gyermekkorban a szakadási pont jobb oldalán (2. ábra, a különbség 90 százalékon szignifikáns);

Ezekon kívül is lehetnek még olyan ismérvek, amelyek szintén eltérnek a szakadási pont két oldalán már az első hullámban, de nem tudjuk őket megfigyelni (és mind az iskolázottsággal, mind pedig a fertilitással korrelálnak). Mit lehet mindezekkel kezdeni? Optimális esetben olyan adatbázist érdemes választani, amiben ilyen torzítások nincsenek (nagy véletlen mintás adminisztratív adatokat). Minden más megoldás csak részlegesen kielégítő, és megfelelő feltevések mellett használható.

Az F.6. táblázat a reform hatását mutatja alacsony és magas családi háttérből származó férfiak, illetve olyan nők között, akiknél kevés vagy sok könyv volt otthon gyermekkorban. Az iskolai évekre kifejtett hatásról tudjuk az irodalomból, hogy nagyobb az alacsony státuszból származó, vagy alacsonyan iskolázott szülők gyermekei között, mert a magas státuszú szülők gyerekei a reformtól függetlenül is tovább járnak iskolába. Valóban ezt látjuk: a reform hatása majdnem kétszer akkora az alacsony, mint a magas státuszból származó férfiak között (0.698 és 0.345). A különbség a reform hatásában a nők között a könyvek száma szerint vett két csoport sokkal kisebb. A gyermekek számára kifejtett hatásban is van különbség, érdekes módon a magasabb státuszú férfiak, és a gyermekkorban több könyvvel rendelkező nők között a reform negatív hatása a gyermekvállalásra sokkal nagyobb. Valószínű tehát, hogy ezek az összetételbeli különbségek a szakadási pont két oldalán okoznak torzítást a becsült hatásokban. Kérdés, hogy ez a torzítás milyen irányú lehet. Ez attól függ, hogy mi nagyobb: a szakadási pont két oldalán lévő változó korrelációja a kimeneti mutatóval, vagy a reform hatásában jelentkező különbség az eltérő csoportok között. A férfiak példájánál maradva, az alacsony státuszból származó gyerekek átlagosan rövidebb ideig járnak iskolába, viszont közöttük a reform hatása nagyobb. A tanulmány utolsó táblázata azt

mutatja, mi történik, ha az eddigi modellekben kontrollálunk a családi háttérre és a könyvek számára (F.7. táblázat). Az alacsony családi háttér a férfiak és a nők között is negatívan korrelál az iskolában töltött évek számával, de ez a korreláció több mint kétszer akkora a férfiak, mint a nők között (-0.803 és -0.360). Érdekes módon a könyvek száma és az iskolában töltött évek száma közötti pozitív összefüggés hasonló a két nem között. A reform becsült hatásának nagysága a nők között némileg csökken, ha kontrollálunk a családi háttérre és a könyvek számára, de a többi becsült hatás nem változik.⁴ A nem megfigyelt változókból eredő fennmaradó torzítás nagyságát olyan módszerekkel lehetne megbecsülni, amelyek valamilyen feltevéssel élnek a nem megfigyelt szelekció mértékéről a megfigyelt szelekció mértékéhez képest (lásd például Manski (1990)). Úgy tűnik azonban, hogy mivel kevés olyan megfigyelt változó van az ELSA-ban, amit nem befolyásolhatott a reform, és ami használható a megfigyelt szelekció mérésére, ezek a módszerek szélsőséges eredményeket adnak, ezért ezek további tárgyalásától eltekintek.

5. Következtetések

A tanulmány azt vizsgálta, befolyásolta-e a kötelező iskolalátogatási korhatár 1947-es emelése Angliában az ELSA adatfelvételben való részvételt, a lemorzsolódást és a válaszadást. A reform az ötödik hullámra szignifikánsan csökkentette a lemorzsolódást, illetve a mintát a negyedik hullámban úgy bővítették, hogy pont a szakadási pont felett született emberek körét bővítették. A megfigyelések száma ezek miatt a negyedik és az ötödik hullám adatfelvétele után ezért nem folytonos a szakadási pontban. A minta összetétele a szakadási pont alatt és felett olyan változóban is eltér, amelyeket nem befolyásolhatott a reform. A szakadási pont felett született férfiak nagyobb valószínűséggel jöttek alacsony családi háttérből, mint a szakadási pont alatt születettek, a nők között pedig abban van különbség, hogy gyermekkorban hány könyvük volt otthon. A szakadási pontban a negyedik-ötödik hullámban nem folytonos elemszám ellenére a szokásos módszerrel végzett becslések a reform iskolázottságra és a gyermekek számára kifejtett hatásáról az egyes hullámok között stabilak maradnak. Bár a tanulmány becslései az irodalomban talált hatásokhoz hasonlóak, valószínű, hogy a mintaösszetétel eltérése a szakadási pont két oldalán torzítást visz a becslésekbe.

⁴ Az F.7. táblázat becsült hatásnagyságai némileg eltérnek a korábbiaktól, mert a gyermekkori családi háttér mutatói csak a minta egy részére elérhetők. A F.7. táblázat két-két modelljét azonos mintán becsültem.

Hivatkozások

- Angrist, Joshua D., és Alan B. Krueger. 1991. „Does compulsory school attendance affect schooling and earnings?” *The Quarterly Journal of Economics* 106 (4): 979–1014. <https://doi.org/10.2307/2937954>.
- Arendt, Jacob. 2005. „Does education cause better health? A panel data analysis using school reforms for identification”. *Economics of Education Review* 24 (2): 149–60. <https://doi.org/10.1016/j.econedurev.2004.04.008>.
- Banks, J., G. David Batty, J. Breedvelt, K. Coughlin, R. Crawford, M. Marmot, J. Nazroo, és mtsai. 2021. „ELSAEnglish Longitudinal Study of AgeingEnglish Longitudinal Study of Ageing: Waves 0-9, 1998-2019”. UK Data Service. <https://doi.org/10.5255/UKDA-SN-5050-24>.
- Banks, James, és Fabrizio Mazzonna. 2012. „The Effect of Education on Old Age Cognitive Abilities: Evidence from a Regression Discontinuity Design”. *Economic Journal* 122 (560): 418–48. <https://doi.org/10.1111/j.1468-0297.2012.02499.x>.
- Brunello, Giorgio, Daniele Fabbri, és Margherita Fort. 2013. „The Causal Effect of Education on Body Mass: Evidence from Europe”. *Journal of Labor Economics* 31 (1): 195–223. <https://doi.org/10.1086/667236>.
- Brunello, Giorgio, Margherita Fort, Nicole Schneeweis, és Rudolf Winter-Ebmer. 2016. „The Causal Effect of Education on Health: What Is the Role of Health Behaviors?” *Health Economics* 25 (3): 314–36. <https://doi.org/10.1002/hec.3141>.
- Brunello, Giorgio, Margherita Fort, és Guglielmo Weber. 2009. „Changes in Compulsory Schooling, Education and the Distribution of Wages in Europe*”. *The Economic Journal* 119 (536): 516–39. <https://doi.org/10.1111/j.1468-0297.2008.02244.x>.
- Chib, Siddharta, és Liana Jacobi. 2016. „Bayesian Fuzzy Regression Discontinuity Analysis And Returns To Compulsory Schooling”. *Journal of Applied Econometrics* 31 (6): 1026–47. <https://doi.org/10.1002/jae.2481>.
- Devereux, Paul J., és Robert A. Hart. 2010. „Forced to Be Rich? Returns to Compulsory Schooling in Britain”. *The Economic Journal* 120 (549): 1345–64. <https://doi.org/10.1111/j.1468-0297.2010.02365.x>.
- Dong, Yingying. 2015. „Regression Discontinuity Applications with Rounding Errors in the Running Variable: RD: ROUNDING ERRORS IN RUNNING VARIABLE”. *Journal of Applied Econometrics* 30 (3): 422–46. <https://doi.org/10.1002/jae.2369>.
- Fang, Hai, Karen N. Eggleston, John A. Rizzo, Scott Rozelle, és Richard J. Zeckhauser. 2012. „The Returns to Education in China: Evidence from the 1986 Compulsory Education Law”. Working Paper. Working Paper Series. National Bureau of Economic Research. <https://doi.org/10.3386/w18189>.
- Fletcher, Jason M. 2015. „New Evidence of the Effects of Education on Health in the US: Compulsory Schooling Laws Revisited”. *Social Science & Medicine* (1982) 127 (február): 101–7. <https://doi.org/10.1016/j.socscimed.2014.09.052>.
- Fort, Margherita, Nicole Schneeweis, és Rudolf Winter-Ebmer. 2016. „Is Education Always Reducing Fertility? Evidence from Compulsory Schooling Reforms”. *The Economic Journal* 126 (595): 1823–55. <https://doi.org/10.1111/eoj.12394>.
- Fort, Margherita, Nicole Schneeweis, és Rudolf Winter-Ebmer. 2016. „Is Education Always Reducing Fertility? Evidence from Compulsory Schooling Reforms”. *The Economic Journal* 126 (595): 1823–55. <https://doi.org/10.1111/eoj.12394>.

- Hahn, Jinyong, Petra Todd, és Wilbert Van der Klaauw. 2001. „Identification and estimation of treatment effects with a regression-discontinuity design”. *Econometrica* 69 (1): 201–9. <https://doi.org/10.1111/1468-0262.00183>.
- Harmon, Colm, és Ian Walker. 1999. „The Marginal and Average Returns to Schooling in the UK”. *European Economic Review* 43 (4–6): 879–87. [https://doi.org/10.1016/S0014-2921\(98\)00101-9](https://doi.org/10.1016/S0014-2921(98)00101-9).
- Imbens, Guido, és Thomas Lemieux. 2008. „Regression discontinuity designs: a guide to practice”. *Journal of Econometrics* 142 (2): 615–35. <https://doi.org/10.1016/j.jeconom.2007.05.001>.
- Lee, David S., és David Card. 2008. „Regression Discontinuity Inference with Specification Error”. *Journal of Econometrics* 142 (2): 655–74. <https://doi.org/10.1016/j.jeconom.2007.05.003>.
- Lee, David S., és Thomas Lemieux. 2010. „Regression Discontinuity Designs in Economics”. *Journal of Economic Literature* 48 (2): 281–355. <https://doi.org/DOI:10.1257/jel.48.2.281>.
- Liwiński, Jacek. 2020. „The Impact of Compulsory Schooling on Hourly Wage: Evidence From the 1999 Education Reform in Poland”. *Evaluation Review* 44 (5–6): 437–70. <https://doi.org/10.1177/0193841X20987104>.
- Manski, Charles F. 1990. „Nonparametric Bounds on Treatment Effects”. *The American Economic Review* 80 (2): 319–23.
- McCrary, Justin. 2008. „Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test”. *Journal of Econometrics* 142 (2): 698–714. <https://doi.org/10.1016/j.jeconom.2007.05.005>.
- Mocan, N., és L. Pogorelova. 2014. „Compulsory Schooling Laws and Formation of Beliefs: Education, Religion and Superstition”. <https://doi.org/10.1016/J.JEBO.2017.07.005>.
- Oreopoulos, Philip. 2006. „Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter”. *American Economic Review* 96 (1): 152–75. <https://doi.org/10.1257/000282806776157641>.
- Roy, Brita, Catarina I. Kiefe, David R. Jacobs, David C. Goff, Donald Lloyd-Jones, James M. Shikany, Jared P. Reis, Penny Gordon-Larsen, és Cora E. Lewis. 2020. „Education, Race/Ethnicity, and Causes of Premature Mortality Among Middle-Aged Adults in 4 US Urban Communities: Results From CARDIA, 1985–2017”. *American Journal of Public Health* 110 (4): 530–36. <https://doi.org/10.2105/AJPH.2019.305506>.
- Schneeweis, Nicole, Vegard Skirbekk, és Rudolf Winter-Ebmer. 2012. „Does schooling improve cognitive functioning at older ages?” NRN working paper 2012–11. The Austrian Center for Labor Economics and the Analysis of the Welfare State, Johannes Kepler University Linz, Austria. https://econpapers.repec.org/paper/jkunrnwps/2012_5f11.htm.
- Vakhitova, Ganna, Tom Coupe, International Labour Office, és ILO Decent Work Technical Support Team and Country Office for Central and Eastern Europe. 2013. *The Relations between Education and Migration in Ukraine*. Budapest: ILO.

Függelék

F.1. táblázat: A kötelező iskolalátogatási korhatár emelését survey adatokon vizsgáló tanulmányok

Tanulmány, folyóirat	Reform, adatok	Kimeneti mutató	Eredmény	Teszteli-e a megfigyelések számának folytonosságát a szakadási pontban?	Teszteli-e a reform hatását a mintában való részvételre/lemorzsolódásra/válaszadásra?	Teszteli-e hogy a minta összetétele eltér-e a szakadási pont két oldalán?	Running variable (születési év/hónap/nap)
Fort, Schneeweis, és Winter-Ebmer (2016), The Economic Journal	9 reform 6 európai országban 1936 és 1975 között SHARE és ELSA	Befejezett fertilitás (gyermek száma 45 éves kor felett, gyermektelenség)	Az iskolázottság csökkentette a fertilitást Angliában, de a többi európai országban nem.	Nem.	Nem.	Nem.	Születési év
Schneeweis, Skirbekk, és Winter-Ebmer (2012), IZA Working Papers	SHARE, 6 reform 6 országban	Időskori kognitív készségek	Oksági kapcsolat az oktatásban eltöltött évek és az azonnali és késleltetett memória között, valamint a nőknél szignifikáns a demenciánál is.	Nem.	Nem.	Nem.	Születési év
Brunello, Fabbri, és Fort (2013), Journal of Labor Economics	SHARE, ELSA, ECHP, British Household Panel Survey, German Socioeconomic Panel, 60-as és 70-es évek kötelező iskolalátogatási korhatár-reformjai	Testtömeg-index	Az iskolázottság csökkenti a testtömeg-index értékét a nőknél 9 európai országban.	Nem.	Nem.	Nem.	Születési hónap, kivéve Angliában, ahol nincs erről információ
(Brunello és mtsai. 2016) Health Economics	SHARE és ELSA, 1947-es reform	Egészségi állapot	Az egészségre káros viselkedésformák rövid- és hosszútávon hatással vannak arra, hogy az oktatás hogyan befolyásolja az egészséges életmódot.	Nem.	Nem.	Nem.	Születési év
Brunello, Fort, és Weber (2009), The Economic Journal	25 reform 12 európai országban 1949 és 1983 között SHARE	Jövedelem	Az iskolázottság növeli a béreket és csökkenti a béregyenlőtlenséget.	Nem.	Nem.	Nem.	Születési év
Fletcher (2014), Social Science and Medicine	NIH/AARP Diet and Health Study, reformok az 50-es években	Egészségi állapot és szubjektív egészségi állapot	Oksági összefüggés lehet az iskolalátogatás és egészségügyi mutatók között, de nem feltétlenül elég nagy a minta.	Nem.	Nem.	Nem.	Születési év

Liwiński (2020), Evaluation Review	1999-os iskolaelhagyási korhatár-emelés, Polish Labour Force Survey	Óránkénti bér és foglalkoztatottsági ráta	13%-kal nőtt a szakmai iskolákat elvégzettek óránkénti bevétele a reform hatására.	Igen.	Nem	Nem.	Születési év
Arendt (2005), Economics of Education review	Több dán iskolalátogatási reform, The Danish National Work Environment Cohort Study	Szubjektív egészségi állapot és testtömeg-index	Összefüggnek a kimeneti mutatók az oktatással, de az okság nem egyértelmű, összességében az eredmény inkonzkluzív.	Nem	Nem.	Nem.	Születési negyedév (birth quarter)
Devereux és Hart (2010), The Economic Journal	1947-es brit reform, GHS, NESPD	Jövedelem	A férfiaknál az iskolában töltött évek száma emeli a jövedelmet.	Igen.	Nem.	Nem.	Születési év
Banks és Mazzonna (2012), The Economic Journal	1947-es brit reform, ELSA	Férfiak memóriája	A kötelező iskolalátogatás meghosszabbítása erős és azonnali hatással volt a 14 évesek csaknem felének kognitív képességeire.	Igen.	Igen.	Nem.	Születési negyedév (birth quarter)
(Fang és mtsai. 2012), NBER Working Paper	1986-es reform Kínában, China Health and Nutrition Survey	Jövedelem	A kötelező iskolalátogatási törvény bevezetése növelte a jövedelmet.	Igen.	Nem.	Nem.	Születési negyedév (birth quarter)
Mocan és Pogorelova (2014), NBER Working Paper	11 európai ország oktatási reformjai, ESS, EVS	Vallásos és babonás viselkedés	Az iskolázottság konzisztensen negatív hatással van a vallásos viselkedésre (pl. imádkozás gyakorisága, szertartásokon való részvétel) és a babonásságra.	Nem.	Nem.	Nem.	Születési év

Forrás: saját gyűjtés.

F.2. táblázat: A RoSLA hatását mérő tanulmányok és eredményeik

Tanulmány, folyóirat	Adatok	Minta	Módszer (running variable/futó változó)	Fő kimeneti mutató	A reform hatása az iskolázottsága (first stage)	Fő eredmény
Oreopoulos (2006), The American Economic Review	General Household Survey (GHS)	1935-1965 között 14 éves nők és férfiak.	Regression discontinuity design (RDD) és IV	éves bér logaritmusa	0.440*** (0.076)	Egy év extra oktatás bérhozáma 14-17%.
Harmon és Walker (1999) European Economic Review	General Household Survey (GHS)	1978-1986-ban 16-64 éves, alkalmazott férfiak. N=34336.	OLS és IV (parametrikus becslés)	bér	Nem közli.	Az oktatás bérhozáma 16%.
Devereux and Hart (2010), The Economic	1979-1998 U.K. General Household	1921-1951 között született nők és férfiak 28-64 éves	RDD (parametrikus becslés)	bér	Férfiak: 0.47** (0.026)	A nőknél nincs bizonyíték az extra oktatás bérhozáma, a

Journal	Survey (GHS)	korban.			Nők: 0.55** (0.044)	férfiaknál 4-7%.
Banks and Mazzonna (2012), The Economic Journal	English Longitudinal Study of Ageing (ELSA)	Az ELSA 1998, 1999, és 2001-es hullámaiból csaknem 12 000 válaszadó.	Fuzzy Regression Discontinuity (FRD) design	kognitív képességek	Férfiak: 0.136 (0.355) Nők: 1.256** (0.307)	Az extra oktatás pozitív és szignifikáns hatással volt az alacsonyabb iskolázottságúak időskori memóriájára.
Fort, Schneeweis, and Winter-Ebmer (2016), The Economic Journal	Survey on Health, Ageing and Retirement in Europe (SHARE) és ELSA	ELSA 2004-2010-ben több, mint 50 éves nők.	IV	befejezett fertilitás (gyermek száma 45 éves kor felett, gyermektelenség)	0.697*** (0.130)	Az extra oktatásnak nincs kimutatható hatása a fertilitásra.
Chib és Jacobi (2016), Journal of Applied Econometrics	UK General Household Survey	1931. október és 1934. szeptember között született férfiak.	FRD design	óránkénti bér logaritmus	Nem közli.	Az extra oktatás hatása kicsi a bérekre, mortalitásra és egészségre.

Forrás: saját gyűjtés. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

F.3. táblázat: A reform hatása az első öt hullámban való részvétel valószínűségére az alapmintán belül

	(1) 1. hullám	(2) 2. hullám	(3) 3. hullám	(4) 4. hullám	(5) 5. hullám
Reform	0.009 (0.022)	0.005 (0.023)	0.020 (0.019)	0.021 (0.022)	0.055* (0.021)
Relatív születési év	0.007* (0.003)	0.011** (0.003)	0.011*** (0.002)	0.013*** (0.003)	0.014*** (0.002)
Relatív születési év *reform	-0.013** (0.003)	-0.014** (0.004)	-0.014*** (0.003)	-0.014*** (0.003)	-0.016*** (0.003)
Konstans	0.689*** (0.016)	0.553*** (0.020)	0.472*** (0.015)	0.420*** (0.017)	0.380*** (0.014)
Megfigyelések száma	9180	9180	9180	9180	9180

Forrás: saját becslés az ELSA adataiból. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Bandwidth: 10-10 születési év a szakadási pont alatt és felett. Lokális lineáris valószínűségi modellek az (5) egyenlet szerint. A minta csak azokat az embereket tartalmazza, akik már az alapmintában is benne voltak. Független változó: az adott személy interjút adott az adott hullámban.

F.4. táblázat: Az alapmintában még nem szereplő, később hozzáadott emberek előfordulási gyakorisága a szakadási pont körül

	(1) 2. hullám	(2) 3. hullám	(3) 4. hullám	(4) 5. hullám
Reform	-0.002 (0.004)	0.004 (0.002)	0.132*** (0.017)	0.126*** (0.017)
Relatív születési év	0.001 (0.001)	0.000* (0.000)	0.011*** (0.002)	0.010*** (0.002)
Relatív születési év *reform	-0.001 (0.001)	0.001* (0.000)	-0.007* (0.003)	-0.006* (0.003)
Konstans	0.005 (0.004)	0.003* (0.002)	0.109*** (0.010)	0.094*** (0.009)
Observations	4814	4160	4555	4181

Forrás: saját becslés az ELSA adataiból. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Bandwidth: 10-10 születési év a szakadási pont alatt és felett. Lokális lineáris valószínűségi modellek az (5) egyenlet szerint. Az egyes hullámok mintája mindenkit tartalmaz, aki az adott hullámban interjút adott. Független változó: az adott személy nem volt benne az alapmintában.

F.5. táblázat: A megfigyelések száma a reform függvényében az ELSA egyes kumulált hullámaiban

	(1)	(2)	(3)	(4)	(5)
	1. hullám	1-2. hullám	1-3. hullám	1-4. hullám	1-5. hullám
Lokális lineáris regresszió, OLS ((1) egyenlet)					
A törés nagysága a szakadási pont körül (β_1 , fő)	11.455 (13.008)	12.327 (12.444)	12.673 (12.416)	59.006*** (13.403)	60.576*** (13.740)
Globális parametrikus modell Dong (2015) alapján, OLS ((2) egyenlet)					
A törés nagysága a szakadási pont körül (β_1 , fő)	53.996 (38.652)	49.355 (38.540)	56.532 (37.185)	94.252* (40.631)	95.585* (40.599)
Kerekítéssel korrigált együttható ($\beta_1 - \frac{1}{2} * \beta_{31} + \frac{1}{6} * \beta_{32} - \frac{1}{30} * \beta_{34}$, fő)	29.577 (25.867)	28.419 (26.288)	35.119 (24.133)	65.117* (27.727)	66.700* (27.230)
Lokális lineáris regresszió, Poisson regresszió ((3) egyenlet)					
A törés nagysága a szakadási pont körül (β_1 , logpont)	0.009 (0.044)	0.011 (0.042)	0.011 (0.042)	0.120** (0.040)	0.124** (0.041)
Globális parametrikus modell Dong (2015) alapján, Poisson regresszió ((4) egyenlet)					
A törés nagysága a szakadási pont körül (β_1 , logpont)	0.163 (0.100)	0.146 (0.099)	0.169 (0.096)	0.259** (0.098)	0.262** (0.098)
Kerekítéssel korrigált együttható ($\beta_1 - \frac{1}{2} * \beta_{31} + \frac{1}{6} * \beta_{32} - \frac{1}{30} * \beta_{34}$, logpont)	0.093 (0.060)	0.088 (0.060)	0.108 (0.055)	0.184** (0.056)	0.188** (0.055)
Megfigyelések száma	20	20	20	20	20

Forrás: saját becslés az ELSA adataiból. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Zárójelben születési év szerint klaszterezett standard hibák. Bandwidth: 10-10 születési év a szakadási pont alatt és felett.

F.6. táblázat: A reform hatása a férfiak és nők egyes almintáiban

	(1)	(2)	(3)	(4)
	Férfiak		Nők	
	Magas családi háttér	Alacsony családi háttér	Kevés könyv otthon	Sok könyv otthon
A reform hatása az iskolában töltött évek számára				
Reform	0.345 (0.321)	0.698*** (0.169)	0.852*** (0.159)	0.795*** (0.166)
Megfigyelések száma	1583	1192	932	2215

A reform hatása a gyermekek számára

Reform	-0.230* (0.093)	-0.154 (0.197)	-0.062 (0.108)	-0.263** (0.071)
Megfigyelések száma	1452	1085	864	2028

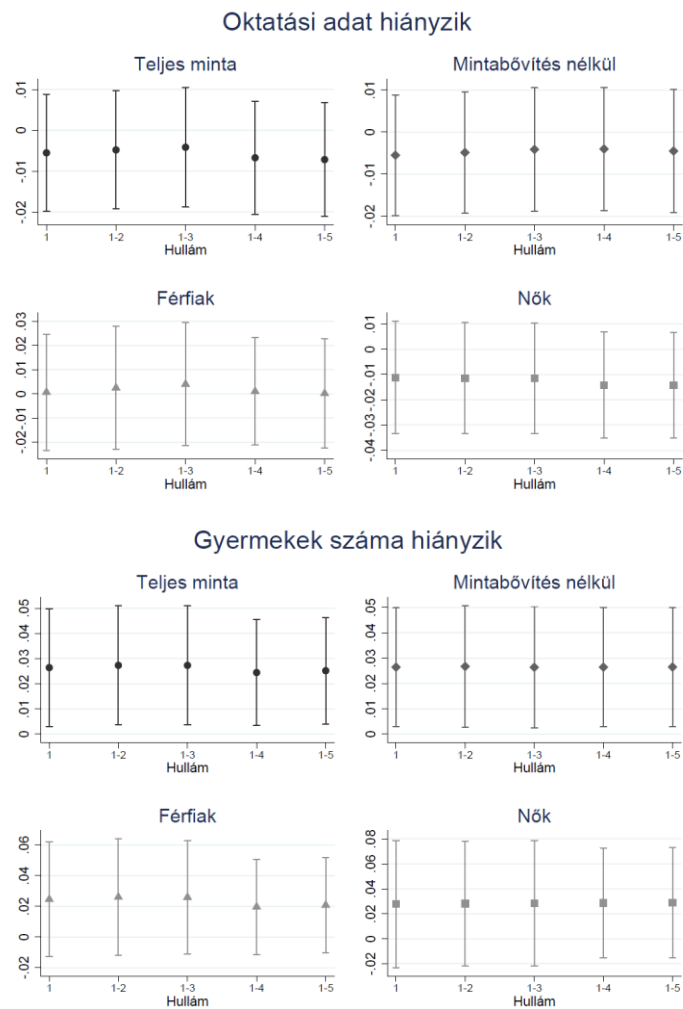
Forrás: saját becslés az ELSA adataiból. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Zárójelben születési év szerint klaszterezett standard hibák. Bandwidth: 10-10 születési év a szakadási pont alatt és felett. (5) egyenlet szerinti becslések.

F.7. táblázat: A reform hatása a családi háttérre történő kontrollálás után

	(1)	(2)	(3)	(3)
	Férfiak	Férfiak	Nők	Nők
A reform hatása az iskolában töltött évek számára				
Reform hatása	0.535 (0.321)	0.549* (0.206)	0.989*** (0.118)	0.813*** (0.120)
Alacsony családi háttér		-0.803*** (0.098)		-0.360*** (0.066)
Könyvek száma otthon 10 éves korban		0.709*** (0.043)		0.734*** (0.048)
Megfigyelések száma	1429	1429	1711	1711
A reform hatása a gyermekek számára				
Reform hatása	-0.309 (0.157)	-0.315 (0.159)	-0.173** (0.056)	-0.166* (0.059)
Alacsony családi háttér		0.138 (0.067)		0.021 (0.063)
Könyvek száma otthon 10 éves korban		-0.006 (0.031)		-0.021 (0.024)
Megfigyelések száma	1311	1311	1579	1579

Forrás: saját becslés az ELSA adataiból. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Zárójelben születési év szerint klaszterezett standard hibák. Bandwidth: 10-10 születési év a szakadási pont alatt és felett. (5) egyenlet szerinti becslések.

F.1. ábra: A reform hatása a válaszadásra az ELSA számára az ELSA kumulatív hullámaiban



Forrás: saját becslés az ELSA-ból. Minden ábrázolt együtttható külön becslésből származik az (5) egyenlet alapján, és azt mutatja meg, hogyan változik az adathiány valószínűsége a szakadási pont körül. A 95 százalékos konfidencia-intervallumok születési év szerint klaszterezett robusztus sztenderd hibákból származnak. Bandwidth: 10-10 születési év a szakadási pont alatt és felett.