

Hova tovább a nemválaszolókkal?

A European Social Survey alapján végzett módszertani kísérlet eredményei

Szeitl Blanka – Tóth István György

szeitl@tarki.hu, toth@tarki.hu

<https://doi.org/10.51624/SzocSzemle.2020.3.5>

Beérkezés: 2020. 03. 19.

Átdolgozott változat beérkezése: 2020. 07. 08.

Elfogadás: 2020. 08. 27.

Összefoglaló: A társadalomkutatási célú empirikus adatfelvételek egyik legnagyobb kihívása a csökkenő válaszadási hajlandóság. A probléma nem csak Magyarországon, nem csak egyik vagy másik vizsgálsorozatban jelentkezik, és nem is egyik vagy másik adatfelvevő céghez kötődik. Mivel a részvétel és az elutasítás nem véletlenszerűen oszlik el a vizsgált populációkban, a megvalósult adatfelvételek interpretációjához és validitásának értékeléséhez elengedhetetlen, hogy többet tudjunk a nemválaszolók társadalmi-gazdasági karakteréről. Cikkünkben egy ezzel kapcsolatos kutatásról számolunk be. Alapkérdésünk az, hogy miképpen jellemezhetők a nem vagy nehezen elérhető csoportok, és hogyan lehet/érdemes változtatni az adatfelvételi eljárásokon ahhoz, hogy jobb legyen az elérés. Kutatásunk Magyarországon hiánypótló, „re-contact” típusú adatfelvétel; a European Social Survey 8. hullámának eredeti, kiinduló címlistájához ismét visszamentünk egy rövid kérdőívvel, és megvizsgáltuk, hogy az ismétlés mennyire hoz más eredményeket az első vizsgálathoz képest. A mindkét felmérésben sikeresen lekérdezett címek aránya 25,7%, a mindkét felmérésben elérhetetlen (no-contact) válaszadók aránya 6,5%, a stabilan válaszmegtadóké 11,3% volt. Kutatásunk egyértelműen bizonyította, hogy egy adott vizsgálat sikertelen eseteit nem szabad „reménytelennek” tekinteni, hiszen az ő körükben is majdnem akkora volt a második körös elérés, mint amekkora magában a fő vizsgálatban volt (38,3 százalék a 41,1 százalékhoz viszonyítva). Az elérhetetlenség és a nemválaszolás meghatározói között a megkérdezettek társadalmi-gazdasági jellemzőit, az adott címekhez tartozó települési-lakókörnyezeti jellemzőket és az adatfelvételi apparátus jellemzőit vizsgáltuk, utóbbi két tényezőcsoport változóit pedig közvetlenül a kérdőív válaszaihoz kapcsoltuk. Eredményeink alapján az újra elérés sikerességét a demográfiai jellemzők mellett terepmunka-specifikus tényezők is befolyásolják, az általános el nem érés esélyeire pedig a települési lejtő mellett a lakókörnyezet fizikai állapota is lényegi hatást gyakorol. Vizsgálatunk utal arra is, hogy a nemválaszolás mintázatait a mintavétel során nem figyelembe vehető pszichológiai tényezők is befolyásolják: ilyen hatásra utal például az objektív anyagi helyzet és annak szubjektív értékelése között talált sajátos interakciós hatás. Cikkünk végén a tanulságok mellett további kutatási irányokra és konkrét adatfelvételi technikákra is javaslatot teszünk.

Kulcsszavak: empirikus felmérések, nemválaszolás, torzítás

Bevezetés

Anélkül, hogy részletesen érvelnénk emellett, adottnak vesszük, hogy az empirikus társadalomtudomány műveléséhez empirikus lakossági adatfelvételekre az adminisztratív adatok és a „big data” alkalmazások széles körű elterjedésének időszakában is szüksége van – és lesz is. Efféle vizsgálatok során mintát kell vennünk a teljes populációból vagy annak a kutatás szempontjából releváns alcsoportjából. Ennek során feltételezzük azt, hogy az általunk használt kiválasztási algoritmusok segítségével reprezentatív mintához jutunk, így az adatokból származó becsléseinket a teljes kiválasztott népeiségre (országosan reprezentatívnak szánt felvételek esetén például a magyar lakosságra) általánosíthatjuk.

Az elmúlt 20 évben azonban folyamatosan és jelentősen növekszik azoknak az aránya, akiket egy adatgyűjtés során a kérdezők nem tudnak elérni, vagy a kiválasztott személy megtagadja a válaszadást. Habár az adatok súlyozási és utólagos rétegezési eljárásokkal javíthatóak, egyre kevésbé beszélhetünk reprezentatív, általánosítható felmérésekről. Noha a szakmai érdeklődés megvan a minta sikertelen részére vonatkozó információk iránt, kevés tényleges, kvantifikálható empirikus ismeretünk van a „kimaradók” társadalmi jellemzőit illetően. Márpedig erre az információra nagy szükség lenne az adatfelvételek érvényességének és megbízhatóságának megítéléséhez is. Az érvényességen egyfelől a kutatási kérdés megfelelő operacionalizálását értjük, vagyis azt, hogy a kérdőív megalkotásával valóban arra a kérdésre kapunk-e választ, amire szeretnénk. A kutatás egészének érvényességi köre pedig arra ad választ, hogy *tekinthető-e az adatfelvétel érvényesnek a teljes lakosságra, vagy az adatfelvételi torzítások, vagy tágabb értelemben a reprezentativitás sérülése miatt szűkebb ez az érvényességi kör.* Az adatok megbízhatósága azt jelezheti, hogy mennyire stabilak az eredményeink, azaz mennyire lehetünk biztosak abban, hogy adott értékek nem a véletlen művei, vagyis egy ismételt vizsgálat eredményei (valamilyen statisztikai valószínűséggel) *ugyanazeket az eredményeket adnák.*

A Tárki Társadalomkutatási Intézet által 2017-ben indított, az NKFIH által támogatott¹ survey-módszertan témájú alapozó pilotkutatás a válaszadás és nem válaszadás meghatározóinak vizsgálatát célozta azzal a reménnyel, hogy ezzel hozzájárulunk az érvényességi kör, illetve az adatok megbízhatóságának elemzéséhez is. A kutatás különlegessége az úgynevezett „re-contact” design, amely egy korábbi kutatás újbóli, megismételt felkeresését jelenti. Módszertani szempontból éppen ez ad lehetőséget arra, hogy az adatfelvételi kérdések állandóságára és stabilitására választ adhassunk, továbbá mélyebben vizsgálhassuk a sikertelenséget befolyásoló tényezőket.²

1 Az NKFI Alapból (OTKA) támogatott kutatás, szerződési azonosítószám: K 125162, a kutatás címe: Megközelíthetetlen csoportok, vizsgálata – Pilot kutatás az empirikus felmérésekből kimaradó elérhetetlen, vagy válaszmegtagadó népeiségről, azok speciális feltérképezésével.

2 A kutatás zárótanulmánya itt található: https://tarki.hu/sites/default/files/2020-03/OTKA_kut_besz_megkozelithetetlen_csoportok.pdf; a módszertani dokumentáció és az adatállomány elérhető a Tárki Adatbankból: www.tarki.hu/adatbank.

Cikkünkben ennek a kutatásnak az eredményeit ismertetjük. Az adatfelvételek megbízhatóságát a kutatás sikerességi eredményeivel, illetve ezek különböző csoportokra vett bontásaival vizsgáljuk meg, majd a nemválaszolók jellemzésére összetett, többváltozós modellt építünk és tesztelünk. Előbb rövid kitekintést adunk a nemzetközi és a magyarországi nemválaszolási trendekről, majd bemutatjuk a kutatás szerkezetét és az adatfelvételi dizájnt. Ezután ismertetjük a kapott eredményeket külön a megbízhatóság és külön a nemválaszolók jellemzésével, majd összegezzük a konklúzióinkat.

Nemválaszolási trendek itthon és külföldön

A „unit-nonresponse” (általános nemválaszolás), vagyis a mintába választott háztartás vagy személy lekérdezésének sikertelensége világszerte növekszik a survey-felmérések körében.³ A unit-nonresponse jellemzően a sikertelen kapcsolatfelvételtől (a háztartás nem érhető el), a válaszmegtagadásból (a háztartást eléri, de megtagadja a felmérésben való részvételt), illetve a válaszképtelenségből (a háztartással sikerül kapcsolatba lépni, válaszolna is a felmérésre, de például egészségügyi problémák miatt nem tud) áll össze (Osier 2016). Az általános nemválaszolás emelkedésével az adatfelvételi folyamatok nehezebbé, lassúbbá, költségesebbé és munkaigényesebbé válnak, a jelenség tehát jelentősen rontja az adatfelvételi procedúrák pontos elvégezhetőségét és hatékonyságát. A nemválaszolás ezen túl az adatok minőségére is hatással van, hiszen a lakosság felmérésre nyitott és elutasító része között alapvető eltérések is lehetnek. Egy magas nemválaszolási eredménnyel záródott kutatás adataira alapozott becslések ezért torzítottak lehetnek, és az általános következtetéseket sem lehet a teljes népességre vetíteni (Osier 2016). Az 1990-es években 3-12 százalékponttal emelkedett a unit-nonresponse – hat amerikai népszámlálási iroda által készített – felmérések során, és 2013-ra a válaszadási arány csaknem 35 százalékra esett (Meyer et al. 2015). Az utóbbi évtizedekben olyan nemzetközi szintű tendencia bontakozott ki, amely alapjaiban ingatja meg a survey-típusú felmérésekből származó adatok segítségével végzett statisztikai becslések pontosságát. Ez pedig az előbb említett nemválaszolás növekedése (Beullens et al. 2018; Blom 2008; Blom et al. 2008; de Heer 1999; de Leeuw – de Heer 2002; de Leeuw et al. 2018; Stoop et al. 2010), amely a nemzetközi felmérések harmonizálhatóságát is veszélybe sodorja (Boros 2020).

Az International Social Survey Programme (ISSP) nemzetközi adatfelvételeinek átlagos válaszadási aránya az 1996–2000 közötti időszakban 69 százalékos volt, ami a következő ötéves periódusokban 58, 53 és 49 százalékra csökkent, 2016-ban pedig mindössze 44 százalékos volt. Az országok közötti szórás természetesen óriási, de a csökkenő trend általánosnak mondható. A European Social Survey rendkívül nagy figyelmet szentel a mintavételre, valamint az adatfelvételi protokollok betartására, és mégis, a csökkenő trend az ESS-vizsgálatok esetében is kimutatható. Bár az ESS 2002 és 2006 között némileg tudta

3 A survey-felmérések esetében használatos teljesülési és sikerességi ráták hivatalos definícióit az AAPOR (American Association for Public Opinion Research) nyilvánosan is elérhető dokumentációi is tartalmazzák: https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

növelni a válaszadási arányát, 2006-tól folyamatos csökkenés figyelhető meg; összességében 60 százalékról 54,5 százalékra csökkent a válaszadási arány 2016-ra (Boros 2020).

Az általános értelemben vett nemválaszolás is, illetve a meg nem találás és a válaszmegtagadás külön-külön is erősödött az elmúlt 20 évben Magyarországon: míg 1996-ban a kiválasztott háztartások 16 százaléka nem volt hajlandó válaszolni egy opcionális blokkra a mikrocenzus során (Havasi 1997), 2016-ban már lényegesen alacsonyabb válaszadási arányokat tapasztaltak⁴. A KSH *Miből élünk?* címmel folytatott háztartási fogyasztási és vagyonzvizsgálatában a 2014-es hullám mintegy 35 százalékos válaszadási arányt eredményezett (Simon–Valentiny 2016), amelyet a következő hullámban körülbelül negyven százalékra sikerült emelni (MNB 2019).⁵ A Tárki Háztartás Monitor vizsgálataiban az 1999-es 58 százalékról 2005-re 48 százalékra, 2012-re 41 százalékra, 2015-re pedig 39 százalékra esett a válaszadási arány (Tárki 2016).

A csökkenő válaszadásból fakadó problémát orvosolhatjuk azzal, ha nagyobb „bruttó” mintát választunk ki az elérni kívánt „nettó” mintaméret teljesítéséhez, illetve azzal, ha speciális súlyozási eljárásokat alkalmazunk. A csökkenő válaszadási arány tendenciája azonban a felmérések reprezentativitása szempontjából is komoly torzításokhoz vezet. A unit-nonresponse csak és kizárólag akkor nem befolyásolja az adatok megbízhatóságát, ha véletlenszerű, tehát ha minden mintába kerülő személy azonos valószínűséggel lesz válaszmegtagadó. Ezt viszont nehéz vizsgálni, mivel általában csak korlátozott információval rendelkezünk a nemválaszolók jellemzőiről, így az összehasonlítás gyakran lehetetlen. Továbbmenve: ha a nemválaszolók alapvető dimenziók – például nem, életkor és földrajzi hely – alapján közel megegyezőnek is látszanak, ez nem jelenti azt, hogy ezek a csoportok hasonlóak más dimenziók mentén is, például a kormányzati programokban való részvételi hajlandóság, a társadalmi hozzáállás, a pártpreferenciák stb. szerint (Meyer et al. 2015). Az eddig rendelkezésre álló eredmények alapján szinte mindig egyértelmű, hogy a unit-nonresponse nem véletlenszerű: 1996-ban az átlagos 16 százalékos nemválaszadási arányhoz képest Budapesten (28 százalék), a leginkább iskolázottak (30 százalék) és a jó anyagi helyzetben élők körében (34 százalék) mérték a legmagasabb nemválaszadási arányt (Havasi 1997). A nemválaszolók leggyakrabban olyan alapvető demográfiai jellemzők mentén írhatóak körül, mint a településméret, a foglalkoztatottság, a háztartásméret, a nem és az életkor⁶. Ezeken felül pedig olyan attitűddimenziók is szóba

4 A mikrocenzus esetében mindez nem könnyen számítható ki. Mindenesetre a rétegződéskiegészítő felvétel mintakeretét adó 88021 címen összesen 55930 háztartás (63,5%) lekérdezése valósult meg. Ezen felül volt a háztartásokon belüli meghíúsulás (12702 személy, lásd KSH 2018: 27), valamint azt is tekintetbe kell venni, hogy magán az (egyébként kötelező) alapmintán belül is volt meghíúsulás (a kiinduló minta 91%-a valósult meg, KSH 2017: 38). Összességében tehát – háztartások szintjén – a 2016-os, rétegződésrel foglalkozó opcionális blokk háztartási szinten számolt megvalósulása 91*63,5=57,8 százalékra tehető. Az egyéni szintű megvalósulás ennél alacsonyabb – de a rendelkezésre álló adatokból nem tudjuk pontosan rekonstruálni.

5 A közlemény szerint „a kérdőbiztosok 15 ezer háztartást kerestek föl, és a felkeresett háztartások közül majdnem 6 ezer választott” (MNB 2019: 2).

6 Korábbi egyszerű, leíró elemzésekkel olyan eredményeket tapasztaltunk, mint hogy a sikertelen megkeresések (válaszmegtagadás és/vagy no-contact) aránya a település méretével folyamatosan növekszik; az átlagosnál nagyobb települések esetén a nem válaszadás mindkét nem esetében nagyjából egyforma arányban fordul elő, de a kisebb települések esetében inkább a nők körében magasabb; a háztartás létszáma (vagy nagysága) jelentősen befolyásolja a no-contact arányát (minél nagyobb a háztartás létszáma, annál nagyobb a valószínűsége, hogy valamelyik háztartástaggal sikerül kapcsolatba lépni), és így a növekvő háztartáslétszám a nem válaszadás arányát is csökkenti (saját számítások a Tárki 2015-ös monitorvizsgálata alapján).

jöhetnek, mint a társadalmi izoláltság⁷ (Goyder 1987; Groves–Couper 1988), az általános bizalmatlanság⁸, az érdektelenség vagy az időhiány.

A kutatás szerkezete

A „re-contact” dizájn azt jelenti, hogy egy korábbi adatfelvételt megismétlünk, és függetlenül attól, hogy az adott kutatásban kijelölt cím sikeres vagy sikertelen megkereséssel zárult, az új hullámban minden korábbi szabályt betartva új felkereséseket rögzítünk. A „re-contact” szerkezetű kutatások tehát tulajdonképpen első és második adatfelvételi hullámra oszthatók, ahol az adatfelvétel első hulláma a kiinduló adatfelvétel, a második hullám pedig az újra felkeresés, azaz maga a „re-contact”. Az ilyen típusú kutatások első számú kutatási kérdése az, hogy mennyiben sikerül a második hullámban reprodukálni az első adatfelvételi hullám eredményeit (válaszadási arányok, a kérdezőbiztosok sikerességi mutatói, a kérdőív kérdéseire adott válaszok megoszlása stb.), azaz mennyire tekinthető megbízhatónak maga az adatfelvételi metódus, illetve a kutatás érdemi eredménye.

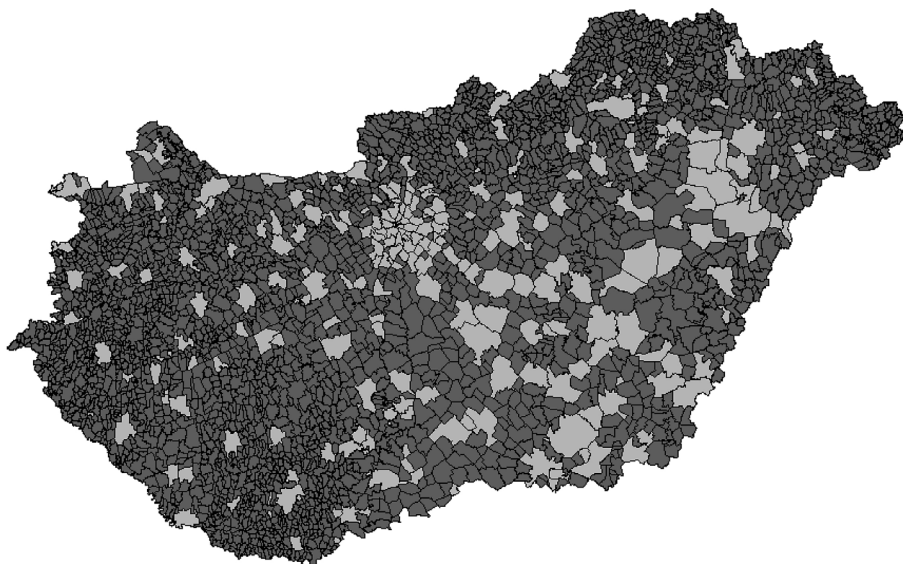
Kutatásunkban első adatfelvételi hullámnak a European Social Survey 8. magyarországi hullámát választottuk, amely egy nagyjából négyezer (pontosan: 3954) kiinduló címmel dolgozó országos reprezentatív kutatás volt⁹. A második hullám során ennek a címbázisnak a 75 százalékát kerestük fel újra, azaz 3000 címet. Az országos lefedettséget az 1. ábra mutatja.

7 A feltételezés szerint azok a személyek, akik jobban be vannak integrálva a társadalomba (azt gondolják, hogy fontos a véleményük, szociálisan aktívak, hatást gyakorolnak a kisebb/nagyobb környezetükre, illetve sokféle dolgot csinálnak), nagyobb valószínűséggel válaszolnak a megkeresésekre.

8 A társadalmi intézményekbe vetett és az általános bizalom állapota két szempont alapján is kapcsolódik a nemválaszolási mintázatokhoz: egyrészt azzal, hogy a kérdezőbiztosnak sikerül-e kapcsolatba lépnie a kijelölt személlyel/háztartással, tehát be tud-e jutni a házba/lakásba, másrészt pedig a vélemény megosztásával kapcsolatban, ami a valós és őszinte válaszok csökkenését jelenti.

9 Maga az adatfelvétel a közvélemény-kutatásokra vonatkozó szabályozásoknak maradéktalanul eleget tett. A válaszadók külön GDPR-nyomtatványon, saját aláírásukkal igazolták azt, hogy önkéntesen vesznek részt a felmérésben, válaszaikat pedig egy – a saját nevüket és személyes adataikat nem tartalmazó – külön adatbázisban kezeltük. A re-contact adatok összekapcsolását az ESS-kutatás eredményeivel kizárólag az anonimizált adatbázisokkal tettük meg, így a válaszadók személyazonossága sem a korábbi, sem a második adatfelvételi hullámban adott válaszaikkal nem beazonosítható.

1. ábra: Az ESS 8. magyarországi hullámának, illetve a „re-contact” felmérésnek a területi lefedettsége (térkép a mintába került településekről – zöld színnel jelölve)



Forrás: saját szerkesztés

A második hullám során az volt a legfontosabb szempont, hogy lehetőleg minden technikai részletet, kérdezői és általános értelemben vett adatfelvételi szabályt pontosan ugyanúgy valósítsunk meg, mint az első hullám esetében tettük, hiszen ez az összehasonlíthatóság miatt kulcsfontosságú. Minden kijelölt címet fel kellett keresni a kérdezőbiztosoknak függetlenül attól, hogy az előző hullámban sikeres vagy sikertelen volt az adott válaszadó¹⁰. A második hullám során az első hullám kérdőívének egy átdolgozott verziójával kérdeztünk: számos releváns demográfiai és attitűdre vonatkozó kérdést megtartottunk, amelyeket az intézményekbe vetett, illetve az általános bizalommal kapcsolatos kérdésekkel egészítettünk ki¹¹.

A sikeresség és a sikertelenség szempontjából a két hullám együttes eloszlását az 1. táblázat tartalmazza.

¹⁰ Az adott válaszadó a korábban kiválasztott konkrét személyt jelenti. Tehát nem csak ugyanazt a háztartást kellett a kérdezőknek felkeresniük, hanem pontosan ugyanazt a személyt. A két felmérés között nagyjából 1 év telt el, a felkeresések száma pedig mindkét kutatás esetében minimum 3 volt (ebből legalább egynek délelőtt és legalább egynek hétvégén kellett megtörténnie).

¹¹ A kutatás részleteiről lásd Szeitl–Tóth 2020.

1. táblázat: Az összes megkeresés együttes megoszlása (%)

	2. hullám, sikertelen minta	2. hullám, sikeres minta	Nem része a 2. hullám mintájának	Összesen
1. hullám, sikertelen minta	27,6	17,1	14,2	58,9
1. hullám, sikeres minta	11,7	19,5	9,9	41,1
Összesen	39,3	36,6	24,1	100,0

Az első hullám mintájának 24,1 százalékát, 954 esetet (költségkorlátok miatt) nem vettük be a második hullámba. Ebből 392 olyan eset, amelyet az első hullámban lekérdeztünk, de a másodikban nem, illetve 562 olyan eset, amelyet valamilyen okból nem találtunk meg az első hullámban, és nem kerestünk a másodikban sem. Ehhez adódik hozzá az az 1091 eset (a teljes mintakeret 27,6 százaléka), amelyet egyik vizsgálatban sem sikerült lekérdezni. A teljes minta 19,5 százalékáról, azaz 772 esetről rendelkezésre áll mindkét lekérdezett kérdőív, valamint az adott adatfelvételekhez tartozó megkeresési lapok információi. Van ezeken kívül 461 olyan eset, ahol csak az első hullám kérdezése zajlott, és 676 olyan eset, amelyről csak a második hullámban használt kérdőív adatai állnak rendelkezésre.

Ki kell emelni: a mi mostani kutatásunk szempontjából mindegyik kombinációnak fontos jelentősége van, különösen azzal együtt, hogy az egyik, a másik vagy mindkét adatfelvétel szerint „sikertelen” esetekre vonatkozóan is rendelkezünk információkkal (a megkeresett lakcímek jellemzői, a települések adatai, a megghiúsulások okai stb.), amelyeket a kérdezőbiztosoknak kellett vezetniük.

A továbbiakban elsőként a felkeresések stabilitását mutatjuk be, majd a nemválaszolók jellemzésére is kitérünk. Az elemzésekben (ha külön ki nem térünk rá) alapvetően a második hullámban is felkeresett személyek, azaz a 3000-es minta szerepel bázisként.

Eredmények a felkeresések stabilitásával kapcsolatban

Ahogy korábban említettük, a válaszadói magatartás megbízhatósága kulcsfontosságú az adatfelvételek minőségének és hatékonyságának szempontjából. Ennek vizsgálatához összesen 7 kategóriába osztottuk a címeteket a két hullám megkeresésének összesített eredményei alapján. Esetszámaikat és a százalékos megoszlásokat a 2. táblázatban közöljük.

2. táblázat: A két felmérés együttes megkeresési eredményei¹²

	N	%
STABIL válaszadó	772	25,7
MEGGYŐZHETŐ válaszmegtagadó	630	21,0
ELÉRHETŐ no-contact	398	13,3
VEGYES válaszadó	109	3,6
STABIL válaszmegtagadó	340	11,3
STABIL no-contact	194	6,5
VEGYES sikertelen	557	18,5
Összesen	3000	100,0

A második hullámban felkeresett minta negyede (25,7 százalék) tekinthető stabil válaszadónak, a válaszmegtagadók tekintetében pedig a minta tizede (11,3 százalék) olyan személy, aki mind a két felmérésben határozottan válaszmegtagadó volt, azaz sikerült kapcsolatba lépni vele, de a kérdező nem tudta rábírnival a válaszadásra. Definiáltunk továbbá a meggyőzhető válaszmegtagadókat is, akiket a két adatfelvételi hullám valamelyikében sikerült válaszadásra bírni. Ebbe a csoportba a minta ötöde (21 százalék) tartozik. No-contact vagy elérhetetlen címek esetében valamivel alacsonyabbak ezek az arányok (6,5 százalék, illetve 11,3 százalék). A továbbiakban a „VEGYES sikertelen” kategória heterogenitása, illetve a hibás címek miatti bizonytalanságok okán ezt a típust kizárjuk az elemzésből.

A stabil válaszadás inkább jellemző a nőkre, mint a férfiakra (30 százalék a 21 százalékkal szemben), illetve az idősebbekre. A stabil válaszmegtagadás és a stabil no-contact inkább a férfiakra jellemző, korcsoport tekintetében a meggyőzhető válaszmegtagadás inkább az idősebbekre, míg az elérhető no-contact inkább a fiatalokra jellemző (a nem változóra és a korcsoport szerinti bontásra is $p=0,00$). A településtípus szerinti bontást a 3. táblázat szemlélteti. Látható, hogy Budapest a legtöbb szempontból speciális helyzetben van, a leginkább szembetűnő különbséget az elérhető no-contact címek alacsony aránya (az átlagos 13 százalékkal szemben 2 százalék), illetve a stabil válaszmegtagadók átlagosnál magasabb aránya (az átlagos 11 százalékkal szemben 23 százalék) jelenti.

12 A különböző csoportok részletes jelentései:

A „STABIL válaszadó” esetében a kérdőívek mind a két adatfelvétel esetén sikerrel zárultak, azaz a megkeresett személy válaszolt a kérdőívre. „MEGGYŐZHETŐ válaszmegtagadó” alatt azokat a megkereséseket értjük, amelyek az egyik felmérés esetében sikerrel, de a másik felmérés esetében válaszmegtagadással zárultak (függetlenül attól, hogy melyik felmérésben, de fontos, hogy csak az egyik felmérésben készült sikeres kérdőív, a másik felmérésben pedig egyértelműen válaszmegtagadást rögzítettünk). Az „ELÉRHETŐ no-contact” kategória esetében az egyik kutatásnál az adott címen sikerült lebonyolítani a kérdést, de a másik adatfelvétel esetében a kérdező no-contactot rögzített, azaz nem talált otthon senkit. „VEGYES válaszadó” kategóriába soroltuk azokat a személyeket, akik a két felmérés közül egyszer válaszoltak a kérdőívünkre, a másik felmérés esetében pedig sikertelen volt a cím (de a sikertelenség nem egyértelműen válaszmegtagadást vagy no-contactot jelentett). A „STABIL válaszmegtagadók” azok a személyek, akik mind a két felmérés esetében visszautasították a válaszadást, a „STABIL no-contactokat” pedig egyik felmérésben sem találta otthon a kérdező, és nem tudott velük kapcsolatba lépni. A „VEGYES sikertelen” kategória esetében válaszmegtagadás, no-contact, hibás cím vagy egyéb sikertelenség szerepel a két felmérés eredménye között.

3. táblázat: A két felmérés együttes megkeresési eredményei településtípus szerint (% , $p=0,00$)

	Budapest	város	község	átlag
STABIL válaszadó	33	25	27	26
MEGGYŐZHETŐ válaszmeztagadó	25	19	21	21
ELÉRHETŐ no-contact	2	16	15	13
VEGYES válaszadó	0	4	4	4
STABIL válaszmeztagadó	23	9	8	11
STABIL no-contact	4	8	6	6
Összesen	100	100	100	100

Lakókörnyezet szempontjából rendelkezésünkre állnak olyan típusú információk, amelyek a lakóépület környékén található szemetességre, rendezettségére, illetve az általános állapotokra vonatkoznak (ezeket a kérdezők külön dokumentumon jelölik minden címről). A 4. táblázatban egy összetett mutató bontásában vizsgáljuk a megkeresési kimenetek stabilitását. Látható, hogy a nagyon rossz és a nagyon jó környékeken kevésbé számíthatunk a megkeresési eredmények változására a második hullámban, illetve a rossz környékeken a stabil elérhetetlenség aránya az átlagosnál valamivel magasabb.

4. táblázat: A két felmérés együttes megkeresési eredményei lakókörnyezet szerint (% , $p=0,01$)

	nagyon rossz	inkább rossz	inkább jó	nagyon jó	átlag
STABIL válaszadó	34	25	27	21	26
MEGGYŐZHETŐ válaszmeztagadó	21	21	19	26	21
ELÉRHETŐ no-contact	9	12	15	14	13
VEGYES válaszadó	4	3	4	3	4
STABIL válaszmeztagadó	11	13	9	12	11
STABIL no-contact	9	6	7	6	6
Összesen	100	100	100	100	100

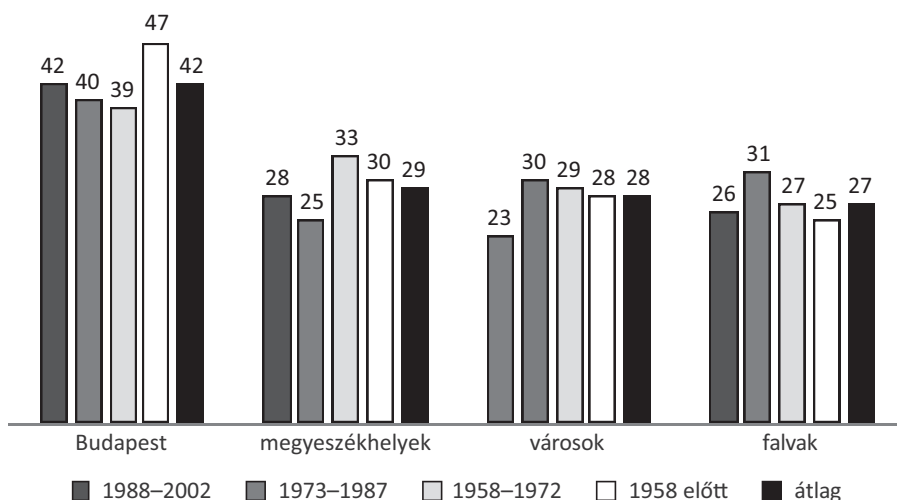
Fontos dimenzió a kérdezői hatás is, amelyet az 5. táblázatban olyan szempontból vizsgálunk, hogy milyen hatást gyakorol az a tény, hogy a két hullámban az adott címen ugyanaz a kérdező dolgozik. Azt találtuk, hogy a kérdezők állandósága növeli a stabil válaszadás arányát (31 százalék), viszont negatív irányba is befolyásol, hiszen a sikertelenségi eredmények is magasabbak (a stabil válaszmeztagadók és no-contactok esetében is). A kérdezők helyismertségének fontosságát támasztja alá az is, hogy a kistélepüléseken a stabil megkeresési eredmények az országos átlagnál nagyobb arányban tapasztalhatóak.

5. táblázat: A két felmérés együttes megkeresési eredményei kérdezőbiztos szerint (% , $p=0,00$)

	más kérdező	ugyanaz a kérdező	átlag
STABIL válaszadó	23	31	26
MEGGYŐZHETŐ válaszmegtagadó	24	14	21
ELÉRHETŐ no-contact	15	10	13
VEGYES válaszadó	4	3	4
STABIL válaszmegtagadó	10	14	11
STABIL no-contact	5	9	6
Összesen	100	100	100

Korábbi tapasztalataink alapján azt várhatjuk, hogy különböző településtípusok esetében más az elérés, amit fontos figyelembe venni a kutatások tervezésénél és a mintavételi dizájn kialakításánál. A 2. és a 3. ábrán külön-külön mutatjuk be a válaszmegtagadók, majd a no-contact, azaz az elérhetetlen címek megoszlását településtípusokon belül a különböző korcsoportok szerinti bontásban¹³. A 2. ábra alapján elmondható, hogy míg a megyeszékhelyeken, a városokban és a falvakban a legidősebb korcsoportban az átlagnál alacsonyabb válaszmegtagadási arány tapasztalható, addig ez Budapesten magasabb, a kisebb településeken (városok és falvak) a fiatalabbakat is könnyebben sikerül válaszra bírni.

2. ábra: A válaszmegtagadó címek aránya a különböző településtípusokon életkori csoportok szerint (%)

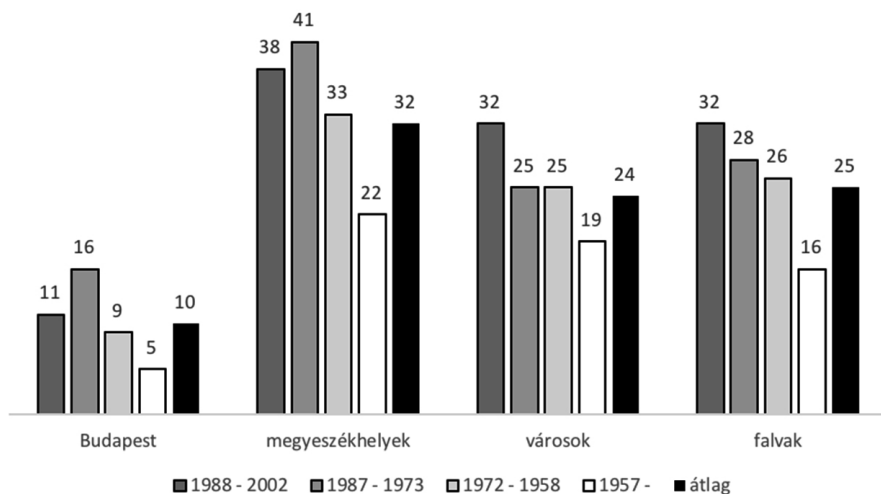


Megjegyzés: Az életkori csoportokat az ábrán szereplő születési évszámok alapján értjük. A különböző településtípusok esetében a khi-négyzet próbákhoz tartozó fontosabb értékek (próbastatisztika értéke és szignifikanciaszint sorrendjében) a következők: Budapesten belül 25, 245 és 0,003; a megyeszékhelyek esetében 41,966 és 0,000003; a városok esetében 40,310 és 0,000007; a községek esetében pedig 37,847 és 0,00001.

¹³ Erre azért van lehetőség, mert a mintakeretbe tartozó címeket bizonyos alapvető demográfiai adatokkal együtt szerezzük be, amilyenek például a nem, illetve a korcsoportba tartozás. Így ezek az adatok azokra a válaszadókra vonatkozóan is a rendelkezésünkre állnak, akik egyik felmérés során sem válaszoltak, vagy egyik felmérés során sem tudtunk kapcsolatba lépni velük.

A no-contact esetében (3. ábra) is érdekes különbségeket tapasztalhatunk: habár mindegyik településtípusra igaz, hogy minél idősebb a kiválasztott személy, annál könnyebben meg lehet találni, de szembeűnő, hogy a kisebb településen (városok és falvak) kisebbek az eltérések a korcsoportokon belül. A budapestiek drasztikusan alacsony kontaktálási esélyei a korcsoport szerinti bontásban is eltérő képet mutatnak (a legidősebb korosztály no-contact aránya [5 százalék] a belső átlagnál [10 százalék] jóval alacsonyabb).

3. ábra: A no-contact címek aránya a különböző településtípusokon életkori csoportok szerint (%)



Megjegyzés: Az életkori csoportokat az ábrán szereplő születési évszámok alapján értjük. A különböző településtípusok esetében a khi-négyzet próbákhoz tartozó fontosabb értékek (próbastatisztika értéke és szignifikanciaszint sorrendjében) a következők: Budapesten belül 25, 245 és 0,003; a megyeszékhelyek esetében 41,966 és 0,000003; a városok esetében 40,310 és 0,000007; a községek esetében pedig 37,847 és 0,00001.

A válaszmegtagadók jellemzése

A továbbiakban kifejezetten a válaszmegtagadók egy csoportjának alaposabb vizsgálatával foglalkozunk (azaz azokkal, akik valamelyik felmérés során kifejezetten megtagadták a választ, míg a másik felmérésben válaszoltak¹⁴), és arra vagyunk kíváncsiak, miben térnek el a válaszmegtagadók jellemzői a többiekétől.¹⁵

14 Valamelyest ingadozó válaszmegtagadóknak is tekinthetjük ezt a csoportot.

15 Az nyilván kutatói döntés kérdése, hogy ilyenkor az elemzést a teljes mintakeret két kategóriája között végezzük el (válaszmegtagadók és mindenki más), vagy csak azok között, akikkel kapcsolatba tudunk kerülni (válaszmegtagadók és válaszadók). Mivel itt a re-contact dizájn miatt volt adatunk az adott vizsgálatban no-contact személyekről is, az előbbi kategorizálás mellett döntöttünk. Természetesen a szűkebb leválogatásra is elkészítettük az elemzést, de a legfontosabb összefüggések tekintetében nagyon hasonló eredményeket kaptunk.

A megtagadást potenciálisan befolyásoló tényezők három csoportját különítjük el. Ezek egyrészt a települési és lakókörnyezeti jellemzők, majd a megkérdezettek demográfiai jellemzői, amelyek közül feltételezésünk szerint a kérdezettek neme, korcsoportja, iskolázottsága és anyagi helyzete lehet fontos¹⁶. A harmadik csoportba tartoznak az adatfelvételi jellemzők, amelyek közül az instruktor és a kérdező kódját¹⁷, valamint a kérdezők iskolai végzettségére, nemére és korára vonatkozó információkat is használjuk. Mindemellett megvizsgáljuk azt is, hogy szerepet játszik-e az a tény, hogy ugyanarra a címre ugyanaz a kérdező, vagy éppen egy másik kérdező ment vissza.

Ezt az elemzést az első hullám (n = 3954) adatain végezzük el¹⁸, és a települési és lakókörnyezeti jellemzőkkel kezdjük¹⁹. Természetesen nem a települések tagadják meg a választ, hanem a személyek, de ettől még elképzelhető, hogy bizonyos közegben lakók inkább, más közegben lakók pedig kevésbé utasítják el azt, hogy kooperáljanak egy adatfelvételben.

A válaszmegtagadás valószínűsége az alapmodellben (minden egyéb, ebben az elemzésben megvizsgált tényezőt változatlanul véve) nem tér el lényegileg a nagyobb és a kisebb településeken, különbség van viszont a „gazdagabb” településeken és Budapesten (6. táblázat). A lakókörnyezet fizikai állapota (vagy legalábbis annak a kérdező által megítélt szintje) lényegi hatást gyakorol a válaszmegtagadás valószínűségére. A „jó vagy nagyon jó” környékeken lényegesen nagyobb, a „rossz vagy nagyon rossz” helyeken pedig lényegesen kisebb a válaszmegtagadás valószínűsége, mint azokon a helyeken, amelyeket a kérdező „megfelelőnek” minősített. A kérdezettek demográfiai összetételének bevonása nem változtat ezeken a hatásokon, kivéve egy esetben: a kérdezettek attribútumaira kontrolláló második modellben a legnagyobb lakosságú településeken szignifikáns lesz a visszautasítás alacsonyabb szintje.

Összességében azt tapasztaljuk, amit várhattunk is: ha bevonjuk a modellbe a megkérdezettek egyéni jellemzőit akkor jelentősen növekszik a magyarázó modelljeink prediktív ereje. (Ezt mutatja a különböző pszeudo R^2 mutatók változása a modell kibővítésének eredményeképpen. A Cox & Snell R^2 értéke 0,03-ról 0,17-re, a Nagelkerke R^2 értéke pedig 0,04-ről 0,24-re növekszik.)

16 Az anyagi helyzetet egy „objektív” (havi nettó jövedelem) és egy szubjektív (a kérdezett értékelése arról, hogy mennyire jól jönnek ki adott jövedelemből) változóval mérjük. A szóban forgó változók közül a nem és a korcsoport a népesség-nyilvántartótól a címhez kapott információ. A többi esetben azt használjuk ki, hogy voltak olyan eseteink, amelyeket kétszer megkérdeztünk, és a kettőből az egyik alkalommal szereztünk róluk iskolázottsági vagy jövedelmi információt. Ilyen esetekben az egyik vizsgálatból a másikba imputáltuk a begyűjtött értéket.

17 Ezeket a többváltozós modellekben fix hatásként vezetjük be.

18 Az, hogy helyenként az ESS, helyenként a Bizalom, helyenként pedig az egyesített adatbázis adatain dolgozunk, attól függ leginkább, hogy milyen változók szerepelnek a modelljeinkben. Mivel helyenként az ismételt (és sikeres) megkeresés révén adatokat tudtunk gyűjteni azokról is, akik egyik vagy másik vizsgálatban nem válaszoltak, vagy elérhetetlennek bizonyultak, egyes változókat (iskolai végzettséget, életkort, jövedelmi szintet stb.) imputáltunk egyik vizsgálatból a másikba. Ezeknek a két adatfelvétel közötti változását kisebb kockázatnak tartottuk, mint amekkora előnyökkel kecsegtetett az, hogy be tudtuk azonosítani (vagyis inkább: közelíteni tudtuk) a meg nem találtak vagy elutasítók jellemzőit is.

19 Fontos, hogy a települési jellemzők a válaszadó lakóhelyére vonatkozóan kerülnek be a modellbe, és ezáltal érhetőek el ilyen típusú adatok a nemválaszolók mintára is.

6. táblázat: A visszautasítás (válaszmeztagadás) valószínűsége és magyarázó tényezői az ESS-vizsgálatban (n=3954)²⁰

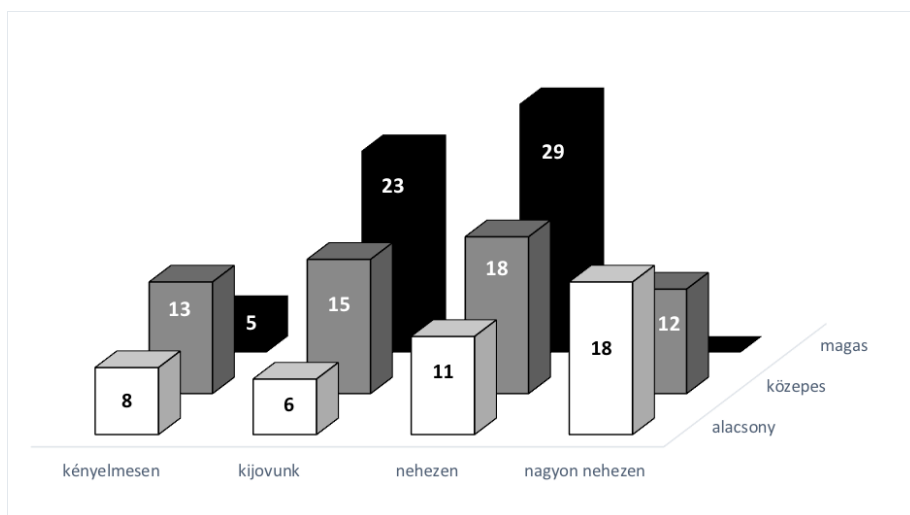
	1. modell: települési jellemzők			2. modell: települési és személyi jellemzők			3. modell: települési, személyi és adatfelvételi jellemzők		
	B	Exp(B)	sig	B	Exp(B)	sig	B	Exp(B)	sig
Magyarázandó (nemválaszolás)									
A település népességszáma 2014-ben (10001–50000)									
0–2000 fő	-0,02	0,98		-0,03	0,97		-0,06	0,94	
2001–10000 fő	0,09	1,10		0,07	1,07		0,03	1,03	
50001–város	-0,18	0,84		-0,36	0,70	***	-0,38	0,69	**
A település fajlagos adóerő-képessége (10000–34999 Ft)									
0–9999 Ft	0,18	1,20		0,21	1,23	*	0,24	1,27	*
35000Ft felett, város	0,52	1,69	***	0,50	1,65	***	0,50	1,64	***
Budapest	0,95	2,59	***	1,18	3,25	***	0,72	2,06	***
A lakókörnyezet fizikai állapota (megfelelő)									
Jó vagy nagyon jó	0,22	1,25	***	0,24	1,27	**	0,19	1,21	*
Rossz vagy nagyon rossz	-0,66	0,52	***	-0,79	0,45	***	-0,71	0,49	***
A kérdezett neme (férfi)				-0,02	0,98		-0,02	0,98	
A kérdezett korcsoportja (születési év, ref.: 1953–1972)									
1988–				-0,30	0,74	***	-0,30	0,74	***
1973–1987				-0,14	0,87		-0,16	0,85	
1953 előtt				0,16	1,18		0,16	1,18	
A kérdezett iskolai végzettsége (legfeljebb szakmunkás)									
Érettségi				-0,28	0,75	**	-0,26	0,77	*
Felsőfok				-0,30	0,74		-0,26	0,77	
Összevont jövedelemkategóriák (havi nettó jövedelem: 181–240 ezer forint közötti)									
125 ezer vagy alatta				-0,18	0,83		-0,15	0,86	
126–180 ezer				-0,30	0,74		-0,21	0,81	
241–300 ezer				0,68	1,97	***	0,69	1,99	***
301–400 ezer				0,90	2,46	***	0,86	2,37	***
401 ezer vagy fölötté				0,49	1,64		0,42	1,52	
Szubjektív anyagi helyzet („kényelmesen megélünk”)									
Kijövünk				0,86	2,35	***	0,78	2,17	**
Nehezen élünk meg				1,12	3,06	***	1,10	3,00	***
Nagyon nehezen élünk meg				1,29	3,65	***	1,36	3,90	***
A kérdező neme (férfi)							0,00	1,00	
Kérdező korcsoportja (születési év: 1953–1972)									
1987 után							-1,00	0,37	**
1973–1987							0,45	1,57	***
1953 előtt							-0,21	0,81	
A kérdező iskolázottsága (végzettség: érettségi)									
Legfeljebb szakmunkás							-0,06	0,94	
Felsőfok							-0,79	0,45	***
Kérdezői stabilitás (nem ugyanaz kérdezte)									
							-0,01	0,99	
Konstans		-0,70			-2,59			-2,04	
-2 Log likelihood		4736			4130			3914	
Cox & Snell R Square		0,03			0,17			0,20	
Nagelkerke R Square		0,04			0,24			0,28	
Jó predikciók aránya		69,70			74,40			76,30	
Ebből a helyesen előrejelzett kitüntetett érték aránya					41,90			48,30	

20 Az ESS kiinduló teljes mintakeretének nagysága 3954 fő, körükből került kiválasztásra a Bizalom kutatás 3000-es mintakerete (lásd 1. táblázat).

Ugyanakkor azt láthatjuk, hogy a személyi (demográfiai) jellemzők tekintetében nem rajzolódik ki határozott mintázat a válaszmegtagadókra. A megtagadás esélye enyhén kisebb a legfiatalabbak (1988 után születettek), mint a középkorúak (1953 és 1972 között születettek) között. A megtagadás kockázata nagyobb a felsőbb jövedelmi kategóriákban (legalábbis a havi nettó 240 és 400 ezer forint közötti jövedelmi csoportban), mint a közepes jövedelműek (havi nettó 181–240 ezer forint közöttiek) kategóriájában (és a tőlük ebből a szempontból nem különböző alacsonyabb jövedelmű kategóriákban). Érdekes viszont, hogy a szubjektív jövedelmi kategóriákat illetően éppen fordítottnak látszik a trend: azok, akik nagyobb anyagi gondokról számolnak be, nagyobb mértékben lesznek válaszmegtagadók, mint azok, akik azt mondják, hogy „kényelmesen kijövünk a pénzünkől”.

Ez utóbbi jelenséget érdemes kicsit alaposabban megvizsgálni. A 4. ábra eredményei alapján az sejthető, hogy van egy sajátos keresztthatás az objektív jövedelmi szint, a szubjektív jövedelemérzet és a megtagadási arányok között.

4. ábra: Megtagadási arányok (adott kategóriában a megtagadók aránya az adott kategóriába tartozó összes „célszemélyhez” viszonyítva) a jövedelemszint és a szubjektív jövedelemszint együttes kategóriáiban (%)



Az egyik tengelyen a szubjektív értékelés („kényelmesen kijövünk” stb.) látszik, a másikon pedig a jövedelmi kategóriák, három csoportba sorolva (legalacsonyabbak, közepesek, magasabbak). Az oszlopok magassága az adott kategórián belül a megtagadások átlagos értéke. Azt látjuk, hogy a legalacsonyabb jövedelmi kategóriában a szubjektív nehézségek növekedésével nő a megtagadási arány is. Ugyanakkor a közepes nehézséget érzékelők között annál nagyobb a megtagadási arány, minél magasabb egyébként az objektív jövedelem. A legmagasabb megtagadási rátát éppen abban a kategóriában látjuk, ahol egyébként a mért „objektív” jövedelmek a legmagasabbak.

Összefoglalás és következtetések

Cikkünkben egy speciális „re-contact” szerkezetű, kifejezetten survey-módszertani felmérés eredményeit mutattuk be, két kérdésre koncentrálnak. Egyrészt azt jártuk körbe, hogy egy első és második hullámra osztott adatfelvétel során mennyire stabilak a megkeresési eredmények, ami alapvető kapcsolatban áll az egyik legfontosabb minőségi kritériummal, a megbízhatósággal. Másik kérdésünk az volt, hogy hogyan tudjuk pontosabban jellemezni azokat a személyeket, akiket nem sikerül elérni a felmérésekben vagy azért, mert megtagadják a válaszadást, vagy pedig azért, mert nem sikerül kapcsolatba lépni velük.

A válaszadási arányok stabilitása és megbízhatósága szempontjából állandóságot figyelhetünk meg az újra felkeresett minta vizsgálatok is, de a belső sikereségi és teljesülési mutatók földrajzi és demográfiai bontásban valamelyest eltérnek. Az első és a második megkeresési hullámban is a standardnak tekinthető, nagyjából 40 százalékos sikereségi arányt tapasztaltunk (az ESS-minta esetében 41,1 százalék, az újra felkereséskor pedig valamivel magasabb, 48% volt ez az arány). Azok, akiket mindkét felmérésben sikerült lekérdezni, a teljes címlista mintegy negyedét tették ki, míg a mindkét felmérésben elérhetetlen (no-contact) válaszadók aránya 6,5 százalék, a stabilan válaszmegtagadóké 11,3 százalék volt. Fontos eredmény, hogy a bizonytalan/meggyőzhető címek majdnem a teljes minta harmadát teszik ki, amiből az következik, hogy az adatfelvételeknek az a szegmense, amelyik adott pillanatban nem elérhető valamilyen okból, nem tekinthető reménytelennek, hiszen egy másik alkalommal sikerült őket kooperációra bírni. Az újra elérés sikerességét terepmunka-specifikus jellemzők is befolyásolják: tapasztalatunk szerint a stabil válaszadók aránya magasabb azon címek esetében, amelyeket mindkét hullám esetében ugyanaz a kérdező próbál lekérdezni, ez a hatás pedig kisebb településtípusok esetében erősödik. A kérdezői hatás negatív irányban is jelentős, hiszen az első hullámban sikertelen címek nagyobb valószínűséggel lesznek a második hullámban is sikertelenek, ha ugyanaz a kérdező. A meggyőzhetőség tekintetében nem tapasztaltunk olyan kapcsolatot, hogy a tapasztaltabb kérdezők nagyobb sikerességgel tudnak teljesíteni. Komoly instruktori hatást tapasztaltunk ugyanakkor a kiosztott minta sikereségi eredményei tekintetében, ami azt mutatja, hogy a terepen teljesítők munkájának standardizálása felé nagyobb figyelmet kell fordítani.²¹

A válaszadási valószínűséget befolyásoló tényezők három nagyobb csoportra oszthatóak, amelyek különböző mértékben befolyásolják a „no-contact” és a válaszmegtagadás valószínűségét.

Területi és lokális dimenziók. Budapesten a meg nem találási esély szignifikánsan alacsonyabb, mint a közepes adóerő-képességű településeknél. A lakókörnyezet fi-

21 A kérdezőkkel, illetve a kérdezői munkával kapcsolatos következtetésekhez természetesen szorosan kapcsolódik az olyan típusú adathiba, amelyet a kérdező szándékosan követ el. Az ilyen típusú hibákat a kérdezői munka rendszeres és szigorú ellenőrzésével tudjuk kontrollálni, de természetesen ez a kontrol sem lehet 100 százalékos. Habár a kérdezői hibák és csalások minimalizálását semmi sem biztosítja, az adatfelvételi rendszerhez kapcsolódó következtetéseket ennek szellemében vagyunk kénytelenek meghozni.

zikai állapota fontos hatást gyakorol az adott cím el nem érésének valószínűségére. Ahol ez a környezet „rossz vagy nagyon rossz” minőségű, ott lényegesen nagyobb eséllyel leszünk képtelenek bárkit is elérni, mint ahol ez a környezet „megfelelő”.

A kiválasztott egyén demográfiai jellemzői. A megtagadás esélye enyhén kisebb a legfiatalabbak (1988 után születettek), mint a középkorúak (1953 és 1972 között születettek) esetében. A megtagadás kockázata nagyobb a felsőbb jövedelmi kategóriákban (legalábbis a havi nettó 240 és 400 ezer forint közötti jövedelmi csoportban), mint a közepes jövedelműek (havi nettó 181–240 ezer forint közöttiek) kategóriájában (és tőlük ebből a szempontból nem különböző alacsonyabb jövedelmű kategóriákban). Érdekes viszont, hogy a szubjektív jövedelmi kategóriákat illetően éppen fordítottnak látszik a trend: azok, akik nagyobb anyagi gondokról számolnak be, nagyobb mértékben lesznek válaszmegtagadók, mint azok, akik azt mondják, hogy „kényelmesen kijövünk a pénzünkből”. Várakozásainknak némiképpen ellentmond az a megfigyelés, hogy a magasabb iskolai végzettségűek között alacsonyabb a megtagadók és a meg nem találtak aránya is, illetve külön érdekes az is, hogy a Budapest-város-falu dimenzióban vizsgálva egy viszonylag konzisztens települési lejtő is kirajzolódik: a megtagadás Budapestről a települési lejtőn lefelé haladva csökken, a nem megtalálás viszont növekszik.

A terepmunka és a kérdezői hálózat. A kérdezői hálózattal kapcsolatos két változó (az instruktor és a kérdezők személye) fontos magyarázó erővel bír, hiszen beemelésük a magyarázó modellbe valamelyest javítja mind a modellek magyarázó erejét, mind pedig azok illeszkedését. Ennek mélyebb megértéséhez további kutatások szükségesek, jelenleg csak a terepmunkával kapcsolatos tapasztalatainkra, illetve az általános intuícióinkra támaszkodhatunk: az instruktorok személye, precizitása, tapasztaltsága, ellenőrzési gyakorlata eltérő, ami minden bizonnyal hatással van az általuk vezetett csapat munkájára, ahogy a csapat kialakulása, a tagok összeválogatása esetében is elképzelhető szelekciós torzítás.

A jövőbeni empirikus adatfelvételek tervezői számára megfogalmazhatunk néhány következtetést a mintavételre és a terepmunka megszervezésére egyaránt.

A mintavételi folyamat megtervezését illetően fontos tapasztalat, hogy a mintába kerülési valószínűségek a csökkenő/fogyó mintás kutatások esetében egy újra felkeresett minta alapján pontosabban kioszthatók, ha a településrétegeken belüli nem és kor (esetleg háztartásméret) mentén az így kialakított cellákon belüli válaszadási valószínűségekkel is számolunk a bruttó minta kialakításakor²². A mintába kerülési valószínűségek továbbá az olyan mintavételi módszerek bizonyos lépéseibe építhetők be, mint a véletlensétás, háztartáson belüli kiválasztáson alapuló módszerek, amelyek a Kish-kulcsos kiválasztással dolgoznak. Ez jelenleg három paramétert vesz

22 Evidensnek tűnhet, de a gyakorlatban a survey felmérések mintavétele jellemzően (néhány speciális felmérést leszámítva) nem követi a releváns válaszmegtagadási mintázatokat demográfiai csoportosításokra nézve. Általában településtípusokon belüli elérési arányok szerint történik a minta allokálása, valamint a címek rendelése is.

figyelembe: nem, (relatív) korcsoport, háztartásméret. Az ezek által kapott kiválasztási valószínűségeket súlyozni lehet a válaszadási valószínűségek stabilitásával.

A terepmunka különböző tényezőinek standardizálása komoly hatást gyakorol az eredményekre. A különböző kérdezői csapatok azonos szinten történő képzése kulcsfontosságú, és ugyanilyen fontos a kérdezőkről való pontos információk begyűjtése, amelyek alapján rendszeres ellenőrzések végezhetőek a kérdezők hatékonyságával kapcsolatban. A kérdezők demográfiai jellemzői, munkatapasztalatuk szintén kulcsfontosságú. Az eredeti és a megismételt adatfelvételben elért és el nem ért esetek részletes elemzése azt mutatja, hogy van értelme több részletben lebonyolítani az empirikus adatfelvételeket. Kutatásunk egyértelműen bizonyította, hogy egy adott vizsgálat sikertelen eseteit nem szabad „reménytelennek” tekinteni, hiszen az ő körükben is majdnem akkora volt a második körös elérés, mint amekkora magában a fő vizsgálatban volt (38,3 százalék a 41,1 százalékhoz viszonyítva). Ha egy adott alkalommal lekérdezzük egy mintát, majd egy év vagy esetleg több idő múlva visszamegyünk hozzájuk, határozottan jobb elérési arányokkal dolgozhatunk, ami akár a kettős adatfelvételi dizájn plusz költségeit is megtérítheti.

Megismételt adatfelvételek esetében érdemes lehet tudatosan allokálni a kérdezőket. Akik a korábbi vizsgálatban sikeresen kerestek meg egy háztartást, ott érdemes visszaküldeni ugyanazt a kérdezőt. Ott azonban, ahol egy cím az egyik hullámban sikertelen volt, érdemes „cserélni”, és esélyt kell adni annak, hogy egy új, másik kérdező sikeresebb lesz.

Vizsgálatunk több ponton is rámutatott, hogy az adatfelvételi apparátus felkészültsége és összetétele rendkívül fontos eleme az empirikus vizsgálatok adatminőségének. Mindezt nemcsak a konkrét megvalósulási eredmények elemzése támasztotta alá, hanem az is, hogy a megkérdezettek egyértelműen jelezték: a kérdező személye mennyire fontos abban, hogy „szóba állnak-e” a közvélemény-kutató céggel a megkérdezni kívánt állampolgárok. Ezért külön kiemeljük, hogy egyszerre van szükség a kérdezők kiválasztásának, képzésének és motivációs szerkezetének a javítására, és legalább ennyire fontos az is, hogy a kérdezői apparátus ellenőrzése és minőségbiztosítása minél magasabb szinten valósuljon meg.

Abstract: One of the major challenges in empirical social research is the decreasing response rate. The problem affects Hungarian and international data collections as well and may lead to bias in most cases. The key question is what are the socio-economic characteristics of non-responders. This article reports a methodological research on survey nonresponse in Hungary. The main question is how non-responders and ineligible households can be characterized, and how data collection procedures can be improved. In the fieldwork we followed “re-contact” design; as a new survey, we re-contacted the original list of selected responders of the 8th wave of the European Social Survey with a short questionnaire, following the same standards regarding the entire fieldwork. The proportion of those, successfully reached in both surveys was 25.7%, proportion of totally ineligible was 6.5%, and the rate of stable non-responders was 11.3%. To examine inaccessibility and non-response, we observed the socio-economic characteristics of the responders, the characteristics of the settlement and environment of the given addresses, the characteristics of the fieldwork team and other external variables of the data collection.

According to our results, besides socio-economic characteristics, fieldwork-specific factors also influence the success rate of re-contacts, and the physical condition of the living environment was also found to be significant. Moreover, our study suggests that non-response rates are also influenced by psychological factors that cannot be taken into account during sampling, such as specific interaction between the objective and subjective evaluation of own living conditions. Besides the conclusions, we also propose additional research directions and potential improvements in specific data collection techniques.

Key words: social research, survey nonresponse, bias

Irodalomjegyzék

- Beullens, K. – Loosveldt, G. – Vandenplas, C. – Stoop, I. (2018): Response Rates in the European Social Survey: Increasing, Decreasing, or a Matter of Fieldwork Efforts? *Survey Methods: Insights from the Field*: 1–12. <https://doi.org/10.13094/smif-2018-00003> (utolsó letöltés: 2021. 02. 01.)
- Blom, A. G. (2008): *Measuring nonresponse cross-nationally*. ISER Working Paper Series, No. 2008-41. Colchester: University of Essex – Institute for Social and Economic Research (ISER). <http://dx.doi.org/10.2866/936989>
- Blom, A. G. – Lynn, P. – Jäckle, A. (2008): *Understanding cross-national differences in unit non-response: the role of contact data*. ISER Working Paper Series, No. 2008-01. Colchester: University of Essex – Institute for Social and Economic Research (ISER).
- Boros K. (2020): A nemzetközi surveyek módszertani diverzitása a nemválaszolás szempontjából. In Szeitl B. – Tóth I. Gy. (szerk.): *Megközelíthetetlen csoportok elérése: hogyan lehet javítani a személyes megkeresésen alapuló empirikus adatfelvételek minőségén? A „Hogyan kezeljük a „nemválaszolók”-at? – Az empirikus kutatások során jellemzően válaszmegtagadó, vagy elérhetetlen csoportok pilot kutatása speciális felméréssel és módszerekkel” című K 125162 számú OTKA Kutatás záró beszámolója*. Budapest: Társi.
- Goyder J. (2019): *The Silent Minority: Nonrespondents on Sample Surveys*. New York: Routledge <https://doi.org/10.4324/9780429314520>
- Groves R., Cialdini R., Couper M. (1992): Understanding The Decision to Participate in a Survey. *Public Opinion Quarterly* 56. <https://doi.org/10.1086/269338>
- Havasi É. (1997): Válaszmegtagadó háztartások. *Statisztikai Szemle*, 75(10): 831–843. http://www.ksh.hu/statszemle_archive/1997/1997_10/1997_10_831.pdf (utolsó letöltés: 2021. 02. 01.)
- de Heer, W. (1999): International Response Trends: Results of an International Survey. *Journal of Official Statistics*, 15(1): 129–142.
- KSH (2017): *Mikrocenzus 2016. 2. A népesség és a lakások jellemzői*. Budapest: KSH. http://www.ksh.hu/docs/hun/xftp/idoszaki/mikrocenzus2016/mikrocenzus_2016_2.pdf (utolsó letöltés: 2020. 02. 01.)

- KSH (2018): *Mikrocenzus 2016. 9. A foglalkozási szerkezet változása és jellemzői Magyarországon*. Budapest: KSH.
https://www.ksh.hu/docs/hun/xftp/idoszaki/mikrocenzus2016/mikrocenzus_2016_9.pdf
 (utolsó letöltés: 2020. 02. 01.)
- de Leeuw, E. – de Heer, W. (2002): Trends in household survey nonresponse: a longitudinal and international comparison. In Groves, R. – Dillman, D.A. – Eltinge, J. L. – Little, R. J. A. (eds.): *Survey nonresponse*. New York: John Wiley, 41–54.
- de Leeuw, E. – Hox, J. – Luiten, A. (2018): International Nonresponse Trends across Countries and Years: An analysis of 36 years of Labour Force Survey data. *Survey Insights: Methods from the Field*. <https://doi.org/10.13094/SMIF-2018-00008>
 (utolsó letöltés: 2021. 02. 01.)
- Meyer, B. D. – Mok, W. K. C. – Sullivan, J. X. (2015): *Household Surveys in Crisis*. NBER Working Paper No. 21399. Issued in July 2015.
<http://dx.doi.org/10.3386/w21399> (utolsó letöltés: 2021. 02. 01.)
- MNB (2019): *Miből élünk? a 2017-es háztartási vagyonszerzés első eredményei*. Budapest: Magyar Nemzeti Bank Statisztikai Igazgatóság.
- Osier, G. (2016): *Unit non-response in household wealth surveys. Experience from the Eurosystem's Household Finance and Consumption Survey*. ECB Statistics Paper No 15. Frankfurt am Main: European Central Bank. July 2016. <https://doi.org/10.2866/844445>
<https://www.ecb.europa.eu/pub/pdf/scpsps/ecbsp15.en.pdf?2f354a8519e343ed4bce974d250d87e5> (utolsó letöltés: 2021. 02. 01.)
- Simon B. – Valentiny Á. (2016): Miből élünk? Az első átfogó hazai háztartási vagyonszerzés bemutatása. *Statisztikai Szemle*, 94(7): 717–735. <http://dx.doi.org/10.20311/stat2016.07.hu0717>
- Stoop, I. – Billiet, J. – Koch, A. – Fitzgerald, R. (2010): Improving survey response: Lessons learned from the European Social Survey. *Revista Espanolade Investigaciones Sociologicas*, 1. 166–170. <http://dx.doi.org/10.1002/9780470688335>
- Szeitl B. – Tóth I. Gy. (2020): *Megközelíthetetlen csoportok elérése: hogyan lehet javítani a személyes megkeresésen alapuló empirikus adatfelvételek minőségén? A „Hogyan kezeljük a „nemválaszolók”-at? – Az empirikus kutatások során jellemzően válszagmentagadó, vagy elérhetetlen csoportok pilot kutatása speciális felméréssel és módszerekkel” című K 125162 számú OTKA Kutatás záró beszámolója*. Budapest: Tárki.
- Tárki (2016): *„Makro sokkok – mikro válaszok: sikeres és sikertelen háztartási alkalmazkodás a válság idején Magyarországon - A Tárki Háztartás Monitor kutatás”*. 1. kötet. A kutatás koncepciója, adatfelvételei és dokumentációja. Budapest: Tárki.
http://www.tarki.hu/hu/research/hm/2016/monitor2015_gyorsjelentes_1_kotet.pdf
 (utolsó letöltés: 2021. 02. 01.)