

Közzététel: 2021. december 14.

A tanulmány címe:

Keverni vagy nem keverni? A hibrid adatfelvételekben rejlő lehetőségek politikai témájú kérdőívek esetén

Szerző:

STEFKOVICS ÁDÁM,

az Eötvös Loránd Tudományegyetem egyetemi tanársegédje, PhD-hallgatója

E-mail: adam.stefkovics@tatk.elte.hu

DOI: <https://doi.org/10.20311/stat2021.12.hu1127>

Az alábbi feltételek érvényesek minden, a Központi Statisztikai Hivatal (a továbbiakban: KSH) Statisztikai Szemle c. folyóiratában (a továbbiakban: Folyóirat) megjelenő tanulmányra. Felhasználó a tanulmány vagy annak részei felhasználásával egyidejűleg tudomásul veszi a jelen dokumentumban foglalt felhasználási feltételeket, és azokat magára nézve kötelezőnek fogadja el. Tudomásul veszi, hogy a jelen feltételek megszegéséből eredő valamennyi kárért felelősséggel tartozik.

1. A jogszabályi tartalom kivételével a tanulmányok a szerzői jogról szóló 1999. évi LXXVI. törvény (Szt.) szerint szerzői műnek minősülnek. A szerzői jog jogosultja a KSH.
2. A KSH földrajzi és időbeli korlátozás nélküli, nem kizárólagos, nem átadható, térítésmentes felhasználási jogot biztosít a Felhasználó részére a tanulmány vonatkozásában.
3. A felhasználási jog keretében a Felhasználó jogosult a tanulmány:
 - a) oktatási és kutatási célú felhasználására (nyilvánosságra hozatalára és továbbítására a 4. pontban foglalt kivétellel) a Folyóirat és a szerző(k) feltüntetésével;
 - b) tartalmáról összefoglaló készítésére az írott és az elektronikus médiában a Folyóirat és a szerző(k) feltüntetésével;
 - c) részletének idézésére – az átvevő mű jellege és célja által indokolt terjedelemben és az eredetihez híven – a forrás, valamint az ott megjelölt szerző(k) megnevezésével.
4. A Felhasználó nem jogosult a tanulmány továbbértékesítésére, haszonszerzési célú felhasználására. Ez a korlátozás nem érinti a tanulmány felhasználásával előállított, de az Szt. szerint önálló szerzői műnek minősülő mű ilyen célú felhasználását.
5. A tanulmány átdolgozása, újra publikálása tilos.
6. A 3. a)–c.) pontban foglaltak alapján a Folyóiratot és a szerző(ke)t az alábbiak szerint kell feltüntetni:
„*Forrás: Statisztikai Szemle c. folyóirat 99. évfolyam 12. számában megjelent, **Stefkovics Ádám** által írt, **'Keverni vagy nem keverni? A hibrid adatfelvételekben rejlő lehetőségek politikai témájú kérdőívek esetén'** című tanulmány (link csatolása)*”
7. A Folyóiratban megjelenő tanulmányok kutatói véleményeket tükröznek, amelyek nem esnek szükségképpen egybe a KSH vagy a szerzők által képviselt intézmények hivatalos álláspontjával.

Stefkovics Ádám

**Keverni vagy nem keverni?
A hibrid adatfelvételekben rejlő lehetőségek
politikai témájú kérdőívek esetén***

**To mix or not to mix?
Potential of mixed-mode data collections
in the case of political surveys**

STEFKOVICS ÁDÁM,
az Eötvös Loránd Tudományegyetem egyetemi tanársegédje, PhD-hallgatója
E-mail: adam.stefkovics@tatk.elte.hu

A világszerte csökkenő válaszadási hajlandóság és más problémák miatt egyre kevésbé képes jó minőségű és megbízható adatokkal szolgálni az empirikus szociológia meghatározó hagyományos módszere, a kérdőívezés. A nehézségek kezelését célzó számos innovatív módszertani kísérlet közül az adatfelvételi módok keverése népszerű megközelítés. Jelen tanulmány célja a hibrid adatfelvételekben rejlő lehetőségek feltérképezése. A szerző – különböző paraméterek szerint – két hibrid mintát hasonlít össze egy egysatosnás telefonos mintával. A minták felvételére a 2018-as magyarországi országgyűlési választások előtt került sor. Az eredmények azt mutatják, hogy a hibrid mérések több szempontból is kisebb hibát eredményeznek, mint az egysatosnás mérés. A párhuzamos, nemvalószínűségi online mintát tartalmazó hibrid mérések nagyon pontosan jelzik előre a választási eredményeket. Ugyanakkor más problémákra (például az alacsony válaszadási arányokra) az adatfelvételi módok keverése nem feltétlenül jelent megoldást. A hibrid módszertan potenciálisan javíthatja a kérdőíves felvételek minőségét és megbízhatóságát, azonban Magyarországon egyelőre alig van ismeretünk arról, hogy mely adatfelvételi konstrukciók lehetnek a leghatékonyabbak, így további kutatások szükségesek a témában.

TÁRGYSZÓ: módszertan, survey, választáskutatás

Due to declining willingness to respond worldwide and other problems, surveying, the main traditional method of empirical sociology, is becoming less capable of providing high quality and reliable data. Mixed-mode data collection is a popular approach among the innovative methodological techniques that aim to tackle these challenges. The goal of this study is to explore the potential of

* A címválasztást *Edith D. de Leeuw* „To mix or not to mix data collection modes in surveys” című tanulmányának címe ihlette (*de Leeuw* [2005]).

using mixed-mode surveys in Hungary. Two mixed-mode samples are compared by the author with a single-mode telephone sample along different parameters. The samples are from surveys that were conducted before the parliamentary elections in Hungary in 2018. The results show that the parallel mixed-mode surveys with a non-probability sample have led to less error compared to the single-mode survey, and predicted the election results very precisely. Mixing modes, however, may not provide solution for other problems (e.g., low response rates). This methodology may improve the data quality and reliability of surveys, nevertheless, very little scientific evidence is available in Hungary as to which designs are the most effective, thus further research is needed on this topic.

KEYWORD: survey, methodology, election research

Az elmúlt évtizedekben egyre több szakember jutott arra a véleményre, hogy az empirikus szociológia válságban van (*Savage-Burrows* [2007]). A szociológia domináns módszere, a hagyományos kérdőívezés marginalizálódásának prognosztizálása jöllehet túlzás lenne, ugyanakkor kétség kívül komoly kihívásokkal kell szembenézni a kérdőíves adatgyűjtés során. Egyrészt, a jó minőségű, megbízható survey adatok gyűjtésének *belső* feltételei, azaz mindazon paraméterek, amelyek egy ilyen kutatás lefolytatásához szükségesek, egyre kevésbé teljesülnek. Itt csak példaként említem az egyik legmarkánsabb problémát, az évek óta világszerte csökkenő válaszadási hajlandóságot (*Luiten-Hox-de Leeuw* [2020]), amely a magyar kutatói környezetet is érinti. A politikai témájú méréseknél a *belső* problémák látványos manifesztációi a választási előrejelzések kudarcai (például a 2016-os amerikai választások vagy a 2019-es magyarországi önkormányzati választások). A kérdőívezés *külső* problémája, hogy az empirikus szociológia nagyon lassan, illetve alig tudott reagálni az adatforradalom (*Kitchin* [2014]) nyomán létrejövő nagymennyiségű, új típusú adatok (Big Data) társadalomtudományi célú használatára. A digitalizáció során létrejövő óriási adattömeg (digitális lábnyomok) elemzése gyakran a hagyományos módszerekkel gyűjtött adatok alternatívájaként, „kihívójaként” értelmeződik. Tegyük hozzá, hogy a lassú reagálást azóta számos jó gyakorlat, kreatív felhasználás követte (*Salganik* [2019]).

A szakemberek számára a *belső* problémák természetesen ismertek, és a módszertankutatók régóta kísérleteznek azok kezelésével. Az egyik ilyen innovatív megoldás a hibrid adatfelvételek használata, azaz az adatfelvételi módok keverése. A hibrid módszer nemzetközileg a 2000-es évekre vált elterjedtté, *Biemer-Lyberg* [2003]-ban fogalmaztak úgy, hogy „a hibrid felmérések jelentik a normát napjainkban” (idézi *Pintér-Kótay* [2010] 724. old.). A hibrid minták népszerűsége abban keresendő, hogy az eljárások kombinálása által az egyik adatfelvételi mód gyengesége a másik mód erősségével kompenzálható, ami összességében csökkenthe-

ti a mérés teljes hibáját (total survey error, TSE) (*de Leeuw–Hox* [2011]). A személyes interjúk költségességét és lassúságát kompenzálhatja a gyors és olcsó online adatfelvétel bevonása, az online mérések magas fedési hibáját, azaz az online mintavételi keretek hiányosságait pedig a telefonos megkeresések; egy online kitöltési opció bevonhat olyan válaszmegtagadókat, akik ódzkodnak a személyes interjúktól. Tegyük hozzá, az adatfelvételi módok keverésének negatív következményei is lehetnek, például az adatminőségre (*Jäckle–Roberts–Lynn* [2006]).

A European Values Survey (Európai Értékrendkutatás – EVS) nemrégiben váltott át egy hibrid módszerre (*Luijckx et al.* [2021]), és a European Social Survey (Európai Társadalomstatistikai Felmérés – ESS) is régóta kísérletezik az átállással (*Messing–Ságvári* [2020]). Magyarországon a KSH (Központi Statisztikai Hivatal) a 2011-es népszámlálást és a 2016-os mikrocenzust is vegyes adatfelvételi eljárással készítette. Mindazonáltal hazánkban egyelőre inkább a bizalmatlanság a meghatározó kutatói attitűd a hibrid adatfelvételekkel szemben (*Lakatos* [2011]), kevésbé népszerűek politikai, társadalmi témájú kutatásoknál; e témáról tudományos publikáció alig akad (kivétel *Pintér–Kótay* [2010]).

Jelen tanulmányom empirikus adalékokkal szolgál a hibrid adatfelvételek alkalmazhatóságáról Magyarországon politikai témájú kérdések esetén. A 2018-as országgyűlési választások előtt készült két hibrid adatfelvétel eredményeit mutatom be, és egy egycsatornás, telefonos adatfelvétel eredményeivel hasonlítom össze. A felvételek több különböző paraméterét vizsgálom, így az elemzésem betekintést nyújthat abba, hogy a hibrid adatfelvétel milyen típusú hibákat/paramétereket képes hatékonyan csökkenteni, javítani, és melyekre nem nyújt megoldást. A nemzetközi tapasztalatok azt mutatják, hogy a vegyes módszertan gyakran jobb, megbízhatóbb adatokat eredményez. Az adott ország kutatói környezetének függvénye, hogy mely adatfelvételi variáció működik hatékonyan. Úgy vélem, fontos a jelen kutatáshoz hasonló empirikus adatok gyűjtése ahhoz, hogy Magyarországon is sikerrel alkalmazzuk ezeket az innovatív módszereket a politikai attitűdök mérésére.

1. A hibrid adatfelvételek módszertani háttere

A fejezet elsőként a hibriditás fogalmát tisztázza, majd bemutatja a hibrid adatfelvételek főbb típusait, keverésük lehetséges előnyeit és veszélyeit.

1.1. Mitől hibrid egy adatfelvétel?

A hibrid módszer lényege, hogy a vizsgálatban nemcsak egy, hanem többféle adatfelvételi eljárást alkalmaznak. Fontos, hogy a módozatok nem csak az adatfelvé-

tel során keveredhetnek (*de Leeuw* [2005] 237. old.). A kutatók és a kérdezőbiztosok a vizsgálat több pontján léphetnek kapcsolatba a mintába kerültekkel. A felkérések, az emlékeztetők, a kérdésfeltevések különbözőképpen valósulhatnak meg. A hibriditás tágan úgy értelmezhető, hogy a vizsgálat bármely szakaszában (beleértve a kontaktálási szakaszt is) többféle módot használnak a kutatók, míg a szűkebb értelmezés szerint a vegyes adatfelvételi módot alkalmazó kutatás hibrid.

1.2. A hibrid adatfelvételek főbb típusai

Előbbieken utaltam rá, hogy létezik egy tágabb értelmezése a hibrid mintáknak, ezeket *kvázi-hibrid adatfelvételeknek* hívjuk. Az adatfelvétel egy csatornán történik, ám a kutatásban több mód is részt vesz; például különböző módon rekrutáljuk a válaszadóinkat. Gyakori az online kitöltés felé terelés, amikor például postai felkérőlevelet kapnak a mintába kerültek egy online kérdőív kitöltésére (lásd például: *Messing–Ságvári* [2020]). A hibrid mérések két leggyakoribb fajtája mégis a *párhuzamos* és a *szekvenciális* adatfelvétel. A párhuzamos változatban egy időpontban egy kérdőívet töltenek ki a válaszadók, azonban egy részük A módon, másik részük B módon stb., azaz a minta hibrid. A szekvenciális típusnál is egy minta és egy kérdőív adott, a válaszadók egy része itt is A módon, egy másik része B módon stb. tölti ki a kérdőívet, ám van egy fő adatfelvételi mód, valamint kiegészítő mód(ok), amelyek tipikusan a nemválaszolók bevonását célozzák. A leggyakoribb, hogy fő adatfelvételi csatornának (például online kérdőív) az olcsóbb lehetőséget, míg a drágább eljárást a nemválaszolók bevonására választják (jó példa erre hazánkban *KSH* [2016]). Nemzetközi vagy regionális kutatásokban nagyon gyakori a párhuzamos adatfelvétel módosított változata, amikor a kérdőív továbbra is azonos, ám több minta is szerepel a vizsgálatban, és az egyes minták válaszadói különböző módokon válaszolnak. Amikor a kérdőív hibrid, továbbra is egy minta válaszol egy időpontban, azonban a kérdőív különböző szakaszait más-más módon töltik ki a válaszadók. A módszert hatékonyan használják szenzitív kérdések esetén. Végül longitudinális vagy panelkutatásoknál sokan kísérleteznek hibrid megoldásokkal (az adatfelvétel különböző szakaszai eltérő módon valósulnak meg).

1.3. Keverni vagy nem keverni? Előnyök és hátrányok

A Confront 2008-as globális jelentése alapján a piackutató cégeknél a hibrid adatfelvétel választását a válaszadási hajlandóság növelése motiválja (idézi *Pintér–Kótyay* [2010] 727. old.). A legjellemzőbb stratégia éppen a nemválaszolók bevonását célozza. Ráadásul arra is van bizonyíték, hogy a válaszadók jelentős részének

vannak preferenciái a válaszadás módjára vonatkozóan (*Groves–Kahn* [1979], *de Leeuw* [1992]). Egyes eredmények azt mutatják, hogy a válaszadási alternatívák felajánlása önmagában is javíthat a válaszadási hajlandóságon, a válaszok minőségén, sőt a kitöltés gyorsaságán is (*Olson–Smyth–Wood* [2012]). Igaz számos más kutatás nem mutatott ki ilyen hatást (*Dillman–West–Clark* [1994], *Manfreda–Vehovar–Batagelj* [2001]).

Az adatfelvételi módok keverése akár a *fedési hiba* csökkenésével is járhat. A hagyományosan nagyon jó lefedettséget biztosító személyes megkérdezés gyakorlati alkalmazhatósága nagyon sokat romlott az elmúlt években, azonban a nagyfokú reprezentativitás változatlan elvárás maradt. Emellett nagy az igény az online mód használatára, ami azonban a legtöbb társadalomban továbbra is jelentős fedési hibát von maga után, hiszen az internetpenetráció nem 100 százalékos. Az online stratégiák lényege az, hogy a mintavétel nem online történik (hanem például lakcím alapján, postai felkérőkkel), és minél több mintatagot próbálunk ösztönözni a webes kitöltésre (*Messing–Ságvári* [2020]). A különböző módok erősségeinek kombinálása összességében csökkentheti a mérés teljes hibáját.

A hibrid minták használata bizonyos esetekben a mérési hibák minimalizálását célozza. Szenzitív (például politikai) témáknál a kérdőíves kutatások egyik régi problémája az ún. társadalmi elvárás szerinti válaszadás (*Tourangeau–Rips–Rasinski* [2000]). Ennek lényege, hogy a válaszadók hajlamosak olyan válaszokat adni, melyek a társadalmi elvárásoknak megfelelőek, és eltitkolni azokat, melyek szembe mennek az aktuális normákkal. Viszonylag konzisztens a szakirodalom a tekintetben, hogy szenzitív témáknál az önkitöltős forma (papíralapú vagy online) őszintébb és pontosabb válaszokat eredményez, mint a kérdezőbiztossal történő válaszadás (*Holbrook–Green–Krosnick* [2003], *Kuran* [1997], *Tourangeau–Smith* [1996], *Tourangeau–Yan* [2007]).¹ Egyes témáknál célravezetőbb lehet egy önkitöltős mód vagy egy olyan hibrid változat, ahol a kérdések egy részére kérdezőbiztossal, egy másik részére önkitöltős formában válaszol a megkérdezett.

A hibrid megoldás választása gyakran a költségek csökkentését vagy éppen az adatfelvételi idő rövidülését is eredményezheti. Az online mód bevonása mögött többségében éppen ezek a fő motivációk. Az, hogy mekkora költségcsökkentési potenciál van az adatfelvételi módok keverésében, főként attól függ, hogy mekkora az eltérés az adatfelvételi módok között. *Villar–Fitzgerald* [2017] gyűjtése az ESS kapcsán jól mutatja, hogy például a személyes adatfelvétel ára 5-6 szorosa is lehet egyes országokban az online adatfelvételének, míg más országokban (Magyarországon is) ez a szorzó „csak” kétszeres. (Lásd a Függelék F1. táblázatát.) Az online adatfelvételek jelentősen gyorsabbak a hagyományos

¹ Magyarországon többek között *Csányi–Dés* [2000] hangsúlyozzák a kérdezőbiztosok képzésének jelentőségét szenzitív témák esetén.

eljárásoknál. Ám olyan kutatásról is tudunk, ahol úgy sikerült érdemben rövidíteni az adatfelvétel idejét, hogy a válaszadók az általuk preferált módon válaszoltak (*Olson–Smyth–Wood* [2012]).

A több adatfelvételi mód szimultán használatának ugyanakkor ára is van. Az ún. módhatás lényege, hogy a mintában megfigyelt különbségek az eltérő adatfelvételi eljárásoknak tulajdoníthatók, s nem a célpopulációban létező valós különbségeket mutatják (*Roberts* [2007]). Bár itt a módhatások negatív következményeiről értekezem, fontos megjegyezni, hogy az adatfelvételi módból fakadó különbség nem feltétlenül hátrány (*De Leeuw–Hox* [2011]). Éppen azért ötvözzük a módokat, mert különböznek. A fedési hiba minimalizálásakor azt használjuk ki, hogy egyes módokon más-más társadalmi csoportok érhetők el hatékonyan. Azonban, ha a módhatások szignifikáns torzításokat okoznak a mérési hibával összefüggésben, az összességében korlátozhatja lehet a hibrid adatfelvételek alkalmazhatóságának.

A módhatásvizsgálatok kevés konzisztens eredménye közül fontos kiemelni a kérdezőbiztos szerepét. Több olyan kutatásról is tudunk, amely alapján a kérdezőbiztos személye meghatározó lehet politikai témák esetében. *Németh–Luksander* [2018] vizsgálata alapján a kérdezőbiztos pártpreferenciája számottevően befolyásolja a válaszadó pártpreferenciáját, valamint *Pillók* [2010] is talált összefüggést egyes kérdéseknél a kérdezőbiztos kora és a válaszok között. Ehhez hasonló kérdezőbiztos-hatásokra korábbi politikai témájú kutatások is rámutattak (*Katz* [1942], *Pickery–Loosveldt* [2002], *Lipps–Lutz* [2010]). A legtöbb korábbi elemzés elhanyagolható módhatásokról számolt be a postai és az online adatfelvételek között (önkitöltős módok), és nagyobb különbségekről, amikor önkítöltős formákat hasonlítottak össze személyes vagy telefonos felvételekkel. Tehát az önkítöltős és kérdezőbiztossal operáló módok kombinálása veszélyt jelenthet az adatminőségre (*de Leeuw–Hox* [2011]), különösen szenzitív témák esetén.

1.4. A magyarországi helyzet

Magyarországon sokáig csak személyesen vettek fel kérdőíveket. Az 1970-es, 1980-as évek adatfelvételeit nagyon magas válaszadási arányok, a kutatásokra kifejezetten nyitott közvélemény jellemezte. Az ezredfordulóra azonban ez olyannyira megváltozott, hogy a klasszikus címlistas megkérdezés mára Magyarországon is ellehetetlenült (*Závecz* [2010]). A válaszadási hajlandóság csökkenése (*Angelusz–Tardos* [2006]), a kérdezőbiztos személye körüli bizonytalanságok (bizalmatlanság) miatt ma már közel sem olyan egyértelmű a személyes felvételek kiemelt státusa. A 1990-es években elterjedő telefonos mintákat erős módszertani szkepszis övezte (*Rudas* [1995]), ennek ellenőre rövid időn belül komoly kihívói lettek a személyes mintáknak (*Kmetty* [2011]). A telefonos interjúk gyorsabbak, olcsóbbak voltak,

és jobban ellenőrizhetővé váltak a kérdezőbiztosok. Ugyanakkor bizonyos módszertani problémák (fedési hiba, alacsony válaszadási arány) még jobban kidomborodtak (Kmetty [2011]). A 2000-es évekre az online kérdőívek is elterjedtté váltak, ma már egyre több akadémiai kutatás is ezzel a módszerrel készül. A legnépszerűbb forma az ún. online access panel, amely nemvalószínűségi mintavételre ad módot (Pillók–Darvas [2001]). Az online minták használata rendkívül vonzó, hiszen a telefonos kutatásoknál is olcsóbbak, egyszerűbbek és gyorsabbak. Azt azonban le kell szögezni, hogy ma Magyarországon nem lehet online adatfelvételt végrehajtani országosan a felnőtt korú magyar lakosságra, még az alap demográfiai jellemzők szerint sem. Ennek fő oka, hogy azok, akik potenciálisan bekerülhetnének egy ilyen panelbe, azaz a rendszeres internetezők aránya Magyarországon 75 százalék².

Angelusz–Tardos már [2009]-es tanulmányukban rávilágítottak arra, hogy a hibrid megoldásokban nagy potenciál lehet. Ennek ellenére, míg szakmai beszélgetések alapján hibrid adatfelvételeket piaci vállalatok rendszeresen használnak, egyéb, például tudományos célra készülő mérésekről alig tudunk. A politikai preferenciákat rendszeresen vizsgáló hazai közvélemény-kutató cégek mind egysatornás, személyes, telefonos vagy online módszereket használnak. Az NRC piackutató cég 2010-es választásokat megelőző online kutatása, amelyet telefonos interjúkkal egészített ki, „meglepően pontos becslést” adott a választás kimenetelére (Kmetty [2011] 60. old.). Nemrégiben az online mód alkalmazhatóságát tesztelte az ESS magyarországi csapata (Messing–Ságvári [2020]). Postai levélben kérték fel a mintatagokat egy online kérdőív kérdéseinek megválaszolására. Azoknak, akik az online kérdőívet nem töltötték ki, egy papíralapú kérdőívet is kipostáztak. Kifejezetten magas, a személyes minták válaszadási arányát megközelítő, 40 százalékos kitöltési hajlandóságot találtak a kutatók. Ebből a válaszadók 30 százaléka online, 10 százaléka papíralapon töltötte ki a kérdőívet.

2. Adatok és módszerek

2.1. Adatok

A 2018. évi áprilisi magyarországi országgyűlési választások előtti egy hónapban két hibrid adatfelvételre (a továbbiakban H1 és H2 jelölést alkalmazom), valamint egy benchmark, egysatornás, számítógéppel támogatott telefonos interjúra (computer-assisted telephone interviewing, CATI) került sor. Az adatokat a Századvég Alapítvány gyűjtötte.

² Forrás: KSH idősoros adatok. https://www.ksh.hu/docs/hun/xstadat/xstadat_eves/i_oni017.html

Vizsgálatunk során a H1-ben egy szekvenciális hibrid felvételt alkalmaztunk, amelyben a fő adatfelvételi mód (első szakasz) a számítógéppel támogatott személyes interjúzás (computer-assisted personal interviewing, CAPI), míg a második szakaszban a kiegészítő adatfelvételi módok (online és telefonos interjúk) célja a visszautasítók bevonása volt. A stratégia arra épült, hogy amennyiben az interjúalany az elsődleges címen visszautasította a kérdezőbiztost, lehetőséget kapott arra, hogy a kérdőívet egy online felületen töltsse ki, vagy egy alkalmas időpontban telefonon adhatott választ a kérdésekre. Amennyiben az interjúalany nyitottnak bizonyult az utóbbi két opció valamelyikére, úgy a kérdezőbiztos felírta az e-mail címét, illetve telefonszámát. Ezek az alanyok kerültek be a rekrutáció második szakaszába; az e-mail címüket megadók egy felkérő levelet kaptak, és a kérdőív online verzióját tölthették ki, a telefonos interjúra nyitott alanyokat pedig a kérdezőbiztosok a rákövetkező napokban több időpontban is felhívták telefonon az azonos tartalmú kérdőív megválaszolására. A két szakasz egy idő után lényegében szimultán futott. Ez fontos következményekkel járt a mintatervezési mechanizmusokra. Az adatfelvétel alatt folyamatosan szükséges volt nyomon követni a másodlagos módokon beérkező kérdőíveket, és ennek megfelelően tervezni az elsődleges mód adatfelvételét. Töblépcsős, arányosan rétegzett, valószínűségi mintavételi eljárást alkalmaztunk. Elsőként a települések kerültek a mintába, majd véletlen séta módszerével választottuk ki a háztartásokat. A háztartáson belül a válaszadó kiválasztását a születésnapos módszer segítette. A mintavételi keretet az érvényes lakcímmel rendelkező állandó magyarországi lakosság képezte. A végső mintaelemszám 800 fő volt.

A H2 párhuzamos hibrid adatfelvétel során már a mintavételezést megelőzően meghatároztuk, hogy mely csoportokat, milyen módon kívánjuk elérni. Szelekciós változóként az életkort határoztuk meg: az 55 éven aluliakat online, az 55 éven felülieket telefonon kérdeztük. Ez lényegében két teljesen független mintavételt jelentett.

Az online mintát a Daedalus Online³ Access Paneljéből vettük fel. A paneltagok rekrutálása nemvalószínűségi eljárásokkal (például hirdetésekben keresztül) történt, csatlakozásuk önkéntes volt. Az adatfelvétel idején 160 ezer tagot számlált a panel. Kvótás mintavételt alkalmaztunk, mely során a 2016-os mikrocenzus adatai alapján nem, életkor és településtípus alapján határoztuk meg a kvótákat. A végső online mintaelemszám 498 fő volt.

Az 55 éves vagy annál idősebb lakosság vizsgálatára telefonos mintát vettünk fel, ennek során egy töblépcsős, arányosan rétegzett, valószínűségi mintavételi eljárást alkalmaztunk. Az adatbázisban egyaránt szerepeltek vezetékes- és mobiltelefonszámok. Az adatfelvétel CATI-módszerrel történt. A végső telefonos mintaelemszám 307 fő volt.

³ <http://hu.daedalusonline.eu/>

A két adatbázis egyszerű egyesítéseként jött létre az a 805 fős adatbázis, amely – bizonyos megszorításokkal ugyan, de – reprezentatívnak tekinthető a teljes magyar lakosságra nem, kor, iskolai végzettség és településtípus szerint.

Végül a benchmarkként szolgáló egycsatornás telefonos minta felvételéhez a H2 telefonos mintánál alkalmazott mintavételi stratégiát hajtottuk végre. Ezen minta végső elemszáma 500 fő volt.

A mintavételből fakadó hibákat mindegyik mérésnél iteratív súlyozás segítségével korrigáltuk.

2.2. Módszerek

A kutatás kérdőíve a válaszadók politikai preferenciáit, attitűdjeit mérte. A kérdések döntő része a korábbi választáskutatások bevett kérdéseit, valamint a Magyar Tudományos Akadémia Társadalomtudományi Kutatóközpont Politikatudományi Intézetében 2018-ban végzett választáskutatás kérdéseit tartalmazta (Böcskei–Szabó [2018]). A programozás során arra törekedtünk, hogy a lehető legkevésbé térjenek el a kérdések a különböző módokon. Ez különösen az online mód esetén jelentett kihívást. Az egységesítés során Dillman unimode konstrukcióját követtük (Dillman–Smyth–Christian [2014]), így például az online kérdőíven is minden alkalommal lehetősége volt a megkérdezetteknek a „nem tudja” választ jelölni, illetve a választ megtagadni, valamint kerültük a nagyobb táblázatok használatát.

Az adatgyűjtés eredményeinek elemzése során a háromféle mérést a következő paraméterek alapján hasonlítottam össze: fedési hiba, válaszadási arány, adatminőség (adathiány, pontosság), költségek és az adatfelvétel ideje. A legtöbb összehasonlítás-hoz χ^2 -, illetve t -próbát használtam.

3. Eredmények

3.1. Válaszási arányok

A válaszadási arányok összehasonlításának fő tanulsága, hogy a válaszadási hajlandóság mindegyik kutatásban meglehetősen alacsony volt. (Lásd az 1. táblázatot.) A legmagasabb válaszadási arányt a szekvenciális hibrid mintában (H1) érték el (32,10%). A H1-felvétel kifejezetten a válaszadási arány javítását tűzte ki célul, azonban az extra törekvések (addicionális módok alkalmazása) a válaszadási arány

növelésére mindössze 2 százalékos emelkedést eredményeztek. A második szakasz során bevont interjúalanyok mindegyike online válaszolt, telefonon nem érkeztek be válaszok. A kérdezőbiztosok beszámolója alapján a visszautasításkor egy új mód felajánlása nagyon kis eséllyel törte volna át a válaszadó által felhúzott „falat”.

1. táblázat

Válaszadók száma és a válaszadási arány a három adatfelvételben
(Number of respondents and response rates in the three surveys)

Adatfelvételi jellemzők	H1	H2-telefonos	H2-online	Egycsatornás telefonos
	adatfelvétel			
	Válaszadók száma (fő)			
Teljes interjú	798	428	–	500
Részleges interjú	2	0	–	1
Visszautasító háztartás	532	243	–	378
Nemkontaktált háztartás	518	210	–	353
Ismeretlen háztartás	610	421	–	765
Ismeretlen egyéb	245	267	–	476
	Válaszadók aránya (%)			
Válaszadási arány első szakasz	30,23	–	–	–
Válaszadási arány végleges	32,10	27,28	–	20,19
Kooperációs arány	58,50	60,71	–	52,60
Kitöltési arány	–	–	51,00	–

Megjegyzés. A válaszadási, kooperációs és kitöltési arányok kiszámításához az AAPOR (American Association of Public Opinion Research – Amerikai Közvélemény-kutató Társaság) aktuális irányelveit követtem. A számítási módokkal kapcsolatban részletesen lásd AAPOR [2016].

A legalacsonyabb válaszadási arányt az egycsatornás telefonos mérés produkálta (20,19%). A párhuzamos hibrid adatfelvételnél a telefonos mintaszeletnél ennél magasabb válaszadási arányt találtunk (27,28%), ami arra utal, hogy az idősebb korosztályoknál magasabb a válaszadási hajlandóság telefonon, mint a fiatalok körében. Nemcsak a válaszadási, hanem a kooperációs arány⁴ is magasabb volt a H2 telefonos almintában (60,71% vs. 52,60%), ez alapján tehát könnyebben elérhetők az idősebb korcsoportok telefonon, és nagyobb valószínűséggel válaszolnak, mint a fiatalabbak. A H2 telefonos kooperációs arány nagyjából azonos a H1-ével. A H2 adatfelvétel

⁴ A kooperációs arány a ténylegesen kontaktált esetek és a megvalósult interjúk hányadosa.

online része 51 százalékos kitöltési arányt eredményezett, ami átlagosnak mondható ilyen panelek esetén.⁵

3.2. Mintaösszetétel – fedési hiba

Jelen kutatásban a fedési hibára a mintaösszetételek közötti különbségek elemzése és egyes általános statisztikák (például internethasználat) alapján tudok következtetni. Az F2. táblázat összesíti a négy legfontosabb szociodemográfiai változó súlyozatlan eloszlása és a mikrocenzus eredményei (KSH [2016]) közötti különbségeket a három adatfelvétel esetén. Fontos leszögezni, hogy a párhuzamos hibrid minta online részénél a rekrutáció során az adatfelvételt végző cég kvótákat alkalmazott a nemre, az életkorra és a településtípusra vonatkozóan. Így a minta ezen szeletének arányai pontosan tükrözik a populációs arányokat, ám nem a véletlennek köszönhetően, hanem a kvótázási eljárás miatt. Ez a körülmény tovább nehezíti az összehasonlítást és annak mérését, hogy vajon a nemvalószínűségi online mintára milyen mértékű fedési hiba jellemző.

Összességében a személyes felmérés (H1) arányai közelítették a legpontosabban a mikrocenzus eredményeit mind a négy változónál. Mindegyik minta igen pontosan leképezte a nemek és a településtípus szerinti megoszlásokat, míg az életkori, és különösképpen az iskolai végzettségbeli arányok helyenként jelentősen eltértek a KSH adataitól. Az életkort tekintve az egycsatornás telefonos minta eredményezte a 60 éven felüliek számottevő felülreprezentáltságát (+17,1%), míg a másik két felvétel jól közelítette a mikrocenzus eredményeit e tekintetben. Az iskolai végzettség szerinti reprezentativitás volt az egyetlen, amelyet a személyes adatfelvétel sem tudott maradéktalanul biztosítani. A H1-nél a legalacsonyabb és a legmagasabb végzettségűek alulreprezentáltsága jelent problémát, míg a H2 és az egycsatornás telefonos kutatásban a maximum 8 általános iskolai végzettséggel rendelkezők alulreprezentáltsága kritikus (–19, illetve –15,6 %). Mindezen jelentős fedési hibák egy része jól magyarázható a mintavételi keret hiányosságaival, különösen az online minta kapcsán. Ismert probléma a valószínűségi és a nemvalószínűségi online paneleknél is, hogy az alacsony végzettségűek jóval nehezebben rekrutálhatók, és így eleve kisebb aránnyal szerepelnek a panelekben.

⁵ A klasszikus értelemben vett, valószínűségi mintákra alkalmazott válaszadási arány kalkulációja nem értelmezhető egy online access panel esetén (Loosveldt–Sonck [2008], Pennay et al. [2018]). Az ilyen típusú online paneleknél a nemválaszolás legalább két lépésben jelenhet meg: a panelben való részvételre és az aktuális kérdőív kitöltésére irányuló hajlandóságnál (Bosnjak et al. [2013]). Becslések szerint jóval kevesebb mint 1 százalék csatlakozik online panelekbe azok közül, akik meghívást kapnak erre (Tourangeau–Conrad–Couper [2013]). Az online access panelek esetén a kitöltési arány a válaszadási hajlandóság leggyakrabban használt indikátora, amely azon paneltagok arányára utal, akik kitöltötték a kérdőívet a kutatásba meghívottak közül (Callegaro–DiSogra [2008]). Szemben a valószínűségi mintákkal, a nemvalószínűségi mintáknál nincs adekvát módja a nemválaszadási hiba becslésének.

3.3. Mérési hiba

3.3.1. Adathiány, részleges nemválaszolás

Mindhárom mérés meglehetősen magas részleges nemválaszolási arányokat eredményezett. (Lásd az F3. táblázatot.) A teljes kérdőívet tekintve a legrosszabb minőséget e tekintetben a két hibrid adatfelvétel produkálta. Amennyiben az átlagokat vetjük össze, a H1 rosszabbul teljesített (átlag = 16,28%, medián = 7,13%, szórás = 19,64%), a medián összehasonlításakor azonban a H2 mutatója bizonyult gyengébbnek (átlag = 14,5%, medián = 12,31%, szórás = 14,29%).⁶ Ehhez képest az egycsatornás telefonos minta mindkét statisztika szerint jobb eredményeket hozott (átlag = 10,58%, medián = 6,15%, szórás = 12,69%). A H2 adatfelvételben az online rész teljesített gyengébben, ebben az összehasonlításban itt volt a legmagasabb a nemválaszolási arány. A telefonos mintaszelet arányai nagyon hasonlóak az egycsatornás telefonos minta arányaihoz.

A részleges nemválaszolási arányok jelentős különbségeket mutatnak a demográfiai és a tartalmi kérdések összehasonlításakor. (Lásd az F4. és F5. táblázatot.) Logikusan, minden mérésnél alacsonyabban alakultak a nemválaszolási arányok a demográfiai változóknál, a medián 0 és 2 százalék között mozog. A magas átlagok oka az, hogy szerepelt egy kérdés a kérdőívek végén a háztartás és a válaszadó havi nettó jövedelmére vonatkozóan, ahol nagyon magas volt a válaszmegtagadás (40–60%-os).⁷ Ha ezektől a kérdésektől eltekintünk, a válaszmegtagadás a demográfiai változóknál minden mérésnél alacsony volt. Nem úgy a tartalmi kérdéseknél. Ebben az összehasonlításban egyértelműen a H1 adatfelvétel eredményezte a legrosszabb minőséget (átlag = 20,51%, medián = 19,48%, szórás = 14,93%) és az egycsatornás telefonos a legjobbat (átlag = 12,27%, medián = 11,38%, szórás = 11,53%). A H2 mérés relatíve jobb eredményei az egycsatornás telefonos mintaszeletnek voltak köszönhetőek, az online rész nemválaszolási arányai igen gyengén alakultak (átlag = 21,29%, medián = 22,66%, szórás = 13,47%).

A 2. táblázat rávilágít az adatfelvételek különbözőségére a tekintetben, hogy milyen a válaszadási hajlandóság a szenzitív kérdéseknél (pártpreferencia és jövedelmi helyzet). A pártpreferenciára vonatkozó kérdésnél hasonló volt a válaszadási hajlandóság a három mérésnél (nemválaszolási arányok: 31–41 százalék, $p > 0,05$ minden párosításnál). A H2 mérésen belül az online minta tagjai 6 százalékkal nagyobb valószínűséggel válaszoltak, mint a telefonon válaszolók. A jövedelmet tudakoló kérdésre adott válaszok ugyanakkor már nagyobb differenciákat mutattak ($p < 0,001$ minden párosításnál). A H1 adatfelvétel messze a legmagasabb visszautasítási arányt hozta (65,38%), és az egycsatornás telefonos mérés a legalacsonyabbat (40,30%). A H2 online része szintén relatíve alacsony visszautasítást produkált

⁶ Az eloszlások erősen ferdek, kiugró értékekkel, így szerencsésebb a medián használata.

⁷ A jövedelmi kérdéseknél egyébként a H1 mérés volt a legkevésbé hatékony.

(44,44%). Összességében tehát szenzitív kérdések esetén a telefonos minta eredményezte a legalacsonyabb nemválaszolást, igaz a kép nem egyértelmű, hiszen a H2 mérésnél az online felvételen sokkal magasabb volt a válaszadási hajlandóság, mint telefonon, mindkét kérdésnél.

2. táblázat

Részleges nemválaszolási arányok a három adatfelvételen a pártpreferenciára és a háztartás havi nettó jövedelmére vonatkozó (szenzitív) kérdéseknél (%)
(Partial non-response rates in the three surveys for [sensitive] questions on party preference and monthly net household income [%])

Nemválaszolás	H1	H2	H2- telefonos	H2- online	Egycsa- tornás telefonos	H1 vs. H2	H1 vs. egycsatornás telefonos	H2 vs. egycsatornás telefonos
	adatfelvétel							
	Pártpreferencia							
Részleges nemválaszolás összesen	31,82	35,24	40,60	34,70	33,00	+3,42	+1,18	-2,24
„Nem válaszol”	14,78	20,09	23,20	17,30	15,30	+5,31***	+0,52	-4,79***
„Nem tudja”	17,04	15,15	17,36	17,40	17,70	-1,89	+0,66	+2,55
	Háztartás havi nettó jövedelme							
Részleges nemválaszolás összesen	65,38	49,45	56,76	44,44	40,30	-15,93***	-25,08***	-9,15***
„Nem válaszol”	53,88	33,70	26,13	38,89	33,83	-20,18***	-20,05***	+0,13
„Nem tudja”	11,50	15,75	30,63	5,56	6,47	+4,25**	-5,03***	-9,28***

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

3.3.2. Pontosság

A pontosság mérésére a három felmérés eredményeit a 2018. áprilisi országgyűlési választások eredményeivel vettem össze.

Az első célváltozó, amelyet becsülni próbáltam, a választáson való részvétel. Egyrészt általános megfigyelés, hogy a közvélemény-kutatások felülmérk a részvételi arányokat, mind a valószínűségi (Kmetty [2011], Angelusz–Tardos [2009]), mind a nemvalószínűségi online mintáknál (Kennedy *et al.* [2016]). Arra számítottam, hogy mindegyik mérés felülbecsülte a részvételt, igaz, feltételezhetünk különbségeket az adatfelvételi módok között. Korábbi kutatások alapján (Duffy *et al.* [2005]) politikailag aktívabbak a rendszeresen internetezők, ami magasabb részvételt jelez a H2 esetében. Egy alternatív magyarázat alapján azonban éppen azok a módok mérik felül a részvételt, amelyek kérdezőbiztost alkalmaznak, mivel a kérdezőbiztos jelenlétében magasabb társadalmi elvárás szerinti választ várunk. Az emberek hajlamosak

szavazási hajlamukat felülbecsülni, mivel a szavazás aktusához mint állampolgári jogaink gyakorlásához, a demokrácia szimbólumához alapvetően pozitív dolgok társulnak; a szavazás egy pozitív aktus (*Voogt–Saris* [2003], *Liao* [2016]).

Ha azok arányát tekintem, akik valamilyen valószínűséget adtak annak, hogy részt vesznek az április választáson⁸, akkor mindhárom mérés alapján jócskán felülbecsültem volna a részvételt. (Lásd az F6. táblázatot.) Jobb a helyzet, ha csak azok arányát veszem, akik biztosra ígérték részvételüket.⁹ Ez esetben a H2 felvétel alapján nagyon pontos, hibahatáron belüli becslést adhattam volna a részvételre (+3,35%, $p > 0,05$), míg a H1 alapján alul- (-19,12%, $p < 0,001$), az egycsatornás telefonos minta alapján pedig felülbecsült lenne a részvétel (+12,74, $p < 0,001$). Annak érdekében, hogy pontosan lássuk az adatfelvételi módok közötti különbségeket, leválogattam a 18–54 évesek válaszait erre a kérdésre; körükben is az egycsatornás telefonos mérés alapján jóval magasabb részvételt becsülhettem volna, mint a többi mérésben.

A pontosság egy másik indikátora, hogy mennyire voltak képesek a mérések pontosan becsülni a választás kimenetelét. A pártpreferenciák mérésére egy fő kérdést¹⁰ és egy másodlagos kérdést¹¹ alkalmaztam. Választási előrejelzéseknél a fő problémát a bizonytalanok és a válaszmegtagadók magas aránya jelentheti. Az már a részleges nemválaszolás tárgyalásakor kiderült, hogy a három adatfelvételnél nagyon hasonló (magas) volt az adathiány a pártpreferenciára vonatkozó kérdésnél (31–35%). A közvélemény-kutatók jellemzően nem a „tisza” pártpreferencia-eloszlásokat használják, hanem ún. szakértői becsléseket készítenek (*Kmetty* [2011] 42. old.). Az egyik stratégia első lépése az ún. biztos szavazó pártválasztók preferenciáinak becslése. Ennek lényege, hogy csak azon megkérdezettek válaszait vesszük figyelembe, akik biztosra ígérték a részvételt, és ki is nyilvánították preferenciáikat; hátránya, hogy tapasztalatok szerint a bizonytalanok vagy válaszmegtagadók egy része később részt vesz a választáson, és érvényes szavazatot ad le, továbbá ilyen nagymértékű hiányzó érték esetén a megbízhatósági szint jelentős mértékben torzít (*Rudas* [1998]). Jelen tanulmányban egy másik módszer, az imputálás mellett döntöttem. Az imputálás lényege, hogy a kutató helyett, hogy figyelmen kívül hagyná, megpróbálja megbecsülni a hiányzó adatokat. Számos imputálási eljárás közül választhatunk, attól függően, hogy milyen jellegű az adathiány és a változóink. Mindezek részletezése jelen tanulmány kereteit meghaladná.¹² Elemzésemben

⁸ Tehát arra a kérdésre, hogy „2018 áprilisában lesznek az országgyűlési választások. Mit gondol, Ön...” a „...biztosan elmegy szavazni” vagy „...valószínűleg elmegy szavazni” választ adták.

⁹ Magyarországon ez a bevett gyakorlat (*Marián* [2006]). Marián ugyanakkor tanulmányában részletesen tárgyalja, hogy mikor és milyen mértékben hibás stratégia pusztán a részvételüket biztosra ígérők preferenciáira alapozni becslésünket.

¹⁰ „(És ha mégis elmenne) melyik pártra szavaz(na) az országgyűlési választáson?”

¹¹ „Van-e olyan párt ma Magyarországon, amely közelebb áll Önhöz, mint a többi? Melyik ez a párt?”

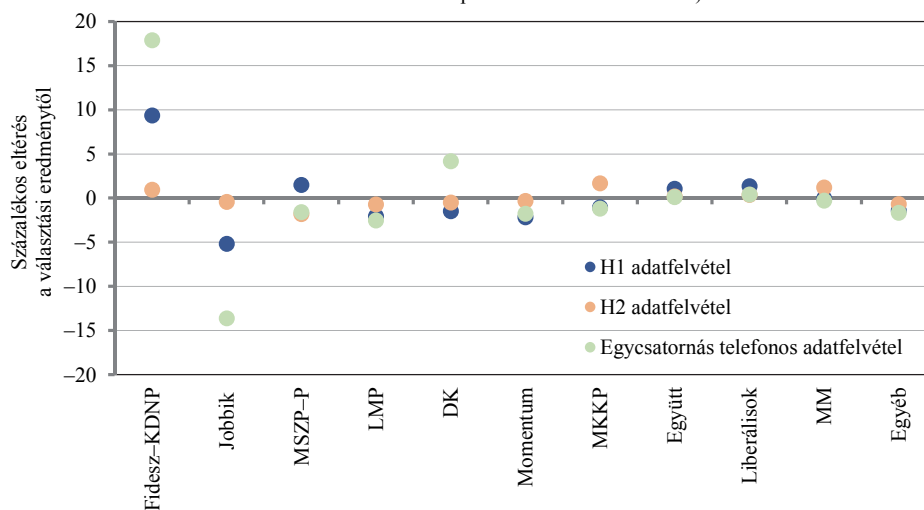
¹² Az imputálási technikákról bővebben lásd például: *Hunyadi–Vita* [2002], *Oravecz* [2008], *Máder* [2005].

hot deck imputálási módszert¹³ alkalmaztam, amelyhez az R program VIM csomagját használtam (Kowarik–Templ [2016]).

Az ábra az imputálás utáni eltéréseket mutatja a három mérés eredményei és a listás választási eredmények között pártonként, a pontos különbségek pedig az F7. táblázatban találhatók. Összességében a legpontosabb a H2 mérés volt, ennél nem akadt olyan párt, amely esetén a becslés és a választási eredmény közötti különbség meghaladta volna a 2 százalékot. Ezt követte a H1 mérés, majd az egysatosornás telefonos mérés egyaránt két szignifikáns eltéréssel. Ha pártonként vizsgáljuk, csak a H2 tudta pontosan előre jelezni a Fidesz–KDNP¹⁴ választási eredményét; mind a H1, mind az egysatosornás telefonos minta alapján jelentősen felülbecsült a kormánypártok szavazati aránya. Ezzel ellentétesen a Jobbik Magyarországért Mozgalom pártját ez a két vizsgálat erősen alulmérté, míg a H2 pontosan becslte. A kisebb pártokat jellemzően inkább alulmérték a kutatások, igaz valamivel kisebbek voltak a tévedések.

Az egyes pártok támogatottsága a három adatfelvételben imputálás után, valamint a választási eredmények különbségei (%)

(Results on the support of parties in the three surveys after imputation, and their differences compared to the election results)



Megjegyzés. KDNP (Kereszténydemokrata Néppárt), MSZP-P (Magyar Szocialista Párt-Párbeszéd), LMP (Lehet Más a Politika), DK (Demokratikus Koalíció), MKKP (Magyar Kétfarkú Kutya Párt), MM (Mindeni Magyarországa).

¹³ A módszer lényege, hogy az eljárás során egy donort, más szóval egy olyan válaszadót keresünk, aki más kérdésekre adott válaszai alapján nagymértékben hasonlít arra a válaszadóra, akinél adathiány van. A donor választását imputáljuk az adathiány helyére. A donor kiválasztását az euklideszi távolságon alapuló legközelebbi szomszéd meghatározásával végzi az algoritmus.

¹⁴ KDNP (Kereszténydemokrata Néppárt).

3.4. A költségek összehasonlítása

A három kutatás költségeinek összehasonlítását a 3. táblázat tartalmazza. A H1 felvétel összességében jóval nagyobb költségű volt, mint a H2 vagy az egycsatornás telefonos mérés. A H1 összköltsége több mint másfélszeresét (1,6-szorosát) tette ki a H2 hibrid mérésnek, és közel 25 százalékkal haladta meg az egycsatornás telefonos mérését. Ezek a különbségek minden költségelemre igazak: a H1-nél a tervezési/szervezési, programozási, valamint az adatfelvételi költségek is magasabbak voltak. A tervezés és a programozás (mivel több felületre kellett elkészíteni a kérdőívet) mindkét hibrid mintánál több költséggel járt, mint az egycsatornás telefonos kutatásnál. Azonban az összköltséget tekintve ezek a költségelemek csak mintegy 2 százaléknyi különbséget eredményeztek a mérések között. Nagyobb a súlya a terepmunkának. Amennyiben kérdőívdarabárát számítunk az egységárból (tehát a tiszta adatfelvételi költséget vizsgáljuk), a személyes adatfelvételi mód több mint kétszer (2,6-szor) drágább az online, és 25 százalékkal drágább az egycsatornás telefonos módnál. Ennek – azaz az online mód szerepeltetésének – köszönhető, hogy a párhuzamos hibrid kutatás összköltsége 12,5 százalékkal alacsonyabban alakult a standard CATI-kutatásénál. A szekvenciális adatfelvételnél extra adatfelvételi költséget jelentett az is, hogy a visszautasítók bevonására tett törekvések miatt magasabb díjazásban részesültek a kérdezőbiztosok, mint egy normál személyes adatfelvételnél. Azt is fontos leszögezni, hogy a tervezés és a programozás fix költségek, a terepmunka azonban elemszám függő. Azaz az előbbieken vázolt árdifferenciák az elemszám emelésével párhuzamosan nőnek. Minél nagyobb mintát szeretnénk, annál jobban megéri az olcsóbb online mód használata – legalábbis anyagi szempontból.

3. táblázat

Relatív költségek 800 fős mintákra szimulálva a három adatfelvételnél
(Relative costs simulated for 800-person samples in the three surveys)

Költségelem	H1 (N = 800)	H2 (N = 805)	H2-telefonos (N = 307)	H2-online (N = 498)	Egycsatornás telefonos (N = 800)
	adatfelvétel				
Tervezés, kérdőívírás	4,0	3,0	1,0	1,0	3,0
Programozás	5,0	5,0	2,5	2,5	3,0
Terepmunka	91,0	54,5	39,0	16,5	69,0
<i>Összesen</i>	<i>100,0</i>	<i>62,5</i>	<i>42,5</i>	<i>20,0</i>	<i>75,0</i>
Kérdőívdarabár	0,11	0,07	0,09	0,04	0,08

Megjegyzés. A telefonos minta 500 fős volt, a táblázatban szereplő adatok egy elméleti kalkuláció eredményei 800 főre.

3.5. A kutatás időigénye

Egyik potenciális előnye lehet a hibrid mintáknak a gyorsabb adatfelvétel, ugyanakkor nem mindegyiktől várható ez az előny. A három mérés közül a H1 készült el a leglassabban (25 nap alatt), ami nem meglepő, hiszen köztudottan ez a legidőigényesebb adatfelvételi mód. A H1 esetében magából a módból is adódott, hogy hosszú volt az adatfelvétel, hiszen a válaszmegtagadók bevonására extra erőfeszítéseket tettünk. Ebből egyrészt az következett, hogy valamelyest megnőtt a tervezési és a programozási idő. Másrészt a kérdezőbiztosok extra erőfeszítései az első adatfelvételi szakaszt is növelték, igaz ez kontaktonként maximum 1-2 percet jelentett, ami a teljes adatfelvételi munka időtartamában nem tekinthető számottevőnek. A második szakasz – tehát azon válaszmegtagadók újbóli felkeresése, akik nyitottnak bizonyultak a válaszadásra egy másik módon –, részben párhuzamosan zajlott a primer adatfelvétellel, így ez összességében csak 2 nappal növelte a teljes adatfelvételi munkát. A H1 tanulsága, hogy jól felkészült csapattal a hibrid jellegből fakadó extra munka ideje hatékonyan minimalizálható – a H1 mérés hosszú adatfelvételi ideje elsősorban a személyes mód adottságából fakadt.

A H2 méréstől már joggal várhattuk, hogy rövid idő alatt elkészüljön, hiszen azt az online módot involválta, amely hagyományosan gyorsabb, mint a többi adatfelvételi mód. Ez a mérés 9 nap alatt készült el, a leggyorsabban a három mód közül, majdnem háromszor gyorsabban (2,77-szer), mint a H1, és 33 százalékkal gyorsabban, mint az egycsatornás telefonos, ami 12 napig tartott volna a 800 fős mintán. Ez az előny valóban az online alminta eredménye, a 376 online válasz mindössze három nap alatt „jött össze”. Ahogy korábban tárgyaltam, a „túl gyors” válaszadás, azaz amikor a minta túlságosan hamar megtelik, gyakori az online paneleknél. Ilyenkor az történik, hogy azok, akik előbb értesülnek a kérdőívről, rövid időn belül kitöltik azt, és egyes csoportoknak (például azoknak, akik csak hétvégén interneteznek) már nincs ideje/esélye csatlakozni a kutatáshoz. Ez akár szisztematikus mintatorzulást is eredményezhet. Ezért a panelt kezelő cégek általában igyekeznek kontrollálni a túl gyors kitöltést. A vizsgálat során nem tapasztaltam komolyabb torzulásokat a minta összetételének elemzésekor.

4. Összegzés

Jelen tanulmány empirikus adalékokkal szolgál a politikai témájú kérdőíves vizsgálatok egy hazánkban kevésbé elterjedt módszertani megközelítéséről, a hibrid adatfelvétel alkalmazhatóságáról. A 2018-as választást megelőző két vegyes adatfel-

vételi mód és az egycsatornás telefonos mérés összehasonlításának eredménye arra utal, hogy a hibrid adatfelvételek használata ígéretes módszertani irány lehet. A három mérés közül a párhuzamos H2 változat adta a legpontosabb közelítést a választási eredményekre úgy, hogy ennek a költségei voltak a legalacsonyabbak, adatfelvételi ideje pedig a legrövidebb.

Az összehasonlítás a hibrid mérésekben rejlő potenciál mellett arra is rávilágít, hogy bizonyos problémákra az adatfelvételi módok keverése nem feltétlenül jelent megoldást. A válaszadási arányok mindhárom mérésnél rendkívül alacsonyak voltak. A H1-nél a válaszadási arány növelését célzó stratégiát lényegében kudarcnak ítéltethetjük; ilyen jellegű hibrid stratégiák alkalmazása – legalábbis jelen eredmények alapján – nem javasolt. Igaz, *Szeidl–Tóth* [2020] nemrégiben lefolytatott „re-contact” kísérletében a visszautasítók újbóli felkeresése sikeres volt. Ez arra utal, hogy a visszautasítás pillanatában történő eredménytelenség ellenére egy későbbi felkérés akár más módokon célravezető lehet. Felvetődik a kérdés, hogy mekkora jelentőséget tulajdonítsunk az alacsony válaszadási arányoknak. Az a tény, hogy a H2 mérés alacsony válaszadási arány mellett is pontos közelítést tudott adni a választási eredményekre, arra enged következtetni, hogy a nemválaszolás nem okozott számottevő torzítást. Ez például azért lehet, mert a nemválaszolás nem szisztematikus a pártválasztás szempontjából. Ennek eldöntésére további kutatások szükségesek.

Kérdés lehet továbbá, hogy a H2-ben előforduló, a nemvalószínűségi online panelhez köthető magas fedési hiba miért nem okozott mérési torzításokat. Az egyik lehetséges magyarázat, hogy a pártválasztás független az internethasználatától, illetve a hagyományos strukturális változóktól. Ez az érvelés ugyanakkor szembemegy a hazai tapasztalatokkal, amelyek – ha nem is erős, de – érdemi különbségekről tanúskodnak mind a digitális tőke (például digitális eszköz-ellátottság), mind a szociodemográfiai jellemzők pártválasztásra gyakorolt hatásairól (*Enyedi–Tardos* [2018], *Makay–Stefkovics* [2021]). A kérdőíves adatok ugyanakkor arra is rámutatnak, hogy a telefonos és az online mintákban jóval nagyobb volt a politika iránt érdeklődők aránya. Ebből az következhet, hogy bár ezek a minták a teljes populációt kevésbé reprezentálják jól, mint a személyes módon felvett minták, mégis pontosabban tudják leképezni a politikailag aktív, a választásokon ténylegesen részt vevő réteget. Ezt erősíti az is, hogy a személyes minta eredményezte a legalacsonyabb részvételt (összhangban *Angelusz–Tardos* [2009] korábbi tapasztalataival), más szóval jóval kevésbé voltak politikailag aktívak a H1 minta tagjai.

A H2 pontosságára magyarázatot adhat az önkitöltős online mód bevonása az adatfelvételbe. Mint arra korábban utaltam, szenzitív kérdéseknél a kérdezőbiztos személye a mérés egy lehetséges hibaforrása, míg az online, önkitöltős kérdéssel akár őszintébb válaszokat kaphatunk. Ráadásul az adatok nem arra utalnak, hogy a kérdezőbiztos személye érdemben elősegítette volna a motiváltabb válaszadást; a részleges nemválaszadási arányok a személyes mintában is nagyon magasak voltak.

Annak sincs számottevő nyoma, hogy az online kérdőíveknél egyébként gyakoribb válaszadói hibák, mint például a nemdifferenciált válaszadás, jellemzőbb lett volna a H2 mérésben, mint a másik két mintában.

Az eredmények azt is demonstrálják, hogy egy hibrid minta számottevő költségcsökkenést és rövidebb adatfelvételi időt jelenthet. A H2 esetében ezek az online mód bevonásának voltak köszönhetők. A költségek csökkenése pedig akár a mintaelemszám növeléséhez és ezáltal megbízhatóbb adatokhoz, mélyebb elemzési lehetőségekhez juttathatja a kutatókat.

Jelen vizsgálat korlátja, hogy az adatfelvételi módok keverésének adatminőségre gyakorolt lehetséges negatív hatásait nem tudtam meghatározni. A mérési módhatások feltárásában a fő probléma az, hogy nehéz elkülöníteni a szelekciósakat a mérési módhatásoktól (*Vannieuwenhuyze–Loosveldt–Molenberghs* [2010]). Önmagában például egy dummy jellegű módváltozó bevonása független változóként egy regressziós modellbe nem mutatná meg, hogy van-e hatása pusztán annak, hogy más módon válaszoltak a kérdezettek. Erre egy olyan kontrollált kísérletben lenne lehetőség, ahol a résztvevőket egy csatornán választják ki, majd véletlenszerűen allokálják különböző adatfelvételi módokhoz.

5. Konklúzió

Mind a nemzetközi tapasztalatok, mind a jelen tanulmányban bemutatott eredmények azt mutatják, hogy a hibrid adatfelvételekben rejlő lehetőségek kihasználása általában javíthatná a hazai politikai mérések minőségét. Nyugat-Európában és az Egyesült Államokban a hibrid adatfelvételek jelentik a normát (*Biemer–Lyberg* [2003]), amiből az következik, hogy – a hivatalos statisztikai adatfelvételeket leszámítva – egyfajta módszertani „versenyhátrányban” van Magyarország. A nyilvánosság számára is látható piaci, illetve akadémiai kísérletek híján azonban nem könnyű meghatározni, hogy Magyarországon mely adatfelvételi változatok lehetnek a leghatékonyabbak.

Jelen kutatás eredményei azt mutatják, hogy a nemválaszoló bevonását célzó, más módokat felajánló szekvenciális stratégiákkal nem számíthatunk komolyabb áttöresre, igaz, a válaszadók felkeresése egy későbbi időpontban akár eredményes lehet (*Szeitl–Tóth* [2020]). A párhuzamos, online mintaszeleteket tartalmazó adatfelvételekkel azonban érdemes lehet kísérletezni, továbbá *Messing–Ságvári* [2020] kutatása azt mutatja, hogy az online típusú hibrid stratégiák kifejezetten hatékonyak lehetnek Magyarországon is. A H2 pontossága azt is demonstrálta, hogy nem minden esetben törvényszerű a nemvalószínűségi online minták gyenge adatminősége. Kétségtől

további kutatások szükségesek annak megértésére, hogy a nemvalószínűségi online mintákat milyen körülmények között érdemes bevonni egy hibrid felvételbe, valamint, hogy azok milyen hibákat involválnak. Mindazonáltal az online mód keverése más módokkal azokban az országokban célszerű, ahol az internetpenetráció már elérte azt a szintet, hogy nagy tömegeket lehet elérni online kérdőívekkel, viszont a teljes országos lefedettséghez további módok bevonása szükséges. Magyarország ilyen hely (Pintér–Kótya [2010]).

Tanulmányomban Magyarországon az elsők között vizsgáltam a hibrid minták alkalmazhatóságát politikai témájú kérdőíves felméréseknél. Nincs ok abban bízni, hogy az adatfelvételek módszertani feltételei a közeljövőben számottevően javulnak, vagy, hogy a csökkenő válaszadási hajlandóság hirtelen növekedésnek indul. Mindezek fényében fontos további kísérleteket folytatni azzal kapcsolatban, hogy az adatfelvételi módok keverése mikor és hogyan juttathatja a kutatókat jobb, illetve megbízhatóbb adatokhoz, és ezáltal a politikai viselkedés pontosabb megértéséhez.

Függelék

F1. táblázat

*Relatív adatfelvételi költségek egy átlagos, 1 000 fős véletlen lakossági mintán,
20 perces kérdőívvel, európai országok szerint, 2017*
(Relative data collection costs for an average random population sample of 1,000,
with a 20-minute questionnaire, by European country, 2017)

Ország	Személyes	Telefonos	Online	Személyes/ telefonos	Személyes/ online	Telefonos/ online
	adatfelvétel					
Ausztria*	100	30	–	3,33	–	–
Belgium*	100	80	62	1,25	1,61	1,29
Belgium**	100	70	25	1,43	4,00	2,80
Ciprus*	100	80	40	1,25	2,50	2,00
Ciprus**	100	50	–	2,00	–	–
Dánia**	100	35	25	2,86	4,00	1,40
Finnország*	100	50	–	2,00	–	–
Írország**	100	65	–	1,54	–	–
Izland**	100	40	15	2,50	6,67	2,67
Lengyelország**	100	–	–	–	–	–
Magyarország**	100	80	50	1,25	2,00	1,60
Németország**	100	43	15	2,33	6,67	2,87

(A táblázat folytatása a következő oldalon)

(Folytatás)

Ország	Személyes	Telefonos	Online	Személyes/ telefonos	Személyes/ online	Telefonos/ online
	adatfelvétel					
Norvégia**	100	50	20	2,00	5,00	2,50
Olaszország**	100	60	50	1,67	2,00	1,20
Oroszország**	100	63	–	1,59	–	–
Portugália**	100	75	–	1,33	–	–
Románia**	100	60	–	1,67	–	–
Spanyolország**	100	60–70	30	1,54	3,33	2,17
Svájc*	100	40	5	2,50	20,00	8,00
Svájc**	100	50	–	2,00	–	–
Svédország**	100	60	–	1,67	–	–
Szlovákia**	100	85	65	1,18	1,54	1,31
Szlovénia**	100	30	20	3,33	5,00	1,50
Ukrajna**	100	60	40	1,67	2,50	1,50
Átlag (szórás)	–	–	–	1,91 (0,62)	4,77 (4,71)	2,34 (1,73)
Medián	–	–	–	1,67	3,67	1,80

* Nemzeti statisztikai hivatalok adatai.

** ESS-t végző cégek adatai.

Forrás: Villar–Fitzgerald ([2017] 285. old.) alapján saját szerkesztés.

F2. táblázat

*Egyes szociodemográfiai változók súlyozatlan megoszlásai
az adatfelvételekben, valamint összehasonlításuk a mikrocenzus eredményeivel (%)*
(Unweighted distributions of some sociodemographic variables in the surveys
and their comparison with the Microcensus 2016 results [%])

Változó	H1	H2	Egycsatornás telefonos	Mikrocenzus 2016
	adatfelvétel			
Nem				
Férfi	47,3	47,5	44,1	46,6
Nő	52,8	52,5	55,9*	53,4
MSE	0,9	1,6	12,5	

(A táblázat folytatása a következő oldalon)

(Folytatás)

Változó	H1	H2	Egycsatornás telefonos	Mikrocenzus 2016
	adatfelvétel			
Kor				
18–29 éves	18,1	18	13,4***	18,1
30–39 éves	18,9	19,5	12,4	19,4
40–49 éves	17,5	17,0	13,2	16,1
50–59 éves	14,5*	14,2*	15,4	17,7
60 év és a feletti	31,0	31,3	45,7***	28,6
MSE	18,2	20,4	355,1	
Legmagasabb iskolai végzettség				
8 általános	19,3***	9,8***	13,2***	28,8
Szakmunkás, szakiskolai	29,1***	19,9	14,6***	22,2
Érettségi	40,6***	51,6***	45,7***	31,3
Felsőfokú	11,0***	18,8	26,5***	17,7
MSE	269,2	779,6	585,9	
Lakhely településtípusa				
Budapest	17,9	17,6	17,6	18,1
Megyeszékhely	19,8	22,2*	20,8	17,9
Egyéb város	34,0	33,4	33,3	34,0
Község, tanya	28,4	26,7*	28,3	30,0
MSE	6,2	30,0	12,0	

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

Megjegyzés. MSE (mean squared error): átlagos négyzetes hiba.

F3. táblázat

Részleges nemválaszolási arányok az adatfelvételekben leíró statisztikai mutatók szerint (%)
(Partial non-response rates in the surveys by descriptive statistical measure [%])

Mutató	H1	H2	H2- telefonos	H2-online	Egycsatornás telefonos	H1 vs. H2	H1 vs. egycsatornás telefonos	H2 vs. egycsatornás telefonos
	adatfelvétel							
	Részleges nemválaszolás összesen							
Átlag	16,28	14,54	11,35	17,34	10,58	–1,74*	–5,70***	–3,96***
Medián	7,13	12,31	6,38	15,89	6,15			
Szórás	19,64	14,29	14,39	15,54	12,69			
Minimum	0,00	0,00	0,00	0,00	0,00			
Maximum	68,09	59,71	56,76	63,49	51,61			

(A táblázat folytatása a következő oldalon)

(Folytatás)

Mutató	H1	H2	H2- telefonos	H2-online	Egycsatornás telefonos	H1 vs. H2	H1 vs. egycsatornás telefonos	H2 vs. egycsatornás telefonos
	adatfelvétel							
	„Nem válaszol”							
Átlag	12,65	8,53	6,23	10,63	4,92	-4,12***	-7,73***	-3,61***
Medián	4,81	5,32	1,86	5,37	1,10			
Szórás	19,62	11,42	10,33	13,01	11,32			
Minimum	0,00	0,00	0,00	0,00	0,00			
Maximum	65,96	56,83	53,95	60,32	51,61			
	„Nem tudja”							
Átlag	9,53	6,01	5,12	6,71	5,66	-3,52***	-3,87***	-0,35
Medián	6,31	1,93	1,33	3,15	2,40			
Szórás	10,04	6,75	7,58	7,88	6,56			
Minimum	0,00	0,00	0,00	0,00	0,00			
Maximum	30,13	22,39	30,63	26,40	22,80			

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

Megjegyzés. Egyes demográfiai kérdéseknél nem volt differenciálva a „nem válaszol” és a „nem tudja” válaszlehetőség. Ezeknél a kérdéseknél a nemválaszolást 50-50 százalékban osztottam meg a két opció között.

F4. táblázat

*Részleges nemválaszolási arányok az adatfelvételekben demográfiai változók esetén
leíró statisztikai mutatók szerint (%)*

(Partial non-response rates for demographic variables in the surveys by descriptive statistical measure [%])

Mutató	H1	H2	H2- telefonos	H2-online	Egycsatornás telefonos	H1 vs. H2	H1 vs. egycsatornás telefonos	H2 vs. egycsatornás telefonos
	adatfelvétel							
Átlag	15,76	9,90	9,26	10,52	7,08	-5,86***	-8,68***	-2,82**
Medián	0,63	1,55	0,93	2,57	2,00			
Szórás	27,08	16,87	17,33	16,85	14,68			
Minimum	0,00	0,00	0,00	0,00	0,00			
Maximum	65,38	49,45	56,76	47,45	40,30			

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

Megjegyzés. Egyes demográfiai kérdéseknél nem volt differenciálva a „nem válaszol” és a „nem tudja” válaszlehetőség. Ezeknél a kérdéseknél a nemválaszolást 50-50 százalékban osztottam meg a két opció között.

F5. táblázat

*Részleges nemválaszolási arányok az adatfelvételekben tartalmi változók esetén
leíró statisztikai mutatók szerint (%)*
(Partial non-response rates for content variables in the surveys by descriptive statistical measure [%])

Mutató	H1	H2	H2- telefonos	H2-online	Egycsatornás telefonos	H1 vs. H2	H1 vs. egycsatornás telefonos	H2 vs. egycsatornás telefonos
Átlag	20,51	17,23	12,56	21,29	12,27	-3,28*	-8,24***	-4,96***
Medián	19,48	16,58	11,17	22,66	11,38			
Szórás	14,93	12,04	12,53	13,47	11,53			
Minimum	1,25	1,99	0,00	3,74	0,00			
Maximum	68,09	59,71	56,58	63,49	51,61			

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

F6. táblázat

Részvételi arányra vonatkozó eredmények az adatfelvételekben (%)
(Results on participation rates in the surveys [%])

Részvételi szándék	H1	H2	Egycsatornás telefonos	Választási eredmény
	adatfelvétel			
Biztosan elmegy szavazni	50,40***	72,87	82,26***	69,73
Valószínűleg elmegy szavazni	29,24	12,29	11,62	
Valószínűleg nem megy el szavazni	4,42	3,86	1,65	
Biztosan nem megy el szavazni	7,56	4,95	1,64	
Nem tudja	5,16	4,01	2,61	
Nem válaszol	3,22	2,02	0,22	

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

Megjegyzés. A pontos kérdés: „2018 áprilisában lesznek az országgyűlési választások. Mit gondol, Ön:...?”

Forrás: A részvételi adat forrása: valasztas.hu

F7. táblázat

*Az egyes pártok támogatottsága az adatfelvételekben imputálás után
és a választási eredmények (%)*
(Survey results on the support of parties after imputation,
and the election results [%])

Párt	H1	H2	Egycsatornás telefonos	Választási eredmény
	adatfelvétel			
Fidesz–KDNP	56,8***	48,3	65,3***	47,4
Jobbik	14,6***	19,4	6,2***	19,8
MSZP-P	13,9	10,6	10,8	12,4
LMP	5,3	6,6	4,8	7,3
DK	4,1	5,1	9,8	5,6
Momentum Mozgalom	1,0	2,9	1,4	3,2
MKKP	0,8	3,5	0,6	1,8
Egyéb párt	0,3	1,0	0,0	2,6
Együtt	1,8	0,9	0,8	0,7
Magyar Munkáspárt	0,3	1,5	0,0	0,3
Liberálisok	1,4	0,4	0,4	0,0
MSE	134,5	10,2	540,4	

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

Megjegyzés. A pontos kérdés: „Melyik pártra szavaz(na) az országgyűlési választáson?/És ha mégis elmenne, melyik pártra szavaz(na) az országgyűlési választáson?”

Forrás: A választási eredmények forrása: valasztas.hu

Irodalom

- AAPOR (AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH) [2016]: *Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys.*
https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf
- ANGELUSZ R. – TARDOS R. [2006]: „Rejtőzködő szavazatok” és a társadalmi-politikai miliő, a látencia probléma vizsgálatának egy megközelítése. In: *Angelusz R. – Tardos R. (szerk.): Mérésről mérésre. A választáskutatás módszertani kérdései.* Századvég Kiadó, Demokrácia Kutatások Magyar Központja Közhasznú Alapítvány. Budapest. 111–134. old.
- ANGELUSZ R. – TARDOS R. [2009]: Demoszkópiai reprezentativitás és demokratikus reprezentáció. Módszertani problémák és tartalmi dilemmák. In: *Enyedi Zs. (szerk.): A népakarat dilemmái.* Századvég Kiadó, Demokrácia Kutatások Magyar Központja Közhasznú Alapítvány. Budapest. 293–327. old.

- BIEMER, P. P. – LYBERG, L. E. [2003]: *Introduction to Survey Quality*. Wiley Series in Survey Methodology. Wiley. Hoboken.
- BÖCSKEI B. – SZABÓ A. (szerk.) [2018]: *Várakozások és valóságok. Parlamenti választás 2018*. Magyar Tudományos Akadémia Társadalomtudományi Kutatóközpont Politikatudományi Intézet, Napvilág Kiadó. Budapest.
- CSÁNYI G. – DÉS F. [2020]. A kérdezőbiztosok képzése nemi alapú erőszak tematikájú kvantitatív adatfelvételekkel. *Statistikai Szemle*. 98. évf. 1. sz. 34–61. old. <http://doi.org/10.20311/stat2020.1.hu0034>
- DE LEEUW, E. D. [1992]: *Data Quality in Mail, Telephone and Face to Face Surveys*. Dissertation. Institute of Education Sciences. Amsterdam. <https://files.eric.ed.gov/fulltext/ED374136.pdf>
- DE LEEUW, E. D. [2005]: To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*. Vol. 21. No. 2. pp. 233–255. <https://www.bebr.ufl.edu/sites/default/files/To%20Mix%20or%20Not%20to%20Mix%20Data%20Collection%20Modes%20in%20Surveys.pdf>
- DE LEEUW, E. D. – HOX, J. J. [2011]: Internet surveys as part of a mixed-mode design. In: *Das, M. – Ester, P. – Kaczmirek, L. (eds.): Social and Behavioral Research and the Internet: Advances in Applied Methods and Research Strategies*. Routledge. London. pp. 45–76. <http://www.joophox.net/publist/DeLeeuwHox2011.pdf>
- DILLMAN, D. A. – SMYTH, J. D. – CHRISTIAN, L. M. [2014]: *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 4th Edition. Wiley. Hoboken.
- DILLMAN, D. A. – WEST, K. K. – CLARK, J. R. [1994]: Influence of an invitation to answer by telephone on response to census questionnaires. *Public Opinion Quarterly*. Vol. 58. No. 4. pp. 557–568. <https://doi.org/10.1086/269447>
- DUFFY, B. – SMITH, K. – TERHANIAN, G. – BREMER, J. [2005]: Comparing data from online and face-to-face surveys. *International Journal of Market Research*. Vol. 47. No. 6. pp. 615–639. <https://www.ipsos.com/sites/default/files/publication/1970-01/comparing-data.pdf>
- ENYEDI ZS. – TARDOS R. [2018]: A magyar választók társadalmi beágyazottsága és ideológiai tagoltsága. In: *Böcskei B. – Szabó A. (szerk.): Várakozások és valóságok. Parlamenti választás 2018*. Magyar Tudományos Akadémia Társadalomtudományi Kutatóközpont Politikatudományi Intézet, Napvilág Kiadó. Budapest. 43–75. old.
- GROVES, R. M. – KAHN, R. L. [1979]: *Surveys by Telephone: A National Comparison with Personal Interviews*. Academic Press. Cambridge.
- HOLBROOK, A. L. – GREEN, M. C. – KROSNIK, J. A. [2003]: Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*. Vol. 67. No. 1. pp. 79–125. <https://doi.org/10.1086/346010>
- JÄCKLE, A. – ROBERTS, C. – LYNN, P. [2006]: *Telephone versus Face-to-Face Interviewing: Mode Effects on Data Quality and Likely Causes: Report on Phase II of the ESS-Gallup Mixed Mode Methodology Project*. ISER Working Paper. Institute for Social and Economic Research. Colchester.
- KATZ, D. [1942]: Do interviewers bias poll results? *Public Opinion Quarterly*. Vol. 6. No. 2. pp. 248–268. <https://www.jstor.org/stable/pdf/2745023.pdf>
- KENNEDY, C. – MERCER, A. – KEETER, S. – HATLEY, N. – MCGEENEY, K. – GIMENEZ, A. [2016]: *Evaluating Online Nonprobability Surveys*. Pew Research Center. Washinton, D.C.

- <https://www.pewresearch.org/methods/2016/05/02/evaluating-online-nonprobability-surveys/>
- KITCHIN, R. [2014]: *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. SAGE Publications Ltd. Los Angeles, London, New Delhi, Singapore, Washington, D.C.
- KMETTY Z. [2011]: A telefonos kutatások speciális problémái. *Statistikai Szemle*. 90. évf. 1. sz. 41–63. old.
- KOWARIK, A. – TEMPL, M. [2016]: Imputation with the R Package VIM. *Journal of Statistical Software*. Vol. 74. No. 7. pp. 1–16. <https://doi.org/10.18637/jss.v074.i07>
- KSH (KÖZPONTI STATISZTIKAI HIVATAL) [2016]: *Elkészült a társadalmi szelfi. Sikeresen zárta a Központi Statisztikai Hivatal a Mikrocenzust*. https://www.ksh.hu/mikrocenzus2016/docs/hirek/20161114_ksh_sikeresen_zarta_a_mikrocenzust.pdf
- KURAN, T. [1997]: *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Harvard University Press. Cambridge.
- LAKATOS ZS. [2011]: Választási előrejelzések. In: *Enyedi Zs. – Szabó A. – Tardos R. (szerk.): Új Képlet. Demokrácia Kutatások Magyar Központja Közhasznú Alapítvány*. Budapest. 241–269. old.
- LIAO, P.-S. [2016]: Social desirability bias and mode effects in the case of voting behavior. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*. Vol. 132. No. 1. pp. 73–83.
- LIPPS, O. – LUTZ, G. [2010]: How answers on political attitudes are shaped by interviewers. Evidence from a panel survey. *Swiss Journal of Sociology*. Vol. 36. No. 2. pp. 345–357. https://serval.unil.ch/resource/serval:BIB_557B73B8C687.P001/REF
- LUIJKX, R. – JÓNSDÓTTIR, G. A. – GUMMER, T. – STÄHLI, M. E. – FREDERIKSEN, M. – KETOLA, K. – REESKENS, T. – BRISLINGER, E. – CHRISTMANN, P. – GUNNARSSON, S. T. – HJALTASON, A. B. – JOYE, D. – LOMAZZI, V. – MAINERI, A. M. – MILBERT, P. – OCHSNER, M. – POLLIEN, A. – SAPIN, M. – SOLANES, I. – VERHOEVEN, S. – WOLF, C. [2021]: The European Values Study 2017: On the way to the future using mixed-modes. *European Sociological Review*. Vol. 37. No. 2. pp. 330–346.
- LUITEN, A. – HOX, J. – DE LEEUW, E. D. [2020]: Survey nonresponse trends and fieldwork effort in the 21st century: Results of an international study across countries and surveys. *Journal of Official Statistics*. Vol. 36. No. 3. pp. 469–487. <https://doi.org/10.2478/jos-2020-0025>.
- MAKAY M. – STEFKOVICS Á. [2021]: Mi oszt meg minket? Politikai törésvonalak vizsgálata Magyarországon. In: *Barthel-Rúzsza Zs. – Fűrész G. – Pillók P. – Stefkovics Á. (szerk.): Magyarország 2021: Társadalom, gazdaság és politika napjainkban*. Századvég. Budapest. 241–277. old.
- MANFREDA, K. Z. – VEHOVAR, V. – BATAGELI, Z. [2001]: Web versus mail questionnaire for an institutional survey. *The Challenge of the Internet*. pp. 1–11. http://www.websm.org/db/12/1497/Bibliography/Web_versus_Mail_Questionnaire_for_an_Institutional_Survey/?menu=1&lst=&q=search_1_1_-1&qdb=12&qsort=1
- MARIÁN B. [2006]: A várható választási eredmények modellezése. In: *Angelusz R. – Tardos R. (szerk.): Mérésről mérésre. A választáskutatás módszertani kérdései*. Demokrácia Kutatások Magyar Központja Közhasznú Alapítvány. Budapest. 135–162. old.

- MATTHIJSE, S. M. – DE LEEUW, E. D. – HOX, J. [2015]: Internet panels, professional respondents, and data quality. *Methodology*. Vol. 11. No. 3. pp. 81–88. <https://doi.org/10.1027/1614-2241/a000094>
- MESSING V. – SÁGVÁRI B. [2020]: *ESS Kutatási Infrastruktúra Magyarország. Éves összefoglaló*. Társadalomtudományi Kutatóközpont. https://tk.hu/uploads/files/ess/ESS_2020-as_beszamolo.pdf
- NÉMETH, R. – LUKSANDER, A. [2018]: Strong impact of interviewers on respondents' political choice: Evidence from Hungary. *Field Methods*. Vol. 30. No. 2. pp. 155–170. <https://doi.org/10.1177/1525822X18769502>
- OLSON, K. – SMYTH, J. D. – WOOD, H. M. [2012]: Does giving people their preferred survey mode actually increase survey participation rates? An experimental examination. *Public Opinion Quarterly*. Vol. 76. No. 4. pp. 611–635. <https://doi.org/10.1093/poq/nfs024>
- PICKERY, J. – LOOSVELDT, G. [2002]: A multilevel multinomial analysis of interviewer effects on various components of unit nonresponse. *Quality and Quantity*. Vol. 36. No. 4. pp. 427–437. <https://doi.org/10.1023/A:1020905911108>
- PILLÓK P. [2010]: *Az elhanyagolt tényező, avagy nem mintavételi hibák a kérdőíves adatfelvételekben*. Doktori disszertáció. Eötvös Loránd Tudományegyetem. Budapest.
- PILLÓK P. – DARVAS P. [2001]: *Online adatfelvételi módszerek és felhasználásuk*. Szakdolgozat. Eötvös Loránd Tudományegyetem. Budapest.
- PINTÉR R. – KÓTAY B. [2010]: A hibrid adatfelvétel módszertani kihívásai. *Statistikai Szemle*. 8. évf. 7–8. sz. 723–738. old.
- ROBERTS, C. [2007]: *Mixing Modes of Data Collection in Surveys: A Methodological Review*. Discussion Paper. Economic and Social Research Council National, Centre for Research Methods. Southampton.
- RUDAS T. [1995]: Vita a telefonos közvélemény-kutatásokról. *Replika*. 19–20. sz. 181–196. old.
- RUDAS T. [1998]: *Hogyan olvassunk közvélemény-kutatásokat?* Új Mandátum Könyvkiadó. Budapest.
- SALGANIK, M. J. [2019]: *Bit by Bit: Social Research in the Digital Age*. Princeton University Press. Princeton.
- SAVAGE, M. – BURROWS, R. [2007]: The coming crisis of empirical sociology. *Sociology*. Vol. 41. No. 5. pp. 885–899. <https://doi.org/10.1177/0038038507080443>
- SZEITL B. – TÓTH I. GY. [2020]: Hova tovább a nemválaszolókkal? A European Social Survey alapján végzett módszertani kísérlet eredményei. *Szociológiai Szemle*. 30. évf. 3. sz. 96–114. old. <https://doi.org/10.51624/SzocSzemle.2020.3.5>
- TOURANGEAU, R. – RIPS, L. J. – RASINSKI, I. K. [2000]: *The Psychology of Survey Response*. Cambridge University Press. Cambridge. <https://doi.org/10.1017/CBO9780511819322>
- TOURANGEAU, R. – SMITH, T. W. [1996]: Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*. Vol. 60. No. 2. pp. 275–304.
- TOURANGEAU, R. – YAN, T. [2007]: Sensitive questions in surveys. *Psychological Bulletin*. Vol. 133. No. 5. pp. 859–883. <https://doi.org/10.1037/0033-2909.133.5.859>
- VANNIEUWENHUYZE, J. – LOOSVELDT, G. – MOLENBERGHS, G. [2010]: A method for evaluating mode effects in mixed-mode surveys. *Public Opinion Quarterly*. Vol. 74. No. 5. pp. 1027–1045. <https://doi.org/10.1093/poq/nfq059>

- VILLAR, A. – FITZGERALD, R. [2017]: Using mixed modes in survey research: Evidence from six experiments in the ESS. In: *Breen, M. I. (ed.): Values and Identities in Europe: Evidence from the European Social Survey. 1st Edition.* Routledge. London. pp. 273–310.
- VOOGT, R. J. J. – SARIS, W. E. [2003]: To participate or not to participate: The link between survey participation, electoral participation, and political interest. *Political Analysis*. Vol. 11. No. 2. pp. 164–179. <https://doi.org/10.1093/pan/mpg003>
- ZÁVE CZ T. [2010]: *A személyes, a telefonos és az online adatfelvétel módszertani összehasonlítása, a hibrid adatfelvétel lehetőségei.* Magyar Választáskutató Program „Részvétel és képviselet” kutatás. Ipsos Zrt. Budapest.