

CHPS

ČESKÉ PANELOVÉ ŠETŘENÍ DOMÁCNOSTÍ

**Dokumentace ke sběru a zpracování dat
Vlna 6 (2023–2024)**

Verze 1.0 (16. 12. 2024)

Tato práce vznikla v rámci projektu NPO „Národní institut pro výzkum socioekonomických dopadů nemocí a systémových rizik“ č. LX22NPO5101, financovaného Evropskou unií – Next Generation EU (MŠMT, NPO: EXCELES).

V této publikaci byla využita data, která byla získána prostřednictvím datových služeb projektu Czech Social Science Data Archive / European Social Survey – Czech Republic (CSDA/ESS-CZ). Projekt výzkumné infrastruktury CSDA/ESS-CZ podporuje MŠMT v rámci grantu č. LM2023046.

Röschová, M., Klusáček, J., Sukhotskaya, A. 2024. České panelové šetření domácností: Dokumentace ke sběru a zpracování dat. Vlna 6 (2023–2024). Praha: Sociologický ústav AV ČR. (Verze 1.0). <https://doi.org/10.13060/dk.2024.2>

OBSAH

SHRNUTÍ	5
1. O VÝZKUMU	7
1.1. Základní informace	7
1.2. Realizátor sběru dat	7
1.3. Etické aspekty výzkumu	8
1.4. Předchozí ročníky	8
2. VÝBĚROVÝ SOUBOR	10
2.1. Cílová populace	10
2.2. Jednotka výběru	10
2.3. Postup výběr původního panelu	11
2.4. Doplnující soubor domácností	12
2.5. Složení dotázaného souboru	16
2.6. Výběr respondentů individuálního dotazníku	22
3. SBĚR DAT	24
3.1. Příprava tazatelů	24
3.2. Tazatelé pracující na sběru dat	25
3.3. Průběh sběru dat	28
3.4. Sběr dat na původním panelu domácností	31
3.5. Sběr dat na doplňujícím souboru domácností	38
3.6. Monitorování práce tazatelů	42
3.7. Administrace dotazovacích nástrojů	44
3.8. Odměny respondentům	45
3.9. Kontrola kvality práce tazatelů	46
4. DOTAZOVACÍ NÁSTROJE	49
4.1. Typy instrumentů	49
4.2. Návratnost na úrovni instrumentů	50
4.3. Délka dotazníku	51
4.4. Podmíněné dotazování	52
4.5. Zpracování obsahu dotazníku	54
4.6. Skriptování	55
4.7. Testování a finální úpravy	55
4.8. Změny ve srovnání s předchozími vlnami	56
4.9. Odlišnosti CAPI a CAWI verze dotazníku	59
4.10. Neplatné („únikové odpovědi“)	60
4.11. Dokumentace dotazovacího nástroje	62
5. DATOVÝ SOUBOR	63
5.1. Obsah a struktura datového souboru	63

5.2.	Identifikační čísla	65
5.3.	Status členů domácnosti a změny ve složení domácností	67
5.4.	Účast v dotazování v minulosti	69
5.5.	Odvozené proměnné	71
5.6.	Odvozené ukazatele socioekonomického postavení	74
5.7.	Geografické (lokalizační) proměnné	76
5.8.	Proměnné příjmu domácnosti a jednotlivce	77
5.9.	Paradata	78
6.	ZPRACOVÁNÍ DAT	82
6.1.	Zpracování hrubých dat	82
6.2.	Kontroly a čištění průběžných dat	82
6.3.	Zpracování finálních dat	87
6.4.	Imputace v datech	90
6.5.	Anonymizace dat	97
7.	POST-STRATIFIKAČNÍ VÁHY	98
	PŘÍLOHA: ZVEŘEJNĚNÉ MATERIÁLY K VÝZKUMU	105
	SEZNAM TABULEK	106

Shrnutí

Cílová skupina

Cílovou populaci výzkumu tvoří domácnosti, které obvykle žijí na území České republiky, a to bez ohledu na občanství a národnost. Domácnost je vymezena jako skupina osob, které společně žijí a společně hospodaří. Do cílové populace spadají domácnosti žijící v bytech a domech určených k bydlení.

Cílovou populaci individuálního dotazníku tvoří neinstitutionalizovaná populace České republiky ve věku 15 a více let.

Jednotka výběru

Hospodařící domácnost

Výběrový soubor

Osloveny byly domácnosti z původního panelu CHPS (České panelové šetření domácností), které se šetření zúčastnily ve čtvrté vlně v roce 2018 (celkem 3 188 domácností). Domácnosti byly osloveny nezávisle na změně bydliště.

Soubor původních panelových domácností byl doplněn o nové domácnosti (tzv. doplňující soubor domácností) tak, aby byla posílena reprezentativita celkového souboru zejména z hlediska věkového složení.

Postup výběru

Původní panel domácností, založený v roce 2015: dvoustupňový stratifikovaný náhodný výběr domácností. Zúčastnit se individuálního dotazování může libovolný člen (členové), neprobíhá náhodný výběr respondentů.

Doplňující soubor: kvótní výběr domácností (typ domácnosti dle struktury členů, kraj, velikost místa bydliště, typ domu, nejvyšší dosažené vzdělání v domácnosti) a jednotlivých respondentů individuálního dotazníku (pohlaví, věk, vzdělání)

Realizátor sběru dat

MEDIAN, s. r. o., a STEM/MARK, a. s.

Termín sběru dat

23. 8. 2023 – 25. 2. 2024

Počet dotázaných domácností a individuálních respondentů

3 297 domácností, 4 459 respondentů individuálního dotazníku, z toho:

1 807 domácností z původního panelu CHPS, 2 762 respondentů individuálního dotazníku

1 490 nových domácností, 1 697 respondentů individuálního dotazníku

Retence domácností z původního panelu

56,7 % z celkového počtu oslovených domácností

61,3 % z celkově oslovených domácností způsobilých k účasti (čistá návratnost)

Způsob (mód) sběru dat

Osobní (face-to-face) rozhovor s tazatelem s využitím počítače („computer assisted personal interviewing“, CAPI): 51,9 % domácností

Samostatné vyplňování dotazníku respondentem na internetu („computer assisted web interviewing“, CAWI): 48,1 % domácností

Dotazovací nástroje

Standardizovaný dotazník: dotazník pro domácnost (pro jednoho představitele domácnosti), dotazník pro jednotlivce (pro členy domácnosti ve věku 15 a více let)

Jednotka analýzy

Domácnost a jednotlivec

Obsah a struktura datového souboru

Datový soubor obsahuje proměnné z dotazníku pro domácnost a individuálního dotazníku, včetně identifikačních a odvozených proměnných, proměnných vztahujících se k průběhu dotazování (paradata) a post-stratifikačních vah

Kontrola dat

Průběžné (týdenní) kontroly struktury dat (unikátnost identifikátorů, rozsah hodnot a filtrování, korektního výpočtu odvozených proměnných užívaných při dotazování), věcné smyslnosti distribucí, logických vazeb mezi základními sociodemografickými proměnnými, změn základních sociodemografických proměnných mezi vlnami. Tyto kontroly proběhly také na finálních datech.

Čištění a zpracování dat

Bylo provedeno začištění struktury dat v souladu se strukturou dotazníku (rozsah hodnot, filtrování) a dále vypořádání nesrovnalostí v základních sociodemografických proměnných (v rámci vlny a mezi vlnami). V datech bylo rovněž provedeno kódování polootevřených a otevřených otázek, harmonizace kódů mezi vlnami, výpočet odvozených proměnných a imputace odpovědí z předchozích vln, pokud nebyla otázka dotazována u opakovaně se účastnícího respondenta.

Vážení dat

Byly vytvořeny post-stratifikační váhy na úrovni domácností, členů domácností a respondentů individuálního dotazníku, vždy pro celý soubor a samostatně pro podsoubor původních panelových domácností. K vážení jednotlivců (členů domácností a respondentů) byla využita kombinace pohlaví, věkových skupin a vzdělání a dále samostatně kraj. K vážení domácností byl využit pouze kraj. K vážení byla využita populační data (balance obyvatel) Českého statistického úřadu z roku 2023 a data z Výběrového šetření pracovních sil 2022.

1. O výzkumu

1.1. Základní informace

České panelové šetření domácností je longitudinální výběrové šetření, v němž jsou od roku 2015 sbírány údaje o reprezentativním souboru domácností žijících na území České republiky. Dotazování se opakovaně účastní tentýž soubor domácností, jedná se tedy o panelový sběr dat. Mezi lety 2015 a 2019 se s ročním odstupem uskutečnilo celkem pět vln šetření. V letech 2020 a 2021 byla na částech vzorku třetí, čtvrté a páté vlny provedena dvě krátká šetření zaměřená na tematiku pandemie nemoci covid-19.

V letech 2023 až 2024 proběhla **šestá vlna šetření**. Jejím cílem bylo získat ke spolupráci co nejvyšší počet domácností, které se zúčastnily čtvrté vlny v roce 2018, tedy **zajistit maximální udržení původního panelového souboru**. Do šetření byly také vůbec poprvé od první vlny v roce 2015 zapojeny nově vybrané domácnosti (tzv. doplňující soubor domácností, „refreshment sample“). Velikost původního panelu se po téměř deseti letech existence snížila a došlo v něm k zestárnutí věkové struktury. Cílem doplňujícího souboru bylo proto navýšit velikost zkoumaného souboru a obnovit jeho složení, zejména z hlediska věku, ale také z hlediska vzdělání a dalších charakteristik. Šestý ročník Českého panelového šetření domácností byl realizován jako součást projektu SYRI (Národní institut pro výzkum socioekonomických dopadů nemocí a systémových rizik).

Pro šetření je využíváno zkrácené označení CHPS, které je odvozeno z anglického názvu Czech Household Panel Survey. Při komunikaci s veřejností je výzkum obvykle označován svým podtitulem Proměny české společnosti. Tímto způsobem je také výzkum představován respondentům.

Tato dokumentace se **vztahuje k datům ze šesté vlny šetření**. Pro data z předchozích ročníků byly v minulosti vydány samostatné dokumentace (souborná dokumentace pro první až čtvrtý ročník, dokumentace pro pátý ročník a dokumentace pro ad hoc covid-19 šetření), které jsou volně dostupné v Českém sociálněvědním datovém archivu.

1.2. Realizátor sběru dat

Jako realizátor sběru dat pro šestý ročník šetření byla ve veřejném zadávacím řízení vybrána agentura STEM/MARK, a. s., která zajišťovala sběr dat již v předchozích ročnících. Na dotazování domácností z původního panelu spolupracovala jako poddodavatel také společnost MEDIAN, s. r. o., spolurealizátor šetření v minulých vlnách. Zapojení realizátorů z předchozích ročníků umožnilo využít k dotazování do značné míry stejné tazatele, kteří na šetření pracovali v minulosti.

Společnost STEM/MARK připravila elektronickou verzi dotazníku pro hlavní sběr, zajistila školení tazatelů, koordinovala terénní práce, prováděla kontrolu kvality práce tazatelů a zpracování dat. Společnost MEDIAN zajistila dotazování části domácností z původního panelu, na dotazování nově vybraných domácností se již nepodílela.

1.3. Etické aspekty výzkumu

Realizace sběru dat šesté vlny Českého panelového šetření domácností byla **schválena Komisí pro etiku výzkumu Sociologického ústavu AV ČR**.

Respondenti obdrželi při osobním dotazování od tazatelů dokument s názvem *Informace o účasti ve výzkumu*, který popisoval obsah výzkumu, náležitosti dotazování, požadavky kladené na respondenty a rovněž uváděl povinné informace ke zpracování osobních údajů. Seznámení s podmínkami a souhlas s účastí ve výzkumu potvrzovali respondenti podpisem formuláře *Souhlas s účastí ve výzkumu a zpracováním osobních údajů*. Zpracována mohly být pouze data respondentů, kteří udělili souhlas (podepsaný formulář se přitom podařilo získat od všech účastníků).

Respondentům, kteří se dotazování zúčastnili online, byl před vstupem do dotazníku zobrazen dokument s informacemi o účasti a formulář k udělení souhlasu s účastí. Respondenti v tomto případě potvrzovali svůj souhlas označením příslušných polí v online formuláři.

1.4. Předchozí ročníky

Ročníky 1 až 4 (2015–2019): zakládající vlny

První čtyři ročníky výzkumu, realizované v letech 2015 až 2018, byly provedeny v rámci projektu Dynamika změny v české společnosti Grantové agentury ČR (GB14-36154G). Výzkum organizoval Sociologický ústav Akademie věd ČR, v. v. i., CERGE-EI (společné pracoviště Centra pro ekonomický výzkum a doktorské studium Univerzity Karlovy a Národohospodářského ústavu Akademie věd ČR) a Fakulta sociálních studií Masarykovy univerzity. Sběr dat zajišťovaly agentury STEM/MARK, a. s., a MEDIAN, s. r. o.

Výběrový soubor pro první vlnu sběru (2015) byl sestaven pomocí vícestupňového stratifikovaného náhodného výběru z Registru sčítacích obvodů a budov Českého statistického úřadu (ČSÚ) tak, aby byl reprezentativní za populaci českých domácností. První vlny šetření se účastnilo 5 159 domácností (návratnost 42 %). V následných třech vlnách v letech 2016 až 2018 byly osloveny všechny domácnosti, které se zapojily do bezprostředně předcházející vlny. Druhé vlny šetření se zúčastnilo 4 147 domácností (retence 80 %), třetí vlny 3 616 domácností (retence 85 %) a čtvrté vlny 3 188 domácností (retence 86 %). Dotazování probíhalo kombinací osobního (CAPI) a online (CAWI) dotazování. Pomocí CAWI bylo každoročně dotázáno přibližně 10 % vzorku.

Ročník 5 (2019): propojení s odběrem krevních vzorků

V roce 2019 proběhla návazná pátá vlna, v níž bylo standardizované dotazování doplněno odběrem krevního vzorku. Pátá vlna se uskutečnila jako součást projektu Kumulativní efekty sociálního znevýhodnění na zdraví a kvalitu života, který financovala Technologická agentura ČR (TL02000190). Pod vedením Sociologického ústavu AV ČR, v. v. i., na jeho realizaci společně pracovaly Fyziologický ústav AV ČR, v. v. i., a agentury MEDIAN, s. r. o., a STEM/MARK, a. s. Aplikačním garantem studie bylo pracoviště preventivní kardiologie Institutu klinické a experimentální medicíny (IKEM). K účasti v pátém ročníku byly osloveny domácnosti, které ve čtvrté vlně předběžně souhlasily s účastí

v lékařském výzkumu spojeném s odběrem krve. V souvislosti s touto změnou designu došlo k poklesu retence a do páté vlny se zapojilo 1 609 domácností (retence 62,5 %). V páté vlně šetření probíhal sběr dat pouze pomocí CAPI, nikoli CAWI.

Ad hoc výzkumy (2020–2021): tematika covid-19

V letech 2020 a 2021 se uskutečnily dva kratší ad hoc sběry dat na domácnostech z Českého panelového šetření domácností, které se týkaly aktuálně probíhající pandemie nemoci covid-19.

Tyto výzkumy byly propojeny s populačními studiemi imunity proti viru způsobujícímu onemocnění covid-19. Zahrnovaly jak rekrutaci účastníků pro tyto populační studie z řad panelistů CHPS, tak samostatná dotazníková šetření.

Hlavním cílem těchto dotazníkových šetření bylo porozumět socioekonomickým, zdravotním a dalším faktorům, které ovlivňují šíření koronaviru v populaci. Oba výzkumy využívaly zjednodušenou metodiku sběru dat ve srovnání s předchozími ročníky a odlišovaly se od nich také potřebou velmi rychlého uspořádání.

Dotazníkového šetření v roce 2020 se zúčastnilo 1 533 respondentů z 1 115 domácností z původního souboru CHPS. V šetření z roku 2021 bylo dotázáno 1 036 respondentů z 871 původních panelových domácností.

2. Výběrový soubor

2.1. Cílová populace

Cílovou populaci šetření tvoří jednotlivci a domácnosti žijící v době sběru dat na území České republiky, a to bez ohledu na občanství či národnost. Do cílové populace nespádají jednotlivci dlouhodobě pobývajících v institucích (vězení, nemocnice, kasárna).

Za způsobilé k účasti byly považovány domácnosti schopné se zúčastnit dotazování, tj. alespoň někteří členové domácnosti hovoří česky nebo slovensky a v účasti v dotazování jim nebrání zdravotní stav.

Cílovou populací individuálního dotazníku v šestém ročníku výzkumu jsou osoby ve věku 15 let a starším bez horní věkové hranice. Šestý ročník se v tomto aspektu **odlišuje od předchozích vln**, v nichž byl individuální dotazník určen dospělým osobám (ve věku 18 let a vyšším).

Minimální velikost dotázaného souboru byla stanovena na 3 100 domácností a 4 000 respondentů individuálního dotazníku.

2.2. Jednotka výběru

Jednotkou výběru je **hospodařící domácnost** definovaná v souladu s metodikou Českého statistického úřadu. Domácnost tvoří jedinci obvykle bydlící v daném bytě či rodinném domku, kteří společně hospodaří, tj. společně hradí základní nezbytné výdaje domácnosti, jako je strava, náklady na bydlení, ostatní provozní výdaje domácnosti (včetně těch osob, které na náklady samy nepřispívají, ale hradí jim je další členové domácnosti, např. děti, osoby bez příjmu).

Pro určení členství v domácnosti je rozhodující obvyklý pobyt, nikoli trvalý pobyt. Studenti dlouhodobě pobývajících mimo rodičovskou domácnost jsou považováni za členy rodičovské domácnosti, pokud nemají žádnou další soukromou adresu, na níž se zdržují častěji než v domácnosti rodičů, a mají finanční vazbu na rodičovskou domácnost.

Předmětem šetření jsou všichni členové vybraných domácností. Za každého člena domácnosti jsou v úvodním dotazníku pro domácnost sepsány základní sociodemografické údaje (pohlaví, měsíc a rok narození, rodinný stav, ekonomické postavení, nejvyšší dosažené vzdělání). Individuální dotazník může vyplnit libovolný člen domácnosti v požadovaném věkovém rozmezí (případně více členů), tedy **neprobíhá náhodný výběr respondentů v rámci domácnosti**. Více o výběru respondentů individuálního dotazníku obsahuje podkapitola 2.5 *Složení dotázaného souboru*.

2.3. Postup výběr původního panelu

Výběrový soubor v šestém ročníku sestával ze dvou částí – domácnosti z původního panelu a doplňující soubor domácností. Domácnosti z původního panelu pocházejí z výběrového souboru, který byl sestaven již pro první ročník výzkumu v roce 2015. Následující část popisuje vznik původního výběrového souboru a práci s panelovými domácnostmi v následujících ročnících.

Výběr souboru pro první ročník

Výběrový soubor domácností pro první vlnu šetření byl sestaven pomocí **dvoustupňového stratifikovaného náhodného výběru**, který byl svým designem srovnatelný s jednostupňovým stratifikovaným skupinkovým výběrem („cluster sampling“). Velikost designového efektu byla dále snížena využitím vysokého počtu menších výběrových bodů. Hrubý výběrový soubor pro první ročník šetření sestával z 1 275 výběrových bodů po 10 adresách, celkem tedy z 12 750 domácností. Tazatelé v prvním ročníku oslovili s nabídkou účasti domácnosti žijící na těchto adresách.

Výběr vzorku pro první vlnu provedly agentury MEDIAN, s. r. o, a STEM/MARK, a. s. Jako opora výběru byl využit aktuální Registr sčítacích obvodů a budov Českého statistického úřadu, který byl pro účely výběru převeden do formy databáze adres (zahrnutí budovy do databáze tolikrát, kolik obsahuje bytů). Podrobnosti k sestavení výběrového souboru pro první ročník výzkumu obsahuje dokumentace k datům z prvního až čtvrtého ročníku. Kompletní dokumentace i data jsou dostupné v Českém sociálněvědním datovém archivu.

Sledování členů panelu ve druhém až pátém ročníku

Ve druhé až páté vlně byly k účasti oslovovány vždy ty domácnosti, které se zapojily do bezprostředně předcházejícího ročníku.

Ve druhé až čtvrté vlně šetření spadaly do výběrového souboru takové domácnosti, v nichž v době sběru dat žila alespoň jedna osoba, která byla členem domácnosti dotázané v bezprostředně předcházejícím ročníku („original sample member“). Pokud tedy domácnost nebyla ve druhém roce dotázána, v dalších letech již její členové k účasti přizvání nebyli a totéž platí pro ukončení účasti v následujících vlnách. Domácnosti s alespoň jedním původním členem byly oslovovány, i pokud se rozdělily na více částí a/nebo se přestěhovaly. Detailní informace k vlnám 1 (2015) až 4 (2018) jsou dostupné v dokumentaci k datům z příslušných vln.

Obdobně i v rámci páté vlny šetření byly součástí výběrového souboru domácnosti, které se účastnily všech čtyř předcházejících šetření (tj. žila v nich aspoň jedna osoba, která byla členem domácnosti dotázané ve čtvrté vlně). K šetření však nebyly přizvány domácnosti, které při předběžném zjišťování ochoty zúčastnit se odběru krve ve čtvrté vlně rozhodně odmítly. Osloveno k účasti bylo 83 % domácností zapojených do čtvrté vlny. V páté vlně byla nabídnuta účast v šetření za podmínky, že alespoň jeden člen zároveň podstoupí odběr krve pro změření vybraných biomarkerů. Tento vyšší nárok na zapojení se do šetření vedl ke sníženému počtu dotázaných domácností oproti předchozím ročníkům (snížení míry retence). Stejně jako v předchozích vlnách šetření byly domácnosti dotazovány, i pokud se od vlny poslední účasti přestěhovaly na jinou adresu.

Sledování členů panelu v ad hoc covid-19 výzkumech

Pro oslovení k účasti v ad hoc výzkumech v letech 2020 a 2021 byla přijata specifická pravidla ovlivněná potřebou rychlého uspořádání těchto sběrů dat v reakci na aktuální stav pandemie nemoci covid-19 a spojením dotazníkových šetření s nábořem účastníků pro populační studie imunity.

K účasti v prvním ad hoc výzkumu byly vyzvány všechny domácnosti, v nichž alespoň jeden člen absolvoval odběr krve v páté vlně šetření v roce 2019 (nebyly tedy osloveny domácnosti, které v páté vlně pouze vyplnily dotazník). Tyto domácnosti byly vybrány, protože byly považovány za dobře vymezenou a dosažitelnou skupinu osob s předpokládanou vyšší ochotou k absolvování populační studie imunity, která zahrnovala odběr krve na jednom z několika testovacích míst v ČR.

V ad hoc šetření v roce 2021 byli k vyplnění dotazníku osloveni respondenti z panelu CHPS, kteří byli dotázáni v ad hoc šetření v roce 2020 a/nebo v páté vlně výzkumu v roce 2019. Dále byly přizvány vybrané domácnosti zapojené ve čtvrté vlně v roce 2018 či třetí vlně v roce 2017, z nichž se však nikdo neúčastnil sběrů v letech 2019 a 2020. Z domácností, které se naposledy zúčastnily v letech 2017 až 2018, byla z kapacitních důvodů oslovena pouze část (78 %) na základě podrobnosti dostupných kontaktních údajů.

Sledování členů panelu v šestém ročníku

K účasti v šestém ročníku šetření byly přizvány domácnosti, které byly **dotázány ve čtvrtém ročníku v roce 2018**, bez ohledu na to, zda se účastnily pátého ročníku šetření (2019) či ad hoc covid sběrů dat (2020 a 2021). Do výběrového souboru šestého ročníku tedy spadaly všechny domácnosti, v nichž v době sběru dat žila alespoň jedna osoba, která byla členem domácnosti dotázané ve čtvrté vlně. Jako referenční ročník byla vybrána čtvrtá a nikoli pátá vlna, protože účast v pátém ročníku byla omezena požadavkem absolvování odběru krevního vzorku. Stejně jako v předchozích letech byly domácnosti oslovovány k účasti, i pokud se rozdělily na více částí a/nebo se přestěhovaly.

2.4. Doplnující soubor domácností

V šestém ročníku šetření byl soubor původních panelových domácností doplněn o soubor nových domácností („refreshment sample“). **Doplnující soubor byl sestaven pomocí kvótního výběru** takovým způsobem, aby se struktura celkového souboru maximálně blížila populační distribuci vybraných charakteristik. Kvótní předpis sestavil realizátor sběru dat (společnost STEM/MARK) ve spolupráci s řešitelským týmem v Sociologickém ústavu AV ČR. Kvótní předpis byl sestaven pro celkový soubor zahrnující jak domácnosti z původního panelu CHPS, tak nové domácnosti.

Kvótní charakteristiky na úrovni domácnosti

Mezi kvótní znaky na úrovni domácnosti patřil typ domácnosti (6 kategorií), nejvyšší dosažené vzdělání členů domácnosti (3 kategorie), kraj (14 kategorií), velikost sídla (5 kategorií) a typ budovy (2 kategorie). Během sběru dat byly průběžně sledovány distribuce kvótních znaků a domácnosti

byly v jednotlivých fázích sběru rekrutovány tak, aby se v maximální míře podařilo dosáhnout požadované struktury celkového souboru.

Složení domácnosti a nejvyšší dosažené vzdělání členů byly stanoveny jako volné kvóty, tedy byly při sběru dat naplňovány nezávisle na ostatních kvótních charakteristikách. Typ domu a velikost sídla byly zkříženy s krajem a byly tedy naplňovány v rámci jednotlivých krajů samostatně.

Jako zdroj kvótního předpisu byla využita následující data:

- **složení domácnosti dle věku a vztahů mezi členy:** Výběrové šetření pracovních sil 2020 (vážená data za druhé čtvrtletí), Český statistický úřad,
- **nejvyšší vzdělání členů domácnosti:** Výběrové šetření pracovních sil 2020 (vážená data za druhé čtvrtletí), Český statistický úřad,
- **kraj a velikost sídla:** populační bilance obyvatelstva ČR za rok 2022 (počáteční stav k 1. 1. 2023), Český statistický úřad,
- **typ budovy:** Sčítání lidu, domů a bytů 2021 (Hospodařící domácnosti podle velikosti a typu domácnosti), Český statistický úřad.

Hodnoty jednotlivých kvótních proměnných jsou zobrazeny v *Tabulce 1*. Složení domácnosti na základě počtu členů domácnosti, jejich věků a vztahů bylo rozlišováno následujícím způsobem:

- **seniorské domácnosti jednočlenné:** domácnost má pouze jediného člena, kterým je osoba ve věku 65 a více let;
- **seniorské domácnosti vícečlenné:** domácnost má více než jednoho člena, všichni členové jsou ve věku 65 a více let;
- **partnerská domácnost s dítětem do 17 let:** v domácnosti se nachází partnerský nebo manželský pár, alespoň jeden člen domácnosti je dítě do 17 let věku;
- **domácnost jednoho rodiče s dítětem do 17 let:** v domácnosti není partnerský či manželský pár, alespoň jeden člen domácnosti je dítě do 17 let věku;
- **domácnost jednotlivce ne-seniorská:** domácnost má pouze jediného člena, kterým je osoba mladší 65 let;
- **partnerská domácnost bez dítěte:** v domácnosti se nachází partnerský nebo manželský pár, žádný člen domácnosti není mladší 18 let;
- **ostatní:** zbývající domácnosti nezahrnuté do kategorií výše.

Kvótní charakteristiky na úrovni jednotlivců

Při sběru dat byla monitorována rovněž struktura respondentů individuálního dotazníku (pohlaví, věk, nejvyšší dosažené vzdělání) a výběr domácností a jejich členů k dotázání byl řízen tak, aby se složení respondentů individuálního dotazníku z hlediska pohlaví, věkových skupin a nejvyššího dosaženého vzdělání neodchyloval od populačního složení o více než 20 %.

Kvóty na úrovni jednotlivců (pohlaví, věkové kategorie, vzdělání) měly převážně kontrolní funkci. Nebyly v nich stanoveny cílové počty, ale pouze povolené minimální a maximální hodnoty, a to vzhledem k předpokládanému konečnému počtu respondentů individuálního dotazníku.

Kvótní předpis byl sestaven na základě následujících zdrojů:

- **pohlaví a věk jednotlivců:** populační balance obyvatelstva ČR za rok 2022 (koncový stav k 31. 12. 2022), Český statistický úřad,
- **vzdělání jednotlivců:** Výběrové šetření pracovních sil 2020 (vážená data za druhé čtvrtletí), Český statistický úřad.

Některé kvótní proměnné na úrovni domácnosti byly logicky provázány s kvótami na úrovni jednotlivců. Jednalo se o kvótu na nejvyšší dosažené vzdělání v domácnosti a vzdělání jednatelce a dále složení domácnosti a věk jednotlivce. Naplňování vybraných kategorií díky tomuto vztahu vytvářelo omezení pro druhou kvótní proměnnou. Například pokud by byly v domácnostech s mladšími členy dotazovány často osoby ve věku 65 a více let, nebylo by možné následně dodržet cílové počty seniorských domácností, aniž by byl přesažen maximální počet respondentů ve věku 65 a více let. Po celou dobu sběru dat byl proto monitorován souběžně stav domácnostních a individuálních kvót, aby bylo možné usměrňovat výběr domácností a respondentů individuálního dotazníku pro dodržení celkového systému kvót.

Postup sběru

Během sběru dat na doplňujícím souboru domácností byly v jednotlivých fázích postupně určovány požadované skupiny pro dotázání na základě dosavadního naplňování kvótního předpisu. Kraj a velikost sídla byly omezeny až v posledních dvou až třech týdnech sběru, kdy bylo možné dotazovat jen domácnosti z vybraných krajů a velikostí sídel, které dosud ve výběrovém souboru chyběly oproti kvótnímu předpisu.

Sběr na doplňujícím souboru se nejprve soustředil pouze na ne-seniorské domácnosti (tedy domácnosti, v nichž je alespoň jeden člen mladší 65 let), protože domácnosti seniorů byly v souboru z původního panelu CHPS nadreprezentovány. Dotazování seniorských domácností (všichni členové jsou ve věku 65 let či vyšším) bylo v doplňujícím souboru zahájeno později. Podobně bylo v první fázi omezeno dotazování domácností s vysokoškolským vzděláním, které opět byly nadproporčně zastoupeny v dotázaném souboru z původního panelu. K této skupině domácností se sběr rovněž později vrátil po rozvolnění kvótních požadavků. Domácnosti, které byly nejprve odmítnuty ve fázi rekrutace, protože neodpovídaly aktuálním výběrovým kritériím, ale později odpovídaly vyhledávaným skupinám, byly znovu osloveny s nabídkou účasti.

Zdroje výběru nových domácností

Pro nábor nových domácností byly použity tři základní postupy:

- oslovení tazatelů, kteří pro realizátora sběru pracují na osobním dotazování (dále „tazatel“) nebo na telefonickém dotazování (dále „operátor“),
- oslovení prostřednictvím členů komerčního online panelu realizátora sběru dat,
- přímé oslovení členů komerčních online panelů realizátora sběru dat.

V případě oslovování tazatelem/operátorem se jedná o běžný postup využívaný při kvótním výběru v sociálněvědních výběrových šetřeních. Při vyhledávání potenciálních účastníků členy komerčního online panelu přebírají rekrutační roli tazatelů členové panelu, zároveň se částečně jedná o metodu sněhové koule („snowball“, tj. doporučování dalších respondentů účastníky výzkumu), protože někteří z online panelistů se mohli sami do výzkumu zapojit. Nakonec u posledního postupu se jedná o cílený výběr vhodných účastníků z registrovaných členů online panelů, které vznikly převážně na základě dobrovolnosti (tzv. „volunteer“ nebo „opt-in“ panely).

Preferovány byly rekrutační kanály mimo přímé oslovování členů komerčních online panelů. K jejich intenzivnějšímu využití bylo přistoupeno v posledním měsíci sběru dat po vyčerpání předchozích možností. Zároveň byl při sběru na doplňujícím souboru domácností stanoven limit na počet respondentů individuálního dotazníku, kteří jsou členové online panelů. Využity byly dva online panely: Český národní panel a Dialog.

V případě přímé rekrutace z online komerčních panelů však nemusel být sám respondent členem jejich členem. Touto otázkou se podrobněji zabývá podkapitola 3.5 *Sběr dat na doplňujícím souboru domácností*.

Tabulka 1: Způsoby rekrutace domácností z doplňujícího souboru

	Domácností	% domácností	Individuálních dotazníků	% individuálních dotazníků
Oslovení tazatelem	472	31,7	523	30,8
Oslovení operátorem	18	1,2	21	1,2
Doporučení členem panelu Dialog	130	8,7	158	9,3
Přímá rekrutace z panelu Dialog	81	5,4	91	5,4
Přímá rekrutace z panelu ČNP	789	53,0	904	53,3
Celkem	1 490	100,0	1 697	100,0

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Využití komerčních online panelů

Český národní panel (ČNP) je společným projektem výzkumných agentur Nielsen Admosphere, NMS Market Research a STEM/MARK. Vznikl v roce 2012 spojením dříve samostatných online panelů těchto tří agentur. Do Českého národního panelu se registrují zájemci o účast na výzkumech trhu a veřejného mínění za finanční odměnu. Při výzkumech na Českém národním panelu jsou oslovováni pouze stávající registrovaní členové panelu. Dotazníky jsou administrovány v uživatelském rozhraní po přihlášení do webové aplikace, jsou tak dostupné pouze stávajícím členům a nemohou být šířeny odkazem po internetu.

Noví členové jsou rekrutováni za pomoci inzerce na webu a sociálních sítích. Současní členové mohou také doporučit další zájemce a získat za úspěšnou rekrutaci nového člena odměnu. Členové panelu jsou rovněž rekrutováni mimo internet, a to během face-to-face a telefonických sběrů dat výzkumných agentur spravujících ČNP. Dle oficiálních údajů zveřejněných ČNP bylo přibližně 30 % registrovaných členů Českého národního panelu rekrutováno mimo internet (telefonicky či face-to-face). Pro nábor nových členů nejsou používány externí databáze. Respondenti jsou za vyplňování dotazníků finančně odměňováni. Finanční odměnu mohou ale také poukázat na charitativní účely.

Do panelu se mohou registrovat lidé starší 15 let trvale žijící na území České republiky. Při registraci je vyžadována unikátní e-mailová adresa a telefonní číslo a vyplnění základních profilových informací. Po vyplnění online registračního formuláře procházejí zájemci ověřovacím telefonním hovorem, při němž potvrzují dříve zadané údaje a také je automatizovaně kontrolováno pohlaví a věk osoby (pomocí hlasové stopy). Pokud je registrace úspěšná, vyplňují noví členové další profilové informace, které jsou využívány k oslovování subpopulací. K vyplnění co nejvíce profilových informací jsou respondenti motivováni finanční odměnou. Profilové údaje je nutné aktualizovat minimálně jednou za devět měsíců.

V roce 2024 měl Český národní panel přibližně 50 000 členů, z toho přibližně 32 000 aktivních členů (pravidelně vyplňují dotazníky). Ze členů panelu lze sestavit kvótně reprezentativní soubor pro českou internetovou populaci.

Respondenti mohou vyplňovat maximálně tři dotazníky týdně a 48 dotazníků ročně. Český národní panel při sběru dat dodržuje kvalitativní standardy ICC/ESOMAR a kvalitativní standardy pro online kvantitativní výzkumy české asociace SIMAR.

Panel Dialog je od roku 2014 nástroj společností STEM/MARK a ČSOB, který slouží k dotazování reprezentativní bankovní populace ČR, ale je využíván také pro jiné výzkumy. Členové panelu jsou oslovováni pro CAPI, CATI a CAWI sběry. Pro rekrutaci, registraci a správu účastníků panelu platí obdobné principy jako u Českého národního panelu.

2.5. Složení dotázaného souboru

Charakteristiky výběrového souboru byly určeny primárně složením souboru domácností z původního panelu CHPS. Tyto domácnosti byly dotazovány v první fázi sběru dat, v níž bylo cílem maximální vytěžení původního panelu. Při plánování sběru dat nebylo předpokládáno, že účast domácností z původního panelu CHPS bude regulována za účelem dosažení určité požadované struktury dotázaného souboru. Dotazování domácností z původního panelu CHPS pak také reálně omezováno nebylo, s výjimkou regulace věkové struktury individuálních respondentů v závěru této fáze sběru dat (podrobnosti jsou obsaženy níže).

Při výběru nových domácností tak byla **východiskem struktura dotázaných domácností z původního panelu**, jejíž specifické tendence bylo třeba korigovat na úrovni domácností či respondentů individuálního dotazníku v doplňujícím souboru.

Věková struktura domácností

Při plánování doplňujícího výběru bylo především nutné počítat se skutečností, že původní panel CHPS od svého založení zestárl. Výběrový soubor byl sestaven v prvním pololetí roku 2015, tedy v roce 2023 byl soubor v průměru o osm let starší než populace.

Na úrovni domácnosti byla věková struktura regulována ve zjednodušené podobě pomocí zastoupení domácností, v nichž všichni členové dosáhli alespoň věku 65 let (tzv. seniorské domácnosti). Zastoupení vícečlenných seniorských domácností bylo o několik procentních bodů vyšší ve srovnání s populací již ve čtvrtém ročníku v roce 2018 a pro šestou vlnu o pět let později bylo nutné předpokládat další nárůst.

Ve skupině domácností z původního panelu je o přibližně šest procentních bodů vyšší zastoupení vícečlenných seniorských domácností než v populaci, zastoupení jednočlenných seniorských domácností je v souladu s populační hodnotou. Výběr doplňujícího souboru domácností se orientoval primárně na jiné než seniorské typy domácností s ohledem na potřebu dotazovat ve druhé fázi sběru především mladší osoby (podrobnější vysvětlení je uvedeno v další sekci Soubor respondentů dotazníku). Ve výsledném souboru domácností jsou **domácnosti seniorů přibližně o šest procentních bodů podhodnoceny** z důvodu nižšího zastoupení jednočlenných domácností v tomto věku.

Velikost domácností

Soubor panelových domácností obsahoval ve všech předchozích ročnících **nižší podíl jednočlenných domácností** ve srovnání s distribucí dle Výběrového šetření pracovních sil. Možné důvody zahrnují nižší šanci zastihnout jednočlenné domácnosti, menší finanční motivace k účasti u menších domácností (finanční odměna se ve všech vlnách odvíjela od počtu členů) či odlišná práce tazatelů/respondentů s definicí hospodařící domácnosti (vnímání více samostatných hospodařících domácností jako jedné domácnosti).

Při sběru dat šestého ročníku byla respektována tato výchozí odchylka v zastoupení jednočlenných domácností přítomná ve všech předchozích letech. Cílové zastoupení jednočlenných domácností bylo proto na začátku sběru stanoveno jako hodnota ležící mezi populační hodnotou (dle Výběrového šetření pracovních sil) a proporcí jednočlenných domácností ve čtvrtém ročníku CHPS, a to na 28 % (dle Výběrového šetření pracovních sil 2022 tvořily jednočlenné domácnosti 36 % všech domácností).

Při sběru dat se počty jednočlenných domácností v obou kategoriích (jednotlivci ve věku 65 a více let, mladší jednotlivci) naplňovaly nejpomaleji a dodavatel sběru dat se setkával s obtížemi tyto domácnosti získat (možnosti byly navíc v této fázi sběru dat omezeny tím, že vyhledávány musely být zejména domácnosti se základním či středoškolským vzděláním bez maturity). Cílové počty jednočlenných domácností byly během sběru proto sníženy. Přihlédnuto bylo rovněž ke skutečnosti, že předností panelového šetření domácností je dotazování více členů jedné domácnosti a vyšší podíl domácností s více členy dává větší potenciál dotazování většího počtu členů na domácnost (ten nebyl nicméně v aktuálním sběru využit, protože v doplňujícím souboru domácností byl dotazován převážně jeden respondent). Ve výsledném souboru jsou **jednočlenné domácnosti podhodnoceny**

o přibližně 13 procentních bodů ve srovnání s daty Výběrového šetření pracovních sil. Více (o přibližně osm procentních bodů) jsou naopak zastoupeny domácnosti partnerů/manželů s dětmi.

Vzdělání domácností

Další výchozí charakteristikou domácností z původního panelu CHPS byla **vyšší úroveň vzdělání členů** ve srovnání s populací. Soubor domácností, který se zúčastnil čtvrté vlny v roce 2018, zahrnoval vyšší podíl domácností s vysokoškolským vzděláním a nižší se středním vzděláním bez maturity či základním vzděláním. Tato tendence se u původních panelových domácností v šestém ročníku dále prohloubila.

Cílová distribuce této kvóty byla na počátku sběru dat stanovena na základě populační hodnoty (dle Výběrového šetření pracovních sil), avšak během sběru dat byla **tato kvóta rozvolněna** vzhledem k velmi komplikovanému získávání domácností s nižším vzděláním (členové dosahují nejvýše středního vzdělání bez maturity nebo základního). Místo cílových počtů pro jednotlivé skupiny byl stanoven minimální požadovaný počet domácností, kde nejvyšší dosažené vzdělání členů je základní nebo střední bez maturity (a naopak maximální počet pro další skupiny).

Výsledné zastoupení domácností, jejichž členové mají nejvýše základní či středoškolské vzdělání bez maturity, je o přibližně deset procentních bodů nižší než v populaci. Naopak zastoupení domácností se středoškolským vzděláním s maturitou je o čtyři procentní body vyšší a podíl domácností s vysokoškolským vzděláním je o šest procentních bodů vyšší než v populaci.

Typ budovy, kraj, velikost sídla

Z hlediska typu domu (rodinný či bytový/jiný) a kraje nebyly v souboru dotázaném ve čtvrtém ročníku CHPS **pozorovány zásadnější odchylky od populační distribuce**. Rozdíl větší než dva procentní body od proporcí v populaci se vyskytl pouze u tří krajů – Hlavní město Praha (nižší zastoupení o přibližně tři procentní body), Středočeský kraj (nižší zastoupení o čtyři procentní body) a Jihočeský kraj (vyšší o čtyři procentní body).

V dotázaném souboru původních panelových domácností se v šestém ročníku mírně zvětšila odchylka u typu domu, ale dosažená distribuce se liší od populační pouze o přibližně jeden procentní bod. U krajů se odchylky celkově zvýšily, ale větší rozdíl než dva procentní body byl zaznamenán jen u pěti krajů (Středočeský, Jihočeský, Ústecký, Královéhradecký, Jihomoravský). Tyto výchozí rozdíly se při sběru dat na doplňujícím souboru domácností promítly do nutnosti dříve uzavřít dotazování v některých krajích, ale nevznikla potřeba speciálních postupů pro vyrovnání distribuce původního souboru a nebyly zaznamenány žádné problémy s naplňováním kvót za kraje.

V souboru domácností ze čtvrtého ročníku CHPS byly o několik procentních bodů **podhodnoceny malé obce do 5 000 obyvatel** a také největší města s 100 000 a více obyvateli. V dotázaném souboru původních panelových domácností se dále o tři procentní body snížilo zastoupení obcí do 5 000 obyvatel, naopak proporce největších měst se vyrovnala s populační. Při rekrutaci nových domácností tak bylo třeba dbát na vyšší zastoupení nejmenších obcí. Dvě nejmenší velikostní kategorie sídel trvalo nejdéle naplnit, ale žádné závažnější problémy při získávání domácností z těchto lokalit se neobjevily.

V níže uvedené tabulce je zobrazeno srovnání populačních distribucí kvótních charakteristik na úrovni domácností s distribucemi v dotázaném souboru, a to za původní domácnosti, doplňující soubor a celý soubor. Rozdíl od populační distribuce je vypočten pouze pro celý soubor. Populační data jsou čerpána ze zdrojů popsaných v části 2.4 *Doplňující soubor domácností*. Zatímco pro sestavení kvót bylo využito Výběrové šetření pracovních sil z roku 2020, pro srovnání s výsledným souborem byla využita novější dostupná data, a to z roku 2022.

Tabulka 2: Distribuce kvótních znaků a srovnání s populací – domácnosti

	% populace	% CHPS původní	% CHPS doplňující	% CHPS celý soubor	Rozdíl p. b. (celý)
Typ (struktura) domácnosti					
Seniorské domácnosti jednočlenné	17,1	17,0	3,7	11,0	-6,2
Seniorské domácnosti vícečlenné	10,2	15,9	3,9	10,5	0,3
Partnerská domácnost s dítětem do 17 let	20,3	23,5	34,6	28,5	8,2
Domácnost jednoho rodiče s dítětem do 17 let	5,3	3,5	6,5	4,9	-0,5
Domácnost jednotlivce ne- seniorská	18,4	9,3	14,1	11,5	-7,0
Partnerská domácnost bez dítěte	23,4	23,1	30,1	26,3	2,9
Ostatní	5,2	7,7	7,1	7,4	2,2
Nejvyšší vzdělání v domácnosti					
Střední bez maturity a nižší	33,8	19,6	29,7	24,2	-9,6
Střední s maturitou	38,0	40,7	43,0	41,7	3,7
Vysokoškolské	28,2	39,7	27,3	34,1	5,9
Typ domu					
Rodinný	44,2	45,4	46,5	45,9	1,7
Bytový a ostatní	55,8	54,6	53,5	54,1	-1,7
Kraj					
Hlavní město Praha	13,7	12,2	15,3	13,6	-0,1
Středočeský kraj	12,7	7,7	18,3	12,5	-0,2
Jihočeský kraj	5,9	9,1	3,0	6,3	0,4
Plzeňský kraj	5,5	4,9	6,4	5,6	0,1
Karlovarský kraj	2,8	1,8	3,0	2,3	-0,5

	% populace	% CHPS původní	% CHPS doplňující	% CHPS celý soubor	Rozdíl p. b. (celý)
Ústecký kraj	7,8	10,6	6,0	8,5	0,7
Liberecký kraj	4,2	3,7	5,1	4,3	0,1
Královéhradecký kraj	5,1	7,2	4,2	5,8	0,7
Pardubický kraj	4,6	3,8	5,4	4,5	-0,1
Kraj Vysočina	4,4	3,9	5,1	4,4	0,0
Jihomoravský kraj	11	13,7	7,4	10,9	-0,1
Olomoucký kraj	5,8	5,9	5,6	5,8	0,0
Zlínský kraj	5,1	4,3	4,7	4,5	-0,6
Moravskoslezský kraj	11,3	11,2	10,5	10,9	-0,4
Velikost sídla					
Do 999 obyvatel	16,3	13,0	17,4	15,0	-1,3
1 000 až 4 999 obyvatel	22,3	17,7	25,9	21,4	-0,9
5 000 až 19 999 obyvatel	17,8	20,3	14,8	17,8	0,0
20 000 až 99 999 obyvatel	20,1	26,0	16,7	21,8	1,7
100 000 obyvatel a více	23,5	23,1	25,2	24,1	0,6

Zdroj: Výběrové šetření pracovních sil 2022 (Český statistický úřad), roční vážená data; Sčítání lidu domů a bytů 2021 (Český statistický úřad); Počet obyvatel v obcích k 1. 1. 2023 (Český statistický úřad), Věkové složení obyvatel k 1. 7. 2023 (Český statistický úřad); České panelové šetření domácností, 6. ročník (2023/2024), 3 297 domácností (z toho 1 807 z původního panelu a 1 490 nových domácností), vlastní výpočty

Soubor respondentů individuálního dotazníku

V předchozích ročnících výzkumu CHPS se složení respondentů individuálního dotazníku více odchylovalo od populační distribuce než v případě složení všech členů domácnosti (tedy všechny osoby uvedené v přehledu členů domácnosti). Mezi respondenty byly více zastoupeny ženy, členové s vysokoškolským vzděláním a starší osoby. Tyto odchylky jsou obecně v souladu s odlišnou pravděpodobností různých skupin obyvatel se účastnit výběrových šetření.

V šesté vlně v roce 2023 bylo navíc nutné počítat se zestárnutím všech členů domácností, tedy s dalším posunem věkové distribuce respondentů individuálního dotazníku. Mezi vlnami byla také zaznamenána větší tendence k opakované účasti u vzdělanějších domácností, tedy pro šestou vlnu bylo nutné očekávat větší zastoupení respondentů s vyšším vzděláním (středoškolské s maturitou a vysokoškolské).

Pro sběr dat nebyly určeny cílové počty v individuálních kvótách, ale minimální a maximální povolené počty. **Při sběru dat na původních domácnostech nebyl výběr respondentů omezován** vzhledem ke stavu naplňování kvót. Pouze od začátku listopadu neměli tazatelé dotazovat osoby ve věku 65 a více let v domácnostech, kde se nacházel alespoň jeden mladší dospělý, a v domácnostech pouze s osobami ve věku 65 a více let (tzv. seniorské domácnosti) měl být dotázán pouze jeden

respondent. Po celou dobu sběru dat na původních domácnostech byli také tazatelé vyzýváni, aby co nejvíce dotazovali mladší osoby v domácnosti.

V předchozích ročnících CHPS **převažovaly mezi respondenty ženy nad muži**. Ve čtvrtém ročníku CHPS byly ženy mezi respondenty individuálního dotazníku zastoupeny o deset procentních bodů více než odpovídalo populaci. V šesté vlně se podařilo dotázat více mužů, odchylka pohlaví od populační distribuce přesto činila sedm procentních bodů. V doplňujícím souboru domácností byly proto k vyplnění individuálního dotazníku vyzýváni častěji muži než ženy, aby bylo vyváжено vyšší zastoupení žen mezi respondenty ve skupině původních panelových domácností. Odchylka výsledného souboru respondentů z hlediska pohlaví představuje přibližně jeden procentní bod.

Soubor respondentů dotázaných ve čtvrtém ročníku se vyznačoval **vyšším zastoupením osob ve věku 50 a více let** (o deset procentních bodů více než v populaci) a naopak podhodnocením osob do 34 let věku. Tyto odchylky dále narostly v šesté vlně, v níž je podíl respondentů individuálního dotazníku ve věku 50 a více let o 13 procentních bodů vyšší než v populaci a naopak podíl nejmladších osob do 34 let o 12 procentních bodů nižší. V doplňujícím souboru domácností se tak dotazování zaměřilo primárně na osoby do 49 let věku. Věková struktura respondentů byla také regulována už v závěru dotazování původního panelu domácností (viz výše). V celkovém souboru dosahují odchylky věkových skupin od populace maximálně 2,3 procentního bodu.

Rovněž v případě vzdělání se soubor respondentů individuálního dotazníku v minulosti lišil od populace. Byli v něm **více zastoupeni lidé se středním vzděláním s maturitou a vysokoškolským vzděláním**. V šestém ročníku tyto odlišnosti v podsouboru původních domácností narostly. Mezi respondenty šestého ročníku z původního panelu bylo o pět procentních bodů více osob se středním vzděláním s maturitou než v populaci a o osm procentních bodů více vysokoškoláků. Domácnosti pro doplňující soubor byly proto vybírány tak, aby v nich bylo možné dotázat především respondenty se základním vzděláním a středním vzděláním bez maturity. V celkovém souboru respondentů individuálního dotazníku se odchylky od populace pohybují do 3,7 procentního bodu. Méně zůstali zastoupeni respondenti se základním a středoškolským vzděláním s maturitou.

Při analýzách jednotlivých částí souboru (původní panelové domácnosti a doplňující domácnosti) a zejména při jejich srovnávání je nutné zohlednit jejich odlišnou sociodemografickou strukturu. **V doplňujícím souboru je vyšší podíl mužů a tato skupina respondentů je mladší a méně vzdělaná ve srovnání s respondenty z původních panelových domácností.**

V níže uvedené tabulce je zobrazeno srovnání populačních distribucí kvótních charakteristik na úrovni jednotlivců (respondentů individuálního dotazníku) s distribucemi v dotázaném souboru, a to za původní domácnosti, doplňující soubor a celý soubor. Rozdíl od populační distribuce je vypočten pouze pro celý soubor.

Populační data jsou čerpána ze zdrojů popsaných v podkapitole 2.4 *Doplňující soubor domácností*. Zatímco pro sestavení kvót bylo využito Výběrové šetření pracovních sil z roku 2020, pro srovnání s výsledným souborem byla využita novější dostupná data, a to z roku 2022.

Tabulka 3: Distribuce kvótních znaků a srovnání s populací – jednotlivci

	% populace	% CHPS původní	% CHPS doplňující	% CHPS celý soubor	Rozdíl p. b. (celý)
Pohlaví					
Muž	48,6	41,3	57,5	47,4	-1,2
Žena	51,5	58,7	42,5	52,6	1,2
Věk					
15 až 34 let	25,8	13,9	40,2	23,9	-1,9
35 až 49 let	27,6	26,4	35,6	29,9	2,3
50 až 64 let	22,3	25,4	15,4	21,6	-0,7
65 a více let	24,3	34,3	8,7	24,6	0,3
Vzdělání					
Základní	10,2	8,3	17,2	11,7	-1,8
Střední bez maturity	33,6	24,2	36,3	28,8	-3,6
Střední s maturitou	35,0	38,9	35,1	37,4	3,7
Vysokoškolské	21,2	28,6	11,4	22,1	1,7

Zdroj: Výběrové šetření pracovních sil 2022 (Český statistický úřad), roční vážená data; České panelové šetření domácností, 6. ročník (2023/2024), 4 459 respondentů individuálního dotazníku (z toho 2 762 z původního panelu a 1 697 z nových domácností), vlastní výpočty

2.6. Výběr respondentů individuálního dotazníku

Domácnosti z původního panelu

V domácnostech z původního panelu CHPS se výběr respondentů individuálního dotazníku řídil stejnými pravidly jako v předchozích ročnících. Vyplnění individuálního dotazníku bylo nabízeno **všem členům v cílovém věku** (v šestém ročníku 15 a více let, v předchozích ročnících 18 a více let). O účasti se rozhodovali sami členové domácnosti, výběr konkrétních osob nebyl prováděn.

Individuální dotazník mohlo v původním panelu vyplnit více osob z jedné domácnosti. Stejně jako v předchozích ročnících bylo i v šesté vlně cílem maximalizovat počet individuálních dotazníků na domácnost. Kritériem pro účast byl pouze věk respondenta, výběr nebyl dalším způsobem omezován. Pouze v poslední fázi sběru dat na původních domácnostech byla přijata opatření pro zvýšení podílu mladších respondentů individuálního dotazníku (podrobnosti jsou uvedeny v předchozí podkapitole).

Domácnosti z doplňujícího souboru

U nových domácností dotázaných pomocí osobního rozhovoru byl postup stejný jako u domácností z původního panelu (účastnit se mohl libovolný člen domácnosti a více členů za domácnost). U

domácností rekrutovaných pro online dotazování byl již **výběr respondentů korigován tak, aby se co nejvíce dosáhlo požadované struktury celkového souboru** respondentů individuálního dotazníku. Při rekrutaci domácností byl tak pomocí automatického algoritmu vybrán pro vyplnění individuálního dotazníku jeden člen, který nejvíce odpovídal aktuálně vyhledávaným skupinám (s výjimkou domácností dotázaných přibližně v prvních dvou týdnech této fáze sběru, kde nebyl počet respondentů individuálního dotazníku omezován, a předem určeného počtu domácností z nadvýběru v rámci doplňujícího souboru, v nichž byli k dotazování vybíráni dva členové).

V nových domácnostech dotazovaných online bylo rovněž zpočátku umožněno vyplnění individuálního dotazníku všem zájemcům ve věku 15 a více let. Z důvodu nižší kontroly nad dotazovací situací a více omezení kladených na výběr respondentů individuálního dotazníku kvůli dodržení kvót mohl v převážné části dotázaných domácností vyplnit individuální dotazník jen jeden člen.

Stále však platilo, že každá domácnost se může zúčastnit výzkumu pouze jednou. Pokud tak bylo u těchto nových domácností dotazovaných pomocí internetu více zájemců o dotazování, nemohl se další zájemce zúčastnit novým vyplněním dotazníku pro domácnost a posléze individuálního dotazníku. Při kontrole dat byla proto speciální pozornost věnována detekci duplicitních domácností. Případné duplicity byly vyloučeny (vyplnění individuálního dotazníku stejnou osobu) či sloučeny do jedné domácnosti (vyplnění individuálního dotazníku více členy domácnosti). Tyto kontroly jsou podrobněji popsány v podkapitole 6.2 *Kontroly a čištění průběžných dat*.

3. Sběr dat

3.1. Příprava tazatelů

Výběr tazatelů pro projekt

Pro práci na projektu byli vybíráni přednostně tazatelé, kteří pracovali na předchozích ročnících Českého panelového šetření domácností. Dále byli vybíráni tazatelé, kteří na projektu v minulosti nepracovali, ale měli značné zkušenosti s dotazováním. Důležitým kritériem pro výběr tazatelů k práci na projektu byla také ochota dojíždět, protože tazatelé měli za úkol oslovit geograficky rozptýlený soubor.

Tazatelé byli oslovováni ke spolupráci na projektu během května a června 2023 prostřednictvím oblastních zástupců realizátora sběru dat.

Školení tazatelů

Pro přípravu tazatelů proběhlo celkem 5 školení, z toho dvě prezenční (Praha, Brno) a tři online školení pro ty tazatele, kteří neměli možnost absolvovat prezenční školení.

Tabulka 4: Přehled jednotlivých školení tazatelů s počty účastníků

Termín školení	Místo (obec/online)	Počet hodin školení	Počet tazatelů: STEM/MARK	Počet tazatelů: MEDIAN
12. 8. 2023	Praha	5	29	8
13. 8. 2023	Brno	5	25	4
23. 8. 2023	online	3,5	17	5
24. 8. 2023	online	3,5	24	5
7. 9. 2023	online	3,5	11	9
Celkem účastníků školení			106	31

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Školením prošlo a bylo řádně proškoleny celkem 137 tazatelů. Za agenturu STEM/MARK bylo proškoleny 106 tazatelů a 31 tazatelů za agenturu MEDIAN.

Jednotlivých školení se účastnil projektový manažer, vedoucí terénního oddělení a jeho zástupce, terénní pracovník zodpovědný za sběr dat v projektu a pracovník pro technickou podporu tazatelů. Obvykle byli přítomni také zástupci Sociologického ústavu AV ČR.

Školení sestávalo ze dvou hlavních částí. V první části seznamovali školitelé tazatele dle připravené prezentace s klíčovými aspekty výzkumu, obsahem dotazníku a postupy sběru dat. Zahrnuta byla následující témata:

- Základní informace o výzkumu (zadavatel, témata)
- Terénní materiály (informační list, informovaný souhlas a smlouva o účasti v šetření, karty k dotazníkům, avizní dopis a leták)
- Metodika sběru dat (pravidla kontaktování domácností, evidence kontaktů)
- Postup dotazování v domácnosti (administrace jednotlivých dotazovacích nástrojů)
- Oslovování a motivování domácností
- Odměny tazatelů
- Obsah dotazníků a klíčové metodické body
- Práce s dotazovacím rozhraním (seznam domácností, vyplňování kontaktního formuláře)

Ve druhé části procházeli tazatelé s pomocí školitelů dotazníkem pro domácnost a individuálním dotazníkem. Školitelé v této části komentovali klíčové metodické a technické prvky dotazníku a demonstrovali správné postupy práce s dotazovací aplikací a dotazníkem. Pro školení byla pro každého tazatele připravena sada testovacích domácností z panelového souboru. Tazatelé museli na školení provést alespoň jeden zkušební průchod dotazníkem.

Tazatelé rovněž obdrželi během školení potřebné terénní materiály (leták k výzkumu, dokument s informacemi o účasti, formulář pro udělení souhlasu s účastí a zpracováním osobních údajů, smlouvu o účasti ve výzkumu pro vyplacení odměny respondentům).

Závěry k důležitým otázkám ze školení byly tazatelům zpřístupněny na interní online platformě s informacemi a materiály k projektu.

Školení tazatelů sloužila k přípravě na dotazování původních panelových domácností. Na dotazování nových domácností posléze pracovali stejní tazatelé, kteří byli vyškoleni k práci na původním panelu domácností. Samostatné školení k dotazování nových domácností neproběhlo.

3.2. Tazatelé pracující na sběru dat

Počet tazatelů

Na sběr dat původních domácností bylo přiděleno všech 137 tazatelů, kteří byli vyškoleni. Těmto tazatelům byl vydán alespoň jeden výběrový bod (tj. skupina domácností ve stejné lokalitě). Na osobním dotazování domácností z doplňujícího souboru poté pracovali stejní tazatelé, žádný nový tazatel nebyl na tuto fázi sběru vyškolen a přidělen.

Na sběru dat se alespoň jednou úspěšně dotázanou domácností podílelo 114 tazatelů (tedy 83 % tazatelů ze všech vyškolených tazatelů). Z toho 106 tazatelů osobně dotázalo alespoň jednu domácnost (77 % všech vyškolených tazatelů) a u dalších 8 tazatelů byla online dotázána alespoň jedna domácnost, která jim byla původně přidělena pro osobní dotazování (tedy úspěšně proběhla konverze na online dotazování).

Domácností na tazatele

Jeden tazatel v průměru dotázal 16 domácností osobním rozhovorem. Tento průměrný počet domácností na tazatele je podobný jako v předchozích ročnících (v první vlně dotázal jeden tazatel v průměru 15 domácností, ve čtvrté vlně připadalo na jednoho tazatele 18 domácností). **Nejčastěji ovšem jeden tazatel dotázal 6 až 10 domácností** (29 tazatelů, tj. přibližně pětina tazatelů s aspoň jedním osobním rozhovorem). Poměrně časté bylo dotázání jedné až pěti domácností (přibližně šestina tazatelů). Naopak část tazatelů vykázala vysokou produktivitu a dotázala více než 25 domácností (celkem 22 tazatelů, tj. více než pětina). Maximální počet osobně dotázaných domácností na tazatele byl 58 (u domácností z původního panelu představovalo maximum 56).

Tabulka 5: Počty dotázaných domácností na tazatele

	Pouze původní domácnosti		Domácnosti celkem	
	Počet tazatelů	Podíl (%)	Počet tazatelů	Podíl (%)
1 až 5 domácností	18	17,0	18	17,0
6 až 10 domácností	33	31,1	29	27,4
11 až 15 domácností	16	15,1	19	17,9
16 až 25 domácností	17	16,0	17	16,0
26 až 35 domácností	12	11,3	12	11,3
36 až 50 domácností	9	8,5	7	6,6
51 až 58 domácností	1	0,9	4	3,8

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Skladba tazatelů

Následující tabulka ukazuje distribuce základních charakteristik tazatelů, kteří úspěšně dotázali alespoň jednu domácnost. **Většina tazatelů pracujících na projektu měla značné zkušenosti s dotazováním** – 84 % tazatelů pracovalo v době začátku sběru dat pro danou agenturu alespoň pět let a téměř 40 % tazatelů mělo více než 15leté zkušenosti. Žádný z tazatelů neměl kratší než roční praxi.

Většinu tazatelů tvořily ženy (79 %). Více než polovina tazatelů (57 %) měla středoškolské vzdělání s maturitou. Druhou nejčastější skupinou byli tazatelé se základním vzděláním nebo středoškolským bez maturity (36 %). Vysokoškolské vzdělání měla méně než desetina tazatelů (7 %).

Mezi tazatele patří zejména muži a ženy vyšších věkových kategorií. Průměrný věk tazatelů dosahoval 60 let a výrazně se nelišil mezi muži a ženami (průměrný věk mužů-tazatelů byl pouze o čtyři roky nižší než u žen-tazatelek). Přibližně čtyři pětiny tazatelů měly v době sběru dat 55 a více let. Naopak necelé desetině tazatelů (8 %) bylo méně než 45 let.

Tabulka 6: Skladba tazatelů pracujících na projektu

Kategorie		
Délka praxe u agentury (při začátku sběru)	Počet	Podíl (%)
0 až 12 měsíců praxe	0	0,0
13 až 24 měsíců praxe	4	3,5
25 až 36 měsíců praxe	4	3,5
37 až 60 měsíců praxe	9	7,9
61 až 120 měsíců praxe	22	19,3
121 až 180 měsíců praxe	30	26,3
181 až 240 měsíců praxe	32	28,1
Více než 240 měsíců praxe	13	11,4
Pohlaví		
Muž	24	21,1
Žena	90	78,9
Vzdělání		
Základní a střední bez maturity	41	36,0
Střední s maturitou	65	57,0
Vysokoškolské	8	7,0
Věk		
Do 34 let	1	0,9
35 až 44 let	8	7,0
45 až 54 let	15	13,2
55 až 64 let	56	49,1
65 a více let	34	29,8
Průměrný věk	Průměr	
Muž	57	
Žena	61	
Celkem	60,3	

Poznámka: Do tabulky je zahrnuto 114 tazatelů, kteří se na sběru podíleli alespoň jednou dotázanou domácností (ať už osobní dotazování, nebo konverze na CAWI).

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

3.3. Průběh sběru dat

Počty dotázaných domácností a jednotlivců

V šestém ročníku výzkumu bylo **úspěšně dotázáno 3 297 domácností**, z toho 1 807 domácností (54,8 %) z původního panelu domácností a 1 490 (45,2 %) nově vybraných domácností z doplňujícího souboru.

Individuálních dotazníků bylo celkem získáno 4 459, z toho 2 762 individuálních dotazníků pochází z původního panelu domácností (61,9 %) a 1 697 individuálních dotazníků bylo vyplněno v nových domácnostech (38,1 %). Na jednu domácnost z původního panelu tak v průměru připadá 1,53 individuálních dotazníků, zatímco u nových domácností je to v průměru 1,14 individuálních dotazníků. Rozdílný poměr dotázaných domácností a získaných individuálních dotazníků u domácností z původního a doplňujícího souboru je důsledek rozdílného přístupu k administraci individuálních dotazníků. Podrobnosti poskytuje podkapitola 3.7. *Administrace dotazovacích nástrojů*.

Tabulka 7: Počty dotázaných domácností a jednotlivců dle statusu domácnosti a módu dotazování

Status domácnosti	Mód dotazování	Domácností		Individuálních dotazníků	
		Počet	Podíl (%)	Počet	Podíl (%)
Původní	-	1 807	54,8	2 762	61,9
Nové	-	1 490	45,2	1 697	38,1
Původní	CAPI	1 623	89,8	2 418	87,5
	CAWI	184	10,2	344	12,5
Nové	CAPI	88	5,9	125	7,4
	CAWI	1 402	94,1	1 572	92,6
Celkem	CAPI	1 711	51,9	2 543	57,0
	CAWI	1 586	48,1	1 916	43,0
Celkem		3 297	100,0	4 459	100,0

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Mód dotazování

V šestém ročníku byla využita kombinace CAPI a CAWI dotazování. **Pro domácnosti z původního panelu bylo preferováno osobní dotazování (CAPI)** z důvodu vyššího zastoupení domácností starších osob a vyšší návratnosti při oslovování tazatelů. Sběr dat v 6. ročníku rovněž potvrdil předpoklad, že domácnosti jsou vázány na své původní tazatele (domácnosti přidělené stejným tazatelům jako v minulosti byly dotázány s větší úspěšností).

Přibližně 10 % úspěšně dotázaných domácností z původního panelu se v šestém ročníku zúčastnilo online (CAWI). Tento podíl odpovídá prvnímu až čtvrtému ročníku CHPS, v nichž bylo rovněž kombinováno CAPI a CAWI dotazování (v pátém ročníku bylo umožněno pouze CAPI dotazování z důvodu propojení dotazníkového šetření a odběru krevních vzorků). Prostřednictvím internetu se v šestém ročníku zúčastnily především domácnosti, které byly online dotázány ve čtvrtém ročníku (tyto domácnosti nebyly oslovovány tazateli, ale byly přímo přizvány k online dotazování). Téměř dvě třetiny (65 %) domácností, které v šestém ročníku vyplnily dotazník online, byly tímto způsobem dotázány také ve čtvrtém ročníku. Dále byl online dotazník nabízen domácnostem, které odmítly osobní účast. Konverze na CAWI byla úspěšná u 64 domácností (přibližně třetina všech původních domácností, které vyplnily dotazník online v šestém ročníku).

Online dotazník bylo možné vyplnit na všech druzích zařízení (počítače, tablety, chytré telefony).

Tabulka 8: Počty zúčastněných původních domácností dle módu dotazování

	Dotázané CAPI ve 4. vlně		Dotázané CAWI ve 4. vlně	
	Počet	Podíl (%)	Počet	Podíl (%)
Dotázané CAPI v 6. vlně	1 623	100,0	0	0,0
Dotázané CAWI v 6. vlně	64	34,8	120	65,2

Poznámka: Podíly představují řádková procenta.

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Naopak u doplňujícího souboru bylo preferováno online dotazování (CAWI) z důvodu jeho větší udržitelnosti v budoucích letech. Většina domácností z doplňujícího souboru tak byla dotázána pomocí CAWI (94 %).

Na rozdíl od předchozích ročníků byly módy sběru dat v šesté vlně kombinovány i u jedné domácnosti. Ve fázi dokončování prací na původním panelu (listopad 2023) byli oslovováni vybraní členové domácností dotázaných tazateli s nabídkou vyplnění online dotazníku. Cílem tohoto postupu bylo navýšit podíl respondentů individuálních dotazníků ve věku do 50 let. Tímto postupem se podařilo získat 42 individuálních dotazníků.

Harmonogram sběru dat

Sběr dat v šestém ročníku CHPS probíhal **od 23. srpna do 25. února 2024**. Po realizátorovi sběru dat bylo vyžadováno minimálně 19 týdnů sběru dat (tento požadavek se přitom odvíjel od délky sběru dat v minulých vlnách a bral v úvahu potřebu získat ke spolupráci a dotázat nové domácnosti).

Sběr dat se uskutečnil ve **dvou hlavních fázích**. V první fázi byly oslovovány a dotazovány domácnosti z původního panelu, ve druhé fázi byly oslovovány a dotazovány nové domácnosti pro doplňující soubor. Tyto dvě fáze částečně probíhaly souběžně.

Tabulka 9: Hlavní mezníky v harmonogramu sběru dat

Datum	Aktivita
Domácnosti z původního panelu	
20. 8. 2023	Vydání adres původních domácností tazatelům
23. 8. 2023	Začátek osobního dotazování původních domácností
25. 8. 2023	Začátek online dotazování původních domácností
9. 12. 2023	Ukončení osobního dotazování původních domácností
12. 12. 2023	Provedení posledních pokusů o kontakt s původními domácnostmi
3. 1. 2024	Ukončení online dotazování původních domácností
Nové domácnosti	
říjen 2023	Začátek rekrutace nových domácností pro osobní dotazování
24. 11. 2023	Začátek osobního dotazování nových domácností
7. 12. 2023	Ukončení osobního dotazování nových domácností
7. 12. 2023	Začátek rekrutace nových domácností pro online dotazování
12. 12. 2023	Začátek online dotazování nových domácností
25. 2. 2024	Ukončení online dotazování nových domácností

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Většina domácností z původního panelu byla dotázána od konce srpna do konce října 2023. Přibližně 60 % domácností z původního panelu bylo dotázáno v srpnu a září, dalších 30 % v říjnu a zbývající desetina v listopadu. Většina domácností z doplňujícího souboru byla naopak dotázána během ledna a února 2024 (50 % v lednu a 26 % v únoru).

Tabulka 10: Podíl (%) domácností dotázaných v jednotlivých měsících sběru dat

	Původní domácnosti			Nové domácnosti			Celkem
	CAPI	CAWI	Celkem	CAPI	CAWI	Celkem	
Srpen 2023	8,0	3,3	7,5	0,0	0,0	0,0	4,1
Září 2023	52,4	52,7	52,4	0,0	0,0	0,0	28,7
Říjen 2023	31,5	16,3	30,0	0,0	0,0	0,0	16,4
Listopad 2023	7,9	20,1	9,1	80,7	0,0	4,8	7,2
Prosinec 2023	0,1	7,1	0,8	19,3	18,8	18,9	9,0
Leden 2024	0,0	0,5	0,1	0,0	53,1	49,9	22,6
Únor 2024	0,0	0,0	0,0	0,0	28,1	26,4	12,0

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), 3 297 domácností (z toho 1 807 z původního panelu a 1 490 nových domácností), vlastní zpracování

3.4. Sběr dat na původním panelu domácností

Přidělení primárního módu sběru dat

K osobnímu dotázání bylo určeno **2 932 domácností**, které se ve čtvrtém ročníku zúčastnily formou osobního rozhovoru nebo se zapojily do páté vlny, která probíhala výhradně osobním dotazováním. Všechny tyto domácnosti měli oslovit tazatelé a pokusit se s nimi provést osobní rozhovor. Pouze v případě neochoty absolvovat osobní dotazování, přestěhování a dalších bariér byla domácnost převedena na online dotazování.

Zbývajících **256 domácností** ze souboru domácností dotázaných ve čtvrtém ročníku bylo primárně určeno pro online dotazování. Tyto domácnosti nebyly oslovovány tazateli, ale byly jim prostřednictvím e-mailu i pošty zasílány pozvánky k účasti ve výzkumu.

Přidělování adres tazatelům

Před začátkem sběru dat byly mezi tazatele rozděleny všechny adresy domácností, které byly určeny k osobnímu dotázání. Tazatelům byly **vydány všechny jim přidělené adresy v jedné dávce, tedy** všechny adresy byly tazatelům k dispozici od začátku sběru dat. Původním záměrem bylo rozdělit vydávání adres do několika dávek (dvou až čtyř dle počtu adres, které byly tazatelům přidělovány), jako tomu bylo v předchozích ročnících, ale po konzultaci s tazateli v návaznosti na tazatelská školení byl postup upraven (preferenční tazatelů směrem k samostatnější organizaci práce).

Pokud to bylo možné, domácnosti byly **přidělovány tazatelům, kteří tyto domácnosti dotazovali už v minulosti**. V případě, že původní tazatel nebyl k dispozici, byl pro domácnost vybrán tazatel žijící v blízkosti. Stejný tazatel jako v poslední vlně účasti (čtvrtý či pátý ročník) byl přidělen u 33 % domácností určených pro osobní dotazování. Zachování stejného tazatele mezi vlnami působilo jako významný faktor retence. Zatímco u domácností, u nichž došlo k výměně tazatele, byla dosažena míra retence 56,8 %, u domácností se stejným tazatelem to bylo 73,4 %. Tento rozdíl se projevuje v rámci agentur provádějících sběr dat (STEM/MARK, MEDIAN), krajů nebo vzdělanostních stupňů domácností.

Avizování výzkumu

Domácnostem byl několik dní před začátkem sběru dat **odeslán avizní dopis a avizní e-mail**. Emailovou informací o nadcházejícím sběru dat byly obeslány všechny domácnosti, na které byl z minulých ročníků k dispozici e-mail. Dopisní verze byla zaslána domácnostem, které měly být osloveny k provedení osobního rozhovoru (CAPI domácnosti). Pokud se e-mailová adresa ukázala nefunkční a rovněž se nepodařilo doručit fyzický avizní dopis, byla dané domácnosti odeslána informační SMS s prosbou o aktualizaci poštovní adresy a/nebo e-mailu pro další komunikaci.

Avizní dopis obsahoval poděkování domácnosti za účast v minulých vlnách, informaci o provádění nového ročníku a prosbu o spolupráci na něm. Dále dopis uváděl, že v nadcházejících dnech osloví domácnost tazatel k provedení rozhovoru, a zahrnoval informaci o délce dotazování, peněžní odměně za účast a možnosti účasti pro více členů domácnosti. Obsahoval rovněž kontaktní údaje na

realizátora sběru a odkaz na web výzkumu. Avizní dopisy jsou spolu s dalšími terénními materiály zveřejněny jako součást dokumentace k datovému souboru.

Před startem sběru dat vydal Sociologický ústav AV ČR tiskovou zprávu k zahájení nového ročníku šetření.

Tazatelé byli během školení seznámeni s obsahem avizního dopisu a měli k dispozici také jeho elektronickou verzi. Při osobních návštěvách na adresách domácností předávali tazatelé (případně zanechávali ve schránkách) **leták**, který obsahoval poděkování za účast v minulých ročnících, výzvu k účasti v nové vlně a vybraná zajímavá zjištění z předchozích let. Stejně jako avizní dopis obsahoval leták kontakt na realizátora sběru dat a odkazoval také na web a sociální sítě výzkumu.

Pravidla kontaktování domácností tazateli

První kontakt s domácností měli tazatelé provést primárně telefonicky, případně e-mailem. Na převážnou většinu panelových domácností bylo z minulé vlny dostupné telefonní číslo (94,5 % domácností). E-mailová adresa byla k dispozici u 58,3 % domácností. Alespoň jeden z uvedených kontaktních údajů byl dostupný u 98,5 % domácností.

Pokud nebyly na domácnost dostupné kontaktní údaje (telefonní číslo, e-mailová adresa) nebo domácnost se tazateli nepodařilo zastihnout telefonicky, měli tazatelé provést osobní návštěvu (ideálně ji spojit s dalšími návštěvami v dané oblasti).

Pro pokus o kontakt byla tazatelům předepsána následující **pravidla**:

- Má být provedeno minimálně 6 pokusů o kontakt, než je domácnost uzavřena jako nezastižená.
- Tři primární kontakty mají být provedeny v srpnu a první polovině září, ideálně v rozpětí dvou týdnů, pokusy mají být rozvrženy do různých dnů v týdnu a do různých denních dob.
- Tři odložené kontakty pro zastižení domácností na dovolených (opět ideálně v rozpětí dvou týdnů se střídáním dnů v týdnu a denních dob).

Jako kontakt se přitom počítal každý pokus domácnost zastihnout, ať už osobní návštěva, telefonát, SMS nebo e-mail. Do evidence kontaktů s domácnostmi, která byla používána k monitorování práce tazatelů, však byly zanášeny jen osobní a telefonické kontakty.

Oslovování domácností tazateli

Tazatelům byly 20. srpna 2023 zpřístupněny adresy přidělených domácností. V tento moment mohli tazatelé začít domácnosti kontaktovat a domlouvat rozhovory. Realizace prvních rozhovorů byla možná od 23. srpna. Tazatelé měli oslovit všechny jim přidělené domácnosti od 20. srpna do 20. září. V tomto období byl dle údajů z evidence kontaktů učiněn alespoň jeden pokus o kontakt u 55 % domácností určených pro osobní dotazování.

Od 21. září do 15. října měli tazatelé za úkol dokončovat oslovování a dotazování panelových domácností. Alespoň jeden pokus o kontakt byl do poloviny října zaznamenán u přibližně 75 % domácností určených k osobnímu dotazování. Ve druhé polovině října bylo prioritou oslovit

domácnosti, u kterých nebyl zatím učiněn pokus o kontakt, a zastihnout domácnosti, které se i přes dosavadní pokusy nepodařilo stále kontaktovat. Do konce října bylo osloveno 93 % domácností určených k osobnímu dotazování. Tazatelé měli dále za úkol dokončit maximum rozhovorů do 15. listopadu, ve výjimečných případech do konce listopadu.

Tazatelé byli při sběru dat průběžně urgováni, aby kontaktovali přidělené domácnosti. Aktivita tazatelů byla monitorována pomocí údajů z kontaktního formuláře, do nějž měli tazatelé zapisovat jednotlivé pokusy o kontaktování domácností. S neaktivními tazateli byly probírány důvody nízké produkce a v případě potřeby byly zbývající domácnosti převáděny na jiné tazatele.

Během listopadu dokončovali kontaktování domácností částečně tazatelé (zejména v případě rozpracovaných a odložených rozhovorů), ale většinu neuzavřených domácností převzala pro další komunikaci centrála realizátora sběru dat. Na přelomu října a listopadu proto probíhaly urgency tazatelů k doplnění zápisů do evidence kontaktů, aby byl jasný status domácností pro další práci s nimi.

Oslovování domácností centrálou

Při kontaktování domácností byla **tazatelům průběžně poskytována rozsáhlá podpora z centrály realizátora sběru dat**. Pokud se tazateli nedařilo určitou domácnost kontaktovat, tyto domácnosti byly osloveny e-mailem či telefonicky z centrály. Pokud se u tazatele objevoval vyšší počet odmítnutí účasti, resp. neschopnosti se zúčastnit, byly tyto domácnosti kontaktovány z centrály pro ověření ochoty k účasti ve výzkumu. Příslušným tazatelům byly poté předány informace o výsledku a v relevantních případech se měli tazatelé znovu pokusit oslovit domácnosti ke spolupráci.

Pokud tazatel či centrála zjistili, že domácnost se přestěhovala z předchozí adresy a není již dosažitelná pro původního tazatele, bylo zajišťováno přidělení domácnosti jinému tazateli, případně byla domácnost převedena na online dotazování.

V polovině října byly e-mailem hromadně osloveny všechny dosud nezastižené domácnosti (také domácnosti přestěhované, s nefunkčním telefonickým kontaktem, domácnosti s nenalezenou či neplatnou adresou) a vyzvány k osobnímu dotazování. Během listopadu **převzala uzavírání dosud rozpracovaných domácností z velké části centrála** realizátora sběru dat namísto tazatelů. V prvním listopadovém týdnu byly dosud nezastižené domácnosti kontaktovány telefonicky a na domácnosti bez telefonního a e-mailového kontaktu byly odeslány dopisy. Pracovníci centrály ve spolupráci s dalšími tazateli rovněž prověřovali a pokoušeli se telefonicky oslovit domácnosti, u kterých byla tazateli zaznačena neexistující či neplatná adresa nebo nefunkční telefonní číslo či se jednalo o přestěhované domácnosti.

Cílem centrály bylo do 20. listopadu navázat pomocí telefonu, e-mailu či osobní návštěvy kontakt s co nejvíce zbývajících domácnostmi a zjistit, zda mají zájem o osobní dotazování či nikoli. Při odmítnutí osobního dotazování bylo nabízeno vyplnění dotazníku na internetu. Během listopadu byly centrálou také znovu oslovovány vybrané domácnosti, které odmítly účast, nebo u nich tazatelé uvedli, že nejsou schopni se zúčastnit dotazování.

Do 12. prosince byly prováděny poslední pokusy o kontakt s nezastiženými domácnostmi, u nichž nebyly dosud splněny minimální požadavky na kontaktování (méně než šest kontaktů, malý rozestup

mezi prvním a posledním kontaktem). Několik původních domácností vyplnilo dotazník online ještě po tomto datu – jedná se o dříve kontaktované domácnosti, které měly od posledního pokusu o kontakt několik týdnů na dokončení dotazníku.

Online dotazování domácností: primárně CAWI domácnosti

Domácnosti určené primárně pro online dotazování (tedy domácnosti dotázané pomocí CAWI ve čtvrtém ročníku) byly osloveny ve čtyřech dávkách, a to:

- 25. srpna 2023: osloveno 10 domácností jako pilot CAWI systému,
- 29. srpna 2023: osloveno 20 domácností jako 2. fáze pilotu CAWI systému,
- 8. září 2023: osloveno 153 domácností,
- 12. září 2023: osloveno 74 domácností.

Zvací e-mail byl připraven po vzoru avizního dopisu, který byl zasílán domácnostem primárně určeným pro online dotazování. Zároveň se zvacím e-mailem byla domácnostem zaslána informační SMS s upozorněním na zasláný e-mail a s kódem, který bylo třeba zadat při vstupu do online dotazníku. Pokud se nepodařilo doručit e-mail ani SMS zprávu (nefunkční kontaktní údaje, zaplněné e-mailové schránky atd.), byl domácnostem odeslán zvací dopis poštou.

Domácnosti byly ve zvacím e-mailu vyzvány, aby dotazník vyplnily během následujícího týdne. Pokud domácnost v této době dotazník nevyplnila, byly jí zasílány přibližně po deseti dnech upomínky. Jestliže domácnosti nereagovaly na první e-mailovou upomínku, byla jí poté zaslána druhá e-mailová upomínka. Posléze byly zasílány třetí a čtvrté upomínky pomocí SMS či telefonu. Třetí a čtvrté upomínky byly zaměřeny především na mladší domácnosti.

E-maily s pozvánkou k výzkumu byly zasílány z různých e-mailových adres, aby se zamezilo padání do spamu či blokování e-mailu schránkou respondenta. Také SMS zprávy byly zasílány z různých čísel kvůli zamezení blokování na telefonu respondenta.

Online dotazování domácností: primárně CAPI domácnosti

U domácností přidělených k osobnímu dotazování nabízeli tazatelé (či při pozdějším oslovování centrála) primárně provedení osobního rozhovoru. **Pokud domácnost odmítla, bylo domácnosti nabídnuto online dotazování.** Reakce domácnosti na nabídku online dotazování byla evidována v kontaktním formuláři.

Domácnosti, které odmítly osobní dotazování, ale souhlasily s vyplněním online varianty dotazníku, byly postupně oslovovány v sedmi dávkách od konce září do 12. prosince. Zasílání upomínek k účasti probíhalo obdobně jako u primárně CAWI domácností. Ze 713 domácností, které odmítly osobní dotazování, souhlasilo s vyplněním dotazníku online 183 domácností (25,7 %). Z nich se poté reálně do šetření zapojilo 64 domácností (9 % ze všech odmítnuvších domácností).

Od poloviny listopadu bylo nabízeno online dotazování vybraným členům domácností, které se již zapojily pomocí osobního rozhovoru s tazatelem. Aby byl navýšen počet respondentů v mladších věkových skupinách, byli osloveni dosud nedotázaní členové domácností do 49 let. Celkem bylo

vytipováno 426 členů domácností k oslovení, z nich se vyplněním individuálního dotazníku zúčastnilo 42 (9,9 %).

Rozdělené domácnosti

V šestém ročníku byla využívána stejná pravidla sledování členů panelu jako v předchozích vlnách. Pokud se jedna domácnost rozdělila na více částí (např. děti se odstěhovaly od rodičů), bylo cílem dotázat všechny takto nově vzniklé domácnosti (dceřiné domácnosti). V šestém ročníku se však u žádné rozdělené domácnosti nepodařilo dotázat více než jednu dceřinou domácnost.

Výsledky kontaktu a návratnost (retence)

V šestém ročníku výzkumu bylo dotázáno celkem 1 807 domácností z domácností, které se zapojily do čtvrtého ročníku výzkumu (3 188 domácností). Protože se u žádné rozdělené domácnosti nepodařilo dotázat více než jednu část, nevznikly rozdělením žádné nové domácnosti (v předchozích vlnách vznikl rozdělením menší počet domácností).

Část domácností z původního souboru již nebyla způsobilá pro účast – všichni členové domácnosti zemřeli nebo jim účast nedovoloval zdravotní stav. U několika domácností nebylo možné dohledat fyzickou adresu nebo nebyly funkční kontaktní údaje. Celkem tak bylo k účasti v šestém ročníku způsobilých 2 946 domácností (92 % z 3 188 domácností zapojených do čtvrté vlny).

Účast v šestém ročníku šetření odmítlo 713 domácností (24,2 % ze způsobilých domácností). Dalších 224 domácností nebylo přes minimálně jeden učiněný pokus o kontakt zastiženo (7,6 %).

Výsledná retence/návratnost (podíl dotázaných ze způsobilých domácností) činí v šestém ročníku **61,3 %**. Hrubá retence nezohledňující způsobilost domácností, tedy podíl dotázaných ze všech vydaných domácností, činí 56,7 %.

Tabulka 11: Distribuce dispozičních kódů v původním panelu

Vydané adresy	3 188	
Dotázáno ve 4. vlně (2018)	3 188	100,0 %
Vzniklé rozdělením	0	0,0 %
Splňuje požadavky na dotázání (způsobilé, „eligible“)	2 946	92,4 %
Dotázáno, ověřeno pomocí kontrol	1 807	61,3 %
Odmítnutí	713	24,2 %
Dotazování nedokončeno nebo domluveno na termín po skončení terénu	65	2,2 %
Nezastižení (celé domácnosti nebo dospělého člena)	224	7,6 %
Domácnost nereagovala na CAWI výzvu (primárně CAWI domácnosti)	94	3,2 %
Domácnost nebyla navštívena (žádný pokus o kontakt domácnosti)	0	0,0 %
Změna adresy, nedotázáno	38	1,3 %
Dotázáno, vyřazeno na základě kontrol	5	0,2 %
Nesplňuje požadavky na dotázání (nezpůsobilé, „ineligible“)	242	7,6 %
Neschopnost se zúčastnit – žádný člen nemluví česky	0	0,0 %
Neschopnost se zúčastnit – zdravotní stav, stáří	177	73,1 %
Nenalezená adresa	10	4,1 %
Nefunkční kontaktní údaje	55	22,7 %

Poznámka: Procenta domácností, které splňují (způsobilé) či nesplňují (nezpůsobilé) požadavky na dotázání, jsou vypočtena z celého počtu vydaných adres. Procenta domácností s jednotlivými dispozičními kódy jsou stanoveny z počtu domácností v příslušné skupině (způsobilé či nezpůsobilé k účasti).

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Míra retence dle charakteristik domácnosti

V následujícím textu je vyhodnocena pravděpodobnost účasti domácnosti v šestém ročníku vzhledem k charakteristikám domácností ve čtvrté vlně v roce 2018. Uvedena je jak **míra návratnosti** (tedy podíl dotázaných domácností ze všech způsobilých), tak **míra způsobilosti** (tedy podíl způsobilých domácností ze všech adres domácností vydaných do terénu). Za způsobilé („eligible“) k dotázání byly v šetření považovány všechny ty domácnosti, u byly zaznamenány jiné výsledky kontaktu (dispoziční kódy) než neschopnost se zúčastnit, nenalezená adresa či nefunkční kontaktní údaje (dle rozpisu v tabulce výše). Přibližně tři čtvrtiny nezpůsobilých domácností v 6. ročníku tvoří domácnosti, u nichž se natolik zhoršil zdravotní stav, že nikdo ze členů se nebyl schopen zapojit do dotazování, případně všichni jejich členové již zemřeli. Vymezení způsobilých a nezpůsobilých domácností v 6. ročníku se řídí konvencemi využívanými v předchozích vlnách.

Napříč typy domácností dle složení je míra retence poměrně vyrovnaná. Nejnižší míra retence se objevila u jednočlenných seniorských domácností (58,2 %) a nejvyšší u domácností jednoho rodiče s dítětem (65,8 %). Více se napříč typy domácností liší podíl způsobilých domácností ze všech

vydaných. Nejnižší míra způsobilosti byla zaznamenána u domácností seniorů. Tento výsledek je v souladu se skutečností, že u těchto domácností je vyšší riziko vážnějších zdravotních komplikací znemožňujících účast či přímo zániku celé domácnosti z důvodu úmrtí všech členů.

Výrazněji se míra retence lišila podle nejvyššího dosaženého vzdělání členů domácnosti. Častěji se do výzkumu znovu zapojily domácnosti, v nichž měl ve čtvrté vlně alespoň jeden člen vysokoškolské vzdělání (65,2 %) nebo středoškolské vzdělání s maturitou (62,6 %). Nižší míra retence se projevila u domácností se středoškolským vzděláním bez maturity (53,3 %) a ještě nižší u domácností se základním vzděláním (47,8 %). Také míra způsobilosti byla nižší mezi domácnostmi s nižším vzděláním (zde se projevuje vyšší podíl domácností s maximálně základním vzděláním či středoškolským bez maturity mezi seniorskými domácnostmi).

Míra retence výrazněji kolísá také mezi kraji. Nejnižší pravděpodobnost opětovné účasti vykázaly domácnosti v Libereckém kraji (48,6 %) a Karlovarském kraji (53,1 %). V tomto kraji byla zaznamenána i nízká míra způsobilosti (77,1 % ze všech vydaných adres v Karlovarském kraji byly domácnosti způsobilé k účasti). U ostatních krajů dosáhla míra způsobilosti nejméně 89 %. Nejvyšší míry retence dosáhl Jihomoravský kraj (71,6 %), Hlavní město Praha (70,3 %) a Ústecký kraj (67 %). Odlišnou míru retence napříč kraji lze vztáhnout k rozdílům mezi podíly domácností, u nichž byl zachován stejný tazatel. Kraje s nejvyšší mírou retence vykazují také jedny z nejvyšších proporcí domácností s přiděleným stejným tazatelem jako v předchozí vlně. Korelační koeficient na úrovni krajů mezi mírou retence a podílem domácností se stejným tazatelem činí 0,6.

V případě velikosti sídla se objevují menší rozdíly v míře retence mezi jednotlivými skupinami. Nejnižší retence byla zaznamenána v sídlech s méně než 5 000 obyvateli (55 %), naopak nejvyšší v městech nad 20 tisíc obyvatel (66 až 67 %).

Tabulka 12: Míra způsobilosti a retence původních panelových domácností (%)

	Počet způsobilých	Počet dotázaných	Způsobilost (%)	Retence (%)
Typ (struktura) domácnosti				
Seniorské domácnosti jednočlenné	337	196	81,4	58,2
Seniorské domácnosti vícečlenné	381	236	87,4	61,9
Partnerská domácnost s dítětem do 17 let	794	494	97,2	62,2
Domácnost jednoho rodiče s dítětem do 17 let	114	75	91,9	65,8
Domácnost jednotlivce ne-seniorská	292	186	96,1	63,7
Partnerská domácnost bez dítěte	791	479	94,3	60,6
Ostatní	237	141	93,3	59,5

	Počet způsobilých	Počet dotázaných	Způsobilost (%)	Retence (%)
Nejvyšší vzdělání v domácnosti				
Základní	67	32	75,3	47,8
Střední bez maturity	598	319	87,4	53,3
Střední s maturitou	1 173	734	93,3	62,6
Vysokoškolské	1 108	722	95,7	65,2
Kraj				
Hlavní město Praha	317	223	95,5	70,3
Středočeský kraj	249	139	91,9	55,8
Jihočeský kraj	286	164	92,3	57,3
Plzeňský kraj	137	89	91,3	65,0
Karlovarský kraj	64	34	77,1	53,1
Ústecký kraj	282	189	91,0	67,0
Liberecký kraj	138	67	96,5	48,6
Královéhradecký kraj	190	130	96,9	68,4
Pardubický kraj	106	68	94,6	64,2
Kraj Vysočina	138	70	92,6	50,7
Jihomoravský kraj	345	247	92,0	71,6
Olomoucký kraj	184	105	88,9	57,1
Zlínský kraj	132	80	91,0	60,6
Moravskoslezský kraj	378	202	93,3	53,4
Velikost sídla				
Do 999 obyvatel	444	246	93,1	55,4
1 000 až 4 999 obyvatel	546	299	91,0	54,8
5 000 až 19 999 obyvatel	607	363	92,0	59,8
20 000 až 99 999 obyvatel	721	478	92,3	66,3
100 000 obyvatel a více	628	421	93,7	67,0

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

3.5. Sběr dat na doplňujícím souboru domácností

Harmonogram

Sběr dat na doplňujícím souboru domácností byl zahájen 24. listopadu 2023. V tento den mohly být provedeny první rozhovory s novými domácnostmi, avšak tazatelé pracující na původním panelu

mohli ještě před tímto datem vyhledávat a oslovovat domácnosti ke spolupráci. Dotazování nových domácností bylo ukončeno 25. února 2024.

Domácnosti z doplňujícího souboru byly dotazovány převážně online (CAWI). Osobní dotazování bylo využito v omezenější míře, převážně pro domácnosti, s nimiž byla domluvena účast během první fáze sběru dat (dotazování původních domácností). Osobní dotazování nových domácností probíhalo do 7. prosince. Z domácností patřících do doplňujícího souboru bylo 6 % dotázáno pomocí osobního rozhovoru, zbylé domácnosti vyplnily online dotazník.

Od 7. prosince probíhaly rekrutace nových domácností pro online dotazování. První online dotazníky byly novými domácnostmi vyplněny 12. prosince. Online dotazování probíhalo do ukončení sběru 25. února 2024.

Rekrutace CAPI

Pro osobní dotazování byly domácnosti vyhledávány pomocí tazatelů dle zadaných kvótních znaků. Ve fázi CAPI dotazování nových domácností rekrutovaných tazateli byly v souboru respondentů individuálních dotazníků silněji zastoupeny respondenti ve věku 65 a více let. Struktura souboru byla proto korigována tím, že tazatelé neměli v domácnostech z doplňujícího souboru dotazovat členy starší 65 let.

Nábor nových domácností pro CAPI dotazování byl zahájen již na přelomu září a října, a to tazateli, kteří již měli dokončené rozhovory s panelovými domácnostmi. Tazatelé byli pro nábor nových domácností vybaveni avizním dopisem a letákem pro předání při osobní návštěvě. Avizní dopis a leták byly odvozeny z materiálů pro původní domácnosti s potřebnými úpravami.

Rekrutace CAWI

Nábor nových domácností pro CAWI dotazování byl prováděn **pomocí několika postupů**. Typ rekrutace je v datovém souboru uveden v proměnné *rekrutace*.

Primárním způsobem rekrutace bylo oslovení tazatelem (na rekrutaci pracovali převážně terénní tazatelé pro osobní dotazování, ale doporučovat domácnosti mohli rovněž tazatelé primárně pracující na telefonickém dotazování, tj. operátoři).

V dalším kroku byli osloveni členové online panelu Dialog, aby doporučili jiné domácnosti ze svého okolí, které budou osloveny s žádostí o účast ve výzkumu.

Nakonec bylo také využíváno přímé oslovení členů Českého národního panelu a členů panelu Dialog s výzvou k zapojení se do výzkumu. Tento poslední způsob byl využíván primárně k rekrutaci domácností a respondentů, kteří byli dosud ve výzkumu méně zastoupeni (zejména jednočlenné domácnosti, osoby se základním vzděláním a středním bez maturity).

Při rekrutaci domácností tazateli (resp. telefonickými operátory), na doporučení členů online panelu Dialog nebo přímým oslovováním panelistů Dialog byl využíván následující postup:

1. Tazatel, operátor či člen panelu dodal kontaktní e-mail na domácnost.

2. Domácnosti byl zaslán odkaz na registrační formulář, v němž byla oslovená domácnost požádána o spolupráci na výzkumu a vyplnění několika základních údajů. Do formuláře bylo rovněž třeba zadat telefonní číslo pro další kontaktování.
3. Dále byla domácnost telefonicky kontaktována pro ověření uvedených informací. Při telefonátu uvedl respondent znovu základní charakteristiky domácnosti (kraj bydliště, velikost obce, typ domu) a jejích členů (věky, pohlaví, vzdělání) a bylo vyhodnoceno, zda domácnost odpovídá vyhledávaným kvótám. Pomocí automatického algoritmu byl vybrán jeden či dva respondenti k dotazování, a to dle aktuálních potřeb sběru dat (dle stavu naplňování kvót). V případě více vhodných členů preferoval algoritmus kategorie osob, které byly dosud v dotázaném souboru méně zastoupeny (muži, mladší ročníky, s nižším vzděláním). Při telefonickém ověřování se dále zjišťovalo, zda je člen domácnosti vybraný k dotazování členem komerčního online panelu (na počet těchto respondentů byla stanovena maximální hranice).
4. V posledním kroku byl domácnosti odeslán unikátní online odkaz k dotazníku a do SMS zprávy heslo pro vstup do dotazníku. Odkazy byly obvykle odesílány do druhého dne.

Při přímém oslovování členů Českého národního panelu byly na základě profilů respondentů předem vytipovány domácnosti odpovídající vyhledávaným skupinám. Domácnostem byl zaslán odkaz na krátký rekrutační dotazník, v němž uvedly základní charakteristiky domácnosti a jednotlivých členů. Pokud domácnost vyhovovala hledaným kritériím, byl jí zaslán unikátní online odkaz k dotazníku a SMS heslo pro vstup do dotazníku.

Při rekrutaci domácností byla kontrolována unikátnost telefonního čísla a e-mailu, aby se předešlo duplicitnímu dotázání stejné domácnosti (ať už účasti stejné osoby v domácnosti, nebo účasti dalších osob ve stejné domácnosti). Postupně byla do zjišťování zařazena i adresa bydliště, protože při kontrolách dat z dotazníků a výplatních údajů bylo odhaleno několik desítek duplicitně dotázaných domácností, které byly z dat vyřazeny (podrobnosti k těmto kontrolám jsou obsaženy v podkapitole 6.2 *Kontrola a čištění průběžných dat*). Respondenti byli z tohoto důvodu na začátku dotazníku upozorněni, že každá domácnost se může výzkumu zúčastnit pouze jednou.

Nadvýběr domácností

V rámci sběru dat na doplňujícím souboru byl ve srovnání s kvótním předpisem **navýšen počet domácností s partnerským/manželským párem** (s dětmi i bez dětí). Tento nadvýběr představuje 172 domácností. Domácnosti s partnerským/manželským párem mezi nimi převažují (76,2 %), ale zastoupeny jsou také další typy domácností (mimo seniorských). Tyto domácnosti byly v rekrutačním screeningu vyhodnoceny jako potenciálně partnerské (screening zjišťoval pouze orientační složení domácnosti), ale dle struktury domácnosti uvedené v dotazníku pro domácnost spadají do jiných typů.

Při sběru této části doplňujícího souboru byl zjednodušen kvótní předpis, sledováno bylo pouze zastoupení individuálních charakteristik (pohlaví, věkové kategorie, vzdělání), nikoli další charakteristiky domácnosti. Výběr domácností pro tuto část souboru byl prováděn přímo z Českého národního panelu (stejným způsobem jako probíhalo přímé oslovení popsané výše). U části

domácností z nadvýběru byly pořízeny dva individuální dotazníky. Domácnosti z nadvýběru v doplňujícím souboru označuje proměnná *boost*.

Návratnost

Na kvótní výběr se nevztahuje standardní koncept návratnosti (podíl dotázaných ze způsobilých jednotek). V této sekci jsou proto uvedeny dostupné informace o úspěšnosti dotázání domácností, které byly identifikovány jako vhodné pro výzkum.

U domácností z doplňujícího souboru, které byly rekrutovány pro osobní dotazování, nebyla pořízena evidence, kolik domácností bylo celkově osloveno s nabídkou účasti.

Při rekrutaci nových domácností pro online dotazování pomocí tazatelů, operátorů, prostřednictvím panelistů Dialog a přímým oslovením panelistů Dialog vyplnilo 1 341 domácností registrační formulář a splňovalo kritéria pro ověřovací telefonát. Z toho 743 domácností prošlo ověřovacím telefonátem a splňovalo kritéria pro dotázání (55 %). Nakonec 613 z těchto domácností vyplnilo dotazník (82,5 % z domácností, které vyplnily registrační formulář).

Při přímé rekrutaci z Českého národního panelu vyplnilo úvodní screeningový dotazník a splňovalo kritéria pro dotázání 1 554 domácností. Z toho 789 (50,7 %) domácností bylo následně úspěšně dotázáno.

Rekrutace respondentů z online panelů

Část doplňujícího souboru byla **rekrutována z dosavadních členů komerčních online panelů**, a to 58 % domácností z celkové velikosti doplňujícího souboru (870 domácností). Z toho 81 domácností bylo rekrutováno z členů panelu Dialog a 789 domácností z Českého národního panelu. Jak je podrobněji popsáno níže, **ne nutně je však v těchto případech respondent domácnostního a/nebo individuálního dotazníku sám členem komerčního online panelu**.

Český národní panel (ČNP) je od roku 2012 společným projektem výzkumných agentur Nielsen Admosphere, NMS Market Research a STEM/MARK. Panel Dialog je od roku 2014 nástroj společností STEM/MARK a ČSOB, který slouží k dotazování reprezentativní bankovní populace ČR, ale je využíván také pro jiné výzkumy. Do obou panelů se mohou registrovat zájemci ve věku 15 a více let žijící na území České republiky. Respondenti jsou rekrutováni pomocí inzerce na internetu, metodou snowball (doporučení nových členů stávajícími) i off-line. Registrace do panelů vyžaduje vyplnění online formuláře s unikátním e-mailem a telefonním číslem. Po vyplnění formuláře probíhá telefonická kontrola identity respondenta a zadaných údajů. Panely využívají další kontrolní mechanismy pro ověření autenticity respondentů, ověření profilových údajů respondenta a vyloučení duplicit. Panely provádějí kontroly odpovědí ve výzkumech a vyhrazují si právo vyloučit panelisty, kteří poskytují nepravdivé či jinak nekvalitní odpovědi. Oba panely umožňují sestavit reprezentativní soubory české online populace. Více informací o těchto online panelech obsahuje podkapitola 2.4 *Doplňující soubor domácností*.

U respondentů rekrutovaných pomocí tazatelů, operátorů a prostřednictvím online panelistů bylo zjišťováno, zda osoba vybraná pro dotazování je členem online komerčního panelu, aby bylo možné sledovat a korigovat podíl těchto respondentů. U domácností přímo oslovených členů ČNP nebylo

při rekrutaci ověřováno členství v panelu, automaticky se předpokládalo. Proměnná *onpanel* v datovém souboru označuje, zda člen domácnosti vybraný pro dotazování byl při rekrutaci označen jako člen komerčního online panelu. Tato proměnná se **vztahuje k celé domácnosti**, ne nutně proto je respondent individuálního dotazníku sám členem komerčního online panelu (v některých domácnostech vyplnilo dotazník více členů či dotazník vyplnil jiný než vyzvaný člen).

Zapojení do komerčních online panelů bylo rovněž kontrolováno v individuálním dotazníku otázkou „Přivydělával(a) jste si v posledních 12 měsících vyplňováním dotazníků na internetu?“. Celkově 28,2 % respondentů individuálního dotazníku uvedlo, že v posledním roce alespoň příležitostně odpovídali na dotazníky na internetu za účelem výdělku (12,3 % pravidelně několikrát do měsíce, 15,9 % příležitostně). V původním panelu se přitom jednalo o 10,7 % respondentů individuálního dotazníku, v doplňujícím souboru se jedná o 56,7 % respondentů).

Tabulka 13: Frekvence přivydělávání si vyplňování dotazníků na internetu

Typ domácnosti	Mód	Způsob rekrutace	% pravidelně	% příležitostně	% nikdy
Původní	CAPI	-	2,2	7,2	90,6
	CAWI	-	3,2	16,3	80,5
	Celkem	-	2,3	8,4	89,3
Nové	CAPI	Oslovení tazatelem	1,6	4,0	94,4
	CAWI	Oslovení tazatelem	5,8	16,4	77,8
	CAWI	Oslovení operátorem	0,0	9,5	90,5
	CAWI	Doporučení členem panelu Dialog	10,8	30,4	58,9
	CAWI	Přímá rekrutace z panelu Dialog	47,3	30,8	22,0
	CAWI	Přímá rekrutace z panelu ČNP	44,4	36,4	19,2
	Celkem	-	28,6	28,1	43,3
Celkem	-	-	12,3	15,9	71,8

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), 4 459 respondentů individuálního dotazníku (z toho 2 762 z původního panelu a 1 697 z nových domácností), vlastní zpracování

3.6. Monitorování práce tazatelů

Monitoring tazatelské činnosti

Práce tazatelů byla během sběru dat nepřetržitě monitorována. Sledováno bylo zapojení tazatelů (zda již začali na projektu pracovat či již ukončili práci, přerušení aktivity), kontaktování domácností a domlouvání schůzek, frekvence nezastižení a odmítnutí. Monitorováno bylo rovněž, zda tazatel

průběžně zasílá z terénu originály administrativních dokumentů (souhlas s účastí, smlouva o účasti) a zda jsou dokumenty úplně a správně vyplněny.

Tazatelé obvykle dostávali zpětnou vazbu ke své práci dvakrát týdně, a to na začátku týdne a poté ve čtvrtek či pátek. S méně aktivními tazateli byly probírány důvody nízké produkce, případě přearažení práce na jiného tazatele.

Tazatelům byla poskytována technická podpora po celou dobu dotazování v době od rána (cca 7:00) do večera (cca 20:00 hod). V těchto hodinách mohli tazatelé rovněž kontaktovat pracovníky centrály s metodickými dotazy (např. vymezení členů domácnosti či jak postupovat v případě úmrtí člena domácnosti).

Evidence kontaktů

Při sběru dat na původních panelových domácnostech byl využíván **elektronický kontaktní formulář** pro záznam jednotlivých pokusů o kontakt s domácnostmi. Kontaktní formulář byl pro tazatele přístupný z dotazovacího rozhraní v detailu jednotlivých domácností. Nejprve zaznamenali tazatelé výsledek kontaktu, jeho datum a čas a v závislosti na konkrétním výsledku byly pokládány podrobnější otázky na průběh kontaktu. Do evidence měly být zaznamenány všechny osobní a telefonické kontakty. Pokusy o kontakt pomocí e-mailu či SMS evidovány nebyly.

Data z evidence kontaktů sloužila jako **klíčový vstup pro monitoring práce tazatelů a sledování postupu sběru dat**. Tazatelé, kteří pokusy o kontakt nezapisovali, byli urgováni ke zpětnému vyplnění dosavadních pokusů o kontakt a řádné evidenci dalších pokusů. Někteří, zejména starší, tazatelé zapisovali pokusy o kontakt do svých terénních poznámek (papírová evidenční tabulka) a při přepisu do elektronické podoby jim byla poskytnuta asistence centrály nebo zaslané poznámky byly do kontaktního formuláře přeneseny přímo pracovníky centrály.

Ve formuláři nebyly používány samostatné kódy pro přechodné výsledky (tj. výsledek posledního kontaktu s domácností, který není považovaný za poslední kontakt) a finální výsledky. Pro zaznamenání výsledku sloužila jedna sada kódů, které dle dalších okolností značily přechodné či konečné stavy domácností. Cílem terénních prací bylo, aby u maximálního počtu domácností byl získán finální výsledek kontaktu. Statusy kódů zobrazuje následující tabulka. Pokud nebylo výsledkem kontaktu plné dotázání domácnosti, byla v závislosti na situaci podnikána různá opatření pro zvýšení návratnosti.

Tabulka 14: Výsledky kontaktu v kontaktním formuláři

Výsledek kontaktu	Status kódu	Kroky pro zvýšení návratnosti
Dotazování domácnosti	Finální	
Domácnost odmítla rozhovor	Přechodný do uzavření tazatelem	Pokusy o konverzi tazatelem a centrálou
Rozhovor započat, ale nedokončen	Přechodný	Opakované kontaktování tazatelem či centrálou

Výsledek kontaktu	Status kódu	Kroky pro zvýšení návratnosti
Domluveno dotazování na jiný termín	Přechodný	Opakované kontaktování tazatelem či centrálou a pokusy domluvit dotazování na dřívější dobu
Neschopnost se zúčastnit (nemocní, staří apod.)	Finální	Prověření centrálou
Domácnost nezastižena	Přechodný do splnění minimálních požadavků na pokusy o kontakty	Pokusy o opakované zastižení tazatelem a centrálou
Adresa neexistuje / nenalezena	Finální	Prověření centrálou
Adresa není byt	Finální	Prověření centrálou
Domácnost přestěhována	Finální	Prověření centrálou, převedení na jiného tazatele nebo online dotazování
Nefunkční telefonní číslo	Finální	Prověření centrálou, e-mailové, případně osobní oslovení tazatelem či centrálou

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

3.7. Administrace dotazovacích nástrojů

Pořadí dotazníků

Při osobním i online dotazování byl **nejprve vyplněn dotazník pro domácnost**. Při osobním dotazování bylo preferováno, aby respondentem dotazníku pro domácnost byl člen domácnost nejlépe obeznámený s finanční a bytovou situací domácnosti. V online prostředí typicky vyplňoval dotazník pro domácnost respondent, který prošel rekrutací.

Po dokončení dotazníku pro domácnost následovalo vyplňování jednotlivých individuálních dotazníků. Respondentem v pořadí prvního individuálního dotazníku byl typicky respondent dotazníku pro domácnost, ale mohlo se jednat o jiného člena domácnosti, protože **pořadí vyplnění individuálních dotazníků bylo libovolné**. Při osobním dotazování měli tazatelé v maximální možné míře dotázat všechny členy při jedné návštěvě.

Počet individuálních dotazníků

Při osobním dotazování panelu původních domácností bylo vyplňování individuálních dotazníků nabízeno všem členům domácnosti ve věku alespoň 15 let. Cílem bylo **získat v domácnosti co nejvíce možných dotazníků**. Z důvodu vysokého zájmu o vyplnění individuálního dotazníku, který vysoce převyšoval návratnost v předchozích vlnách, byl počet individuálních dotazníků

v domácnosti od druhé poloviny září regulován a vyplnění dotazníku třetího a vyššího pořadí již nebylo umožňováno.

V online dotazování nového souboru domácností bylo až na počáteční dny této fáze sběru dat omezeno provádění druhých dotazníků z důvodu nižší kontroly nad dotazovací situací než v případě osobního dotazování a nutnosti regulovat strukturu respondentů individuálního dotazníku.

Kombinace módů sběru dat

Jako jedno z opatření na zvýšení návratnosti individuálních dotazníků mezi mladšími věkovými skupinami v původním panelu byli během listopadu 2023 zpětně osloveni členové již dotázaných domácností do 50 let s prosbou o vyplnění online dotazníku. V rámci jedné domácnosti tak mohlo dojít ke kombinaci módů sběru dat (domácnost se nejprve zúčastnila osobního dotazování, posléze byl doplněn další individuální dotazník vyplněný online).

Přístup do online dotazníků

Přístup do CAWI dotazníku byl zajištěn tak, aby do dotazníku mohla vstoupit jen domácnost, pro kterou byl dotazník určen. Domácnosti, které měly být dotázány online, byly obeslány zvacím e-mailem s pokyny pro vstup a vyplnění dotazníku. E-mail obsahoval link na online dotazník, který byl unikátní pro každou oslovovanou domácnost. Souběžně s e-mailem byla odeslána na mobilní číslo domácnosti SMS s přihlašovacím kódem, který byl náhodně vygenerován spolu s odkazem na dotazník. Po kliknutí na odkaz z e-mailu vstoupil respondent na úvodní stránku s výzvou k přihlášení, k němuž musel být využit jedinečný kód z SMS zprávy.

Pokud respondent přerušil vyplňování dotazníku, případně bylo vyplňování přerušeno kvůli technickým problémům, do dotazníku bylo možné opětovně vstoupit pomocí odkazu a přihlašovacího kódu. Dotazník se načetl v místě, kde vyplňování skončilo.

3.8. Odměny respondentům

Za vyplnění individuálního dotazníku náležela respondentům **odměna ve výši 700 Kč**. Za vyplnění dotazníku pro domácnost nebyla poskytována samostatná odměna. Odměnu vyplácel respondentům Sociologický ústav AV ČR na základě smlouvy o účasti ve výzkumu.

Při osobním dotazování respondent vyplnil a podepsal smlouvu o účasti ve výzkumu. Tazatel zároveň zadal základní platební údaje ze smlouvy do elektronického formuláře v závěrečné části individuálního dotazníku. Před předáním těchto dat Sociologickému ústavu AV ČR k proplacení odměny kontroloval realizátor sběru dat údaje zadané do elektronického formuláře vůči snímkům smluv, které pořizovali tazatelé a odesílali do centrály (případně vůči již doručeným papírovým originálům). V případě rozporu mezi údaji v papírovém originále smlouvy a údaji v elektronickém formuláři či neúplného vyplnění smlouvy byl kontaktován respondent a nesrovnalosti byly před proplacením vypořádány. Papírové originály smluv odesílali tazatelé průběžně do centrály, kde byly smlouvy evidovány a na závěr předány do Sociologického ústavu AV ČR.

Při online dotazování vyplňovali respondenti jednotlivé údaje potřebné pro smlouvu, jejíž obsah jim byl zobrazen v závěru individuálního dotazníku. Respondenti museli zároveň vyjádřit souhlas s obsahem smlouvy. Případné nesrovnalosti v údajích zapsaných do elektronického formuláře byly opět řešeny pomocí telefonických konzultací s respondenty před předáním Sociologickému ústavu k vyplacení.

Před odesláním odměn probíhaly v Sociologickém ústavu AV ČR **další kontroly údajů** (např. shoda mezi jménem a příjmením zadaným v současné vlně a údaji z minulých vln, unikátní e-mailové adresy, telefonní čísla, bankovní účty, platnost adres bydliště). Zjištěné nesrovnalosti byly předávány realizátorovi ke zjištění dalších informací a vypořádání.

3.9. Kontrola kvality práce tazatelů

Kontrola nahrávek rozhovorů

Z osobních rozhovorů byly pořizovány nahrávky vybraných úseků dotazníků. Nahrávána byla obrazovka s výčtem členů domácnosti v dotazníku pro domácnost a baterie položek na postoje k sociální spravedlnosti v individuálním dotazníku. Rozhovor byl nahráván pouze v případě, že respondent udělil s nahráváním souhlas (souhlas s nahráváním dotazníku pro domácnost a jednotlivých individuálních dotazníků byl zjišťován zvlášť).

S nahráváním nesouhlasilo 99 respondentů individuálního dotazníku (4 %). Část pořízených nahrávek byla poškozena a nebylo možné je použít k náslechové kontrole (5 % z CAPI individuálních dotazníků). Přibližně desetina (11 %) naslechnutých nahrávek obsahovala technické problémy (zejména tichá nahrávka nebo výrazný šum), které neumožnily vyhodnotit kvalitu rozhovoru.

Naslechnuto a vyhodnoceno tak bylo celkem **2 039 nahrávek rozhovorů (80,2 %)**. Z toho 97,5 % nahrávek neobsahovalo žádný problém. U 51 rozhovorů byl při náslechu zaznamenán věcný nedostatek (vznikly pochybnosti o správném postupu tazatele), tyto rozhovory byly proto dále podrobeny zpětné kontrole. Zpětná kontrola byla také provedena u všech rozhovorů, z nichž byly nahrávky poškozené nebo měly technické problémy.

Při náslechové kontrole byla ověřována přítomnost respondenta při rozhovoru, korektní zajištění souhlasu s nahráváním, dodržování pravidel standardizovaného dotazování (zejména doslovné čtení otázek) a správný záznam odpovědí respondenta.

Průměrná délka naslouchaných nahrávek činila pět minut. Nahrávky byly kontrolovány průběžně při sběru dat, typicky do dvou dnů, kdy tazatel odeslal nahrávky na server.

Tabulka 15: Kontrola nahrávek individuálních rozhovorů

	Počet	% z individuálních dotazníků
Individuálních dotazníků (rozhovorů)	2 543	-
Bez souhlasu s nahráváním	99	3,9
Poškozená nahrávka	124	4,9
Technické problémy nahrávky	281	11,0
Věcné nedostatky v rozhovoru	51	2,0
Naslechnuto bez problémů	1 988	78,2

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Zpětné kontroly

Část získaných rozhovorů a online dotazníků byla podrobena zpětné terénní kontrole. Zpětné kontroly probíhaly telefonicky a pomocí e-mailu průběžně během sběru dat. Kontrola byla dle informací od realizátora sběru dat dokončena (tj. s respondentem se podařilo navázat kontakt a vyhodnotit kontrolní kritéria) u přibližně **62 % individuálních dotazníků z osobního rozhovoru a 54 % individuálních dotazníků vyplněných online**. Část respondentů kontaktovaných ve věci zpětné kontroly se nepodařilo dosáhnout nebo provedení kontroly odmítli. Kontrolu se podařilo dle realizátora sběru dat dokončit u 95 % respondentů individuálního dotazníku, u nichž byl podniknut pokus o kontrolu.

Rozhovory a online dotazníky byly ke kontrole **přidělovány převážně náhodně**. Záměrně byly ke kontrole přiřazeny osobní rozhovory, u nichž byla nahrávka poškozena nebo obsahovala technické či věcné problémy. Pokud byl respondent kontaktován kvůli vyřešení nesrovnalosti v údajích pro vyplacení odměny nebo neúplně vyplněným administrativním dokumentům (informovaný souhlas, smlouva o účasti ve výzkumu), byla zároveň při kontaktu s respondentem provedena zpětná kontrola. Ke zpětné kontrole byly také přidělovány rozhovory produktivnějších tazatelů (typicky pět a více dotázaných domácností).

Při zpětné kontrole osobních rozhovorů byla ověřena přítomnost tazatele, doba trvání rozhovoru a obsah dotazníku. Zjišťováno také bylo, zda nebyla součástí návštěvy i nějaká komerční činnost tazatele. Zpětná kontrola online dotazníků obsahovala dotaz na počet osob v domácnosti, jejich věky a vzdělání, termín vyplňování dotazníku a jeho délku.

Z dat bylo na základě zpětných kontrol a kontrol nahrávek **vyřazeno** dohromady pět domácností.

Tabulka 16: Dokončené zpětné kontroly u respondentů individuálního dotazníku

Status domácnosti	Mód dotazování	Telefonická kontrola	E-mailová kontrola	Celkem kontrol	% indi. dotazníků
Původní	CAPI	197	1 315	1 512	61,5
	CAWI	20	170	190	62,7
Nové	CAPI	18	61	79	62,7
	CAWI – tazatel, operátor	6	289	295	70,6
	CAWI – pomocí panelu	11	92	103	65,2
	CAWI – přímé oslovení z panelu	57	360	417	41,9
Celkem	CAPI	215	1 376	1 591	61,5
	CAWI	94	911	1 005	53,6

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

4. Dotazovací nástroje

4.1. Typy instrumentů

Data byla v šestém ročníku CHPS sbírána pomocí standardizovaného elektronického dotazníku, který byl administrován buď při osobním rozhovoru respondenta s tazatelem (CAPI), nebo pomocí online dotazníku (CAWI).

Využívány byly dva dotazníky, a to **dotazník pro domácnost** a **dotazník pro jednotlivce**, označovaný jako individuální dotazník. Dotazník pro domácnost byl vyplňován za každou zúčastněnou domácnost pouze jednou, individuální dotazník mohlo vyplnit více členů domácnosti ve věku alespoň 15 let.

Ve srovnání s předchozími ročníky byla **skladba instrumentů zjednodušena** vzhledem k novým výzkumným cílům a rovněž potřebě zjednodušit metodiku výzkumu a přizpůsobit ji pro významnější zastoupení online dotazování. V šestém ročníku nebyly administrovány časové snímky pro dospělé a děti (využívané v prvních čtyřech vlnách šetření) a dále doplňující papírové dotazníky pro samostatné vyplnění (využívané ve všech předchozích pěti vlnách).

Tabulka 17: Dotazovací nástroje využité v jednotlivých vlnách šetření CHPS

	Vlna 1	Vlna 2	Vlna 3	Vlna 4	Vlna 5	Vlna 6
Dotazník pro domácnost	ano	ano	ano	ano	ano	ano
Dotazník pro jednotlivce	ano	ano	ano	ano	ano	ano
Proxy dotazník	ano	ano	ano	ano	ano	-
Papírový dotazník pro samostatné vyplnění – dospělí	ano	ano	ano	ano	ano	-
Papírový dotazník pro samostatné vyplnění – děti	ano	ano	ano	ano	-	-
Časový snímek – dospělí	ano	ano	ano	ano	-	-
Časový snímek – děti	ano	ano	ano	ano	-	-

Zdroj: České panelové šetření domácností, 1. až 6. ročník (2015–2023), vlastní zpracování

První část dotazníku pro domácnost zahrnovala soupis členů domácnosti. V této části bylo třeba uvést jednotlivé členy domácnosti bez ohledu na věk a k nim doplnit základní sociodemografické charakteristiky (pohlaví, měsíc a rok narození, rodinný stav, ekonomické postavení, nejvyšší dosažené vzdělání). V domácnostech z původního panelu byl zobrazen seznam členů ze všech předchozích vln a respondenti u jednotlivých členů aktualizovali, zda tyto osoby stále patří k domácnosti, a doplňovali či aktualizovali základní informace o těchto členech. Druhá část

dotazníku pro domácnost obsahovala otázky na charakteristiky celé domácnosti (hospodaření, bydlení, péče o děti a další).

Po dokončení dotazníku pro domácnost se v dotazovacím rozhraní otevřela možnost zahájit individuální dotazník pro jednotlivé členy domácnosti ve věku 15 let a vyšším. V případě vyplňování více individuálních dotazníků v domácnosti bylo pořadí vyplnění za jednotlivé členy libovolné.

4.2. Návratnost na úrovni instrumentů

Domácnost byla vyhodnocena v šestém ročníku CHPS jako úspěšně dotázaná, pokud v ní byl vyplněn dotazník pro domácnost a alespoň jeden individuální dotazník. **Individuální dotazník byl vyplněn za celkem 56,5 % členů domácností ve věku 15 a více let.**

Domácnosti z původního panelu

V domácnostech z původního panelu bylo vyplňování individuálního dotazníku nabízeno všem členům v cílovém věku (s omezením dotazníků třetího a vyššího pořadí od 5. týdne sběru). Individuální dotazník v původních domácnostech vyplnilo 60,5 % členů ve věku 15 a více let. V průměru připadá na dotázanou domácnost z původního panelu 1,53 individuálního dotazníku. To výrazně předstihuje poměr CAPI individuálních dotazníků na domácnost v první až čtvrté vlně šetření, kdy vyplnění tohoto instrumentu bylo rovněž nabízeno všem dospělým členům domácnosti (1,38 v první vlně a mezi 1,26 až 1,28 ve druhé až čtvrté vlně).

Vyšší počet individuálních dotazníků na domácnost v šestém ročníku lze dát do souvislosti s odlišným schématem odměn pro respondenty. V šesté vlně byl odměňován samostatně každý vyplněný individuální dotazník, zatímco v první až čtvrté vlně bylo nutné dosáhnout alespoň minimálního počtu dotazníků za domácnost, aby domácnosti byla uznána účast a mohla získat odměnu jako celek.

Domácnosti z doplňujícího souboru

V domácnostech z doplňujícího souboru dotazovaných pomocí CAPI bylo nabízeno vyplnění individuálního dotazníku dvěma členům domácnosti. V těchto domácnostech je individuální dotazník k dispozici u 68,9 % členů od 15 let (1,43 dotazníku na domácnost). V domácnostech z doplňujícího souboru dotazovaných pomocí CAWI dotazníku bylo v prvních dvou týdnech dotazování (první polovina prosince) umožněno vyplnit individuální dotazník libovolnému počtu členů. Poté byl počet individuálních dotazníků omezen na jeden za domácnost. V těchto domácnostech dosahuje míra vyplnění individuálního dotazníku 50 %. To představuje v průměru 1,12 individuálních dotazníků na domácnost.

Následující tabulka shrnuje úroveň návratnosti individuálního dotazníku dle statusu domácnosti, jejího složení (struktura členů) a módu dotazování domácnosti. Pro zachycení struktury domácnosti je využita proměnná *hhtype*.

Tabulka 18: Míra vyplnění individuálního dotazníku dle statusu a typu domácnosti a módu sběru v %

	Původní domácnosti			Nové domácnosti			Celkem
	CAPI	CAWI	Celkem	CAPI	CAWI	Celkem	
Celkem	60,3	61,7	60,5	68,9	50,0	51,0	56,5
Domácnost jednotlivce	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Partnerská/ manželská domácnost	78,3	71,4	77,5	83,3	55,9	56,6	68,9
Rodič s dospělými dětmi	68,0	87,5	68,6	60,0	52,7	53,5	62,6
Rodič s dětmi do 18 let	86,7	100,0	88,2	80,0	68,0	69,3	76,4
Rodiče s dospělými dětmi	61,0	62,4	61,2	35,0	34,6	34,6	50,2
Rodiče s dětmi do 18 let	70,8	78,9	71,9	76,6	46,4	48,1	59,0
Ostatní	54,3	60,0	54,8	40,0	30,3	30,8	41,0

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), 3 297 domácností (z toho 1 807 z původního panelu a 1 490 nových domácností), vlastní zpracování

4.3. Délka dotazníku

Délka vyplňování dotazníků

Při přípravě dotazníku bylo cílem, aby průměrná délka dotazování jednoho respondenta **nepřesáhla jednu hodinu** (tedy aby kombinace dotazníku pro domácnost a individuálního dotazníku byla v průměru kratší než jedna hodina). Tento cíl se podařilo naplnit.

Délky vyplňování se lišily podle způsobu dotazování a podle toho, zda se jednalo o původní CHPS domácnost nebo novou domácnost (viz následující tabulka). Délka dotazníku pro domácnost se při zohlednění módu sběru dat neliší mezi původními a novými domácnostmi (přibližně 12 minut u CAPI a 15 až 16 minut u CAWI). Individuální dotazník v průměru vyplňovali déle respondenti z původních domácností, ale pouze v případě CAPI (přibližně o 4 minuty). U CAWI tomu bylo naopak, ale zde je rozdíl menší (2 minuty). Ačkoli tak pro respondenty z původních domácností byly dotazníky z obsahového hlediska kratší (předvyplnění části údajů k ověření, nedotazování retrospektivních charakteristik), na průměrných časech vyplňování se tento rozdíl v designu dotazníku neprojevil. Respondenti z původních domácností jsou v průměru starší a délka vyplňování dotazníků obecně s věkem roste, tedy efekt designu dotazníku je pravděpodobně zakryt rozdílnou věkovou strukturou respondentů.

Nejvýraznější odlišnost lze pozorovat u délky dotazníku pro domácnost mezi módy dotazování. Domácnosti dotazované pomocí CAWI vyplňovaly tento dotazník o 3 až 4 minuty v průměru déle. Jedno z možných vysvětlení je větší náročnost úvodní části (výčet členů domácnosti v tabulkové podobě, vyplnění vzájemných vztahů mezi členy domácnosti), pokud s ní nepomáhá tazatel.

Tabulka 19: Délka vyplňování dotazníků dle statusu domácnosti a módu dotazování

Status domácnosti	Mód dotazování	Dotazník pro domácnost		Dotazník pro jednotlivce	
		Medián	Průměr	Medián	Průměr
Původní	CAPI	11,3	12,2	49,5	50,7
	CAWI	13,9	15,3	47,3	48,6
Nové	CAPI	11,2	11,9	46,0	46,5
	CAWI	14,6	16,4	48,7	51,2
Celkem	CAPI	11,3	12,1	50,4	50,6
	CAWI	14,5	16,3	48,9	50,7
Původní	-	11,4	12,5	50,4	50,5
Nové	-	14,3	16,2	48,9	50,8
Celkem		12,6	14,1	49,9	50,6

Poznámka: U dotazníků pro domácnost byly ze zpracování vyřazeny dotazníky delší než 60 minut, u individuálních dotazníků vyřazeny dotazníky delší než 120 minut.

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Čas strávený tazatelem v domácnosti

Celkový čas strávený tazatelem v domácnosti zahrnuje nejen vyplňování dotazníků, ale také další úkony při sběru dat (komunikace s respondentem, vyplňování administrativních dokumentů a další). Do celkové doby, kterou tazatelé strávili v domácnosti, se kromě počtu vyplňovaných individuálních dotazníků promítl také věk a další charakteristiky dotazovaných členů. Standardní doba strávená dotazováním domácnosti se dvěma individuálními dotazníky představovala **100 až 120 minut**. Celkový čas dotazování jedné domácnosti také ovlivňovalo, zda byly dotazníky vyplněny při jedné návštěvě nebo během více návštěv.

4.4. Podmíněné dotazování

Stejně jako v předchozích ročnících výzkumu byly v dotazníku pro šestý ročník aplikovány postupy podmíněného dotazování. Při podmíněném dotazování jsou odpovědi z předchozích vln využívány k přizpůsobení znění či filtrování otázek či k zobrazení kontrolních výzev. Cílem podmíněného dotazování je snížit zátěž kladenou na respondenta, zefektivnit dotazování a omezit chybovost dat.

Využití podmíněného dotazování bylo **ve srovnání s předchozími vlnami redukováno** v souvislosti se zjednodušením dotazníku (zejména vypuštění bloku otázek na meziroční zaměstnaneckou historii) a kvůli delšímu časovému období, které uplynulo od posledního dotazování domácností.

V předchozích ročnících výzkumu bylo podmíněné dotazování využíváno čtyřmi následujícími způsoby:

1. předvyplnění odpovědi hodnotou z minulého dotazování. Nová odpověď má být vybrána, pokud došlo ke změně nebo hodnota byla při minulém dotazování chybně zaznamenána,
2. připomenutí odpovědi z předchozího dotazování v textu otázky. Respondent v otázce uvádí, zda došlo či nedošlo ke změně stavu,
3. připomenutí odpovědi z předchozího dotazování v textu otázky. Respondent na otázku odpovídá, jaká je situace nyní,
4. hodnota je využita při filtrování, nikoli připomenuta respondentovi.

V šestém ročníku byly **uplatněny postupy 1, 2 a 4**. U domácností z původního panelu byly v dotazníku zobrazeny základní sociodemografické charakteristiky (pohlaví, měsíc, rok narození, nejvyšší dosažené vzdělání) a vztahy mezi členy domácnosti z poslední vlny dotazování. Respondenti byli v dotazníku vyzváni ke kontrole a v případě potřeby k opravě či aktualizaci těchto informací. Informace o rodinném stavu a nejvyšším dosaženém vzdělání v kombinaci s odpovědí z minulé vlny byly využity k filtrování u vybraných otázek. U respondentů, kteří v poslední vlně účasti patřili mezi pracující, byl zobrazen popis práce z minulé vlny. Respondent mohl buď potvrdit, že tento popis stále odpovídá současnému zaměstnání, nebo zadat aktuálně platný popis práce.

Tabulka 20: Využívání podmíněného dotazování

Proměnná	Typ podmíněného dotazování	Obsah proměnné
<i>name</i> *)	1	Křestní jméno člena domácnosti
<i>sex</i>	1	Pohlaví člena domácnosti
<i>birthm</i>	1	Měsíc narození člena domácnosti
<i>birthy</i>	1	Rok narození člena domácnosti
<i>mstat</i>	4	Rodinný stav člena domácnosti
<i>educ</i>	1, 4	Nejvyšší dosažené vzdělání člena domácnosti
<i>relN(x)</i>	1	Vztah člena domácnosti k osobám <i>pno</i> = x
<i>isco08</i>	2	Pracovní činnost respondenta zakódovaná dle klasifikace CZ-ISCO

*) Z důvodu anonymizace dat nezařazeno do datového souboru.

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Otázky týkající se retrospektivních či stabilních charakteristik respondenta (předchozí práce, rodina původu, příchod do ČR, občanství, národnost, rok sňatku a začátku společného žití s partnerem) byly dotazovány pouze u respondentů, kteří v minulosti nevyplnili individuální dotazník zahrnující tyto otázky. Průchod individuálním dotazníkem se tak odvíjel od vyplnění tohoto dotazníku v minulosti.

Proměnná *capi* v datech určuje, zda osoba vyplnila individuální dotazník v některém z předchozích ročníků.

Odpovědi respondentů uvedené v předchozích vlnách byly při zpracování finálního datového souboru doplněny do příslušných proměnných, aby odpovědi na tyto retrospektivní otázky byly dostupné u všech relevantních respondentů. V tomto aspektu zpracování se soubor z šestého ročníku odlišuje od datových souborů z předchozích vln (v nich byly z minulých dat doplňovány pouze charakteristiky současné práce, které respondenti v daném ročníku potvrzovali). Podrobnosti poskytuje podkapitola 6.4. *Imputace v datech*.

4.5. Zpracování obsahu dotazníku

Obsah dotazníku pro šestý ročník výzkumu CHPS byl připravován v prvním pololetí roku 2023. Přípravu koordinovali členové projektu SYRI z Českého sociálněvědního datového archivu (ČSDA) Sociologického ústavu AV ČR, kteří pracovali již na předchozích ročnících výzkumu. Na vývoji dotazníku se dále podíleli členové několika výzkumných skupin projektu SYRI.

Protože šestý ročník šetření byl realizován v rámci nového výzkumného projektu, bylo třeba obsah dotazníku **aktualizovat pro potřeby řešení nových výzkumných témat**. Zároveň byl důraz kladen na zachování kontinuity s předchozími vlnami v harmonizaci měření a opakování co nejvyššího počtu dosavadních otázek. Při vývoji dotazníku pro šestý ročník byl proto nejprve připraven **jádrový modul (core)** sestávající z klíčových proměnných z předchozích ročníků výzkumu. Tento modul byl navržen realizačním týmem ČSDA ve spolupráci s výzkumníky ze Sociologického ústavu AV ČR, kteří pracovali na původním Českém panelovém šetření domácností. Pro jádrový modul byly zvoleny klíčové proměnné z hlavních tematických oblastí dřívějších ročníků: hospodaření domácnosti, zdraví a životní styl, politické, společenské, občanské postoje a hodnoty, politické chování, individuální sociodemografické charakteristiky, péče o děti, práce.

Při volbě proměnných pro jádrový modul byly preferovány proměnné, které byly opakovány v co nejvíce předchozích vlnách, a dále standardní proměnné převzaté z jiných národních nebo mezinárodních výzkumů a proměnné s doloženou kvalitou měření. Pomocí těchto kritérií byly identifikovány i další významné proměnné z předchozích ročníků, jejichž opětovné zařazení bylo diskutováno s výzkumníky SYRI. Dotazník byl sestavován tak, aby k zodpovězení výzkumných otázek byly **využívány existující proměnné** a nové otázky byly zařazovány v případě, že pro daný koncept nebyla dosud ve výzkumu použita žádná proměnná. Ve výsledném dotazníku pro šestý ročník byla **více než polovina** všech otázek dotázána v alespoň jedné z předchozích vln.

Obsah dotazníku byl připravován také s ohledem na délku dotazování. Dotazník pro domácnost a jednotlivce byl sestaven tak, aby odhadovaná délka obou dotazníků nepřesáhla ve svém součtu jednu hodinu.

4.6. Skriptování

Dotazník byl naprogramován jako **jeden dotazovací skript**, který zohledňoval odlišnosti dle módu dotazování (CAPI, CAWI) a dle statusu domácnosti (původní či nová domácnost). Nebylo tedy připraveno více samostatných verzí dotazníku, ale jeden souhrnný dotazník, v němž byly mód dotazování a status domácnosti využity jako podmínky, podle nichž se lišila část parametrů a obsahu dotazníku. Toto řešení bylo zvoleno s cílem dosáhnout **maximální harmonizace** osobního a online dotazování a dotazování domácností z původního panelu a z doplňujícího souboru. Odlišnosti CAPI a CAWI dotazníku spočívaly většinou v technických parametrech: CAPI dotazník obsahoval navíc odkazy na karty, otázku na připojení domácnosti k internetu, CAWI dotazník naopak zahrnoval čtyři testy pozornosti. CAPI a CAWI mód se rovněž lišil v práci s neplatnými odpověďmi („nevím“, odmítnutí). Podrobnosti k těmto odlišnostem mezi módy poskytuje podkapitola 4.9. *Odlišnosti CAPI a CAWI verze dotazníku*.

Nejprve byl připraven dotazník zohledňující pouze původní domácnosti (kombinující CAPI a CAWI, filtrování dle módu dotazování). Odlišnosti pro nové domácnosti byly do skriptu zapracovány během příprav sběru dat na nových domácnostech v říjnu 2023.

Nejobtížnějším úkolem při skriptování byla příprava importu existujících údajů o členech původních domácností do otázek, v nichž se tyto odpovědi poté zobrazovaly jako předvyplněné a respondent je ponechal či aktualizoval. Údaje pro import připravil Sociologický ústav AV ČR na základě odpovědí z předchozích let. Rozsah načítání údajů byl ve srovnání s předchozími ročníky omezen.

4.7. Testování a finální úpravy

První verze dotazníku byla připravena na začátku července 2023 a do poloviny srpna probíhalo postupné testování a odlaďování dotazníku. Testování probíhalo u dodavatele sběru dat a realizačním týmem v Sociologickém ústavu AV ČR. Testování bylo prováděno formou volného vyplňování dotazníku a formou vyplňování **dle připravených scénářů**, jejichž cílem bylo otestovat všechny možné průchody dotazováním, zejména pak korektní filtrování v závislosti na vstupních informacích o respondentovi z dotazníku pro domácnost (pohlaví, věk, rodinný stav, ekonomické postavení, vztahy k dalším členům domácnosti).

Ze školení tazatelů byla získána testovací data (každý tazatel měl za úkol absolvovat alespoň jeden zkušební průchod dotazníkem), která byla podrobena podrobné strukturní kontrole. Na základě testovacích dat byly provedeny drobné opravy skriptu.

Na základě zpětné vazby ze školení tazatelů bylo provedeno několik finálních úprav dotazníku. Zejména bylo na začátek individuálního dotazníku doplněno ověření rodinného stavu, ekonomického postavení a nejvyššího dosaženého vzdělání respondenta a možnost tyto údaje upravit, pokud nebyly v dotazníku pro domácnost zadány správně. V CAPI variantě byl v sekci individuálního dotazníku, která se týká partnerských vztahů (od otázky *rlsat* po otázku *breakup*), doplněn pokyn k předání počítače respondentovi k samostatnému vyplnění.

4.8. Změny ve srovnání s předchozími vlnami

V šestém ročníku bylo provedeno několik **změn v administraci dotazování, které si vyžádaly úpravy ve znění či dalších parametrech otázek**. Individuální dotazník byl v předchozích letech administrován většinou pomocí osobního rozhovoru, zatímco v šestém ročníku bylo významnou měrou zastoupeno také online dotazování. Dále byl v šestém ročníku využíván jeden elektronický dotazník (pro osobní či online dotazování), zatímco v předchozích letech bylo dotazování jednotlivců rozděleno mezi osobní rozhovor a samostatné vyplňování dotazníku na papíře (část otázek zařazená v šestém ročníku tak byla v minulosti vyplňována respondenty samostatně v papírovém dotazníku).

Úprava ve znění otázek a pokynů

U některých otázek opakovaných z předchozích ročníků bylo proto mírně upraveno zadání či pokyny tak, aby otázky odpovídaly jak osobnímu, tak online dotazování. Tyto změny měly formální charakter, neměnily význam otázky. Typicky tyto změny zahrnují odstranění odkazů na osobu tazatele (např. „Řekněte mi...“) či na kartu v textu otázky (např. „použijte prosím tuto kartu“) nebo větší upřesnění otázky, aby nevyžadovala případné dovysvětlení tazatelem.

Úprava kategorií odpovědí

U menšího počtu opakovaných otázek byla aktualizována nabídka odpovědí, aby **odpovídala podmínkám v době dotazování**. V otázkách na zemi narození (*cborn*) a národnost (*nation*) byla upravena nabídka možností tak, aby reflektovala aktuální zastoupení zemí původu a národností v České republice (dle údajů ze Sčítání lidu, domů a bytů 2021). Při zpracování dat byly kódy možností **harmonizovány** s předchozími vlnami. Podobně u otázky na stranické preference (*partywv*) byla nabídka možností upravena vzhledem k aktuálním stranám a kódy odpovědí byly mezi vlnami harmonizovány při zpracování dat.

Nabídka **vztahů mezi členy domácnosti** byla ve srovnání s předchozími vlnami **zjednodušena**. Předchozí ročníky ukázaly, že volba vztahů mezi všemi členy domácnosti patří k náročnějším pasážím dotazníku, zejména pokud jsou vztahy zadávány zcela nově. Ponechány byly proto kategorie pro hlavní příbuzenské vztahy (včetně rozlišení biologických, nevlastních a adoptivních dětí či dětí v pěstounské péči) a méně časté příbuzenské vztahy v domácnostech byly spojeny do širší kategorie „Jiný příbuzenský vztah“. Nová sada proměnných ke vztahům v domácnosti nese příponu „N“.

Dále byly upraveny kategorie v doplňujících otázkách na velikost příjmů domácnosti a jednotlivce (pro respondenty, kteří nevěděli nebo odmítli uvést číselný příjem). Tyto kategorie je možné mapovat na starší kategorie s úpravami, proto tyto proměnné nesou v kódu příponu „N“.

Práce s neplatnými odpověďmi

Zároveň bylo třeba přizpůsobit práci s neplatnými odpověďmi („nevím“, odmítnutí) v otázkách aktuálním módům dotazování. Podrobnosti k těmto odpovědím zahrnuje podkapitola 4.10 *Neplatné („únikové odpovědi“)*.

Tabulka 21: Otázky, u nichž došlo k úpravě parametrů mezi současnou a předchozími vlnami

Kód otázky	Změna oproti předchozím ročníkům výzkumu (1. až 5. vlna)
<i>relN(x)</i>	Úprava kategorií proměnné (sloučení méně častých příbuzenských vztahů do souhrnné kategorie „Jiný příbuzenský vztah“, sloučení adoptivních a pěstounských dětí/rodičů do jedné kategorie, sloučení více typů sourozenců do jedné kategorie)
<i>hallow</i>	Změna z položky v baterii na pobírání dávek domácností na samostatnou otázku
<i>ahincome</i>	Přidání instrukce pro respondenta: „Pokud nevíte přesně, odhadněte prosím alespoň přibližnou částku.“; přidání výzvy „Tato otázka je pro výzkum velmi důležitá. Pokud neznáte přesnou částku, stačí uvést přibližný odhad.“ při výběru odpovědi „nevím“; viditelná možnost nevím/odmítnutí u CAWI
<i>hincome</i>	Úprava znění otázky („do některé z možností na kartě“ změněno na „do některé z kategorií“); přidání instrukce pro respondenta („Pokud nevíte přesně, odhadněte prosím alespoň přibližnou částku.“); změna kategorií odpovědí (15 kategorií příjmu místo 18); viditelná možnost „nevím“/odmítnutí u CAWI
<i>exutil, exrent, exmort, exrep, exmed</i>	Úprava znění otázky („tento byt/dům“ upraveno na „byt/dům, ve kterém Vaše domácnost žije“); úprava instrukce (z „Pokud respondent nezná...“ na „Pokud neznáte...“); viditelná možnost „nevím“/odmítnutí u CAWI
<i>sfield</i>	Změna znění kategorie 1 („Všeobecné vzdělávací programy“ nahrazeno „Bez hlavního oboru či zaměření, jde/šlo o všeobecný studijní program“) a její zařazení na konec nabídky; vyřazení kategorie 2 „Osobní dovednosti“; vyřazení kategorie 25 „Nezačlenitelné programy“; přeformulování dřívější instrukce pro tazatele do části otázky („Pokud v současnosti studujete...“)
<i>jobtp</i>	Přeformulování instrukce pro tazatele na instrukci pro respondenta („Respondentů, kteří dělali placenou práci...“ upraveno na „Pokud jste dělal(a)...“)
<i>wrkps</i>	Rozdělení původně jedné otázky na dvě části (název práce a popis pracovních činností) a související úprava znění; přidání příkladů dostatečně podrobných popisů do instrukce; rozdělení na tři samostatné sady proměnných (C1: respondenti s individuálním dotazníkem z minulosti, C2: respondenti vyplňující individuální dotazník poprvé a mající v současnosti práci, L: respondenti vyplňující individuální dotazník poprvé s prací v minulosti)
<i>svise, subor, empl, empls</i>	Rozdělení na dvě samostatné proměnné (C: respondenti se současnou prací, L: respondenti s prací v minulosti); přidání upřesnění „ve své současné práci“/„ve své poslední práci“ do textu otázky
<i>fsize</i>	Rozdělení na dvě samostatné proměnné (C: respondenti se současnou prací, L: respondenti s prací v minulosti); přidání upřesnění „v současnosti / ve své poslední práci“; úprava upřesnění dříve v závorce („Myslí se počet zaměstnanců v místě Vašeho pracoviště.“)
<i>contr</i>	Úprava znění otázky (přidáno upřesnění „ve své současné práci“)
<i>hours, adhours</i>	Přidání instrukce pro respondenta: „Vypište číslem počet hodin. Zaokrouhlete na celé hodiny.“

Kód otázky	Změna oproti předchozím ročníkům výzkumu (1. až 5. vlna)
<i>jbsat</i>	Úprava znění otázky („použijte prosím tuto kartu“ změněno na „použijte prosím tuto stupnici“)
<i>cmtime</i>	Přeformulování instrukce pro tazatele na instrukci pro respondenta („Pokud respondent pracuje z domova...“ upraveno na „Pokud pracujete z domova...“)
<i>swlose, swsafe, swhour</i>	Úprava filtru v souvislosti s přesunutím otázky z papírového dotazníku k samostatnému vyplnění v předchozích vlnách do individuálního dotazníku
<i>mary</i>	Úprava znění otázky („Ve kterém měsíci a roce“ změněno na „Ve kterém roce“)
<i>livey</i>	Úprava znění otázky („Ve kterém měsíci a roce jste“ změněno na „Ve kterém roce“); přeformulování instrukce pro tazatele do části otázky („Pokud je datum shodné se svatbou, uveďte ho znovu.“)
<i>cohy</i>	Úprava znění otázky (dovětek „V jakém roce to bylo?“)
<i>chlaund až chmeal</i>	Úprava instrukce pro respondenta („jen na respondenta a jeho manžela/manželku“ změněno na „jen na Vás (respondenta) a Vašeho/Vaši manžela/manželku“)
<i>pool</i>	Úprava znění otázky („Vyberte prosím z karty tu odpověď“ změněno na „Vyberte prosím tu odpověď“)
<i>dichores až didrink, breakup</i>	Úprava filtru v souvislosti s přesunutím otázky z papírového dotazníku k samostatnému vyplnění v předchozích vlnách do individuálního dotazníku
<i>pheart až pref</i>	Úprava znění otázky („Zaškrtněte“ změněno na „Vyberte“)
<i>hchegp až hchegyn</i>	Změna znění otázky („Kdy naposledy jste...“ změněno na „Kdy naposledy jste navštívil(a)...“)
<i>wvote</i>	Úprava znění možnosti odpovědi 5 („Nemá volební právo“ změněno na „Nemám volební právo“)
<i>partywv</i>	Aktualizována nabídka stran; viditelná možnost „nevím“/odmítnutí u CAWI
<i>lfrg</i>	Úprava znění otázky („Použijte tuto kartu a řekněte mi, kam byste se zařadil(a)“ změněno na „Kam byste se zařadil(a)“); viditelná možnost „nevím“/odmítnutí u CAWI
<i>efigual až efisay</i>	Úprava znění otázky („Použijte tuto kartu. Řekněte mi prosím“ změněno na „Uveďte prosím“)
<i>demo</i>	Úprava znění otázky („výroků na kartě souhlasíte“ změněno na „výroků souhlasíte nejvíce“)
<i>pdisf</i>	Úprava znění otázky („Řekněte mi prosím ještě“ změněno na „Řekněte prosím ještě“)
<i>psapet až psaint</i>	Úprava znění otázky („Prosím řekněte“ změněno na „Prosím uveďte“). V šesté vlně bylo využito stejné referenční období a nabídka odpovědí jako ve druhé

Kód otázky	Změna oproti předchozím ročníkům výzkumu (1. až 5. vlna)
	vlně, kdy byla baterie zařazena poprvé (nikoli upravené referenční období a nabídka z dalších vln)
<i>lipar</i>	Přeformulování instrukce pro tazatele na instrukci pro respondenta („Jestliže se rodinné uspořádání do 16 let věku respondenta“ upraveno na „Jestliže se rodinné uspořádání do Vašich 16 let“)
<i>festat, mestat</i>	Přeformulování instrukce pro tazatele na instrukci pro respondenta („Není důležité, zda v době, kdy respondentovi“ změněno na „Pro otázku není důležité, zda s Vámi“); přidání vysvětlení do instrukce („Svobodným povoláním se myslí podnikání bez živnostenského oprávnění.“)
<i>fwrkps, mwkrp</i>	Rozdělení otázky na dvě části (název práce a popis pracovních činností) a související úprava znění otázky a instrukce; přidání příkladů dostatečně podrobných popisů do instrukce; přesun instrukce pro případ více povolání z úvodu k sekci do otázky; viditelná možnost „nevím“ u CAWI
<i>aincome</i>	Úprava znění otázky (přidání dovětky „a další příjmy (např. výživné).“); přeformulování instrukcí pro tazatele na instrukce pro respondenta (vysvětlení čisté mzdy, možnost využít při odpovídání dokumenty, postup, pokud respondent nemá žádný příjem); viditelná možnost „nevím“/odmítnutí u CAWI
<i>income</i>	Úprava znění otázky („do některé z možností na kartě“ změněno na „do některé z kategorií“, přidání dovětky „a další příjmy (např. výživné)“); změna kategorií odpovědí (17 kategorií příjmu místo 15); viditelná možnost „nevím“/odmítnutí u CAWI
<i>cborn</i>	Aktualizována nabídka zemí a jejich pořadí
<i>nation</i>	Aktualizována nabídka národností a jejich pořadí
<i>relig</i>	Aktualizována nabídka církví

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

4.9. Odlišnosti CAPI a CAWI verze dotazníku

Dotazník byl připraven tak, aby si verze pro osobní (CAPI) a online (CAWI) dotazování v co nejvyšší míře odpovídaly. Cílem bylo, aby obě verze obsahovaly až na výjimky stejné znění otázek a instrukcí. Texty některých otázek a instrukcí z předchozích vln prošly proto mírnými úpravami, aby vyhovovaly jak CAPI, tak CAWI dotazování (shrnutí úprav poskytuje výše uvedená tabulka).

Verze dotazníku pro osobní a online dotazování se tedy odlišují převážně v technických parametrech:

- Verze pro osobní dotazování obsahovala v instrukcích pro tazatele informaci, že k dané otázce náleží karta („Karta č. X“). CAWI verze tuto informaci nezahrnovala.
- CAWI verze zahrnovala čtyři testy pozornosti v individuálním dotazníku pro kontrolu kvality dat.
- CAPI verze obsahovala navíc otázku na připojení domácnosti k internetu v dotazníku pro domácnosti.

- V CAWI verzi byly neplatné („únikové“) varianty odpovědí až na výjimky skryté, dokud se respondent nepokusil otázku přeskočit bez uvedení odpovědi. Podrobnosti k těmto typům odpovědí obsahuje následující podkapitola.

4.10. Neplatné („únikové odpovědi“)

U všech otázek s výjimkou základních sociodemografických údajů a proměnných používaných pro sledování struktury souboru bylo respondentům umožněno uvést, že odpověď na otázku neznají či si ji nepřejí sdělit. V ošetření těchto typů odpovědí bylo třeba přihlídnout k rozdílům mezi módy dotazování (CAPI a CAWI), zároveň bylo dbáno na maximální srovnatelnost s předchozími ročníky výzkumu.

CAPI verze dotazníku

V CAPI verzi byly jako součást nabídky odpovědí zahrnuty varianty „Nevím“ a „Odmítl(a)“. Tyto možnosti neměl tazatel respondentům nabízet, byly proto také vizuálně a graficky oddělené. Výjimkou byly otázky, které byly v minulosti součástí papírového dotazníku pro samostatné vyplnění, v němž obsahovaly explicitní varianty pro „nevím“ a/nebo „odmítnutí“ (*symods až symstan, psapet až psaint*). V těchto případech byly uvedené možnosti součástí explicitní nabídky (tedy byly uvedené na kartách či je tazatel předčítal s ostatními možnostmi). Druhou výjimkou byly otázky, v nichž byla odpověď „nevím“ významově plnohodnotnou součástí nabídky odpovědí (*materho, prodho, chplan, chtwo, chmany*). V těchto případech byla odpověď „nevím“ uvedena na kartách a předčítána, možnost „odmítnutí“ nebyla respondentům nabízena (pouze zakódována tazatelem, jestliže respondent takto spontánně odpověděl).

CAWI verze dotazníku

V CAWI nebyly až na výjimky neplatné („únikové“) odpovědi zobrazeny – respondentům se zobrazily až poté, kdy se pokusili otázku přeskočit. Stejně jako v CAPI variantě byly výjimkou otázky, které byly v minulosti součástí papírového dotazníku pro samostatné vyplnění respondentem, v němž obsahovaly explicitní varianty pro „nevím“ a/nebo „odmítnutí“. Tyto možnosti odpovědí byly v online verzi dotazníku přímo zobrazeny. Dalším případem, kdy byly únikové odpovědi přímo zahrnuty do nabídky v CAWI dotazování, byly vybrané citlivé otázky (příjmy, politické preference) a otázky, v nichž respondenti reálně nemusí znát odpověď (např. povolání rodičů v dětství respondenta). Poslední výjimkou byly otázky, v nichž byla odpověď „nevím“ významově plnohodnotnou součástí nabídky. V těchto případech byla odpověď „nevím“ přímo zobrazena, možnost „odmítnutí“ se objevila až po pokusu otázku přeskočit.

Tabulka 22: Neplatné („únikové“) odpovědi s odlišnostmi od standardního zacházení

Kód proměnné	Popis	Práce s proměnnou
CAPI		
<i>newre, newy, leftre, lefty</i>	household grid (noví a bývalí členové)	Varianty pro „nevím“ a odmítnutí jsou součástí standardní nabídky
<i>materho, prodho, chplan, chtwo, chmany</i>	home office, plány mít děti	Varianta pro „nevím“ je součástí nabídky (tazatel nabízí, zahrnuto na kartách). Odmítnutí je pojato jako standardní úniková odpověď (tazatel nenabízí, není uvedeno na kartách)
<i>symods až symstan, psapet až psaint</i>	politické sympatie, politická angažovanost	Varianta pro „nevím“ je součástí nabídky (tazatel nabízí, zahrnuto na kartách). Odmítnutí je pojato jako standardní úniková odpověď (tazatel nenabízí, není uvedeno na kartách)
CAWI		
<i>newre, newy, leftre, lefty</i>	household grid (noví a bývalí členové)	Varianty pro „nevím“ a odmítnutí jsou součástí standardní nabídky
<i>ahincome, hincomeN</i>	příjem domácnosti	„Nevím“ a odmítnutí jsou zobrazeny spolu s ostatními možnostmi
<i>exutil, exrent, exmort, exrep, exmed</i>	výdaje domácnosti	„Nevím“ a odmítnutí jsou zobrazeny spolu s ostatními možnostmi
<i>materho, prodho, chplan, chtwo, chmany</i>	home office, plány mít děti	Varianta pro „nevím“ je zobrazena spolu s ostatními možnostmi. Odmítnutí je zobrazeno až po pokusu otázku přeskočit
<i>vaxflu až vaxhez</i>	očkování	Varianta pro „nevím“ je zobrazena spolu s ostatními možnostmi. Odmítnutí je zobrazeno až po pokusu otázku přeskočit
<i>screbre až screcol</i>	screening na onemocnění	Varianta pro „nevím“ je zobrazena spolu s ostatními možnostmi. Odmítnutí je zobrazeno až po pokusu otázku přeskočit
<i>partyvt21, pvote23r1, pvote23r2</i>	volená strana či kandidát	„Nevím“ a odmítnutí jsou zobrazeny spolu s ostatními možnostmi
<i>partywv</i>	preferovaná strana	„Nevím“ a odmítnutí jsou zobrazeny spolu s ostatními možnostmi
<i>lfrg</i>	škála levice-pravice	„Nevím“ a odmítnutí jsou zobrazeny spolu s ostatními možnostmi
<i>symods až symstan, psapet až psaint</i>	politické sympatie, politická angažovanost	Varianta pro „nevím“ je zobrazena spolu s ostatními možnostmi. Odmítnutí je zobrazeno až po pokusu otázku přeskočit

Kód proměnné	Popis	Práce s proměnnou
<i>fwrkps1/2, fsvise, fempl, fempls, mwrkps1/2, msvise, mempl, mempls</i>	práce otce a matky v 16 letech respondenta	Varianta pro „nevím“ je zobrazena spolu s ostatními možnostmi. Odmítnutí je zobrazeno až po pokusu otázku přeskočit
<i>aincome, incomeN</i>	příjem jednotlivce	„Nevím“ a odmítnutí jsou zobrazeny spolu s ostatními možnostmi

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

4.11. Dokumentace dotazovacího nástroje

Dotazník byl distribuován pomocí osobního rozhovoru s tazatelem za pomoci počítače (CAPI) a online (CAWI), měl tedy pouze elektronickou podobu. Pro dokumentaci dotazovacího nástroje byl připraven dokument s podkladem pro skriptování dotazníku, do nějž byly během zpracování finálních dat doplněny předpisy pro odvozené proměnné, popis paradat a další informace ke zpracování výsledných proměnných. Tato podoba dotazníku se graficky liší od zpracování dotazníků v první až páté vlně šetření, obsahuje však stále všechny klíčové parametry otázek.

Dokument zahrnuje jak CAPI, tak CAWI variantu dotazníku (odlišnosti mezi módy jsou opatřeny podmínkami, které specifikují příslušný mód, stejně jako v elektronickém skriptu) a rovněž obsahuje jak dotazník pro původní domácnosti, tak pro nové domácnosti z doplňujícího souboru (odlišnosti mezi verzemi pro původní a nové domácnosti jsou opět opatřeny podmínkami).

Dokument s předlohou dotazníku obsahuje následující parametry:

- kód či kódy příslušných proměnných,
- znění otázky,
- znění pokynů pro tazatele,
- varianty odpovědí a jejich číselné kódy,
- formulaci filtrační podmínky (kromě zápisu pomocí kódu proměnných a matematických/logických operátorů obsahuje také slovní popis podmínky),
- další technické parametry (povolené rozsahy u numerických proměnných, podmínky pro zobrazení kontrolních výzev atd.),
- značení odvozených proměnných a paradat,
- stručný popis konstrukce odvozených proměnných,
- specifikaci, pokud daný parametr platí jen pro CAPI či CAWI variantu,
- specifikaci, pokud daný parametr platí jen pro původní domácnosti či domácnosti z doplňujícího souboru.

Dotazník obsahuje také barevné rozlišení otázek zařazených poprvé v šestém ročníku a barevné značení odlišností v parametrech otázek od předchozích vln.

5. Datový soubor

5.1. Obsah a struktura datového souboru

Řazení proměnných

Datový soubor z šestého ročníku obsahuje proměnné z dotazníku pro domácnost a z individuálního dotazníku. Proměnné jsou v souboru řazeny dle logicky využití v datech z předchozích ročníků. Řazení dotazovaných otázek odpovídá řazení v dotazníku. Řazení odvozených proměnných a paradat je v některých případech změněno oproti řazení po skupinách proměnných v dotazníku a dokumentaci kvůli větší přehlednosti datového souboru.

Na začátku datového souboru jsou umístěny identifikátory a proměnné z dotazníku pro domácnost, poté následují proměnné dotazované v individuálním dotazníku. Paradata a váhy jsou umístěny na konci datového souboru. Odvozené proměnné jsou obvykle umístěny za primární proměnné, z nichž jsou vypočítány. Odvozené proměnné socioekonomického statusu a odvozené geografické (lokalizační) proměnné jsou umístěny po bloku otázek z individuálního dotazníku.

Kódy a popisky proměnných

Jako kódy (názvy) proměnných jsou využity **akronymy** vytvořené z klíčových slov otázek v anglickém jazyce. Pro otázky dotazované v předchozích vlnách jsou využity původní kódy (např. lfsat, smoke).

Popisky proměnných v datovém souboru zkráceně vystihují obsah jednotlivých otázek/položek. U proměnných opakovaných z minulých vln byly použity stejné popisky jako v datech z předchozích ročníků.

Na rozdíl od dat z předchozích ročníků není součástí popisku číselný kód otázky (např. v datech z předchozích vln značil kód „I3.“ třetí otázku v individuálním dotazníku). Číslování otázek nebylo v šestém ročníku využíváno (při sběru a zpracování dat byly využívány pouze akronymy, které na rozdíl od číselných označení zůstávají mezi vlnami stabilní).

Kódování možností odpovědí

Možnosti odpovědí jsou u otázek opakovaných z předchozích ročníků kódovány stejnými číselnými hodnotami. Otázky s možností více odpovědí (multiple-choice) jsou na rozdíl od předchozích ročníků kódovány pomocí 1 pro „vybráno“ a pomocí kódu 0 (nula) pro „nevybráno“, zatímco v předchozích letech byly užívány hodnoty 1 pro „vybráno“ („ano“) a 2 pro „nevybráno“ („ne“).

Struktura souboru

Datový soubor obsahuje **řádek za každého současného či bývalého člena** domácností, které se účastnily šestého ročníku šetření. Ze souboru byly vyřazeny domácnosti, u nichž nebyl uznán ani jeden individuální dotazník. Pro každého člena domácnosti jsou dostupné údaje z dotazníku pro domácnost (na úrovni domácnosti a na individuální úrovni). Hodnoty proměnných na úrovni

domácnosti jsou všem členům domácnosti společné. Výjimku představují osoby, které v době dotazování už nebyly členy dané domácnosti. U nich se u většiny proměnných na úrovni domácnosti objevuje neplatný kód 97 / 999997 (podrobnosti k bývalým členům uvádí podkapitola 5.3. *Status členů domácnosti a změny ve složení domácnosti*).

Odpovědi z individuálního dotazníku jsou připojeny pouze u respondentů tohoto dotazníku. **Ostatní členové domácnosti** mají u proměnných z individuálního dotazníku **prázdnou hodnotu** (system missing). Respondenty individuálního dotazníku lze identifikovat pomocí hodnoty 1 v ukazateli *indi*. Strukturu datového souboru ilustruje následující tabulka.

Tabulka 23: Struktura datového souboru

hid	pid	indi	Dotazník pro domácnost, úroveň domácnosti	Dotazník pro domácnost, úroveň jednotlivce	Individuální dotazník
1	11	1	1	11	11
2	21	1	2	21	21
2	22	1	2	22	22
2	23	0	2	23	.
3	31	1	3	31	31
3	32	0	3	32	.
4	41	1	4	41	41
5	51	1	5	51	51
5	52	0	5	52	.

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Proměnné nezařazené do datového souboru

Z veřejného datového souboru byly vyřazeny některé proměnné z dotazníků. Vynechány byly proměnné poskytující detailnější informace o respondentovi (např. křestní jméno, které při dotazování sloužilo k rozlišování členů domácnosti nebo slovní popis pracovní činnosti). Tyto proměnné nezařazené do datového souboru shrnuje následující tabulka.

Podrobnější lokalizační proměnné (okres a obec) mohou být poskytnuty v samostatném souboru na žádost podanou Českému sociálněvědnímu datovému archivu (ČSDA).

Tabulka 24: Proměnné nezařazené do datového souboru

Proměnná	Popis
<i>name</i>	Křestní jméno člena domácnosti sloužící pro identifikaci jednotlivých členů během dotazování
<i>wrkps1</i>	Název práce respondenta
<i>wrkps2</i>	Popis hlavních pracovních činností respondenta
<i>fwrkps1</i>	Název práce otce respondenta v době, kdy respondentovi bylo 16 let
<i>fwrkps2</i>	Popis hlavních pracovních činností otce respondenta v době, kdy respondentovi bylo 16 let
<i>mwrkps1</i>	Název práce matky respondenta v době, kdy respondentovi bylo 16 let
<i>mwrkps2</i>	Popis hlavních pracovních činností matky respondenta v době, kdy respondentovi bylo 16 let
<i>county</i>	Okres bydliště domácnosti
<i>munic</i>	Obec bydliště domácnosti
<i>chcounty</i>	Okres bydliště respondenta v dětství (16 let)
<i>chmunic</i>	Obec bydliště respondenta v dětství (16 let)
<i>wcount</i>	Okres pracoviště respondenta
<i>wmunic</i>	Obec pracoviště respondenta

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

5.2. Identifikační čísla

Datový soubor obsahuje identifikační čísla respondentů (*pid*) a domácností (*hid*), která umožňují rozlišení jednotlivých domácností a jejich členů. Identifikační číslo domácnosti je společné všem členům domácnosti (včetně bývalých členů), tedy proměnná s identifikátorem domácnosti nemá v datovém souboru unikátní hodnoty.

Ve výzkumu CHPS jsou identifikační čísla domácností a jednotlivců **vytvářena nově pro každý ročník šetření**. Aby bylo možné propojit data z jednotlivých vln, jsou doplněna také identifikační čísla ze všech předchozích vln. Vlnu, k níž náleží daná sada identifikačních čísel, označuje číslice X v názvu identifikátoru (*wX_hid*, *wX_pid*). Soubor z šestého ročníku zahrnuje rovněž identifikační čísla domácností a jednotlivců ze dvou krátkých ad hoc covid-19 šetření z let 2020 a 2021. Tato identifikační čísla jsou označena *cX_hid* a *cX_pid*, kde X značí pořadí šetření (1 = 2020, 2 = 2021).

V jednotlivých vlnách je identifikační číslo osoby vytvořeno jako kombinace identifikačního čísla domácnosti (*hid*) a pořadí osoby v rámci dané domácnosti (*pno*). Proměnná s pořadím osoby v domácnosti (*pno*) v datovém souboru se vztahuje k šestému ročníku.

Pouze v ad hoc šetřeních z let 2020 a 2021 byly identifikátory domácností a jednotlivců sestrojeny odlišným způsobem. V ad hoc šetření 2020 byla jako identifikační čísla domácností a jednotlivců

využita čísla z páté vlny. V ad hoc šetření 2021 byla pro identifikátory využita čísla z vlny, v níž se domácnost naposledy zúčastnila.

Tabulka 25: Identifikační čísla domácností a jednotlivců

Proměnná	Popis
<i>w(x)_hid</i>	Unikátní identifikační číslo domácnosti ve vlně (x)
<i>w(x)_pid</i>	Unikátní identifikační číslo osoby ve vlně (x)
<i>c(x)_hid</i>	Unikátní identifikační číslo domácnosti v ad hoc covid-19 šetření (x)
<i>c(x)_pid</i>	Unikátní identifikační číslo osoby v ad hoc covid-19 šetření (x)
<i>pno</i>	Pořadí osoby v domácnosti
<i>pidhr</i>	Identifikační číslo osoby, která vyplnila dotazník pro domácnost

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Jednotlivým členům domácností jsou v datovém souboru přiřazena identifikační čísla dalších osob v domácnosti, které mají k danému členovi určitý vztah. K dispozici jsou identifikátory pro partnery, děti, rodiče a prarodiče.

Tabulka 26: Identifikační čísla dalších osob v domácnosti

Proměnná	Popis
<i>pidp</i>	Identifikační číslo člena domácnosti, který je manžel/manželka (<i>relN</i> = 1), stálý partner/stálá partnerka (<i>relN</i> = 2), registrovaný partner/registrovaná partnerka (<i>relN</i> = 3) osoby
<i>pidf</i>	Identifikační číslo muže (<i>sex</i> = 1) v domácnosti, který je biologický (<i>relN</i> = 7), nevlastní (<i>relN</i> = 8) nebo adoptivní otec nebo pěstoun (<i>relN</i> = 9) osoby
<i>pidm</i>	Identifikační číslo ženy (<i>sex</i> = 2) v domácnosti, která je biologická (<i>relN</i> = 7), nevlastní (<i>relN</i> = 8) nebo adoptivní matka nebo pěstounka (<i>relN</i> = 9) osoby
<i>pidgf</i>	Identifikační číslo muže (<i>sex</i> = 1) v domácnosti, který je dědeček (<i>relN</i> = 12) osoby
<i>pidgm</i>	Identifikační číslo ženy (<i>sex</i> = 2) v domácnosti, která je babička (<i>relN</i> = 12) osoby
<i>pidch(x)</i>	Identifikační číslo člena domácnosti, jehož <i>pno</i> je (x) a který je biologické (<i>relN</i> = 4), nevlastní (<i>relN</i> = 5) nebo adoptivní dítě osoby nebo jeho dítě v pěstounské péči (<i>relN</i> = 6)

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Osobám, které mají v domácnosti nevlastního rodiče stejného pohlaví jako pohlaví biologického rodiče, je v identifikátoru otce (*pidf*) a matky (*pidm*) přiřazeno identifikační čísla partnera/partnerky biologického rodiče.

5.3. Status členů domácnosti a změny ve složení domácností

Zjišťování statusu člena domácnosti v dotazníku

V původních domácnostech byla v úvodní části dotazníku pro domácnost vypsána křestní jména členů domácnosti společně s jejich aktuálním věkem. Uvedeny byly všechny osoby, které v minulosti patřily k domácnosti kromě zemřelých členů (tj. *leftre* = 1 v některé z předchozích vln). Na této obrazovce přitom nebylo rozlišeno, zda daná osoba byla členem v poslední vlně dotazování domácnosti. Respondent zde u jednotlivých osob do proměnné *member* označoval, zda aktuálně jsou či nejsou členové domácnosti.

Do přehledu členů dále tazatelé a respondenti dopisovali všechny nové členy, kteří dosud v žádném roce k domácnosti nepatřili. U těchto osob se členství nezjišťovalo, automaticky vyplývalo z logiky přidání nového člena. Odpovědi jsou tak u proměnné *member* dostupné pouze u původních domácností a jejich členů, kteří k domácnosti patřili již v předchozích letech. Noví členové domácností z původního panelu a členové domácností z doplňujícího souboru mají v této proměnné kód pro filtr (97). **Status člena domácnosti pro všechny osoby** v datovém souboru obsahuje odvozená proměnná *hhmem6*.

V šestém ročníku šetření nebyla na rozdíl od předchozích vln využívána proměnná *rstat* (status člena domácnosti), v níž byl rozlišen typ pobytu člena domácnosti. Jedna z kategorií této proměnné označovala osoby, které již nejsou členy dané domácnosti (*rstat* = 5). V šestém ročníku byl status člena domácnosti zjednodušen na dvě kategorie (je či není členem domácnosti) v proměnné *member*.

Práce s bývalými členy v datovém souboru

Do datových souborů ze všech ročníků CHPS jsou vždy kromě současných členů domácnosti zahrnuti také **všichni bývalí členové** (včetně zemřelých, kteří nebyli zobrazováni během dotazování). V důsledku této skutečnosti dochází při vyšetření obou částí rozdělených domácností k **duplicitám**. Členové těchto rozdělených domácností jsou v takovém případě v datech obsaženi dvakrát, a to jednou jako bývalí členové svých původních domácností a jednou jako soušní členové svých nových domácností. Identifikace osoby (*pid*) ze současného ročníku je vždy unikátní (díky odvození z identifikace domácnosti *hid*), avšak identifikátory z předchozích let nikoli.

Pro rozlišení „primárního výskytu“ (tj. výskytu v domácnosti, jejímž je osoba aktuálním členem) slouží proměnná *primary*. Proměnná *primary* se vztahuje vždy jen k aktuální vlně. Příklad se sekundárním výskytem v dané vlně může být v následující vlně opět primární. V šestém ročníku nebylo u žádné rozdělené domácnosti dotázáno více částí, **tyto duplicity tak nevznikají** a všichni respondenti mají v proměnné *primary* hodnotu 1.

Bývalým členům domácnosti je v proměnných z dotazníku pro domácnost přidělena platná hodnota (při splnění ostatních filtračních podmínek) v následujících proměnných: *w(x)_hid*, *w(x)_pid*, *pno*, *asame*, *new*, *newn*, *sex*, *birthm*, *birthy*, *age*, *agecat*, *primary*, *hhmem(x)*, *capi(x)*, *hhsz*, *leavers*, *newcomers*, *rejoiners*, *leftre*, *leftreo*, *lefty* a *hchange*. Ostatní proměnné z tohoto dotazníku nabývají

pro bývalé členy filtračních hodnot (97 / 999 997). V proměnných z ostatních instrumentů mají bývalí členové prázdné hodnoty (system missing), jsou jim však přiděleny hodnoty proměnných vztahujících se k průběhu a podmínkám dotazování (paradata). V proměnných k indikaci vyplněných dotazníků (*hh_hh*, *hh_indi*, *indi*) mají bývalí členové hodnotu 0 a nejsou pro ně spočteny váhy.

Ukazatele pro status člena domácnosti a změny ve složení domácnosti

Pro domácnosti z původního panelu jsou připraveny proměnné zachycující současné složení domácnosti a změny ve složení domácnosti mezi vlnami.

Proměnná *hhmem(x)* určuje, zda byla daná osoba členem domácnosti v ročníku, který odpovídá číselnému indexu na konci proměnné. Hodnotu 1 (Ano) mají v proměnné *hhmem(x)* osoby, které byly v dané vlně označeny jako rezidenti nebo dočasně nepřítomní členové domácnosti (v první až páté vlně *rstat* = 1/2/3/4/88/99, či v šesté vlně *member* = 1). Hodnotou 2 (Ne) jsou v příslušné proměnné *hhmem(x)* značeny dva typy osob. Jednak jsou to bývalí členové, tedy osoby, které před dotazováním v dané vlně domácnost opustily (*rstat* = 5, resp. v šesté vlně *member* = 2). Do kategorie 2 v proměnných *hhmem(x)* zadruhé spadají osoby, které se členy domácnosti staly až v budoucích vlnách (narození, přistěhování atd.), tedy v současné vlně zatím k domácnosti nepatří. Tyto osoby mohou být rozpoznány pomocí kódu 97 v individuálním identifikátoru *pid* z příslušné vlny.

Proměnné *leaver*, *newcomer* a *rejoiner* se vztahují k statusu člena domácnosti v šestém ročníku. V proměnné *leaver* jsou hodnotou 1 označeni odchozí (bývalí) členové, tedy osoby, které v poslední vlně účasti domácnosti (čtvrtá či pátá vlna) patřily k současným členům, avšak v šestém ročníku u nich bylo zvoleno, že již mezi členy nepatří. V proměnné *newcomer* jsou hodnotou 1 označeni zcela noví členové domácností (tj. osoby, které nikdy v minulosti ke členům domácnosti nepatřily). Nakonec v proměnné *rejoiner* jsou označeny osoby, které v poslední vlně účasti domácnosti patřily mezi bývalé členy, ale v šesté vlně bylo uvedeno, že jsou opět současnými členy (tzv. navrátníci se bývalí členové). Proměnné *leavers*, *newcomers* a *rejoiners* pak obsahují počty jednotlivých kategorií respondentů v rámci domácnosti.

Odvozené proměnné *new* a *newn* udávají, zda je v domácnosti alespoň jeden nový či navrátní se bývalý člen a jejich počet. Nakonec proměnná *hchange* rozlišuje domácnosti, u nichž se změnilo složení (alespoň jeden nový, navrátní se bývalý či odchozí člen) či nezměnilo.

Proměnné *hhmem(x)*, *primary*, *leavers*, *new*, *newn* a *hchange* jsou dostupné také v datech z minulých ročníků (s výjimkou ad hoc covid-19 šetření z let 2020 a 2021, v nichž nebyl prováděn výčet jednotlivých členů domácnosti a zjišťování jejich statusu). Ostatní níže uvedené proměnné byly vytvořeny nově pro šestou vlnu.

Tabulka 27: Proměnné popisující změnu složení domácnosti

Proměnná	Popis
<i>member</i>	Zda je osoba z původního panelu domácností současným členem domácnosti v šestém ročníku
<i>primary</i>	Výskyt osoby ve vzorku: primární (člen nerozdělených domácností, aktuální člen rozdělené domácnosti) nebo sekundární (bývalý člen rozdělené domácnosti)
<i>hhmem(x)</i>	Zda osoba byla ve vlně x současným členem domácnosti
<i>hhmem_last</i>	Vlna, v níž byla osoba naposledy současným členem domácnosti
<i>allmem</i>	Status člena domácnosti pro zahrnutí do výčtu členů při dotazování (nezahrnutí členové, kteří před poslední vlnou účasti zemřeli, tj. před realizací šesté vlny bylo úmrtí známé)
<i>o_member</i>	Zda byla osoba z původního panelu členem domácnosti v posledním ročníku účasti domácnosti
<i>leaver</i>	Bývalý/odchozí člen domácnosti (od minulé vlny přestala být osoba členem)
<i>newcomer</i>	Nový člen domácnosti (osoba se nikdy v minulosti nenacházela mezi členy domácnosti)
<i>rejoiner</i>	Navrátilší se bývalý člen domácnosti (v poslední vlně účasti domácnosti nebyla osoba současným členem, ale alespoň v jedné z přechozích vln patřila mezi současné členy a zároveň nyní opět patří mezi členy)
<i>leavers</i>	Počet bývalých členů domácnosti (domácnost opustili mezi aktuální vlnou a poslední vlnou účasti dané domácnosti)
<i>newcomers</i>	Počet nových členů domácnosti (nikdy v minulosti se nenacházeli mezi členy domácnosti)
<i>rejoiners</i>	Počet navrátilších se bývalých členů domácnosti
<i>new</i>	Přítomnost nových členů v domácnosti (včetně navrátilších se bývalých)
<i>newn</i>	Počet nových členů v domácnosti (včetně navrátilších se bývalých)
<i>hchange</i>	Indikace změny ve složení domácnosti (tj. přítomnost nových nebo bývalých členů domácnosti)

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

5.4. Účast v dotazování v minulosti

Datový soubor obsahuje pomocné ukazatele pro rozlišení domácností a jednotlivců, kteří se v minulosti účastnili dotazování.

Ve druhé až páté vlně výzkumu byly vždy osločovány domácnosti dotázané v bezprostředně předcházející vlně. V šestém ročníku však byly k účasti přizvány všechny domácnosti zapojené do čtvrté vlny **bez ohledu na účast v páté vlně**. Účast v páté vlně dotazování označuje proměnná *hhpart5*. Pro první až čtvrtou vlnu není ukazatel *hhpart* vytvořen, protože domácnosti zapojené ve druhé až páté vlně představovaly podmnožinu domácností z předcházející vlny (účastnily se tedy

všech dřívějších ročníků). Do datového souboru jsou rovněž doplněny ukazatele účasti domácnosti v ad hoc covid-19 šetřeních z let 2020 a 2021 jako proměnné $hhpartc(x)$, kde x označuje pořadí šetření (1 = 2020, 2 = 2021). K účasti v ad hoc šetření 2020 byly přizvány domácnosti, které se zapojily do páté vlny v roce 2019. Účast v šetření z roku 2021 bylo nabídnuto domácnostem, které se naposledy zúčastnily ve třetí, čtvrté či páté vlně nebo v ad hoc šetření 2020. Podrobnosti obsahuje dokumentace k ad hoc covid-19 šetřením zveřejněná v online katalogu Českého sociálněvědního datového archivu¹.

Proměnné $capi(x)$ udávají, zda člen domácnosti vyplnil ve vlně x individuální dotazník. Tímto dotazníkem je myšlen individuální CAPI dotazník, který byl do páté vlny určen dospělým členům domácnosti (18 a více let), nikoli papírový dotazník k samostatnému vyplnění pro dospělé. Souhrnná proměnná $capi$ označuje členy domácnosti, kteří tento individuální dotazník vyplnili v alespoň jedné z předcházejících vln. Tento ukazatel představuje **klíčovou proměnnou, která ovlivňovala průchod dotazníkem v šesté vlně**. Respondentům, kteří individuální dotazník již v minulosti vyplnili, nebyly pokládány otázky na retrospektivní charakteristiky (popis poslední práce, rodina původu, narození v ČR atd.). Zároveň jsou do dat doplněny ukazatele $capi_first$ a $capi_last$, které určují vlny, v nichž respondent vyplnil individuální dotazník poprvé a naposledy.

Nakonec proměnné $adhoc(x)$ udávají, zda daná osoba vyplnila dotazník v ad hoc covid-19 šetření příslušného pořadí (1 = 2020, 2 = 2021).

Tabulka 28: Ukazatele účasti v dotazování v minulosti

Proměnná	Popis
$hhpart5$	Ukazatel, zda se domácnost zúčastnila páté vlny
$hhpartc(x)$	Ukazatel, zda se domácnost zúčastnila ad hoc covid-19 šetření (x)
$capi(x)$	Ukazatel, zda člen domácnosti vyplnil individuální dotazník ve vlně (x)
$adhoc(x)$	Ukazatel, zda člen domácnosti vyplnil dotazník v ad hoc covid-19 šetření (x)
$capi$	Ukazatel, zda člen domácnosti vyplnil individuální dotazník v alespoň jednom z předchozích ročníků (v první až páté vlně)
$capi_first$	Vlna, ve které respondent poprvé vyplnil individuální dotazník
$capi_last$	Vlna, ve které respondent naposledy vyplnil individuální dotazník

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

¹ Röschová, M., Klusáček, J., Sukhotskaya, A. 2024. České panelové šetření domácností: Dokumentace ke sběru a zpracování dat. Ad hoc covid-19 výzkumy 2020 a 2021. Praha: Sociologický ústav AV ČR. (Verze 1.0). <https://doi.org/10.13060/dk.2024.1>

5.5. Odvozené proměnné

Datový soubor zahrnuje proměnné, které nebyly přímo dotazovány jako otázky, ale byly odvozeny z hodnot jiných proměnných. Část proměnných byla vytvářena přímo během dotazování a sloužila k filtrování v následujících otázkách. Ostatní proměnné byly konstruovány až při zpracování finálního datového souboru. Odvozené proměnné byly vytvořeny jak na úrovni domácnosti, tak na úrovni jednotlivce.

K odvozeným proměnným patří i další proměnné popisované v jiných sekcích této dokumentace (např. ukazatele pro status členů domácnosti a změny ve složení domácnosti).

Proměnné na úrovni domácnosti

Odvozené proměnné na úrovni domácnosti mají shodnou hodnotu pro všechny členy v domácnosti. Výjimkou jsou bývalí členové domácnosti, kterým je přidělena platná hodnota jen u vybraných proměnných (podrobnosti jsou dostupné v podkapitole 5.3 *Status členů domácnosti a změny ve složení domácností*, oddíl *Práce s bývalými členy v datovém souboru*).

První skupinu odvozených proměnných na úrovni domácnosti představují **počty členů domácnosti** celkově a v jednotlivých věkových kategoriích. Při výpočtu všech těchto proměnných byli vyloučeni bývalí členové domácnosti (*hmem6* = 2), tedy uvedené proměnné vypovídají jen o počtech současných členů domácnosti.

Tabulka 29: Odvozené proměnné na úrovni domácnosti

Proměnná	Popis
<i>hhsiz</i>	Počet osob v domácnosti
<i>nborn</i>	Počet osob v domácnosti ve věku do 1 roku (<i>age</i> < 1)
<i>ch0to5</i>	Počet osob v domácnosti ve věku do 5 let včetně (<i>age</i> ≤ 5)
<i>ch6to11</i>	Počet osob v domácnosti ve věku od 6 do 11 let včetně (<i>6</i> ≥ <i>age</i> ≤ 11)
<i>ch10to17</i>	Počet osob v domácnosti ve věku od 10 do 17 let včetně (<i>10</i> ≥ <i>age</i> ≤ 17)
<i>ch0to17</i>	Počet osob v domácnosti ve věku do 17 let včetně (<i>age</i> ≤ 17)
<i>p0to26</i>	Počet osob v domácnosti ve věku do 26 let včetně (<i>age</i> ≤ 26)
<i>adults</i>	Počet dospělých osob v domácnosti (<i>age</i> ≥ 18)
<i>nown</i>	Zda se domácnost stala od minulého dotazování vlastníkem bydlení (získala vlastnické nebo družstevní bydlení)

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Druhou skupinu proměnných na úrovni domácnosti představují proměnné týkající se **struktury členů**. V souladu s předchozími ročníky výzkumu byla vytvořena souhrnná proměnná *hhtype* charakterizující typ domácnosti dle složení členů a dále jednotlivé ukazatele přítomnosti různých skupin osob. Proměnné *hhtype2* až *hhtype6* značí přítomnost daných skupin osob v domácnosti bez

ohledu na přítomnost jakýchkoli dalších osob, zatímco kategorie proměnné *hhtype* značí přítomnost právě jen popsanych skupin osob. Například manželský pár žijící s matkou manžela (tchyně manželky) bude mít v *hhtype* hodnotu 7 (ostatní) a hodnotu 1 (ano) v proměnných *hhtype2* a *hhtype3*. Podrobnosti ke konstrukci těchto proměnných obsahuje dokumentace k prvnímu až čtvrtému ročníku výzkumu CHPS.

V šestém ročníku byla do dat doplněna nová odvozená proměnná vztahující se ke složení domácnosti, a to typ domácnosti (dle počtu, věků a vztahů členů) využívaný pro kvótní výběr domácností z doplňujícího souboru (proměnná *hhtype_q*). Jednotlivé kategorie této proměnné jsou podrobněji popsány v části 2.4. Doplňující soubor domácností, Kvótní charakteristiky na úrovni domácnosti.

Klasifikace v proměnných *hhtype* a *hhtype_q* se shodují v identifikaci jednočlenných domácností (*hhtype_q* navíc rozlišuje dva typy jednočlenných domácností podle věku členů), u ostatních kategorií se již liší, protože kvótní proměnná *hhtype_q* nebere v úvahu přítomnost dalších skupin v domácnosti (domácnost partnerů s rodičem jednoho z partnerů tak zařadí jako partnerskou, zatímco klasifikace *hhtype* ji zařadí mezi ostatní typy). V proměnné *hhtype* jsou navíc v souladu s přípravou dat v minulých ročnících považovány domácnosti exmanželů s dětmi za rodičovské, zatímco proměnná *hhtype_q* nevyhodnocuje bývalé manžele jako páry.

Tabulka 30: Ukazatele struktury domácnosti

Proměnná	Popis
<i>hhtype</i>	Typ domácnosti dle struktury
<i>hhtype2</i>	V domácnosti je alespoň jeden manželský / partnerský pár
<i>hhtype3</i>	V domácnosti je alespoň jeden rodič bez manžela/partnera jen s dospělými dětmi
<i>hhtype4</i>	V domácnosti je alespoň jeden rodič bez manžela/partnera s dětmi, z nichž alespoň jedno je mladší 18 let
<i>hhtype5</i>	V domácnosti je alespoň jeden rodičovský pár pouze s dospělými dětmi
<i>hhtype6</i>	V domácnosti je alespoň jeden rodičovský pár s dětmi, z nichž alespoň jedno je mladší 18 let
<i>hhtype_q</i>	Typ domácnosti dle struktury – kvótní proměnná

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Proměnné na úrovni jednotlivce

Individuální proměnné odvozené z proměnných v dotazníku pro domácnost mají hodnotu pro všechny členy domácnosti. Výjimkou jsou bývalí členové domácnosti, kterým je přidělena platná hodnota jen u vybraných proměnných (podrobnosti jsou dostupné v podkapitole 5.3 *Status členů domácnosti a změny ve složení domácností, oddíl Práce s bývalými členy v datovém souboru*).

Odvozené proměnné vycházející zcela či částečně z proměnných v individuálním dotazníku mají hodnoty pro respondenty, kteří vyplnili tento dotazník.

Tabulka 31: Odvozené proměnné na úrovni jednotlivce

Proměnná	Popis
<i>age</i>	Věk v letech odvozený od měsíce a roku narození (<i>birthm</i> , <i>birthy</i>) a měsíce a roku vyplnění dotazníku pro domácnost (<i>hstartm</i> , <i>hstarty</i>)
<i>agecat</i>	Věk kategorizovaný po 10 letech
<i>livesp</i>	Zda v domácnosti žije manžel/manželka (<i>relN</i> = 1) nebo registrovaný partner/registrovaná partnerka (<i>relN</i> = 3) osoby
<i>livewith</i>	Zda v domácnosti žije stálý partner/stálá partnerka (<i>relN</i> = 2) osoby
<i>livefa</i>	Zda v domácnosti žije muž (<i>sex</i> = 1), který je biologický (<i>relN</i> = 7), nevlastní (<i>relN</i> = 8), adoptivní otec nebo pěstoun (<i>relN</i> = 9) osoby
<i>livemo</i>	Zda v domácnosti žije žena (<i>sex</i> = 2), která je biologická (<i>relN</i> = 7), nevlastní (<i>relN</i> = 8), adoptivní matka nebo pěstounka (<i>relN</i> = 9) osoby
<i>livebifa</i>	Zda v domácnosti žije muž (<i>sex</i> = 1), který je biologický otec (<i>relN</i> = 7) osoby
<i>livebimo</i>	Zda v domácnosti žije žena (<i>sex</i> = 2), která je biologická matka (<i>relN</i> = 7) osoby
<i>livegfa</i>	Zda v domácnosti žije muž (<i>sex</i> = 1), který je dědeček osoby (<i>relN</i> = 12)
<i>livegmo</i>	Zda v domácnosti žije žena (<i>sex</i> = 2), která je babička osoby (<i>relN</i> = 12)
<i>bipar</i>	Zda alespoň jeden člen domácnosti je biologické dítě (<i>relN</i> = 4) osoby
<i>child</i>	Počet členů domácnosti, kteří jsou biologické (<i>relN</i> = 4), nevlastní (<i>relN</i> = 5) nebo adoptivní děti osoby nebo jeho/její děti v pěstounské péči (<i>relN</i> = 6)
<i>childy</i>	Počet členů domácnosti mladších 18 let (<i>age</i> < 18), kteří jsou biologické (<i>relN</i> = 4), nevlastní (<i>relN</i> = 5) nebo adoptivní děti osoby nebo jeho/její děti v pěstounské péči (<i>relN</i> = 6)
<i>childa</i>	Počet dospělých (<i>age</i> ≥ 18) členů domácnosti, kteří jsou biologické (<i>relN</i> = 4), nevlastní (<i>relN</i> = 5) nebo adoptivní děti osoby nebo jeho/její děti v pěstounské péči (<i>relN</i> = 6)
<i>etran</i>	Změna ekonomického postavení mezi vlnami
<i>relup</i>	Změna rezidenčního partnera mezi vlnami
<i>sfield_wave</i>	Vlna, z níž je čerpána informace o oboru studia respondenta
<i>job_info</i>	Práce popisovaná v proměnných <i>isco08</i> , <i>svise</i> , <i>subor</i> , <i>empl</i> , <i>empls</i> , <i>fsvise</i> (současná nebo minulé práce)
<i>job_wave</i>	Vlna šetření, ze které jsou čerpány informace o práci respondenta (<i>isco08</i> , <i>svise</i> , <i>subor</i> , <i>empl</i> , <i>empls</i> , <i>fsize</i>)
<i>mary_wave</i>	Vlna, z níž je čerpána informace o roku sňatku
<i>mary_info</i>	Zdroj informace o roku sňatku
<i>livey_wave</i>	Vlna, z níž je čerpána informace o roku začátku společného soužití manželů
<i>livey_info</i>	Zdroj informace o roku začátku společného soužití manželů
<i>cohy_wave</i>	Vlna, z níž je čerpána informace o roku začátku nesezdaného soužití partnerů
<i>cohy_info</i>	Zdroj informace o roku začátku nesezdaného soužití partnerů

Proměnná	Popis
<i>part</i>	Zda má respondent individuálního dotazníku rezidenčního (<i>livesp</i> = 1 or <i>livewith</i> = 1) nebo nerezidenčního partnera (<i>date</i> = 1) nebo je ženatý/vdaná či v registrovaném partnerství (<i>mstat</i> = 1 or <i>mstat</i> = 2)
<i>sumchild</i>	Celkový počet biologických (<i>reIN</i> = 4) a adoptivních (<i>reIN</i> = 6) rezidenčních a nerezidenčních (<i>nrechn</i>) dětí respondentů individuálního dotazníku. Výpočet této proměnné nelze zcela harmonizovat s předchozími ročníky, protože v minulých vlnách byly v kategoriích vztahů v domácnosti rozlišovány adoptivní děti a děti v pěstounské péči, zatímco v šestém ročníku jsou tyto děti sdruženy do jedné kategorie.

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Ukazatele přítomnosti členů domácnosti s vybranými vztahy k dané osobě (proměnné *live-* a *bipar*) jsou odvozeny ze vztahů mezi členy domácnosti (proměnné *reIN*) a nabývají pouze hodnot 1 (ano) a 2 (ne). Tyto proměnné tedy nevypovídají o tom, zda mimo domácnost žije někdo, kdo má k osobě daný vztah (např. hodnota „ne“ v proměnné *livesp* znamená, že nikdo v domácnosti není manžel/manželka osoby, nikoli že osoba má manžel/manželku, který/která žije mimo domácnost).

U osob, které mají v domácnosti nevlastního rodiče stejného pohlaví jako pohlaví biologického rodiče, nabývá u *livefa* a *livemo* hodnoty „ano“ také proměnná vážící se k druhému pohlaví, než je pohlaví biologického rodiče.

5.6. Odvozené ukazatele socioekonomického postavení

V datovém souboru se generují dva typy stratifikačních proměnných – International Socio-Economic Index (ISEI) a European Socio-economic Classification (ESeC). V obou případech se jedná o sociální klasifikace založené na informacích o zaměstnání respondenta (occupation-based social classifications).

ISEI má určené skóre pro každé zaměstnání, které bylo spočítáno na základě průměrné úrovně vzdělání a průměrných příjmů držitelů zaměstnání. Nejnovější verze indexu (ISEI-08) vznikla na základě dat z ISSP z let 2002 až 2007, z informací o zaměstnání více než 200 000 respondentů ze 42 zemí². Pro každou hodnotu klasifikace zaměstnání ISCO je určené skóre socio-ekonomického indexu ISEI.

ESeC je evropská verze Erikson-Goldthorpe-Portocarero třídního schématu (EGP). EGP bylo vytvořeno pro podmínky Velké Británie a je méně vhodné pro použití v kontinentálních evropských státech³. Z potřeby mít evropskou socio-ekonomickou klasifikaci vznikla ESeC. Klasifikace obsahuje

² Ganzeboom, H. B. (2010). A new International Socio-Economic Index (ISEI) of occupational status for the International Standard Classification of Occupation 2008 (ISCO-08) constructed with data from the ISSP 2002–2007. In annual conference of international social survey programme, Lisbon (Vol. 1).

³ Wirth, H. (2023). EU-SILC Tools: European Socioeconomic Classification - ESeC88 and ESeC08. (GESIS Papers, 2023/01). Köln: GESIS - Leibniz-Institut für Sozialwissenschaften. <https://doi.org/10.21241/ssar.83962>

devět tříd určených na základě pozice respondenta na trhu práce a typu zaměstnaneckého poměru. Prvním krokem je rozlišení na osoby samostatně výdělečně činné (třída 4 – živnostníci, 5 – drobní podnikatelé v zemědělství) a na zaměstnance. Zaměstnanci jsou dále rozděleni podle toho, jestli smlouva mezi zaměstnavatelem a zaměstnancem má charakter služební (zaměstnanec zastává určitou pozici a za to je placen) nebo výkonový (zaměstnanec je placen za vykonanou práci), nebo jestli je kombinací těchto dvou.

Smlouva o služebním vztahu se týká zejména nemanuálně pracujících, práce v těchto zaměstnáních je obtížná na kontrolu a k jejich vykonávání je potřeba mít specifické kompetence. Smlouva o vykonávané práci se týká převážně manuálně pracujících, kontrolování výkonu těchto zaměstnání je jednodušší a práce nevyžaduje tolik specifických znalostí nebo schopností. Mezi servisními (služebními) třídami (1 – vyšší kvalifikovanější salarát, 2 – nižší salarát) a pracujícími (výkonovými) třídami (7 – rutinní pracovníci ve službách, 8 – kvalifikovaní dělníci, 9 – polokvalifikovaní nebo nekvalifikovaní dělníci) je několik mezilehlých tříd, které jsou kombinací služebního a výkonového zaměstnaneckého poměru (3 – bílé límečky vyššího řádu, 6 – modré límečky v řídicích pozicích/mistři).

Pro výpočet ESeC v datech jsou použity proměnné o ekonomickém postavení (*estat*), zaměstnávání nebo vedení dalších osob (*empl, svise*), počtu zaměstnanců (*empls*) a hodnota klasifikace zaměstnání ISCO-08 (*isco08*), a to pro současnou i poslední práci. Nevýhodou ESeC klasifikace je, že v podmínkách moderních rozvinutých ekonomik první dvě třídy často tvoří až 50 % pracovní síly⁴.

Tabulka 32: Přehled stratifikačních proměnných

Proměnná	Popis
<i>ISEIisco08</i>	Socioekonomický index – respondent
<i>ISEIfisco08</i>	Socioekonomický index – otec
<i>ISEImisco08</i>	Socioekonomický index – matka
<i>ESeC08</i>	ESeC 9 tříd – respondent
<i>ESeC08_6</i>	ESeC 6 tříd – respondent
<i>fESeC08</i>	ESeC 9 tříd – otec
<i>fESeC08_6</i>	ESeC 6 tříd – otec
<i>mESeC08</i>	ESeC 9 tříd – matka
<i>mESeC08_6</i>	ESeC 6 tříd – matka

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

⁴ Ganzeboom H. B. G. (2018). "QUESTIONS AND ANSWERS ON ISEC-08, EGP AND ESEC." Dostupné z: <http://www.harryganzeboom.nl/isco08/qa-isec-08.htm>.

5.7. Geografické (lokalizační) proměnné

V šestém ročníku byla stejně jako v předchozích vlnách při sběru dat zaznamenávána obec bydliště domácnosti. V předchozích ročnících vycházela informace z adresy domácnosti ve výběrovém souboru, případně z aktualizovaného údaje od tazatele, pokud se domácnost přestěhovala. V šestém ročníku byla obec bydliště zjišťována jako **otázka v dotazníku pro domácnost** z důvodu vyššího zastoupení CAWI dotazování. Aby bylo možné unikátně identifikovat název obce, byl nejprve dotázán okres bydliště. Poté respondent vypisoval název obce a vybíral příslušnou obec z našeptavače, který podle prvních zapsaných písmen nabízel relevantní obce ve zvoleném okrese. Stejným způsobem byla zjišťována obec bydliště v dětství respondenta a obec současné či minulé práce.

Z okresu bydliště byl odvozen kraj (NUTS 3) dle aktuální územní struktury České republiky. Na základě údajů o počtu obyvatel v obcích k 1. 1. 2023 (Český statistický úřad) byla z názvu obce odvozena její velikost.

U bydliště v dětství bylo nejprve zjišťováno, zda se jedná o stejnou obec jako obec současného bydliště (proměnná *chplace*). Pokud ano (*chplace* = 1), otázka na obec nebyla položena. Při závěrečném zpracování dat byl u těchto respondentů **doplněn kód obce současného bydliště do kódu obce bydliště v dětství**. Analogicky byla zpracována proměnná s obcí pracoviště.

Proměnné s obcemi a okresy nejsou zahrnuty do datového souboru. Přístup k nim lze získat na základě žádosti podané Českému sociálněvědnímu datovému archivu.

Tabulka 33: Geografické (lokalizační) proměnné

Proměnná	Popis
<i>munic</i>	Obec současného bydliště domácnosti (LAU 2)
<i>county</i>	Okres současného bydliště domácnosti (LAU 1)
<i>region</i>	Kraj současného bydliště domácnosti (NUTS 3)
<i>msize</i>	Velikost obce současného bydliště domácnosti (k 1. 1. 2023)
<i>chmunic</i>	Obec bydliště respondenta v dětství (LAU 2)
<i>chcounty</i>	Okres bydliště respondenta v dětství (LAU 1)
<i>chregion</i>	Kraj bydliště respondenta v dětství (NUTS 3)
<i>chmsize</i>	Velikost obce bydliště respondenta v dětství (k 1. 1. 2023)
<i>wmunic</i>	Obec pracoviště respondenta (LAU 2)
<i>wcounty</i>	Okres pracoviště respondenta (LAU 1)
<i>wregion</i>	Kraj pracoviště respondenta (NUTS 3)
<i>wmsize</i>	Velikost obce pracoviště respondenta (k 1. 1. 2023)

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

5.8. Proměnné příjmu domácnosti a jednotlivce

V dotazníku pro domácnost byl zjišťován celkový čistý měsíční příjem všech členů domácnosti (*ahincome*). Pokud respondent neznal nebo odmítl uvést odpověď, byl dotázán na kategorii měsíčního příjmu (*hincome*). Při zpracování finálních dat byly do proměnné *hincome* doplněny kategorie odpovídající částce uvedené v otázce na numerický příjem (*ahincome*). Zdroj hodnoty ve výsledné proměnné *hincome* je uveden v pomocné proměnné *hincome_type* (zda je kategorie odvozená z uvedené částky či respondent neuvedl částku, ale až kategorii).

Kategorie proměnné *hincome* byly ve srovnání s předchozími ročníky upraveny, aby v souvislosti se zvýšením příjmů lépe diferencovaly ve vyšších příjmových skupinách. Proměnná *hincome* ze šestého ročníku má proto pro odlišení v datech příponu „N“ (*hincomeN*), protože ji není možné přímo srovnávat s hodnotami z předchozích vln. Pro účely srovnání s hodnotami z předchozích ročníků byla vytvořena proměnná *hincome_h*, kterou lze namapovat na kategorie z předchozích ročníků (nejdříve je však nutné sloučit některé kategorie proměnné ze starších ročníků).

Analogicky byl v datech zpracován individuální čistý měsíční příjem dotazovaný v individuálním dotazníku. U částky individuálního příjmu byly navíc převzaty částky příjmu domácnosti u jednočlenných domácností, v nichž byl dotazován pouze příjem domácnosti, nikoli již individuální příjem (protože se jedná o stejnou informaci).

Tabulka 34: Příjmové proměnné a práce s nimi

Proměnná	Popis
<i>ahincome</i>	Čistý měsíční příjem domácnosti – částka
<i>hincomeN</i>	Čistý měsíční příjem domácnosti – kategorie Dotazováni pouze respondenti, kteří neznali nebo odmítli uvést částku. V datech doplněny kategorie uvedených částek.
<i>hincome_type</i>	Typ kategorického příjmu domácnosti (rekódované z částky nebo uvedené přímo respondentem)
<i>hincome_h</i>	Čistý měsíční příjem domácnosti v kategoriích pro harmonizaci s kategoriemi příjmu v předchozích vlnách
<i>aincome</i>	Čistý měsíční příjem jednotlivce – částka Dotazováni pouze respondenti žijící v domácnosti, která má alespoň dva členy. U respondentů z jednočlenných domácností doplněna odpověď z proměnné <i>ahincome</i> .
<i>incomeN</i>	Čistý měsíční příjem jednotlivce – kategorie Dotazováni pouze respondenti, kteří neznali nebo odmítli uvést částku. V datech doplněny kategorie uvedených částek.
<i>income_type</i>	Typ kategorického příjmu jednotlivce (rekódované z částky nebo uvedené přímo respondentem, případně u jednočlenných domácností odvozeno z příjmu domácnosti)

Proměnná	Popis
<i>income_h</i>	Čistý měsíční příjem jednotlivce v kategoriích pro harmonizaci s kategoriemi příjmu v předchozích vlnách
<i>eqval1</i>	Počet spotřebních jednotek, modifikovaná stupnice OECD
<i>eqval2</i>	Počet spotřebních jednotek, původní stupnice OECD
<i>ahincome_i</i>	Čistý měsíční příjem domácnosti – částka. U domácností, které neuvedly částku, ale pouze kategorii příjmu, byla doplněna střední hodnota (průměr) odpovědí v příslušné kategorii u domácností, které částku uvedly
<i>ahincome_e1</i>	Ekvivalizovaný příjem domácnosti, modifikovaná stupnice OECD (hodnota <i>ahincome_i</i> vydělená <i>eqval1</i>)
<i>ahincome_e2</i>	Ekvivalizovaný příjem domácnosti, původní stupnice OECD (hodnota <i>ahincome_i</i> vydělená <i>eqval2</i>)

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Do dat byl dále spočten **ekvivalizovaný příjem domácnosti**, tedy příjem, který je vztažen k velikosti a složení domácnosti pomocí metody spotřebních jednotek. Při výpočtu spotřebních jednotek jsou členům domácnosti přiřazeny rozdílné váhy, které umožňují zohlednit sdílení výdajů v domácnosti a rozdílnou spotřebu členů domácnosti dle věku.

Pro konstrukci počtu spotřebních jednotek v domácnosti byla využita jak běžná, tak modifikovaná stupnice OECD. Proměnná *eqval1* je vypočtená dle modifikované stupnice spotřebních jednotek OECD. První dospělé osobě v domácnosti je přiřazena váha 1, další osoby starší 13 let obdrží váhu 0,5 a děti do 13 let včetně váhu 0,3. Pro výpočet proměnné *eqval2* je využita původní stupnice OECD, kde je pro první dospělou osobu v domácnosti využita váha 1, další osoby starší 13 let mají váhu 0,7 a děti do 13 let včetně váhu 0,5.

Ekvivalizovaný příjem domácnosti byl vypočten v obou variantách. Proměnná *ahincome_e1* vznikla jako numerický příjem domácnosti vydělený počtem spotřebních jednotek dle modifikované stupnice OECD (*eqval1*), proměnné *ahincome_e2* využívá počet spotřebních jednotek dle původní stupnice (*eqval2*). Pro ekvivalizaci byla využita proměnná *ahincome_i*, v níž byla domácnostem, které uvedly pouze kategorii příjmu, dosazena střední hodnota (průměr) dané příjmové kategorie.

5.9. Paradata

Základní paradata

Do datového souboru je zahrnuto několik proměnných, které se vztahují k procesu dotazování (tzv. paradata). První skupinu představují základní paradata, která se týkají módu dotazování, rozlišují realizátora sběru dat a tazatele a vypovídají o vyplnění individuálního dotazníku.

Zahrnuty jsou ukazatele pro mód dotazování, a to proměnná *qmode* značící mód vyplňování dotazníku pro domácnost a proměnná *iqmode* s módem vyplňování individuálního dotazníku. V datech z předchozích ročníků je zahrnuta pouze proměnná *qmode* odpovídající módu dotazování

domácnosti, protože v dřívějších vlnách nebyla využívána kombinace více módů sběru dat u jedné domácnosti. V šesté vlně naproti tomu bylo umožněno dodatečné vyplnění online dotazníků u vybraných členů domácností, které se již zúčastnily osobního dotazování.

Do datového souboru je doplněna také proměnná *qmode_origin*, která označuje primární mód dotazování domácnosti (tj. zda byla domácnost na začátku sběru dat přidělena k osobnímu či online dotazování). Tento ukazatel rovněž nebyl v předchozích ročnících k dispozici.

Data dále obsahují ukazatel agentury, která dotázala danou domácnost a identifikační číslo tazatele. Číslo tazatele je uvedeno u všech domácností, které byly před začátkem sběru dat určeny k osobnímu dotazování (tedy i u domácností, které byly převedeny na CAWI dotazování). V datech z předchozích ročníků bylo číslo tazatele uvedeno pouze u domácností dotázaných osobně.

Novým ukazatelem v šestém ročníku je *indistat*, který rozlišuje status vyplnění individuálního dotazníku včetně dokončených, ale neuznaných dotazníků. Nakonec je obsažena proměnná *iorder* s pořadím vyplnění individuálního dotazníku v domácnosti. Pořadí vyplnění individuálního dotazníku je odvozeno z času ukončení individuálního dotazníku.

Proměnné *h_indi* a *hp_indi* udávají počet vyplněných (uznaných) individuálních v domácnosti a podíl vyplněných individuálních dotazníků ze všech možných (dle počtu členů ve věku 15 a víc let).

Tabulka 35: Základní paradata

Proměnná	Popis
<i>qmode_origin</i>	Primární mód dotazování domácnosti (pouze pro domácnosti z původního panelu), tj. mód sběru dat přidělený před začátkem sběru dat
<i>qmode</i>	Mód dotazování dotazníku pro domácnost
<i>iqmode</i>	Mód dotazování individuálního dotazníku
<i>agen</i>	Agentura, která provedla dotazování domácnosti
<i>intid</i>	Číslo tazatele (tentýž tazatel má stejné číslo v datech ze všech let výzkumu)
<i>indistat</i>	Status vyplnění individuálního dotazníku
<i>iorder</i>	Pořadí vyplnění individuálního dotazníku v domácnosti
<i>h_indi</i>	Počet vyplněných individuálních dotazníků v domácnosti
<i>hp_indi</i>	Podíl vyplněných individuálních dotazníků v domácnosti (ze všech možných)

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Časové značky

Druhou skupinu paradat představují časové značky obsahující datum a čas začátku a konce dotazování. Časové značky jsou k dispozici zvlášť za dotazník pro domácnost a individuální dotazník. Datum a čas začátku dotazování odpovídá vstupu do daného instrumentu, časové značky konce odpovídají odchodu z poslední obrazovky dotazníku. Časové značky v šestém ročníku jsou

připraveny **ve formátu časové proměnné obsahující datum, hodin, minuty a vteřiny**, nikoli jako tři samostatné numerické proměnné pro den, měsíc a rok, jako tomu bylo v předchozích ročnících.

Časové značky k dotazníku pro domácnost jsou uvedeny u všech členů dané domácnosti, včetně bývalých členů. Časové značky k individuálnímu dotazníku jsou k dispozici u respondentů tohoto dotazníku (*indi* = 1), ostatní členové domácností mají hodnoty proměnných *istart* a *iend* prázdné. U jednoho respondenta nebyl z důvodu technického problému zaznamenán čas začátku individuálního dotazníku, v tomto případě je časová značka prázdná.

Tabulka 36: Časové značky

Proměnná	Popis
<i>hstart</i>	Datum a čas zahájení dotazníku pro domácnost
<i>hend</i>	Datum a čas ukončení dotazníku pro domácnost
<i>istart</i>	Datum a čas zahájení individuálního dotazníku
<i>iend</i>	Datum a čas ukončení individuálního dotazníku

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Domácnosti z doplňujícího souboru

První skupinu představují proměnné týkající se dotazování domácností z doplňujícího souboru. Proměnná *hhstatus* rozlišuje mezi domácnostmi z původního panelového souboru a z doplňujícího souboru domácností. Proměnná *rekrutace* udává způsob výběru domácností z doplňujícího souboru (podrobnosti k výběru nových domácností poskytují podkapitoly 2.4 *Doplňující soubor domácností* a 3.5. *Sběr dat na doplňujícím souboru domácností*). Dále proměnná *boost* označuje domácnosti z doplňujícího souboru, které byly vybrány pro navýšení počtu domácností s manželským/partnerským párem.

Nakonec proměnná *onpanel* určuje, zda při procesu rekrutace byla osoba vybraná pro dotazování označena jako člen online komerčního panelu. Hodnota této proměnné přitom nemusí z několika důvodů vypovídat o reálném respondentovi individuálního dotazníku. V některých domácnostech vyplnilo dotazník více členů, či dotazník vyplnila jiná osoba než člen domácnosti vybraný při rekrutaci. U přímé rekrutace z panelu ČNP také nebylo členství v panelu u osoby vybrané pro dotazování zjišťováno (automaticky se předpokládalo), ačkoli pro vyplnění dotazníku mohl být vybrán jiný člen domácnosti než ten, který je v panelu zaregistrován a vyplňuje v něm dotazníky.

Tabulka 37: Paradata vztahující se k doplňujícímu souboru domácností

Proměnná	Popis
<i>hhstatus</i>	Status domácnosti – domácnost z původního panelu CHPS nebo domácnost z doplňujícího souboru
<i>rekrutace</i>	Způsob rekrutace domácností z doplňujícího souboru
<i>boost</i>	Označení nadvýběru v rámci doplňujícího souboru
<i>onpanel</i>	Zda osoba vybraná pro dotazování v domácnosti z doplňujícího souboru (CAWI) byla při rekrutaci označena jako člen komerčního online panelu

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

6. Zpracování dat

6.1. Zpracování hrubých dat

Během sběru dat zpracovával realizátor pravidelně datové soubory s průběžnými daty k vlastním kontrolám a pro předání do Sociologického ústavu AV ČR. Po většinu sběru dat byla data předávána s týdenní periodicitou, v posledním měsíci sběru (únor 2024) byla předávána dvakrát týdně.

Pro zpracování hrubých dat využíval realizátor software IBM SPSS Statistics. Nejprve byla z dotazovacího systému exportována aktuální data, a to vždy od úplného počátku dotazování včetně nekompletních rozhovorů. Pro potřeby převodu dat do cílového souboru ve formátu SAV byla použita sada syntaxí.

Proces převodu dat do cílového souboru byl rozdělen na dvě části. Zvlášť byla zpracována data z dotazníku pro domácnost a zvlášť data z individuálního dotazníku. Na závěr byly údaje z individuálního dotazníku připojeny k příslušným členům domácnosti z datového souboru za domácnost.

Pro zpracování byly vždy vybrány pouze kompletně vyplněné dotazníky a zkontrolována unikátnost a správná konstrukce identifikátorů. Při zpracování dat byly vybrané hodnoty překódovány tak, aby odpovídaly cílové datové mapě (zejména byly převáděny vlastní standardní kódy pro odpovědi typu „nevím“, odmítnutí, odfiltrováno, bez odpovědi, jiný atd.).

Data z dotazníku pro domácnost bylo nutné transformovat z původního širokého formátu exportovaného z dotazovacího systému (jednu domácnost reprezentuje jeden řádek, hodnoty za jednotlivé členy jsou uloženy v samostatných proměnných) na dlouhý formát (samostatný řádek pro každého člena domácnosti). Dále byly vytvářeny vybrané odvozené proměnné.

6.2. Kontroly a čištění průběžných dat

Kontroly zpracované do dotazníku

Základní kontrolní mechanismy týkající se struktury dat byly zpracovány přímo do dotazovacího nástroje. U všech numerických proměnných byl nastaven **povolený rozsah zadávaných hodnot**. V případě zápisu hodnoty mimo tento rozsah nebylo možné pokračovat na další obrazovku a zobrazila se výzva, že je nutné zadat odpověď v daném rozmezí. Dále byla většina numerických otázek opatřena **kontrolami odlehlých pozorování**. Pokud byla zadána odpověď mimo předem určené pravděpodobné rozmezí, byla v dotazníku zobrazena výzva k ověření zadané odpovědi. Na rozdíl od kontrol povoleného rozsahu hodnot však bylo možné odpověď ponechat a pokračovat dál.

Otázky na vztahy jednotlivých členů v domácnosti byly opatřeny **sadou logických kontrol**. Při navolení vzájemně neslučitelných vztahů mezi členy domácnosti či zvolení vztahů, které neodpovídaly věku členů (dítě starší než rodič, právnouče starší než prarodič), byla zobrazena výzva k opravě hodnot a nebylo možné pokračovat dál, než došlo k vyřešení nekonzistence. Další sada

kontrol se zaměřovala na nepravděpodobné kombinace vztahů v domácnosti a věku a rodinného stavu členů (např. přítomnost více než dvou rodičů jedné osoby v domácnosti). V těchto případech byl tazatel/respondent vyzván ke kontrole a případně k opravě, než přejde na další obrazovku.

Testy pozornosti a doba vyplňování dotazníků

Do CAWI verze dotazníku byly zařazeny celkem čtyři testy pozornosti. Jednalo se o otázky, v nichž byl respondent vyzván, aby zvolil danou možnost z nabídky odpovědí. Testy pozornosti byly umístěny po celé délce dotazníku tak, aby na ně vždy odpovídali všichni respondenti. Dva testy měly formát samostatné otázky s možností výběru jedné odpovědi, dva testy byly zapracovány jako položky dvou různých baterií.

Testy pozornosti byly v kombinaci s délkou vyplňování používány k určení nekvalitně vyplněných individuálních dotazníků. K vyřazení byly vybrány individuální dotazníky podle následujících kritérií:

- Tři nebo čtyři testy pozornosti vyplněny chybně bez ohledu na délku vyplňování dotazníku
- Jeden či dva testy pozornosti jsou chybně a délka individuálního dotazníku je kratší než 50 % mediánu délky vyplňování individuálního dotazníku v CAWI
- Délka vyplňování individuálního dotazníku je kratší než 35 % mediánu v CAWI bez ohledu na testy pozornosti

Respondenti, jejichž dotazníky splňovaly výše uvedená kritéria, byli informováni, že jejich dotazníky neprošly kontrolou kvality a dostali možnost dotazník vyplnit znovu.

Při zpracování finálního datového souboru byly individuální dotazníky splňující výše uvedené podmínky vyřazeny z datového souboru. Vyloučeny byly také všechny dotazníky pro domácnost u domácností, u nichž došlo k vyřazení všech dostupných individuálních dotazníků. Celkem bylo z uvedených důvodů vyřazeno 16 individuálních dotazníků vyplněných online a 11 domácností.

Kontroly průběžných dat

Kontrola průběžných datových souborů probíhala na straně realizátora sběru dat i týmem v Sociologickém ústavu AV ČR. Průběžná kontrola byla nejprve prováděna s týdenní periodicitou. V posledním měsíci sběru dat (únor 2024) byla kontrola průběžných dat prováděna dvakrát týdně.

V rámci průběžných kontrol bylo prováděno **automatizované vyhodnocení shody mezi strukturou dat a strukturou dotazníku**. Cílem těchto kontrol bylo identifikovat hodnoty mimo povolený rozsah ukazatele, chybějící hodnoty v případech, u nichž měla být dle filtračních podmínek odpověď, nebo naopak odpovědi tam, kde otázka neměla být položena. Tyto průběžné kontroly strukturní konzistence dat neodhalily žádný systematický problém ve skriptování dotazníku (např. chybně specifikované filtrační podmínky). Nejčastějším zdrojem problémů ve struktuře dat bylo opakované vstupování do dotazníku nebo změna módu dotazování mezi dotazníkem pro domácnost a jednotlivce. V těchto případech u některých dotazníků došlo k narušení vazby mezi filtračními proměnnými z dotazníku pro domácnost a otázkami v individuálním dotazníku a respondent byl dotazován také na otázky, které pro něj nebyly určeny.

Chybějící hodnoty ve finálním datovém souboru tak nejsou obvykle způsobeny problémy v dotazníku, ale vznikly v důsledku čištění finálního datového souboru. U některých domácností a jednotlivců byla při čištění změněna hodnota filtrační proměnné tak, že po úpravě patří do cílové populace určité otázky, zatímco při dotazování nikoli. Například chybějící údaje v proměnných *up6pre* až *up11lon* ohledně péče o děti z dotazníku pro domácnost jsou způsobeny dodatečným přidáním dětí do 15 let mezi uvedené členy domácnosti.

Při průběžných kontrolách byly rovněž **kontrolovány distribuce** jednotlivých proměnných a jejich křížení (logika a zjevná validita distribucí a křížení).

V Sociologickém ústavu AV ČR byla navíc prováděna **logická kontrola** dat zaměřující se na změny základních sociodemografických údajů mezi vlnami (včetně změn ve vztazích mezi členy), konzistenci vztahů v domácnosti, konzistenci mezi základními charakteristikami (např. věk a ekonomické postavení). Prováděné logické kontroly jsou popsány v textu níže (popsány jsou případy, které kontrola identifikovala jako problematické a byly dále prověřovány). Tyto kontroly vycházejí z informací o jednotlivých členech domácnosti v dotazníku pro domácnost a nebyly prováděny pouze pro respondenty individuálního dotazníku, ale pro všechny členy domácnosti (pro větší srozumitelnost je níže používán výraz „respondent“, ale míněn je jím obecně člen domácnosti).

Srovnání odpovědi v současné vlně s minulou vlnou

- Křestní jméno se změnilo od minulé vlny
- Pohlaví se změnilo od minulé vlny
- Měsíc narození se změnil od minulé vlny
- Rok narození se změnil od minulé vlny
- Vztah člena domácnosti k respondentovi se změnil od minulé vlny

Logika vztahů v domácnosti

- Respondent má 3 a více rodičů
- Respondent má 2 a více partnerů/manželů
- Rodič respondenta je starší než respondent o méně než 18 let
- Prarodič respondenta je starší než respondent o méně než 36 let
- Respondent a partner respondenta se věkem liší o více než 20 let
- Sourozenec respondenta je o 18 a více let starší
- Respondent a partner respondenta jsou stejného pohlaví
- Respondent a partner respondenta jsou jiného pohlaví, ale žijí v registrovaném partnerství
- V přehledu vztahů v domácnosti jsou nesymetrické vztahy (vztah člena A ke členovi domácnosti B je odlišný od vztahu člena B ke členovi A)
- Respondent žije v manželství či v registrovaném partnerství, ale manžel(ka) či registrovaný partner(ka) není uveden(a) v přehledu členů domácnosti
- Respondentův rodinný stav je rozvedený(á)/vdovec/vdova/svobodný(á), ale dle přehledu členů domácnosti má manžela/manželku
- Respondent žije v manželství/registrovaném partnerství, ale v přehledu členů domácnosti má partnera/partnerku místo manžela/manželky nebo registrovaného partnera/partnerky

- Respondent(ka) byl(a) v předchozích vlnách ženatý (vdaná), v registrovaném partnerství, rozvedený(á) nebo ovdovělý(á), ale v současné vlně má uvedený rodinný stav svobodný(a)

Vztah mezi vzděláním a věkem respondenta

- Respondentovi je sedm nebo více let a má uvedeno, že se nikdy nevzdělával
- Respondentovi je osmnáct nebo více let a má uvedeno neúplné základní vzdělání
- Respondentovi je třináct a méně let a má dokončené základní vzdělání
- Respondentovi je sedmnáct a méně let a má středoškolské nebo vysokoškolské vzdělání
- Respondent měl v předchozí vlně vzdělání vyššího stupně, než je současné vzdělání

Vztah mezi ekonomickým postavením a dalšími charakteristikami

- Respondentovi je šedesát nebo méně let a je ve starobním důchodu
- Respondent má ekonomické postavení „Student, učeň, žák“, ale v současnosti nestuduje
- Respondent nemá v domácnosti děti mladší sedmi let, ale má ekonomický stav „Na mateřské (rodičovské) dovolené“
- Respondent nemá v domácnosti žádné děti, ale má ekonomický stav „Trvale v domácnosti“

Informace o nových a odchozích členech

- Respondent je mladší osmnácti let, ale jako důvod opuštění domácnosti má uvedeno „Odešel(a) kvůli rozchodu či rozvodu“, „Odešel(a) do domova důchodců nebo domu s pečovatelskou službou“, „Odstěhoval(a) se za prací do jiné části ČR“, „Odstěhoval(a) se za prací do zahraničí“, „Odešel(a) bydlet s partnerem/partnerkou“, „Vdala se/oženil se“ nebo „Osamostatnil(a) se a založil(a) nový domov“
- Respondentovi je padesát a méně let a má uvedeno jako důvod odchodu z domácnosti „Zemřel(a)“
- Není uveden důvod odchodu člena z domácnosti
- Není uveden rok odchodu člena z domácnosti
- Respondent je mladší osmnácti let, ale jako důvod příchodu do domácnosti má uvedeno „Nový partner/partnerka, manžel/manželka“, „Návrat z důvodu ukončení zaměstnání v předchozím bydlišti“, „Návrat z důvodu ztráty bydlení“, „Přistěhoval(a) se z pracovních důvodů“, „Přistěhoval(a) se, aby zajistil(a) péči o některého z členů domácnosti“ nebo „Nový spolubydlící“
- Respondent je starší pěti let, ale jako důvod příchodu do domácnosti má uvedeno „Narození, adopce“
- Není uveden důvod příchodu nového člena domácnosti
- Není uveden rok příchodu nového člena domácnosti

Další kontroly

- Všichni původní členové jsou označeni jako bývalí, domácnost sestává pouze ze zcela nových osob
- Uvedená obec bydliště domácnosti je jiná než v minulé vlně, přičemž v individuálním dotazníku bylo uvedeno, že se domácnost od minulé vlny nepřestěhovala

Kontroly unikátnosti domácností a přiřazení individuálních dotazníků

Během sběru dat bylo průběžně analyzováno základní složení domácností (křestní jména, měsíce a roky narození, pohlaví členů), aby byly zachyceny případné **duplicitně dotázané domácnosti**. Tyto kontroly byly posíleny ve fázi sběru dat na nových domácnostech, u nichž byl kvůli převaze online dotazování nižší dohled nad dotazovací situací. Při kontrolách unikátnosti byly využívány rovněž údaje poskytnuté pro opětovné kontaktování respondentů v budoucnu (adresa, plné jméno a příjmení, další kontaktní údaje).

Mezi domácnostmi z původního panelu byla identifikována pouze jedna duplicita, která nebyla zachycena v předchozích ročnících. Dále se tři domácnosti zúčastnily nejprve jako původní panelové a později se objevily v souboru nových domácností. Duplicitní výskyty těchto domácností byly z dat při finálním zpracování vyřazeny.

Mezi domácnostmi z doplňujícího souboru bylo identifikováno několik desítek duplicitně dotázaných domácností. Obvykle se jednalo o případy, kdy se výzkumu postupně zúčastnili dva různí členové domácnosti (každý ze zúčastněných členů prošel celým dotazováním, tj. za stejnou domácnost byl dvakrát vyplněn dotazník pro domácnost a daná domácnost tak figurovala v datech dvakrát s individuálním dotazníkem vždy za jiného člena). V těchto případech byl v datech ponechán pouze jeden dotazník pro domácnost (obvykle první v pořadí) a k němu byly do příslušného člena převedeny odpovědi z individuálního dotazníku. V jednotkách případů se výzkumu opakovaně zúčastnil tentýž člen domácností, duplicitní domácnost pak byla z dat zcela vyřazena.

Pomocí údajů od respondentů získaných pro opětovné kontaktování v budoucnu a provádění zpětných kontrol (plné jméno a příjmení, datum narození) bylo rovněž kontrolováno, že individuální dotazník je přiřazen ke správnému členovi domácnosti. Pokud se údaje o respondentovi v přehledu členů domácnosti neshodovaly s údaji zadanými na konci dotazníku, byl respondent kontaktován a bylo prověřováno, který člen domácnosti dotazník reálně vyplnil. V domácnostech z původního panelu tyto záměny téměř nenastávaly, avšak v domácnostech z doplňujícího souboru se v několika desítkách případů stalo, že individuální dotazník byl vyplněn jiným členem domácnosti, než kterému byl v datech přiřazen. U těchto domácností obvykle dotazník reálně vyplňoval jiný člen domácnosti, než který byl pro vyplňování vybrán při rekručním screeningu. Po ověření s respondenty byla data z individuálního dotazníku přeřazena ke správnému členovi domácnosti.

Čištění průběžných dat

Nesrovnalosti ve struktuře a logice dat odhalené během průběžných kontrol byly evidovány do formuláře sdíleného mezi realizátorem sběru dat a týmem ze Sociologického ústavu AV ČR. Dle povahy problému byly jednotlivé záležitosti **předávány příslušným oddělením realizátora sběru**. Při odhalení problému ve struktuře dat bylo prověřováno správné skriptování dotazníku a procedury transformace dat exportovaných z dotazovacího systému. Logické nekonzistence (dle seznamu kontrol v předchozím textu) byly předávány terénnímu oddělení, které mělo za úkol kontaktovat příslušného respondenta a ověřit reálný stav. Dle zjištěných informací byly průběžně upravovány příslušné proměnné (např. opraveny vztahy mezi členy domácnosti). Tyto úpravy byly prováděny realizátorem sběru dat.

6.3. Zpracování finálních dat

Finální soubor s kompletními daty byl po ukončení terénních prací předán do Sociologického ústavu AV ČR, kde proběhlo další zpracování. V předaném datovém souboru bylo od realizátora sběru dat zapracováno čištění logických nesrovnalostí v datech dat popsané výše.

Čištění dat

Data prošla **automatizovaným začištěním struktury**, aby odpovídala struktuře dotazníku. V případě, že se u daného pozorování objevila odpověď u otázky, která neměla podle filtračních podmínek položena, byla hodnota změna na kód pro filtr. Pokud naopak hodnota vzhledem k filtračním podmínkám chyběla, byl doplněn chybový kód (98/999998). Tyto úpravy byly třeba spíše výjimečně, zejména pak v návaznosti na další čištění, které změnilo hodnotu filtrační proměnné.

Dále byly provedeny další individuální ad hoc opravy dat pro vypořádání zbývajících logických nesrovnalostí (situace, kdy byly informace z terénu získány až po předání výsledných dat). V případě potřeby byly dle textových odpovědí v otevřených či polootevřených otázkách upraveny související proměnné, zejména ekonomické postavení dle popisu povolání (menší počet respondentů v popisu povolání upřesnil, že reálně nemá práci).

Kontrola odlehlých hodnot a konzistence roků

U numerických proměnných byla zkontrolována odlehlá pozorování. Menší počty extrémních hodnot u několika proměnných byly nahrazeny chybovým kódem (98/99998). U proměnných s roky sňatku a začátku společného soužití byla provedena kontrola konzistence odpovědí mezi partnery. U jednotek případů s velkými neshodami byla odpověď v šestém ročníku nahrazena chybovým kódem (999998).

Pro čištění odlehlých pozorování byl zvolen **konzervativní přístup** tak, aby byly v maximální míře zachovány původní odpovědi respondentů a vyřazeny byly pouze vysoce nepravděpodobné odpovědi, které by mohly vychylovat agregované statistiky. Analytikům pracujícím s daty je **doporučeno** provádět u numerických proměnných **detekci odlehlých pozorování** a případné extrémy ovlivňující výsledky vyloučit.

Tabulka 38: Proměnné s vyřazenými odlehlými či nekonzistentními hodnotami

Proměnná	Popis proměnné	Úroveň	Počet odlehlých či nekonzistentních hodnot
<i>exutil</i>	Výdaje na spotřebu energií	Domácnost	9
<i>exrent</i>	Výdaje na nájem	Domácnost	7
<i>exmort</i>	Výdaje na hypotéku	Domácnost	0
<i>exrep</i>	Výdaje na údržbu a opravy	Domácnost	12
<i>exmed</i>	Výdaje na léky, zdravotní péči	Domácnost	12
<i>ahincome</i>	Příjem domácnosti – částka	Domácnost	1

Proměnná	Popis proměnné	Úroveň	Počet odlehlých či nekonzistentních hodnot
<i>aincome</i>	Příjem jednotlivce – částka	Jednotlivec	0
<i>height</i>	Tělesná výška	Jednotlivec	8
<i>weight</i>	Tělesná váha	Jednotlivec	8
<i>mary</i>	Rok sňatku	Jednotlivec	3
<i>livey</i>	Rok začátku soužití (manželé)	Jednotlivec	3
<i>cohy</i>	Rok začátku soužití (partneři)	Jednotlivec	1

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

U proměnných s příjmy jednotlivců a domácností byla také zkontrolována konzistence odpovědí. U menší části domácností se objevuje nesoulad mezi uvedeným příjmem domácnosti a individuálními příjmy. U 433 domácností (13 %) tvoří součet individuálních příjmů (za respondenty individuálního dotazníku) částku, která je vyšší než příjem domácnosti. Průměrný rozdíl částek u domácností s vyšším součtem individuálních příjmů, než je uvedený příjem domácnosti, činí přibližně 10 400 Kč, medián rozdílu pak přibližně 6 500 Kč.

Kontrola chybějících odpovědí

U individuálních dotazníků i dotazníků pro domácnost byl zkontrolován podíl odpovědí „nevím“ a odmítnutí ze všech položených otázek. Cílem této analýzy bylo identifikovat dotazníky, v nichž respondenti poskytli nízký počet odpovědí a tyto dotazníky vyřadit jako nekvalitní.

U dotazníků pro domácnost dosahoval průměrný podíl chybějících odpovědí 1,1 %. Pouze dvacet respondentů dotazníku pro domácnost zvolilo odpovědi „nevím“ či odmítnutí u 20 a více % otázek. Maximální podíl chybějících odpovědí činil 28 %.

U individuálního dotazníku činil průměrný podíl chybějících odpovědí 1,8 %. Alespoň 20 % chybějících odpovědí se objevilo pouze u čtyř respondentů individuálního dotazníku, alespoň 10 % chybějících odpovědí u 52 respondentů. Maximální podíl chybějících odpovědí představoval 32 %.

Tyto podíly jsou stále poměrně nízké a i dotazníky s největším podílem chybějících odpovědí stále obsahují většinu dat. Žádné dotazníky nebyly proto z tohoto důvodu vyřazeny.

Výpočet odvozených proměnných

Do dat byly dále doplněny odvozené proměnné v souladu s postupy v předchozích ročnících. Vedle odvozených proměnných vytvářených již pro předchozí vlny bylo připraveno několik nových odvozených proměnných. Odvozené proměnné obsažené v datovém souboru jsou podrobněji popsány v kapitole 5. *Datový soubor*.

Harmonizace kódů odpovědí mezi vlnami

U vybraných otázek byla v dotazníku šestého ročníku upravena nabídka odpovědí ve srovnání s minulými ročníky, aby lépe odpovídala aktuálním podmínkám v době sběru dat. Jednalo se o

následující proměnné: *sfield* (obor studia), *partywv* (stranické preference), *cborn* (země narození), *nation* (národnost), *relig* (náboženské vyznání), *hincome* (kategorie příjmu domácnosti), *income* (kategorie příjmu jednotlivce). Při zpracování dat byly upraveny původní číselné kódy možností tak, aby odpovídaly kódům stejných variant odpovědí z předchozích vln. Novým možnostem odpovědí byly přiděleny číselné kódy navazující na dosavadní číselnou řadu kódů, aby nedocházelo k duplicitnímu zakódování více možností.

Kódování otevřených a polootevřených otázek

Textové odpovědi v polootevřených otázkách byly v relevantních případech převedeny do již existujících možností. Zbývající textové odpovědi byly upraveny tak, aby neobsahovaly pravopisné a gramatické chyby a aby popis stejné skutečnosti byl textově sjednocen (např. různé varianty popisu téhož náboženství či církve v otázce na náboženskou příslušnost *relig* byly upraveny na jedno znění).

Proměnné s popisem povolání respondenta a jeho rodiče byly zakódovány do klasifikace CZ-ISCO, která odpovídá mezinárodní klasifikaci zaměstnání ISCO-08. Data nejprve prošla finálním zakódováním jedním kódem a poté byly všechny kódy zkontrolovány druhým kódem a v případě potřeby byl navržen vhodnější kód.

Nastavení chybějících odpovědí

Odpovědi „Nevím“ a „Odmítnutí“ jsou kódovány pomocí hodnot 88 / 888 888, respektive 99 / 999 999. Tyto hodnoty jsou v SPSS souboru nastaveny jako user-missing. Výjimkou jsou otázky, v nichž byly odpovědi typu „Nevím“ začleněny do standardní nabídky (*materho*, *prodho*, *chmany*).

Odpovědi chybějící v důsledku technické ztráty dat nebo čištění odlehlých či nekonzistentních pozorování jsou kódovány pomocí hodnoty 98 / 999 998. Tyto hodnoty jsou v SPSS souboru nastaveny jako user-missing.

Případy, kdy respondent nespadal do populace otázky, a nebyl na ni proto dotázán, jsou kódovány pomocí hodnot 97 / 999 997. Tyto hodnoty jsou v SPSS souboru nastaveny jako user-missing.

Ostatní případy, kdy se otázka nebo odvozená proměnná na respondenta z různých důvodů nevztahuje či nebylo možné provést doplnění odpovědi v nedotazované otázce (podrobnosti jsou uvedeny v části 6.4. *Imputace v datech*), jsou kódovány pomocí hodnot 96 / 999 996. Tyto hodnoty nejsou v SPSS souboru nastaveny jako user-missing s výjimkou kódů pro označení případů, kdy nebylo možné provést imputaci.

6.4. Imputace v datech

Doplnění údajů ze současného ročníku

Do vybraných proměnných byly doplněny (imputovány) hodnoty na základě odpovědí v jiných otázkách. Jedná se převážně o proměnné vztahující se k rodině původu v 16 letech respondenta z individuálního dotazníku, které nebyly dotazovány u respondentů mladších 18 let.

Dále byly do proměnných s kategoriemi příjmu, které byly dotazovány v případě, že respondent neznal nebo odmítl sdělit číselný příjem, doplněny odpovídající kategorie uvedené částky příjmu. Tato úprava byla provedena u kategorií příjmu domácnosti (*hincomeN*) a kategorií individuálního příjmu (*incomeN*). Do částky individuálního příjmu (*aincome*) byla doplněna částka příjmu domácnosti u respondentů z jednočlenných domácností, u nichž nebyla otázka na individuální příjem pokládána.

Všechny tyto imputace jsou **popsány v příslušných proměnných v dokumentu s dotazníkem**. Imputace je popsána v bloku se světle oranžovým podbarvením. Samostatně je popsáno doplňování údajů ze současného ročníku šetření a z předchozích vln. Uvedeno je, ve kterých případech byla imputace prováděna, a které údaje pro ni byly využity.

Tabulka 39: Proměnné, u nichž byly doplňovány hodnoty dle jiných proměnných z šesté vlny

Proměnná	Obsah proměnné	Popis doplnění (imputace) hodnoty z šestého ročníku výzkumu CHPS
<i>hincomeN</i>	Kategorie příjmu domácnosti	U domácností, které uvedly číselný příjem (<i>ahincome</i>), byla do kategorické proměnné <i>hincomeN</i> doplněna odpovídající kategorie uvedené částky.
<i>aincome</i>	Příjem jednotlivce	U jednočlenných domácností byla do individuálního příjmu doplněna částka příjmu domácnosti (respondenti individuálního dotazníku z jednočlenných domácností nebyli na příjem dotazováni).
<i>incomeN</i>	Kategorie příjmu jednotlivce	U respondentů, kteří uvedli číselný příjem (<i>aincome</i>), byla do kategorické proměnné <i>incomeN</i> doplněna odpovídající kategorie uvedené částky.
<i>pbooks</i>	Počet knih doma v dětství	U respondentů mladších 16 let byla doplněna hodnota proměnné <i>books</i> z dotazníku pro domácnost.
<i>feduc</i>	Vzdělání otce	U respondentů žijících s otcem v jedné domácnosti byla doplněna rekódovaná hodnota proměnné <i>educ</i> otce respondenta z dotazníku pro domácnost.
<i>meduc</i>	Vzdělání matky	U respondentů žijících s matkou v jedné domácnosti byla doplněna rekódovaná hodnota proměnné <i>educ</i> matky respondenta z dotazníku pro domácnost.

Proměnná	Obsah proměnné	Popis doplnění (imputace) hodnoty z šestého ročníku výzkumu CHPS
<i>festat</i>	Ekonomické postavení otce v dětství	U respondentů mladších 18 byla doplněna hodnota proměnné <i>estat</i> otce respondenta z dotazníku pro domácnost.
<i>fisco08</i>	Kód povolání otce v dětství	U respondentů mladších 18 let byla doplněna hodnota proměnné <i>isco08</i> otce respondenta z individuálního dotazníku.
<i>fsvise</i>	Zodpovědnost otce za vedení zaměstnanců v dětství	U respondentů mladších 18 let byla doplněna hodnota proměnné <i>svise</i> otce respondenta z individuálního dotazníku.
<i>fempl</i>	Zaměstnávání dalších osob otcem v dětství	U respondentů mladších 18 let byla doplněna hodnota proměnné <i>empl</i> otce respondenta z individuálního dotazníku.
<i>fempls</i>	Počet zaměstnanců otce v dětství	U respondentů mladších 18 let byla doplněna rekódovaná hodnota proměnné <i>empls</i> otce respondenta z individuálního dotazníku.
<i>mestat</i>	Ekonomické postavení matky v dětství	U respondentů mladších 18 byla doplněna hodnota proměnné <i>estat</i> matky respondenta z dotazníku pro domácnost.
<i>misco08</i>	Kód povolání matky v dětství	U respondentů mladších 18 let byla doplněna hodnota proměnné <i>isco08</i> matky respondenta z individuálního dotazníku.
<i>msvise</i>	Zodpovědnost matky za vedení zaměstnanců v dětství	U respondentů mladších 18 let byla doplněna hodnota proměnné <i>svise</i> matky respondenta z individuálního dotazníku.
<i>mempl</i>	Zaměstnávání dalších osob matkou v dětství	U respondentů mladších 18 let byla doplněna hodnota proměnné <i>empl</i> matky respondenta z individuálního dotazníku.
<i>mempls</i>	Počet zaměstnanců matky v dětství	U respondentů mladších 18 let byla doplněna rekódovaná hodnota proměnné <i>empls</i> matky respondenta z individuálního dotazníku.
<i>chplace</i>	Bydliště v 16 letech	U respondentů mladších 18 let byla doplněna hodnota 1.
<i>chcounty, chmunic</i>	Okres a obec bydliště v 16 letech	U respondentů mladších 18 let byly doplněny hodnoty proměnných <i>county</i> a <i>munic</i> z dotazníku pro domácnost. U respondentů, kteří žili v dětství ve stejné obci jako v současnosti, byly doplněny hodnoty proměnných <i>county</i> a <i>munic</i> z dotazníku pro domácnost.
<i>chbtype</i>	Typ zástavby v 16 letech	U respondentů mladších 18 let byla doplněna hodnota proměnné <i>btype</i> z dotazníku pro domácnost.

Proměnná	Obsah proměnné	Popis doplnění (imputace) hodnoty z šestého ročníku výzkumu CHPS
<i>charsafe</i> , <i>charrep</i> , <i>charatt</i>	Okolí bydliště v 16 letech	U respondentů mladších 18 let byly doplněny hodnoty proměnných <i>arsafe</i> , <i>arrep</i> , <i>aratt</i> z dotazníku pro domácnost.
<i>wcounty</i> , <i>wmunic</i>	Okres a obec současné či minulé práce	U respondentů, kteří pracují/pracovali ve stejné obci jako v současnosti, byly doplněny hodnoty proměnných <i>county</i> a <i>munic</i> z dotazníku pro domácnost.

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Doplnění údajů z přechozích ročníků

Respondenti, kteří vyplnili individuální dotazník již v některé z předchozích vln, nebyli v šesté vlně dotazováni na stabilní či retrospektivní charakteristiky, o kterých vypovídali již v minulosti. Tyto odpovědi z minulých vln byly při zpracování dat **doplněny do příslušných proměnných**, aby datový soubor obsahoval odpovědi všech respondentů z cílové populace otázky.

V tomto aspektu zpracování se datový soubor z šestého ročníku odlišuje od datových souborů z předchozích vln. V nich nebyly z minulosti imputovány hodnoty stabilních či retrospektivních charakteristik, dotazovaných pouze u nových respondentů individuálního dotazníku (odpovědi v těchto proměnných v datech z minulých vln mají pouze respondenti, kteří v dané vlně vyplňovali individuální dotazník poprvé).

Řada retrospektivních údajů byla doplněna z vlny, v níž respondent poprvé vyplňoval individuální dotazník. Tato vlna je identifikována v proměnné *capi_first*. V menším množství případů neposkytli respondenti při prvním vyplňování individuálního dotazníku odpověď na danou otázku (byla zvolena odpověď „nevím“ či odmítnutí). U těchto respondentů byla doplněna odpověď z prvního proxy dotazníku, pokud byl za respondenta v minulosti vyplněn. Tyto případy však byly výjimečné.

Doplňující informace o imputaci do proměnných týkajících se práce respondenta (*isco08* až *fsize*) obsahují proměnné *job_info* a *job_wave*. V proměnné *job_info* je rozlišeno, zda se uvedené informace týkají respondentovy současné nebo poslední práce (respondenti, kteří v době dotazování nepracovali, byli dotazováni na svoji poslední práci). Proměnná *job_wave* informuje, ze které vlny byly údaje v těchto proměnných čerpány. Podobným způsobem jsou připraveny pomocné ukazatele k proměnným *mary*, *livey* a *cohy*.

Všechny tyto imputace jsou **popsány v příslušných proměnných v dokumentu s dotazníkem**. Základní informace k popisu imputací v dotazníku jsou obsaženy v předchozí sekci.

Tabulka 40: Proměnné, do kterých byla doplněna hodnota ze starší vlny

Proměnná	Obsah proměnné	Popis doplnění (imputace) hodnoty z předchozích ročníků CHPS
<i>sfield</i>	Obor studia	Byla imputována hodnota z vlny, v níž byl respondent naposledy dotázán na obor studia (ať už současného studia, nebo nejvyššího dosaženého vzdělání). Tato vlna je identifikována v odvozené proměnné <i>sfield_wave</i> .
<i>isco08</i>	Kód povolání z klasifikace CZ-ISCO	Byla doplněna hodnota z poslední vlny, v níž byl aktuálně nepracující respondent dotazován na popis pracovní činnosti (ať už na současnou, nebo minulou práci). Tato vlna je identifikována v odvozené proměnné <i>job_wave</i> . Byla doplněna hodnota z poslední vlny, v níž byl aktuálně pracující respondent dotazován na popis pracovní činnosti, pokud tento popis v současné vlně potvrdil.
<i>svise</i>	Zodpovědnost za vedení zaměstnanců	U aktuálně nepracujících respondentů (<i>job_info</i> = 2) byla doplněna hodnota z vlny, z níž byl přejat popis práce (<i>job_wave</i>).
<i>subor</i>	Počet podřízených	U aktuálně nepracujících respondentů (<i>job_info</i> = 2) byla doplněna hodnota z vlny, z níž byl přejat popis práce (<i>job_wave</i>).
<i>empl</i>	Zaměstnávání jiných osob	U aktuálně nepracujících respondentů (<i>job_info</i> = 2) byla doplněna hodnota z vlny, z níž byl přejat popis práce (<i>job_wave</i>).
<i>empls</i>	Počet zaměstnanců	U aktuálně nepracujících respondentů (<i>job_info</i> = 2) byla doplněna hodnota z vlny, z níž byl přejat popis práce (<i>job_wave</i>).
<i>fsize</i>	Velikost firmy	U aktuálně nepracujících respondentů (<i>job_info</i> = 2) byla doplněna hodnota z vlny, z níž byl přejat popis práce (<i>job_wave</i>).
<i>mary</i>	Rok sňatku	Byla doplněna hodnota z vlny, v níž respondent poprvé odpovídal na individuální dotazník, pokud má v současné šesté vlně za manžela/manželku (registrovaného partnera / registrovanou partnerku) stále stejnou osobu jako v dané vlně (případně se manžel/manželka aktuálně nenachází v domácnosti). U opakovaně dotazovaných respondentů byla informace odvozena z odpovědí v otázkách na meziroční zněny rodinného stavu ve druhé až šesté vlně. Pokud nebyla a odpověď k dispozici u respondenta (respondent neodpovídal na otázku, nevěděl či odmítl), byla doplněna odpověď partnera. Typ imputace je značen v proměnné <i>mary_info</i> . Vlna, z níž byl údaj čerpán, je zaznamenána v proměnné <i>mary_wave</i> .

Proměnná	Obsah proměnné	Popis doplnění (imputace) hodnoty z předchozích ročníků CHPS
<i>livey</i>	Rok začátku soužití (sezdání)	Byla doplněna hodnota z první vlny (ve vlnách 2 až 5 nebyla otázka dotazována), pokud má respondent v současné šesté vlně za manžela/manželku (registrovaného partnera / registrovanou partnerku) stále stejnou osobu jako v první vlně (případně se manžel/manželka aktuálně nenachází v domácnosti). V dalších vlnách odvozováno z odpovědí v otázkách na meziroční změny struktury domácností (<i>newy a newre</i>). Pokud nebyla a odpověď k dispozici u respondenta (respondent neodpovídal na otázku, nevěděl či odmítl), byl doplněn údaj respondenta z příbuzných proměnných (<i>cohy</i>) nebo údaj partnera (<i>livey, cohy, newy a newre</i>). Typ imputace je značen v proměnné <i>livey_info</i>. Vlna, z níž byl údaj čerpán, je zaznamenána v proměnné <i>livey_wave</i>.
<i>cohy</i>	Rok začátku soužití (nesezdání)	Byla imputována hodnota z první vlny (ve vlnách 2 až 5 nebyla otázka dotazována), pokud má respondent v současné šesté vlně za partnera/partnerku stále stejnou osobu jako v první vlně (případně se manžel/manželka aktuálně nenachází v domácnosti). V dalších vlnách odvozováno z odpovědí v otázkách na meziroční změny struktury domácností (<i>newy a newre</i>). Pokud nebyla a odpověď k dispozici u respondenta (respondent neodpovídal na otázku, nevěděl či odmítl), byl doplněn údaj respondenta z příbuzných proměnných (<i>newy</i>) nebo údaj partnera (<i>cohy, newy</i>). Typ imputace je značen v proměnné <i>cohy_info</i>. Vlna, z níž byl údaj čerpán, je zaznamenána v proměnné <i>cohy_wave</i>.
<i>lipar</i>	Vyrůstání s biologickými rodiči	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník (případně z prvního proxy dotazníku, pokud odpověď nebyla k dispozici v individuálním dotazníku).
<i>pbooks</i>	Počet knih doma v dětství	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník (případně z prvního proxy dotazníku, pokud odpověď nebyla k dispozici v individuálním dotazníku).
<i>feduc</i>	Vzdělání otce	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník (případně z prvního proxy dotazníku, pokud odpověď nebyla k dispozici v individuálním dotazníku).
<i>meduc</i>	Vzdělání matky	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník (případně z prvního proxy dotazníku, pokud odpověď nebyla k dispozici v individuálním dotazníku).

Proměnná	Obsah proměnné	Popis doplnění (imputace) hodnoty z předchozích ročníků CHPS
<i>festat</i>	Ekonomické postavení otce v dětství	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník (případně z prvního proxy dotazníku, pokud odpověď nebyla k dispozici v individuálním dotazníku).
<i>fisco08</i>	Kód povolání otce v dětství	Byla doplněna odpověď z vlny, z níž byla čerpána imputovaná hodnota proměnné <i>festat</i> .
<i>fsvise</i>	Zodpovědnost otce za vedení zaměstnanců v dětství	Byla doplněna odpověď z vlny, z níž byla čerpána imputovaná hodnota proměnné <i>festat</i> .
<i>fempl</i>	Zaměstnávání dalších osob otcem v dětství	Byla doplněna odpověď z vlny, z níž byla čerpána imputovaná hodnota proměnné <i>festat</i> .
<i>fempls</i>	Počet zaměstnanců otce v dětství	Byla doplněna odpověď z vlny, z níž byla čerpána imputovaná hodnota proměnné <i>festat</i> .
<i>mestat</i>	Ekonomické postavení matky v dětství	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník (případně z prvního proxy dotazníku, pokud odpověď nebyla k dispozici v individuálním dotazníku).
<i>misco08</i>	Kód povolání matky v dětství	Byla doplněna odpověď z vlny, z níž byla čerpána imputovaná hodnota proměnné <i>mestat</i> .
<i>msvise</i>	Zodpovědnost matky za vedení zaměstnanců v dětství	Byla doplněna odpověď z vlny, z níž byla čerpána imputovaná hodnota proměnné <i>mestat</i> .
<i>mempl</i>	Zaměstnávání dalších osob matkou v dětství	Byla doplněna odpověď z vlny, z níž byla čerpána imputovaná hodnota proměnné <i>mestat</i> .
<i>mempls</i>	Počet zaměstnanců matky v dětství	Byla doplněna odpověď z vlny, z níž byla čerpána imputovaná hodnota proměnné <i>mestat</i> .
<i>cborn, corno</i>	Země narození	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník.
<i>cmcr</i>	Rok příchodu do ČR	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník.
<i>citcr, citcb, citoc, citot, citdk, citref</i>	Občanství	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník.
<i>nation, nationo</i>	Národnost	Byla doplněna odpověď z vlny, v níž respondent poprvé vyplňoval individuální dotazník.

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Případy, kdy doplnění údajů nebylo možné

V menším počtu případů nebylo možné u respondentů doplnit údaje z jiných otázek v současné vlně či ze stejné proměnné v minulých vlnách.

U 33 respondentů z jednočlenných domácností nebylo možné převzít částku individuálního příjmu z příjmu domácnosti, protože v dotazníku pro domácnost byla poskytnuta pouze odpověď ve formě příjmové skupiny, kterou nebylo možné převést na kategorie individuálního příjmu (nabídka kategorií u příjmu domácnosti a jednotlivce se v dotaznících lišila). U těchto respondentů je v datech uveden kód 96 s popisem „Není k dispozici“.

Imputace údaje o rodičích u respondentů mladších 18 let nebyla v některých případech možná, protože daný rodič se nenacházel v domácnosti či nevyplnil individuální dotazník. V těchto případech je v datech uveden kód 96 / 999 996 s popisem „Údaj nelze doplnit“.

Údaje z minulosti nebylo možné převzít v případech, kdy respondent vyplnil poprvé individuální dotazník v pátém ročníku, v němž nebyly některé otázky položeny (zejména dotazy na rodinu původu).

Počty respondentů, u nichž nebylo možné v jednotlivých otázkách údaje doplnit, jsou uvedeny v následující tabulce.

Tabulka 41: Počty případů, kdy nebylo možné provést imputaci ze současné či minulé vlny

Proměnná	Popis proměnné	Počet	Důvod
<i>sfield</i>	Obor studia	25	Individuální dotazník vyplněn poprvé v páté vlně, v níž otázka nebyla zařazena
<i>mary</i>	Rok sňatku	15	Více možných specifických důvodů (zejména individuální dotazník vyplněn poprvé v první vlně, kdy se manžel/ manželka zároveň nenacházel/a v domácnosti, a otázka <i>mary</i> nebyla proto položena)
<i>livey</i>	Rok začátku společného soužití	13	Viz výše důvody u proměnné <i>mary</i> výše
<i>lipar</i>	Žil(a) s biologickými rodiči do 16 let	29	Individuální dotazník vyplněn poprvé v páté vlně, v níž otázka nebyla zařazena
<i>pbooks</i>	Počet knih v domácnosti v 16 letech	37	Individuální dotazník vyplněn poprvé v páté vlně, v níž otázka nebyla zařazena
<i>feduc</i>	Vzdělání otce	27	Individuální dotazník vyplněn poprvé v páté vlně, v níž otázka nebyla zařazena
<i>meduc</i>	Vzdělání matky	24	Individuální dotazník vyplněn poprvé v páté vlně, v níž otázka nebyla zařazena
<i>festat</i>	Ekonomické postavení otce v 16 letech respondenta	68	Otec respondenta nežije v domácnosti
<i>fisco08</i>	Popis práce otce v 16 letech respondenta	146	Otec respondenta nežije v domácnosti nebo nevyplnil individuální dotazník

Proměnná	Popis proměnné	Počet	Důvod
<i>fsvise</i>	Vedení jiných zaměstnanců otcem v 16 letech respondenta	106	Otec respondenta nežije v domácnosti nebo nevyplnil individuální dotazník
<i>fempl</i>	Zaměstnávání dalších lidí otcem v 16 letech respondenta	19	Otec respondenta nežije v domácnosti nebo nevyplnil individuální dotazník
<i>mestat</i>	Ekonomické postavení matky v 16 letech respondenta	41	Matka respondenta nežije v domácnosti
<i>misco08</i>	Popis práce matky v 16 letech respondenta	145	Matka respondenta nežije v domácnosti nebo nevyplnila individuální dotazník
<i>msvise</i>	Vedení jiných zaměstnanců matkou v 16 letech respondenta	104	Matka respondenta nežije v domácnosti nebo nevyplnila individuální dotazník
<i>mempl</i>	Zaměstnávání dalších lidí matkou v 16 letech respondenta	5	Matka respondenta nežije v domácnosti nebo nevyplnila individuální dotazník
<i>incomeN</i>	Příjem jednotlivce v kategoriích	33	Jednočlenná domácnost, příjem domácnosti uveden pouze jako kategorie, která není převoditelná na kategorie individuálního příjmu

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

6.5. Anonymizace dat

Během sběru dat byly zpracovávány vybrané osobní údaje respondentů za účelem kontaktování domácností, provádění kontrol práce tazatelů (náslechové kontroly, zpětné terénní kontroly) a vyplacení odměny. Tyto údaje byly zpracovávány **odděleně od dat z dotazníku** a přístup k nim měli jen pověřeni pracovníci. Předávání souborů s osobními údaji mezi realizátorem sběru dat a Sociologickým ústavem AV ČR probíhalo pomocí zabezpečených kanálů. Respondenti byli seznámeni s podmínkami zpracování osobních údajů dle platné legislativy a udělili s ním souhlas.

Identifikační čísla domácností a jednotlivců používaná při sběru dat byla v datovém souboru **nahrazena novými identifikačními čísly**, která byla náhodně vygenerována bez spojitosti s původními identifikačními čísly. Stejně tak byla **nahrazena identifikační čísla tazatelů** (pro stejného tazatele je však stále používán stejný identifikátor v datech ze všech dosavadních vln).

Z datového souboru byla dále odstraněna část ukazatelů. Vyřazeny byly proměnné poskytující podrobnější informace o domácnostech a jejich členech, které by v kombinaci s dalšími proměnnými mohly potenciálně sloužit jako nepřímé identifikátory. Seznam těchto proměnných je poskytnut v podkapitole 5.1. *Obsah a struktura datového souboru.*

7. Post-stratifikační váhy

Systém vah

Pro data z šestého ročníku CHPS byly připraveny post-stratifikační váhy, které korigují distribuci souboru vzhledem k základním geografickým a sociodemografickým ukazatelům. Za účelem **základních deskriptivních analýz** celého souboru nebo podsouboru původních panelistů CHPS je doporučeno tyto váhy využít.

Post-stratifikační váhy byly vypočteny **jak pro celkový soubor, tak samostatně pro podsoubor** původních panelových domácností. Datový soubor tak obsahuje dvě skupiny vah – váhy pro celý dotázaný soubor a váhy pro podsoubor domácností z původního panelu CHPS (přípona „p“).

Pro analýzu na úrovni domácností (do analýzy vstupuje každá domácnost jen jednou) jsou připraveny váhy pro domácnosti (W_{hh_hh} , $W_{hh_hh_p}$). K této analýze mohou být použity proměnné z domácnostního dotazníku na úrovni domácnosti, tak agregace individuálních charakteristik na domácnost. Další váhy slouží pro analýzu všech členů domácností, tedy nejen respondentů individuálního dotazníku (W_{hh_indi} , $W_{hh_indi_p}$). Využit je možné individuální charakteristiky jednotlivců z dotazníku pro domácnost či otázky týkající se domácnosti převedené na úroveň jednotlivců. Analýzy pracující s proměnnými z individuálního dotazníku pak vyžadují váhy pro respondenty individuálního dotazníku (W_{indi} , W_{indi_p}).

Váhy byly sestaveny obdobným způsobem jako v předchozích ročnících. Datový soubor opět **obsahuje několik vážících proměnných**, protože dotazování se skládalo z více instrumentů s různými cílovými populacemi a skupinami respondentů. V souvislosti se zjednodušením struktury dotazovacích instrumentů v šestém ročníku však zároveň došlo ke **zjednodušení systému vah**.

Binární proměnné určující vyplnění instrumentů

Pro výběr jednotlivých pozorování, u nichž je vyplněný daný instrument a vztahují se k nim váhy pro tento instrument, slouží proměnné hh_hh (dotazník pro domácnost, proměnné na úrovni domácnosti), hh_indi (dotazník pro domácnost, proměnné na úrovni jednotlivce) a $indi$ (individuální dotazník). Tyto proměnné určují, zda člen domácnosti vyplnil daný instrument (resp. jsou pro něj dostupné tyto údaje) – pokud ano, proměnné nabývají hodnoty 1, všechny ostatní případy mají přiřazenou hodnotu 0. Váhy jsou spočteny jen pro pozorování s hodnotou 1 v příslušném binárním ukazateli. Uvedené proměnné vypovídají o celém souboru domácností a jednotlivců. Analogicky jsou připraveny proměnné s příponou „p“, které se váží k souboru původních domácností (u jednotlivců a domácností z doplňujícího souboru nabývají vždy hodnoty 0).

Přehled binárních indikátorů a příslušných vah obsahuje následující tabulka. Bývalí členové domácnosti mají v ukazatelích vyplnění instrumentu hodnotu 0 a vážící proměnné pro ně nejsou spočteny.

Tabulka 42: Přehled ukazatelů vyplnění jednotlivých dotazníků a příslušných vah

Proměnná	Popis
<i>hh_hh</i>	Člen domácnosti, který vyplnil dotazník pro domácnost – celý soubor
<i>W_hh_hh</i>	Váha pro domácnosti – celý soubor
<i>hh_hh_p</i>	Člen domácnosti, který vyplnil dotazník pro domácnost – původní panel CHPS
<i>W_hh_hh_p</i>	Váha pro domácnosti – původní panel CHPS
<i>hh_indi</i>	Člen dotázané domácnosti (tj. všichni členové kromě bývalých) – celý soubor
<i>W_hh_indi</i>	Váha pro členy domácností – celý soubor
<i>hh_indi_p</i>	Člen dotázané domácnosti (tj. všichni členové kromě bývalých) – původní panel CHPS
<i>W_hh_indi_p</i>	Váha pro členy domácností – původní panel CHPS
<i>indi</i>	Člen domácnosti, který vyplnil individuální dotazník – celý soubor
<i>W_indi</i>	Váha pro respondenty individuálního dotazníku – celý soubor
<i>indi_p</i>	Člen domácnosti, který vyplnil individuální dotazník – původní panel CHPS
<i>W_indi_p</i>	Váha pro respondenty individuálního dotazníku – původní panel CHPS

Zdroj: České panelové šetření domácností, 6. ročník (2023/2024), vlastní zpracování

Postup vážení

Vážení bylo provedeno v programu R funkcí *rake* (survey package), která pomocí iterativní procedury vyrovnává vážené marginální distribuce výběrového souboru a známé populační distribuce. Výsledné váhy přesahující hodnotu 4 byly změněny na hodnotu 4 (funkce *trim*). Váhy vyšší než 4 se v šestém ročníku objevily u proměnných *W_indi* (váha pro individuální dotazník v celém souboru) a *W_indi_p* (váha pro individuální dotazník v souboru původních panelistů).

Pro vážení členů domácnosti (dotazník pro domácnost, úroveň jednotlivce) a respondentů individuálního dotazníku byla využita cross-klasifikace pohlaví, vzdělání a věku a samostatně kraj bydliště (na úrovni jednotlivců). Pro vážení dotazníků pro domácnost byl využit pouze kraj bydliště (na úrovni domácností).

Kontrolní distribuce pro vážení byly získány z následujících zdrojů:

- Výběrové šetření pracovních sil 2022, roční vážná data (cross-klasifikace pohlaví, vzdělání a věku), využity byly vážené distribuce z ročního souboru,
- Sčítání lidu, domů a bytů (SLBD) 2021 (kraj na úrovni domácností),
- údaje Českého statistického úřadu o věkovém složení obyvatel k 1. červenci 2023 (kraj na úrovni jednotlivců).

Pro vážení dat z dotazníku pro domácnost na úrovni jednotlivce byly využity distribuce pro populaci bez věkového omezení. Pro vážení dat z individuálního dotazníku byly využity distribuce pro populaci ve věku 15 a více let.

Výsledky vážení

U všech skupin (domácnosti, členové domácností, respondenti individuálního dotazníku) platí, že **celkový soubor vykazuje převážně velmi nízké rozdíly** oproti populačním distribucím a vyžaduje poměrně malé korekce pomocí post-stratifikačních vah. Největší rozdíly se u celkového souboru projevují u rozdělení věkových kategorií členů domácnosti. Oproti populaci jsou v rozsahu dvou až čtyř procentních bodů nadhodnoceny děti do 17 let a naopak méně jsou zastoupeny osoby ve věku 55 a více let. Tyto odchylky vznikají v důsledku zaměření doplňujícího souboru domácností na domácnosti do 55 let a rodiny s dětmi.

Zejména v podsouboru původních panelových domácností jsou největší korekce post-stratifikačními vahami vyžadovány u respondentů individuálního dotazníku, jejichž rozdělení se od populačního odlišuje více, než je tomu u domácností a všech jejich členů. Tato tendence se projevovala i u všech předchozích ročníků šetření.

Na úrovni **domácností** korigují postratifikační váhy v podsouboru původních domácností zejména nižší zastoupení Středočeského kraje (rozdíl 5 procentních bodů) a naopak vyšší zastoupení Jihočeského (3,3 procentního bodu), Ústeckého (2,8 procentního bodu) a Jihomoravského kraje (2,7 procentního bodu). Doplňující domácnosti byly vybírány tak, aby vyrovnávaly regionální strukturu mezi všemi dotázanými domácnostmi, v celém souboru tak odchylky nevážených dat nepřesahují jeden procentní bod a vliv post-stratifikačního vážení je pro celý soubor poměrně nízký.

Tabulka 43: Distribuce proměnných pro vážení (%) – dotazník pro domácnost, úroveň domácnosti

Proměnná	Populace	Celý soubor			Původní panel		
		Neváženě	Rozdíl	Váženě	Neváženě	Rozdíl	Váženě
Kraj (region)							
Hlavní město Praha (1)	13,7	13,6	-0,1	13,7	12,2	-1,5	13,7
Středočeský kraj (2)	12,7	12,5	-0,2	12,7	7,7	-5,0	12,7
Jihočeský kraj (3)	5,9	6,3	0,4	5,9	9,1	3,2	5,9
Plzeňský kraj (4)	5,5	5,6	0,1	5,5	4,9	-0,6	5,5
Karlovarský kraj (5)	2,8	2,3	-0,5	2,8	1,8	-1,0	2,8
Ústecký kraj (6)	7,8	8,5	0,7	7,8	10,6	2,8	7,8
Liberecký kraj (7)	4,2	4,3	0,1	4,2	3,7	-0,5	4,2
Královéhradecký kraj (8)	5,1	5,8	0,7	5,1	7,2	2,1	5,1
Pardubický kraj (9)	4,7	4,5	-0,2	4,7	3,8	-0,9	4,7
Kraj Vysočina (10)	4,4	4,4	0,0	4,4	3,9	-0,5	4,4

Proměnná	Populace	Celý soubor			Původní panel		
		Neváženě	Rozdíl	Váženě	Neváženě	Rozdíl	Váženě
Jihomoravský kraj (11)	11	10,9	-0,1	11,0	13,7	2,7	11,0
Olomoucký kraj (12)	5,8	5,8	0,0	5,8	5,9	0,1	5,8
Zlínský kraj (13)	5,1	4,5	-0,6	5,1	4,3	-0,8	5,1
Moravskoslezský kraj (14)	11,2	10,9	-0,3	11,2	11,2	0,0	11,2

Zdroj: Věkové složení obyvatel k 1. 7. 2023, Český statistický úřad (cílová distribuce kraje); České panelové šetření domácností, 6. ročník (2023/2024), 3 297 domácnost v celém souboru, z toho 1 807 domácností z původního panelu, vlastní zpracování

U **současných členů dotázaných domácností** se v neváženém souboru projevují z hlediska krajů stejné tendence jako v souboru domácností. Celkový soubor vykazuje rozdíly do jednoho procentního bodu, zatímco v podsouboru původních panelistů je přibližně o pět procentních bodů méně zastoupen Středočeský kraj a naopak více zastoupeny jsou Jihočeský, Jihomoravský a Ústecký kraj. Vliv vážení je tak výraznější u členů původních domácností.

Zastoupení mužů a žen v celém souboru se pouze minimálně liší od populačního, změna distribucí způsobená vahami je tak minimální. V podsouboru původních panelových domácností korigují váhy mírně vyšší zastoupení žen (o dva procentní body ve srovnání s populací).

Výraznější odchylky od populačního rozdělení se ve skupině současných členů domácností projevují u vzdělání a věku. V podsouboru původních panelistů jsou o šest procentních bodů více zastoupeny osoby s vysokoškolským vzděláním, naopak skupiny se středoškolským vzděláním bez maturity a nižším vzděláním jsou oproti populaci podhodnoceny každá o tři procentní body. V celém souboru je u vzdělání třeba nižší korekce (rozdíly zde dosahují maximální výše 1,3 procentního bodu).

V podsouboru původních panelistů upravují váhy vyšší zastoupení osob ve věku 55 a více let (odchylka 6,4 procentních bodů) a naopak nižší u dospělých od 18 do 34 (5,5 procentních bodů). Tyto rozdíly jsou logickým důsledkem postupného stárnutí původního panelu domácností, který byl založen v roce 2015. V celkovém souboru jsou mezi všemi současnými členy domácností více zastoupeny děti do 17 let (sběr na doplňujících domácnostech se soustředil zejména na rodiny s dětmi). Naopak méně jsou v celkovém souboru členů domácností přítomni lidé nad 55 let (3,8 procentních bodů).

Tabulka 44: Distribuce proměnných pro vážení (%) – dotazník pro domácnost, úroveň jednotlivce

Proměnná	Populace	Celý soubor			Původní panel		
		Neváženě	Rozdíl	Váženě	Neváženě	Rozdíl	Váženě
Kraj (region)							
Hlavní město Praha (1)	12,4	12,3	-0,1	12,4	11,9	-0,5	12,4
Středočeský kraj (2)	13,3	12,5	-0,8	13,3	8,0	-5,3	13,3
Jihočeský kraj (3)	6,0	6,8	0,8	6,0	10,1	4,1	6,0
Plzeňský kraj (4)	5,6	5,6	0,0	5,6	5,2	-0,4	5,6
Karlovarský kraj (5)	2,7	2,2	-0,5	2,7	1,6	-1,1	2,7
Ústecký kraj (6)	7,5	8,3	0,8	7,5	9,6	2,1	7,5
Liberecký kraj (7)	4,2	4,0	-0,2	4,2	3,4	-0,8	4,2
Královéhradecký kraj (8)	5,1	5,7	0,6	5,1	7,0	1,9	5,1
Pardubický kraj (9)	4,9	4,2	-0,7	4,9	3,2	-1,7	4,9
Kraj Vysočina (10)	4,8	5,1	0,3	4,8	4,6	-0,2	4,8
Jihomoravský kraj (11)	11,2	11,7	0,5	11,2	14,1	2,9	11,2
Olomoucký kraj (12)	5,9	6,0	0,1	5,9	6,0	0,1	5,9
Zlínský kraj (13)	5,4	4,8	-0,6	5,4	4,7	-0,7	5,4
Moravskoslezský kraj (14)	11,0	10,8	-0,2	11,0	10,7	-0,3	11,0
Pohlaví (sex)							
Muž (1)	49,3	49,2	-0,1	49,3	47,4	-1,9	49,3
Žena (2)	50,7	50,8	0,1	50,7	52,6	1,9	50,7
Vzdělání (educ)							
Bez vzdělání, základní vzdělání (0, 1, 2)	27,3	27,5	0,2	27,3	24,2	-3,1	27,3
Středoškolské (3 až 7)	55,5	54,2	-1,3	55,5	52,7	-2,8	55,5
Vyšší odborné a vysokoškolské (8 až 11)	17,1	18,3	1,2	17,1	23,1	6,0	17,1
Věk (age)							
Do 17 let	19,1	21,5	2,4	19,1	19,0	-0,1	19,1
18 až 34 let	18,2	18,0	-0,2	18,2	12,7	-5,5	18,2
35 až 54 let	29,9	31,5	1,6	29,9	29,1	-0,8	29,9
55 a více let	32,8	29,0	-3,8	32,8	39,2	6,4	32,8

Zdroj: Věkové složení obyvatel k 1. 7. 2023, Český statistický úřad (cílová distribuce kraje); Výběrové šetření pracovních sil 2022, roční vážená data, Český statistický úřad (cílové distribuce pohlaví, vzdělání, věku); České panelové šetření domácností, 6. ročník (2023/2024), 8 481 členů domácností v celém souboru, z toho 4 406 členů z původního panelu, vlastní zpracování

V případě **respondentů individuálního dotazníku** platí stejně jako u souboru domácností a současných členů domácností, že korekce některých kategorií pomocí post-stratifikačních vah jsou třeba zejména u podsouboru původních panelových domácností, zatímco celkový soubor je poměrně vyvážený a výraznější vážení nevyžaduje.

U distribuce krajů a jejich korekce vahami platí stejné tendence jako u souboru domácností a členů domácností (méně je zastoupen Středočeský kraj, více pak kraje Jihočeský, Ústecký, Jihomoravský).

V podsouboru původních panelových domácností odpovídaly na individuální dotazník častěji ženy a jejich proporce je ve srovnání s populací o 7,6 procentních bodů vyšší. Dotazování v doplňujícím souboru domácností se proto zaměřilo výrazněji na muže, v celkovém souboru odchylka u pohlaví činí pouze 1,5 procentního bodu.

Mezi původními panelisty se individuálního dotazování účastnily výrazně častěji osoby s vysokoškolským vzděláním (rozdíl 8,2 procentních bodů). Naopak nižší šance účasti se projevila u středoškoláků s maturitou (rozdíl tři procentní body) a ještě nižší u osob se základním vzděláním a středoškolským bez maturity (5,1 procentních bodů). V celém souboru je stále vyšší podíl osob s vysokoškolským vzděláním než v populaci a méně osob s maximálně středoškolským vzděláním bez maturity, ale rozdíly jsou poměrně nízké (1,7 procentního bodu).

Nejvýraznější odchylky nevážených dat se u respondentů individuálního dotazníku projevují u věkových kategorií. Osoby ve věku 55 a více let jsou přibližně o 11 procentních bodů nadhodnoceny, mladší dospělí do 34 let jsou naopak stejnou měrou podhodnoceni. Efekt post-stratifikačního vážení je tak největší v případě věkových kategorií.

Tabulka 45: Distribuce proměnných pro vážení (%) – individuální dotazník

Proměnná	Populace	Celý soubor			Původní panel		
		Neváženě	Rozdíl	Váženě	Neváženě	Rozdíl	Váženě
Kraj (region)							
Hlavní město Praha (1)	12,4	13,2	0,8	12,5	12,2	-0,2	12,4
Středočeský kraj (2)	13,3	11,7	-1,3	13,0	7,9	-5,1	13,0
Jihočeský kraj (3)	6,0	7,5	1,5	6,1	10,0	4,0	6,0
Plzeňský kraj (4)	5,6	5,6	0,0	5,6	5,2	-0,4	5,6
Karlovarský kraj (5)	2,7	2,2	-0,5	2,7	1,7	-1,0	2,7
Ústecký kraj (6)	7,5	8,7	1,1	7,6	10,4	2,8	7,6
Liberecký kraj (7)	4,2	3,9	-0,2	4,2	3,3	-0,8	4,1
Královéhradecký kraj (8)	5,1	6,3	1,1	5,2	7,4	2,2	5,2
Pardubický kraj (9)	4,9	4,4	-0,5	4,9	3,8	-1,1	4,9
Kraj Vysočina (10)	4,8	4,6	-0,2	4,8	4,3	-0,5	4,8
Jihomoravský kraj (11)	11,2	11,4	0,2	11,2	13,5	2,3	11,2
Olomoucký kraj (12)	5,9	5,8	-0,1	5,9	5,6	-0,3	5,9

Proměnná	Populace	Celý soubor			Původní panel		
		Neváženě	Rozdíl	Váženě	Neváženě	Rozdíl	Váženě
Kraj (region)							
Zlínský kraj (13)	5,4	4,5	-0,9	5,4	4,4	-1,0	5,4
Moravskoslezský kraj (14)	11,0	10,3	-0,8	11,1	10,2	-0,9	11,1
Pohlaví (sex)							
Muž (1)	49,3	47,4	-1,5	48,9	41,3	-7,6	48,9
Žena (2)	50,7	52,6	1,5	51,1	58,7	7,6	51,1
Vzdělání (educ)							
Bez vzdělání, základní vzdělání (0, 1, 2)	27,3	11,7	-1,7	13,3	8,3	-5,1	13,4
Středoškolské (3 až 7)	55,5	66,2	0,1	66,2	63,1	-3,0	66,1
Vyšší odborné a vysokoškolské (8 až 11)	17,1	22,1	1,7	20,5	28,6	8,2	20,4
Věk (age)							
Do 17 let	19,1	23,9	-1,4	25,4	13,9	-11,4	25,3
18 až 34 let	18,2	37,5	1,9	35,7	35,7	0,1	35,6
35 až 54 let	29,9	38,6	-0,5	39,0	50,4	11,3	39,1
55 a více let	32,8	13,2	0,8	12,5	12,2	-0,2	12,4

Zdroj: Věkové složení obyvatel k 1. 7. 2023, Český statistický úřad (cílová distribuce kraje); Výběrové šetření pracovních sil 2022, roční vážená data, Český statistický úřad (cílové distribuce pohlaví, vzdělání, věku); České panelové šetření domácností, 6. ročník (2023/2024), 4 459 respondentů individuálního dotazníku, z toho 1 697 respondentů z nových domácností, vlastní zpracování

Práce s chybějícími hodnotami

Kontrolní data z Výběrového šetření pracovních sil 2022 obsahovala několik chybějících hodnot v proměnné vzdělání. Data z výběrového souboru neobsahovala ve vážících proměnných žádné chybějící hodnoty.

Pro práci s chybějícími hodnotami byla použita stejná pravidla jako pro předchozí ročníky (podrobně jsou tato pravidla popsána v dokumentaci k prvnímu až čtvrtému ročníku šetření). V souladu s těmito pravidly byly chybějící hodnoty v kontrolních datech zanedbány (za předpokladu MCAR). Podíl chybějících hodnot v kontrolních datech nepřesahoval 0,01 %.

Příloha: Zveřejněné materiály k výzkumu

K 6. ročníku Českého panelového šetření domácností byl dosud (do data vydání této dokumentace) zveřejněn datový soubor, seznam proměnných, dotazník a karty, terénní materiály (dopisy a letáky, poučení účastníků, informovaný souhlas a smlouva o účasti) a podklady a skript k vážení.

Soubor se **seznamem proměnných** obsahuje přehled proměnných dostupných v datovém souboru z 6. ročníku se značením, ve kterých předchozích vlnách byly zařazeny. Soubor dále obsahuje pro **kódovací klíč** (přehled popisků proměnných a jejich hodnot a další parametry proměnných).

Všechny materiály jsou zveřejněny v online datovém katalogu Českého sociálněvědního datového archivu. Datový soubor bude dostupný od začátku roku 2026 (do této doby k němu mají přístup pouze členové projektu SYRI).

Tabulka 46: Zveřejněné materiály k 6. ročníku CHPS

Datové soubory
Datový soubor (spojená data z dotazníku pro domácnosti a individuálního dotazníku)
Seznam proměnných v datovém souboru, kódovací klíč (codebook)
Dokumentace
Dokumentace ke sběru a zpracování dat
Dotazníky
Předloha dotazníku s komentáři ke zpracování dat
Karty k dotazníku pro domácnost a individuálnímu dotazníku
Terénní materiály
Dopis pro původní domácnosti
Leták pro původní domácnosti
Dopis pro nové domácnosti
Leták pro nové domácnosti
Informace k účasti ve výzkumu (poučení účastníka)
Souhlas s účastí ve výzkumu a zpracováním osobních údajů
Vážení
R script pro post-stratifikační vážení
Populační (kontrolní) distribuce pro vážení

Seznam tabulek

Tabulka 1: Způsoby rekrutace domácností z doplňujícího souboru.....	15
Tabulka 2: Distribuce kvótních znaků a srovnání s populací – domácnosti	19
Tabulka 3: Distribuce kvótních znaků a srovnání s populací – jednotlivci.....	22
Tabulka 4: Přehled jednotlivých školení tazatelů s počty účastníků.....	24
Tabulka 5: Počty dotázaných domácností na tazatele	26
Tabulka 6: Skladba tazatelů pracujících na projektu.....	27
Tabulka 7: Počty dotázaných domácností a jednotlivců dle statusu domácnosti a módu dotazování	28
Tabulka 8: Počty zúčastněných původních domácností dle módu dotazování.....	29
Tabulka 9: Hlavní mezníky v harmonogramu sběru dat.....	30
Tabulka 10: Podíl (%) domácností dotázaných v jednotlivých měsících sběru dat.....	30
Tabulka 11: Distribuce dispozičních kódů v původním panelu.....	36
Tabulka 12: Míra způsobilosti a retence původních panelových domácností (%)	37
Tabulka 13: Frekvence přivydělávání si vyplňování dotazníků na internetu	42
Tabulka 14: Výsledky kontaktu v kontaktním formuláři.....	43
Tabulka 15: Kontrola nahrávek individuálních rozhovorů	47
Tabulka 16: Dokončené zpětné kontroly u respondentů individuálního dotazníku	48
Tabulka 17: Dotazovací nástroje využité v jednotlivých vlnách šetření CHPS	49
Tabulka 18: Míra vyplnění individuálního dotazníku dle statusu a typu domácnosti a módu sběru v %	51
Tabulka 19: Délka vyplňování dotazníků dle statusu domácnosti a módu dotazování	52
Tabulka 20: Využívání podmíněného dotazování	53
Tabulka 21: Otázky, u nichž došlo k úpravě parametrů mezi současnou a předchozími vlnami	57
Tabulka 22: Neplatné („únikové“) odpovědi s odlišnostmi od standardního zacházení	61
Tabulka 23: Struktura datového souboru	64
Tabulka 24: Proměnné nezařazené do datového souboru.....	65
Tabulka 25: Identifikační čísla domácností a jednotlivců	66
Tabulka 26: Identifikační čísla dalších osob v domácnosti	66
Tabulka 27: Proměnné popisující změnu složení domácnosti.....	69
Tabulka 28: Ukazatele účasti v dotazování v minulosti	70
Tabulka 29: Odvozené proměnné na úrovni domácnosti	71
Tabulka 30: Ukazatele struktury domácnosti	72
Tabulka 31: Odvozené proměnné na úrovni jednotlivce.....	73
Tabulka 32: Přehled stratifikačních proměnných.....	75
Tabulka 33: Geografické (lokalizační) proměnné	76
Tabulka 34: Příjmové proměnné a práce s nimi	77
Tabulka 35: Základní paradata.....	79
Tabulka 36: Časové značky.....	80
Tabulka 37: Paradata vztahující se k doplňujícímu souboru domácností	81
Tabulka 38: Proměnné s vyřazenými odlehlými či nekonzistentními hodnotami.....	87
Tabulka 39: Proměnné, u nichž byly doplňovány hodnoty dle jiných proměnných z šesté vlny	90
Tabulka 40: Proměnné, do kterých byla doplněna hodnota ze starší vlny	93
Tabulka 41: Počty případů, kdy nebylo možné provést imputaci ze současné či minulé vlny.....	96
Tabulka 42: Přehled ukazatelů vyplnění jednotlivých dotazníků a příslušných vah	99

Tabulka 43: Distribuce proměnných pro vážení (%) – dotazník pro domácnost, úroveň domácnosti	100
Tabulka 44: Distribuce proměnných pro vážení (%) – dotazník pro domácnost, úroveň jednotlivce.....	102
Tabulka 45: Distribuce proměnných pro vážení (%) – individuální dotazník.....	103
Tabulka 46: Zveřejněné materiály k 6. ročníku CHPS	105

Mgr. Michaela Röschová, Ph.D.

Mgr. Jan Klusáček

Bc. Aliona Sukhotskaya

Sociologický ústav AV ČR, v. v. i.

chps.soc.cas.cz

Verze 1.0 (16. 12. 2024)

Röschová, M., Klusáček, J., Sukhotskaya, A. 2024. České panelové šetření domácností: Dokumentace ke sběru a zpracování dat. Vlna 6 (2023–2024). Praha: Sociologický ústav AV ČR. (Verze 1.0). <https://doi.org/10.13060/dk.2024.2>