

3

Valundersökningen 2022

Valforskningsprogrammets studier av val och väljarbeteende har alltid kunnat byggas upp kring ambitiösa och omfattande datainsamlingar, främst frågeundersökningar med väljare. I den här boken använder vi oss av data från många olika källor för att ge en bakgrund till valet, beskriva hur väljarna upplevde medievalrörelsen och valkampanjen och för att analysera hur de svenska väljarna bar sig åt för att komma fram till ett röstningsbeslut. Hit hör till exempel valstatistik, publicerade mätningar från olika opinionsinstitut, nationella undersökningar från SOM-institutet, publikmätningar från MMS, sökningar i mediearkivet Retriever och offentlig statistik om energipriser.

Vår viktigaste datakälla är dock förstas Valundersökningen 2022. Studien är den 30:e nationella valundersökningen i Sverige sedan det göteborgska Valforskningsprogrammet inledde väljarstudier under 1950-talet. Fram till 2014 genomfördes undersökningarna som en kombination av besöks- och telefonintervjuer. Sedan 2014 genomförs studien med en mix av postala enkäter och webbenkäter. I Valundersökningen 2022 var det ungefär 30 procent av de svarande som valde att svara på webben. Det stora urvalet på 30 000 personer hämtades från registret för totalbefolkningen RTB och är representativt för den till riksdagsvalet röstberättigade befolkningen boende i Sverige, det vill säga svenska medborgare äldre än 18 år. Bruttosvarsfrekvensen varierar mellan 36–47 procent i undersökningens olika delar (Karlsson & Cederholm Lager, 2023). För Valundersökningen 2022 användes totalt sju olika frågeformulär som tillsammans omfattar 92 sidor med enkätfrågor (!): två förvalseditioner, två eftervalseditioner, en uppföljningsenkät, en edition med länderjämförande frågor samt en minienkät som användes i slutet av fältperioden. Förvalsundersökningen gick i fält redan den 8 augusti och stängde på valdagen den 11 september. Eftervalsundersökningen gick i fält dagen efter valet den 12 september och fältarbetet avslutades den 25 november.

Fyra faktorer gör de svenska Valundersökningarna unika i förhållande till väljarstudier i de flesta andra länder. 1) Vi använder sedan valet 1973 rullande mellanvalspaneler som gör det möjligt att följa samma väljare mellan två val. 2) Vi har en kampanjpanel bestående av en förvalsstudie och en uppföljningsstudie efter valet som gör det möjligt att följa samma väljare före och efter valet. Förvalseditionerna i Valundersökningarna kan användas för att lösa en del av de problem som gängse renodlade eftervalsundersökningar lider av: Eftersom väljarbedömningar och överväganden ofta tenderar att ändras, ibland till och med dramatiskt, när valutgången blivit känd, är det vanskligt att dra slutsatser om orsak-verkan när alla data är insamlade vid samma tidpunkt och i efterhand. 3) Sedan 1956 har vi kontrollerat urvalets valdeltagande mot offentliga röstlängder vilket validerar de viktiga utfallsvariablerna som valdeltagande och partival. 4) Vi hämtar in allt fler kompletterande bakgrundsuppgifter från befolkningsregister för att kunna ha

kortare enkäter och därmed lätta svarsbördan för respondenterna, bredda analyserna och för att kunna vikta och kalibrera resultaten. Sammantaget är den långa serien av svenska valundersökningarna ett av de bästa datamaterialen i världen för studier av politiskt beteende i samband med val. Datamaterialet görs tillgängligt för forskning och utbildning via Svensk Nationell Datatjänst (www.snd.gu.se) så snart det är möjligt med hänsyn till gällande lagstiftning.

I den här boken förekommer flitigt hänvisningar till olika delkomponenter av Valundersökningen 2022 (se figur 3.1). Analyserna som fokuserar på rörelser under mandatperioden använder sig av valundersökningarnas rullande *mellanvals-panel* där hälften av de som ingick i urvalet 2018 kontaktades igen 2022. De urval som ingår i inkommande eller utgående paneler randomiseras i två grupper som får antingen förvalsenkäten (etapp A) eller eftervalsenkäten (etapp B). Analyser som fokuserar på korttidsförändringar under valrörelsen använder sig av valundersökningarnas *kampanjpanel* där alla som besvarade förvalsenkäten före den 11 september (etapp A) fick en uppföljningsenkät efter valet. En särskild eftervalsenkät (etapp C) rymmer de frågeinstrument som ingår i det länderjämförande datainsamlingsprojektet The Comparative Study of Electoral Systems (CSES).

Vi vill gärna göra läsarna uppmärksamma på att det statistiska underlaget för analyserna i boken beror på vilka frågeinstrument som ingått i de olika enkäterna. Vissa nyckelfrågor ställs i samtliga delundersökningar och då kan antalet svarspersoner som analyserna vilar på landa på drygt tiotusen. Andra frågor kan ha ställts i enbart en enskild edition av valundersökningen och då kan antalet svarande vara nedåt 2 000. Detaljer om antalet svarande och internbortfall på enskilda frågor finns nedtecknat i tabell- och figurkommentarer.

Figur 3.1 En schematisk bild över komponenter i Valundersökningen 2022.



Kommentar: En starkt förkortad enkät – här kallad en minienkät – erbjöds de som ingick i ettapp A och B vid den sista påminnelsen i slutet av fältperioden. Totalt användes sju olika enkäter i Valundersökningen 2022: två förvalsenkäter (ettapp A), en uppföljningsenkät, tre eftervalsenkäter (ettapp B och C) samt en minienkät.

Metodologin bakom att genomföra storskaliga samhällsvetenskapliga frågeundersökningar befinner sig under ständig utveckling. I Valundersökningen 2022 intro-

ducerade vi flera nyheter i syfte att effektivisera datainsamlingen, bibehålla jämförbarheten och stärka datakvaliteten. De i urvalet som anslutit sig till en digital brevlåda (som t ex *Kivra* eller *Min myndighetspost*) – hela 66 procent! – fick inbjudningar, information och påminnelser skickat till sig i digital form. För huvudutskicken med enkät, svarskuvert, information om undersökningen och inloggningsuppgifter användes dock postala utskick till alla.

Incitament – det vill säga att svarspersonerna får en liten belöning eller present som tack – är viktiga för att försöka höja utvalda personers motivation att delta i en frågeundersökning. På basis av resultat från omfattande fältexperiment genom åren har vi en god bild av vilka typer av incitament som fungerar för att höja svarsfrekvenserna i den här typen av samhällsundersökningar. Huvuddelen av Valundersökningen (etapp A och B) genomfördes med endast ett ovillkorat incitament som medföljde huvudutskicket (penna). I delar av etapp C fick pennan sällskap av en godisbit. I etapp C användes också ett villkorat incitament: De som svarade fick en Trisslott som tack för hjälpen.

Olika erbjudanden till särskilt upptagna respondenter att svara på en förkortad version av Valundersökningens besöksintervju har använts sedan 1979. I den nya postal-webbmiljön lät vi den förkortade intervjun motsvaras av en *minienkät* med bara fem frågor som skickades ut tillsammans med den allra sista påminnelsen till de respondenter som ännu inte svarat. Minienkäten kunde endast fyllas i digitalt via en webbenkät. De fem frågorna handlade om partival 2022 och 2018, intresse för politik, förtroende för politiker och ideologisk vänster-högerplacering.

Svarsfrekvenserna i frågeundersökningar riktade till medborgare har minskat dramatiskt över tid sedan valundersökningarna inleddes på 1950-talet. I den första valundersökningen 1956 var svarsfrekvensen hela 95 procent. Den bekymmersamma nivån på bortfallet föranledde pionjärerna Jörgen Westerståhl och Bo Särilvik att genomföra omfattande och ambitiösa bortfallsanalyser (Westerståhl & Särilvik, 1957). Sedan dess har bortfallet i samhällsvetenskapliga undersökningar ökat. Sverige och valundersökningarna var länge förskonade från de dramatiskt fallande svarsfrekvenser som noterades i andra länder. Ända fram till 2010 var svarsfrekvenserna i valundersökningarna sjuttio procent eller högre.

Sedan 2010 har även valundersökningarna drabbats av fallande svarsfrekvenser. Vid de senaste undersökningarna 2018 och 2022 har bruttosvarsfrekvensen fallit en bit under 50 procent. Precis som i liknande undersökningar som bygger på befolkningsurval är det yngre och utlandsfödda som tillhör de mest svårnådda grupperna. Kontaktgraden – andelen personer i urvalet som man framgångsrikt lyckas kontakta med en förfrågan om deltagande – har minskat avsevärt bara under det senaste decenniet. Ett illustrativt exempel är att andelen i ett slumpmässigt befolkningsurval som SOM-institutet haft någon slags kontakt med under omkring 100 dagars fältperiod minskat från 82 till 57 procent mellan 2010 och 2020 (Falk, Sandelin & Weissenbilder, 2021). Bortfallsanalyser för valundersökningarna ger mycket likartade resultat som SOM-institutet finner i sina metodstudier (se Bergquist, Falk & Weissenbilder, 2023).

Bruttosvarsfrekvensen – det vill säga antalet svarande i relation till antalet i det ursprungliga befolkningsurvalet – för Valundersökningen 2022 var i genomsnitt 41,1 procent. Bakom detta genomsnitt döljer sig att olika delkomponenter av studien hade olika hög svarsfrekvens: Etapp A 36,2 procent, etapp B 38,2 procent

och etapp C 47,4 procent. Skillnaden mellan etapp A och etapp B beror i huvudsak på fältperiodernas olika längd (förvalsundersökningen hade en naturlig brytpunkt vid valdagen den 11 september). För Etapp C användes villkorade incitament och en något längre fältperiod än för etapp B.

Nettosvarsfrekvensen beräknas genom att ta hänsyn till naturligt bortfall, det vill säga räkna bort personer i urvalet som var förhindrade att delta i undersökningen på grund av fysiska eller psykiska skäl, att de är avlidna, har flyttat utomlands, saknar adress eller inte talar svenska. Med detta sätt att räkna (AAPOR6) varierar valundersökningarnas nettosvarsfrekvens mellan de olika etapperna på följande sätt: Etapp A 36,7 procent, Etapp B 39,0 procent och Etapp C 48,9 procent. Värt att notera är att trots nästan tio procentenheters skillnad i svarsfrekvens är fördelningen av partivalsvariabeln i Etapp B och etapp C identisk ($\chi^2=,63$). När det gäller precision på den huvudsakliga beroende variabeln partival spelar det alltså ingen roll att det skiljer tio procentenheter i svarsfrekvens.

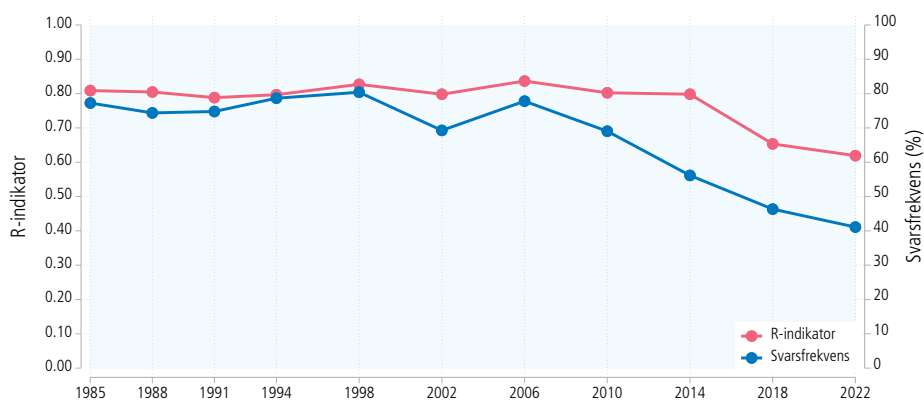
Många analyser i boken tar till vara de unika panelkomponenterna i valundersökningarna. Då är uppgifter om så kallad paneldödlighet viktig. Av dem som valts ut att delta i valundersökningarna *mellanvalspanel* 2018–2022 var det 28 procent som med fyra års mellanrum medverkade i båda undersökningarna. För *kampanjpanelen* är paneldödligheten lägre: Av dem som svarade på etapp A i Valundersökningen 2022 var det 74 procent som också fyllde i eftervalsenkäten efter valet.

Det finns många felkällor i frågeundersökningar (se t ex Biemer (2010) för en introduktion till ramverket *Total Survey Error*), varav de potentiellt mest problematiska ofta är kopplade till sättet att ställa frågor eller hur skalor och svarsalternativ är utformade eller hur data analyseras. Bortfall hör numera till de problem som både är mest överhängande och växande. Svarsfrekvens är dock inte något universellt mått på datakvalitet. Effekterna av ett bortfall kan vara mycket små för vissa typer av skattningar men desto större för andra typer av skattningar. Ett samlande mått på hur representativa de svarande är för den population som undersöks, och som vi tycker borde användas mer som kvalitetsmått, är den så kallade representativitetsindikatorn (R). För att kunna beräkna R-indikator för en studie behöver man ha tillgång till uppgifter för hela urvalet, alltså både de som svarat och de som inte gjort det. I det här fallet (se figur 3.2) har vi beräknat R-indikatorn för Valundersökningarna 1985–2022 med hjälp av en logistisk regression där förklaringsvariablerna kön, födelseår, inkomst och valdeltagande använts för att predicera sannolikhet att svara på undersökningen. I en studie med maximal representativitet är det helt slumpmässigt vem som svarar eller inte och då beräknas R-indikatorn till 1. Ju större avvikelser vad gäller sammansättningen av de svarande och de icke-svarande, desto lägre R-indikator.

Graden av representativitet i valundersökningarna har aldrig varit perfekt trots historiskt mycket höga svarsfrekvenser. Den demografiska profilen för de grupper som inte går att nå eller inte vill svara har tvärtom varit mycket stabil över tid, med en R-indikator omkring .80 under perioden 1985–2014. Valundersökningen 2014 hade klart lägre svarsfrekvens än tidigare studier men representativiteten bland de svarande upprätthölls ändå. Det var först vid bytet av datainsamlingsmetoder från besöksintervjuer till postala-webbundersökningar som även R-indikatorn började falla och ligger vid de två senaste undersökningarna på en klart lägre nivå, strax över .60.

Förutom den klassiska bortfallsproblematiken med yngre och svagt integrerade utrikes födda finns en allt starkare tendens att personer som inte är så intresserade av politik eller som inte röstar i riksdagsvalen väljer bort deltagande i Valundersökningarna. Bland de svarande i Valundersökningen 2022 var det hela 96,3 procent som röstade i riksdagsvalet 2022, enligt de offentliga röstlängder vi kontrollerar mot. Förr fanns det fler politiskt ointresserade och icke-röstare som deltog i valundersökningarna. Idag är dessa grupper så små att det är en stor utmaning att undersöka dem exempelvis i studier av icke-röstares egna motiveringar till varför de inte deltar eller i studier av politiskt alienerade grupper.

Figur 3.2 Svarsfrekvenser och representativitet i de svenska Valundersökningarna 1985–2022 (andel och R-indikator)



Källa: Valforskningsprogrammet.

Kommentar: Svarsfrekvenserna är bruttosvarsfrekvenser för Valundersökningarna rapporterade i enlighet med organisationen AAPORs-standard AAPOR1 (AAPOR, 2015). R-indikator visar graden av representativitet bland de svarande och beräkningarna är baserade på kön, födelseår, inkomstnivå samt valdeltagande i riksdagsvalet. Nettosvarsfrekvensen för tidigare Valundersökningar har varit 1956: 95 procent, 1960: 92 procent, 1964: 92 procent, 1968: 88 procent, 1970: 85 procent, 1973: 82 procent, 1976: 74 procent, 1979: 81 procent och 1982: 82 procent (se Karlsson & Cederholm Lager, 2023).

Ett ytterligare mått på övergripande datakvalitet handlar om träffsäkerhet. Träffsäkerheten i valundersökningarna har under lång tid varit mycket tillfredsställande när det gäller den viktigaste beroende variabeln: partival. Medelfelet per parti, som är ett vanligt mått på träffsäkerhet, har totalt sett varierat mellan 0,3 och 2,1 under perioden 1956–2022 (Karlsson & Cederholm Lager, 2023). Under perioden 1970 fram till Millenieskiftet var träffsäkerheten så pass bra, oftast lägre än 1,0 i medelfel, att ansvariga forskare därför inte såg någon större anledning att ägna sig åt viktningar: fördelningen av partivalsvariabeln bland de svarande låg tillräckligt nära de faktiska valresultaten.

Men med den allt större gruppen SD-väljares kroniska underrepresentation bland de svarande även i valundersökningen har behovet av att justera marginalfördelningarna efter valresultaten vuxit. De vikter som används i våra restaurerade dataserier (se tabell 3.1) har för Sverigedemokraternas del legat påfallande stabilt omkring 1,4 i valen 2006–2022. Det betyder att en svarande SD-väljare behöver

vägas upp med 1,4 för att andelen SD-väljare bland samtliga svarande ska likna valresultatet. En komma fyra tycks vara ett slags konstant i sammanhanget. Även i SVTs vallokalsundersökningar ligger SDs vikter nära denna nivå (Holmberg m fl, 2023). Och vill man vikta upp de nationella SOM-undersökningarnas resultat för partival från valhöstarna så att de liknar valresultatet har det behövts en vikt mellan 1,35 och 1,45 under perioden 2010–2022. Oavsett om Sverigedemokraterna varit ett litet eller stort parti har ungefär samma viktningskvot varit nödvändig. De vikter som presenteras i tabell 3.1 gäller valundersökningarna men är bra riktmärken för hur över- eller underrepresenterade olika partiers väljare har varit i liknande studier under samma period.

Tabell 3.1 Vikter för partival i Valundersökningarna 1956–2022.

Undersökningsår	Parti										
	V	S	C	FP/L	M	KD	MP	SD	FI	NyD	Ovr
1956	3.32	0.91	1.08	0.91	1.15	–	–	–	–	–	–
1960	2.26	0.92	0.89	1.06	1.16	–	–	–	–	–	1.25
1964	1.88	0.92	0.89	1.01	1.25	1.06	–	–	–	–	1.17
1968	1.97	0.91	0.90	1.06	1.22	1.01	–	–	–	–	4.33
1970	1.24	0.92	0.89	1.14	1.30	1.26	–	–	–	–	1.85
1973	1.23	0.94	0.96	1.15	1.14	0.89	–	–	–	–	4.44
1976	1.18	0.93	1.03	0.94	1.15	0.99	–	–	–	–	9.62
1979	1.05	0.96	1.09	0.96	1.02	0.87	–	–	–	–	1.90
1982	1.12	0.98	1.05	0.98	1.01	0.78	0.93	–	–	–	1.32
1985	1.05	1.02	0.96	0.86	1.07	–	0.87	–	–	–	4.33
1988	1.08	0.95	0.97	1.04	1.11	1.01	0.91	–	–	–	1.84
1991	1.28	0.99	0.97	0.99	1.01	0.95	0.95	–	–	0.94	1.39
1994	0.94	0.98	0.91	0.92	1.10	1.00	0.96	–	–	1.16	2.25
1998	1.05	0.97	0.98	0.95	0.99	1.07	1.03	–	–	–	0.97
2002	0.98	1.03	1.06	0.83	1.12	0.99	0.71	–	–	–	1.15
2006	1.06	1.00	1.00	1.01	0.99	0.92	0.88	1.38	0.87	–	1.50
2010	1.09	0.98	0.98	0.94	0.98	1.09	0.82	1.61	0.82	–	3.10
2014	0.85	1.00	0.90	0.93	1.03	0.93	0.79	1.38	0.87	–	1.35
2018	0.91	0.95	0.90	0.80	1.04	0.98	0.82	1.36	1.04	–	1.27
2022	1.04	0.89	0.86	0.81	1.05	0.95	0.72	1.44	–	–	1.81

Källa: Valforskningsprogrammet.

Kommentar: Vikterna för partival i riksdagsvalet gäller för hela fördelningen och alla svarande. I en del korstabeller och flervariabelanalyser i boken tas särskilda vikter fram för att kompensera för internbortfall. Partival är röstlängdskontrollerat för valdeltagande. Vikter som är större än 1 indikerar att vi har lägre andel svarande som röstat på partiet än vad de officiella valresultaten visar. Vikter som är mindre än 1 indikerar att vi har högre andel svarande som röstat på partiet än vad de officiella valresultaten visar.

För vissa analyser i den här boken har vi valt att vikta data från valundersökningarna mot de officiella valresultaten. Om vikter har använts framgår det i så fall alltid av information i tabell- och figurkommentarer. Viktning mot valresultat spelar oftast inte särskilt stor roll för punktskattningar och aldrig för våra huvudslutsatser. För analyser över tid har vi tyckt att det är viktigt att ta hänsyn till de skevheter som uppstått främst på grund av SD-väljarnas underrepresentation. Det samma gäller för analyser av fenomen där vi ser att skattningar för hela väljarkåren påverkas av att Sverigedemokraterna numera är ett stort parti. Många av de faktorer som styr benägenheten att rösta på SD är också viktiga för att förklara varför vissa väljargrupper inte vill vara med i valundersökningarna, som till exempel intresse, kunskap, utbildningsnivå och graden av politiskt förtroende.

I den här boken har vi inte använt oss av demografiska viktningar efter exem-

pelvis egenskaper som kön, ålder, bostadsort, utbildning eller inkomst. Demografiska vikter betraktas ibland som mer statistiskt renlärigt än att mekaniskt rikta upp utfallsvariabeln partival, särskilt i analyser där partival inte är inblandat. I praktiken är det dock inte säkert att exempelvis en åldersvikt förbättrar representativiteten eller gör att vi kommer närmare ett sant populationsvärde. Viktning är inget universellt botemedel. Vikter kan göra mer skada än nytta om vi exempelvis kraftigt viktar upp de unga människor som deltar i undersökningen och låter dem utan vidare representera unga människor som inte deltagit.

Allt valundersökningarna inte perfekt speglar den röstberättigade befolkningen är en insikt som är viktig att internalisera i de analyser som genomförs i boken. Det är viktigt att tolkningar av förändringar över tid eller skillnader mellan grupper tar höjd för de osäkerheter som finns till följd av att vi inte lyckas nå unga och utrikesfödda i den utsträckning som vi skulle önska. Till sist vill vi påminna läsaren om att nästan alla tal du ser i boken är skattningar. Alla skattningar som bygger på slumpmässigt urval, och därmed på statistisk teori, har felmarginaler. Tabell 3.3 nedan visar vilka felmarginaler som gäller för olika procentskattningar vid olika storlekar på grupperna. Ju mindre grupp, desto större felmarginal, och detta gäller särskilt för skattningar runt 50 procent.

Tabell 3.2 Lathund för procentskattningar vid olika gruppstorlekar (felmarginaler och procentenheter, 95 procents signifikansnivå)

Skattning	Gruppstorlek						
	50	100	200	400	800	1 000	2 000
50 procent	±13,9	±9,8	±6,9	±4,9	±3,5	±3,1	±2,2
30 procent	±12,7	±9,0	±6,4	±4,5	±3,2	±2,8	±2,0
10 procent	±8,3	±5,9	±4,2	±2,9	±2,1	±1,9	±1,3
6 procent	±6,6	±4,7	±3,3	±2,3	±1,6	±1,5	±1,0
4 procent	±5,4	±3,8	±2,7	±1,9	±1,4	±1,2	±0,9
2 procent	±3,9	±2,7	±1,9	±1,4	±1,0	±0,9	±0,6

Kommentar: Tabellen visar hur stora felmarginalerna är för olika procentskattningar och gruppstorlekar. Nittiofem procents signifikansnivå innebär att vi kan förvänta oss att procenttalet för hela populationen befinner sig inom konfidensintervallet i 95 fall av 100 i en tänkt situation med många upprepade mätningar. Att vara medveten om felmarginalen gör att vi lättare undviker att dra slutsatser om förändringar eller gruppskillnader som enbart beror på slumpen.

Vi behöver också påminna om att alla slags beräkningar av felmarginaler bygger på ett ofta outtalat antagande om att bortfallet i en undersökning är helt slumpmässigt fördelat – ett antagande som vi vet säkert inte är sant. Även efter viktning tar våra beräkningar av konfidensintervall bara hänsyn till fel som är associerade med själva urvalsdragningen. Vi tar inte alls hänsyn till skevheter som beror på att vissa grupper har en större benägenhet att inte delta i frågeundersökningar. Konventionen inom forskningen är att ändå trots allt använda sedvanliga statistiska signifikanstest som ett disciplinerande verktyg för att hjälpa oss att skilja signaler från brus.

Valundersökningen 2022 är den 30:e nationella frågeundersökningen med väljare i samband med val som genomförs i Sverige

Valundersökningarnas inbyggda mellanvals- och valrörelsepaneler möjliggör noggranna analyser av väljarrörlighet

Mer än 60 000 röstberättigade medborgare har deltagit i de svenska Valundersökningarna sedan 1956. De äldsta åldersgrupperna i materialet föddes på 1870-talet.

Nettosvarsfrekvensen för de olika delkomponenterna i Valundersökningen 2022 varierar mellan 37 och 49 procent.