

**I SONDAGGI
POLITICI**

**LE BUONE
PRATICHE**

2019



A CURA
DEL
COMITATO
STANDARD
DI
QUALITA'

INTRODUZIONE

ASSIRM svolge un ruolo fondamentale nella qualificazione dell'attività dei suoi aderenti, nonché nella diffusione della cultura della ricerca in Italia.

L'obiettivo è quello di dettare i canoni di qualità ed individuarne i criteri di determinazione e di verifica, contribuendo a fornire agli stakeholder dei parametri chiari di identificazione delle aziende che operano con serietà all'interno del settore, siano esse o meno associate.

Un ruolo primario nel raggiungimento di tali obiettivi è svolto dal **Comitato Standard di Qualità Assirm**, che ha il compito di:

- aggiornare le Norme di Qualità ed il Codice Etico dell'Associazione;
- rilasciare il Certificate of Conformity agli standard qualitativi dell'Associazione;
- promuovere la redazione e la divulgazione di norme specifiche e best practice che servano ad individuare il perimetro di competenza all'interno del quali le aziende associate possono operare.

È in questa logica che si colloca la redazione de "I sondaggi politici. Le buone pratiche" curata da Stefano Draghi, docente di *Nuove tecnologie digitali* all'Università IULM di Milano, su incarico del Comitato Standard di qualità Assirm e sotto la supervisione del Prof. Edoardo Lozza, professore ordinario di *Psicologia del marketing e dei consumi* e di *Psicologia economica* all'Università Cattolica del Sacro Cuore di Milano.

Un particolare ringraziamento va al Dottor Maurizio Pessato per il contributo che, grazie alla sua esperienza sul tema, ha dato alla realizzazione di questo documento.

I componenti del Comitato Standard di Qualità Assirm:

Luciana Donnarumma (Presidente)

Edoardo Lozza (Editor scientifico del Codice delle Ricerche Assirm)

Chiara Castelli

Massimo Cesaretti

Lino Coscione

Rosapia Farese

Piergiorgio Rossi

Vilma Scarpino

INDICE

1. OGGETTO, OBIETTIVO E RICHIAMO NORME DI QUALITÀ ASSIRM	1
2. IL QUESTIONARIO: NUMERO, FORMA E SUCCESSIONE DELLE DOMANDE	2
2A. Numero di domande e durata del questionario	2
2B. La formulazione delle domande	2
2C. La successione delle domande e l'effetto contaminazione	5
3. DIGRESSIONE SUI CONCETTI DI NOTORIETÀ, POPOLARITÀ E CONSENSO	6
4. IL CAMPIONAMENTO	6
4A. La numerosità del campione	7
4B. Campioni probabilistici e non, multicanalità	8
4C. La stratificazione del campione	9
4D. La ponderazione dei dati	9
5. I SONDAGGI SULLE INTENZIONI DI VOTO	10

A differenza degli antichi *straw polls* e della ricerca di mercato contemporanea, il sondaggio politico è parte rilevante della sfera pubblica e componente vitale del processo democratico. Le sue funzioni fondamentali di legittimazione degli attori politici, di formazione dell'agenda politica e soprattutto di formazione del consenso¹ sollecitano la definizione di standard di qualità più elevati a maggior tutela del ruolo che l'opinione pubblica ha nei sistemi liberal-democratici e della reputazione degli istituti di ricerca.

1. OGGETTO, OBIETTIVO E RICHIAMO NORME DI QUALITÀ ASSIRM

Oggetto delle buone pratiche proposte in questo documento sono tutti i sondaggi condotti sugli orientamenti politici e di voto degli elettori e più in generale su argomenti politici. Non riguardano metodi e tecniche di anticipazione dei risultati elettorali realizzate con interviste all'uscita dei seggi (*exit-polls*) o a urne chiuse sullo spoglio delle schede in un campione di sezioni (proiezioni elettorali). Né metodi di previsione dell'esito delle elezioni, anche se basati esclusivamente su sondaggi politici.

Obiettivo delle buone pratiche è quello di rendere evidenti e verificabili ai competenti e al pubblico le differenze tra sondaggi professionali e non professionali e garantire il massimo della qualità, compatibilmente con le risorse disponibili, nelle seguenti fasi della ricerca

- impostazione del questionario, ordine e forma delle domande
- campionamento
- elaborazione dei dati

Valgono anche per i sondaggi politici tutte le vigenti "Norme di qualità Market and Social Research" adottate da ASSIRM². Le buone pratiche qui proposte vanno quindi intese come aggiuntive. Molte delle pratiche qui proposte si possono estendere a tutta la ricerca sociale e di mercato, ma assumono particolare rilievo nel caso dei sondaggi politici.

Le proposte riguardano tutti i sondaggi politici condotti con metodo campionario e tramite interviste individuali (CAPI, CATI, CAMI, CAWI, ecc.), sia riservati che destinati alla pubblicazione. Le regole a cui si dovrebbero attenere i media nella pubblicazione dei sondaggi politici non sono comprese in questo documento, ma sarebbe assai importante arrivare alla stesura di un codice condiviso almeno per la presentazione dei risultati, come avviene ad esempio nei paesi anglosassoni³. Ovviamente si estende al settore politico il "divieto di condurre attività di promozione e comunicazione pubblicitaria di beni e servizi, di marketing diretto e telemarketing, di commercializzazione, distribuzione e vendita"

¹ Così P. Ceri, "A che cosa servono i sondaggi politici", in *id.*, *Politica e sondaggi*, Rosenberg & Sellier, Milano, 1997.

² ASSIRM, *Le norme di qualità Market and Social Research*, ultima revisione approvata dall'Assemblea degli Associati del 10 Dicembre 2015. Aggiornamento del giugno 2018.

³ Cfr. ad esempio BBC, *Editorial Guidelines, Sect. 10, Politics, Public Policy and Polls*; ABC News *Polling Methodology and Standards*; New York Times *Polling Standards*.

(Statuto ASSIRM, art.2), che potrebbero assumere la forma di *push polling* o di *negative campaigning*.

Per ciascun sondaggio politico l'Istituto compila una Scheda di Qualità ASSIRM (SQA) in cui registra tutte le operazioni effettuate per assicurare il massimo di conformità del sondaggio alle buone pratiche ASSIRM per i sondaggi politici. La scheda SQA è inserita e conservata nella cartella di lavoro del progetto di ricerca.

Le SQA (o una loro sintesi) possono essere rese disponibili sul sito dell'Istituto o dell'Assirm senza limitazioni di accesso. La possibilità di accedere alla SQA andrebbe comunque garantita se il sondaggio politico è destinato alla pubblicazione.

2. IL QUESTIONARIO: NUMERO, FORMA E SUCCESSIONE DELLE DOMANDE

2A. NUMERO DI DOMANDE E DURATA DEL QUESTIONARIO

Un problema assai frequente che riguarda i sondaggi politici è che sono spesso condotti su temi di stretta attualità ("*instant poll*") e il committente impone tempi di esecuzione della ricerca molto limitati e spesso anche un budget assai ridotto. In questi casi ne può risentire anche fortemente la qualità dei risultati, perché tempi stretti per l'impostazione del questionario impediscono l'esecuzione di test pilota e un numero ridotto di domande non permette di articolare il questionario in modo tale da irrobustire la validità e l'attendibilità⁴ delle risposte, ad esempio con domande filtro e eventuali percorsi di domande diversificate⁵. Il rischio, elevato, è quello di registrare opinioni molto volatili, superficiali o che nascono al momento stesso della domanda, e di raccogliere molti "non so" o molte risposte "incerto/indeciso" o orientamenti (come ad esempio le intenzioni di voto) di cui non è possibile valutare la solidità e l'intensità e di cui è incerta o indefinita la probabilità di trasformarsi in comportamenti conseguenti. Le stesse considerazioni valgono se una/poche domande politiche vengono inserite in questionari di tipo *omnibus*.

In questi casi l'Istituto, per garantire lo standard di qualità della rilevazione, proporrà al Committente in fase di *brief* un questionario e un test pilota in grado di limitare il rischio che gli errori di misurazione (derivanti dall'impossibilità di verificare validità e attendibilità delle domande) possano produrre errori sistematici di entità non valutabile. L'Istituto segnalerà inoltre al Committente il rischio che tali errori possano sommarsi agli errori di campionamento, portando l'errore totale della rilevazione a livelli inaccettabili, e comunque non valutabili. Se il Committente ritiene di non dover accettare le proposte dell'istituto, può dichiararlo in modo esplicito. Tale dichiarazione è inserita nella scheda SQA del progetto di ricerca.

2B. LA FORMULAZIONE DELLE DOMANDE

I sondaggi politici "riguardano quasi sempre argomenti sui quali gli intervistati possono avere un livello molto diverso di conoscenza e di interesse, opinioni e punti di vista in via di

⁴ Validità = la domanda rileva effettivamente il concetto che si intende rilevare (grado di sovrapposizione semantica). Attendibilità = grado di correttezza e stabilità delle risposte.

⁵ Tempi ristretti e numero di casi ridotto influenzano anche altre fasi della ricerca, come ad esempio la rilevazione, perché si accentua il *bias* (errore sistematico) di autoselezione del campione.

formazione, inconsistenti, provvisori e soggetti a cambiare”⁶. Si tratta di opinioni, motivazioni, orientamenti, intenzioni, aspettative, sentimenti, valutazioni, giudizi e valori che riguardano oggetti a volte complessi e sensibili. Per queste ragioni in un questionario su temi politici va tenuta in particolare considerazione la pur ovvia affermazione che il modo in cui una domanda viene formulata può influenzare la risposta. I manuali di metodologia della ricerca riportano innumerevoli esempi di come possono cambiare - e di molto - i risultati con modifiche anche piccole nel *wording* delle domande. Se non è possibile formulare regole precise, si possono però indicare principi generali a cui attenersi⁷.

- Usare un linguaggio semplice, accessibile a tutti, per evitare risposte a caso di chi non capisce un termine o non conosce l'argomento
- Formulare domande concise, non troppo lunghe.
- Minimizzare il numero di alternative di risposta
- Evitare termini ambigui e espressioni gergali
- Non utilizzare parole cariche di emotività, soprattutto se negative
- Evitare domande con doppia negazione
- Preferire le formulazioni di segno positivo a quelle negative⁸
- Evitare domande “doppie” o a risposta non univoca⁹
- Evitare domande tendenziose, a risposta pilotata o che segnalano un'aspettativa di risposta da parte dell'intervistatore¹⁰

⁶ ESOMAR-WAPOR-Guideline-on-Opinion-Polls-and-Published-Surveys-August-2014, p. 5.

⁷ Riprendo parte delle indicazioni da P. Corbetta, *La ricerca sociale: metodologia e tecniche*, Il Mulino, Bologna, 2015, vol. II.

⁸ È questa una delle principali fonti di ambiguità e di possibile forzatura dei risultati. Infatti, l'intervistato che approva l'oggetto della domanda (comportamento, personaggio, ...) presentato negativamente deve esprimersi negativamente, mentre chi disapprova dovrebbe approvare la negazione. “In altre parole, l'oggetto proposto attira l'attenzione dell'intervistato, che esprime il suo accordo o disaccordo nei suoi confronti, e non verso l'affermazione. Il problema non si pone se si utilizzano solo affermazioni positive” (*ibidem*).

⁹ Altra fonte comune di inattendibilità e di distorsione dei risultati.

¹⁰ Tra i numerosissimi esempi che si possono fare di domande di questo tipo ne riporto qui uno solo, in omaggio a Elio Brusati, uno dei “padri” italiani del sondaggio ben fatto.

“Citiamo il caso rilevante di una domanda effettivamente rivolta in un sondaggio, effettuato durante una delle passate crisi di governo: “Se il Presidente della Repubblica nominasse un Capo del Governo al di sopra delle parti, lei pensa che, per il bene del paese, si dovrebbe votare dopo aver fatto le riforme più urgenti, oppure si dovrebbe votare subito?”. Evidentemente oggetto del dibattito in corso al momento del sondaggio erano due tesi contrapposte: formare un governo di attesa con obiettivi limitati, oppure andare subito alle elezioni anticipate. Ma nel testo della domanda una delle due tesi (quella del governo d'attesa) appariva indorata da ben tre connotazioni di tipo “virtuoso”: 1) un Capo del Governo al di sopra delle parti; 2) per il bene del Paese; 3) (che farà / che saprà fare) le riforme più urgenti. Invece l'altra tesi (quella delle elezioni anticipate subito) era liquidata con poche parole, senza alcuna connotazione virtuosa, salvo forse quella - peraltro appena percettibile - nascosta nell'avverbio subito, che agli amanti della sbrigatività efficiente potrebbe anche piacere. È evidente che nella domanda c'è un grave squilibrio: manca quella par condicio, tra l'esposizione delle due tesi a confronto,

- Non dare per scontati comportamenti o conoscenze che non lo sono, utilizzando opportune domande “filtro”
- Preferire le domande concrete a quelle astratte

In merito alla formulazione delle domande altri punti sono rilevanti e in particolare:

- i. il problema della mancanza di opinioni
- ii. la misura dell'intensità degli atteggiamenti
- iii. la desiderabilità sociale delle risposte

i. Nei sondaggi politici agli intervistati vengono spesso sottoposte domande su temi complessi, la cui conoscenza e comprensione può richiedere competenze specialistiche e comunque non banali, come quelle che fanno riferimento a questioni economiche o giuridiche. O domande che richiamao eventi, problemi o personalità che l'intervistato semplicemente non conosce o di cui non è al corrente. In questi casi la dinamica dell'intervista può creare una pressione a rispondere che spinge gli intervistati a esprimere un giudizio anche senza avere un'opinione in proposito.

In un sondaggio di qualità le domande di tipo valutativo sono rivolte solo a coloro che dimostrano di conoscere l'oggetto della domanda, che sarà pertanto preceduta da una domanda-filtro di tipo cognitivo. Le risposte “non so”, “non conosco” sono componenti naturali e legittime dell'opinione pubblica e devono essere ben visibili nella presentazione dei risultati. Evitando di mischiare tra le risposte valide quelle motivate e competenti con quelle date a caso o che nascono al momento stesso della domanda, si otterranno risultati meno volatili e più robusti.

ii. Analogo ragionamento può essere fatto per quanto riguarda l'intensità degli atteggiamenti, che “è importante saper cogliere [...] perché è quella che poi determina il comportamento [...] Anche in questo caso la rilevazione dell'intensità di un atteggiamento non può essere affidata semplicemente ad una più appropriata formulazione della domanda, ma necessita di un approfondimento attraverso domande ulteriori”¹¹.

iii. Meno rilevante per i sondaggi politici, ma da tenere comunque presente, è la questione della desiderabilità sociale delle risposte. Il controllo sociale che agisce nell'interazione tra intervistato e intervistatore può spingere quest'ultimo a dare risposte che ritiene più

che avrebbe potuto evitare che parte dei soliti indecisi (naturalmente i “già decisi”, per l'uno o l'altro partito, non sono sospetti di lasciarsi influenzare dalle sfumature della domanda) scegliessero la tesi che sentivano come la meglio “argomentata”, solo perché tale. [...] Eppure i sostenitori della tesi delle elezioni anticipate avrebbero ben saputo proporre le due o tre “connotazioni virtuose”, con cui indorare anche la loro, se quell'autore avesse avuto la pazienza di ascoltare anche la loro voce, com'era suo dovere, per evitare quel vizio che oggi viene detto del doppiopesismo (usare “due pesi e due misure”). Nei sondaggi va piuttosto usato quell'altro neologismo, non importa se altrettanto brutto, che è il cerchiobottismo (dare “un colpo al cerchio ed uno alla botte”). E. Brusati, Come si fanno i sondaggi, in Induzioni, n. 28, 2003 reperibile su <http://matematica-old.unibocconi.it/statistica/SONDAGG I.htm>.

¹¹ P. Corbetta, op.cit., p. 161.

“accettabili” da parte del suo interlocutore¹². Nell’*issue polling* questo effetto si può manifestare in domande che riguardano temi sociali o sensibili di rilevanza politica, ma tende a produrre un *bias* di modesta entità se si tiene presente che il sondaggio ha proprio lo scopo di misurare la forza delle correnti che investono l’opinione pubblica.¹³

2C. LA SUCCESSIONE DELLE DOMANDE E L’EFFETTO CONTAMINAZIONE

Anche la posizione di una domanda all’interno di un questionario può condizionare le risposte. In aggiunta alle regole generali che vanno tenute presenti e riguardano qualsiasi tipo di intervista con questionario¹⁴, va evitato il cosiddetto *effetto contaminazione* che si determina quando una domanda è preceduta da una o più altre domande che possono influenzare la risposta. È possibile in questo modo “pilotare” le risposte a una domanda e mascherare l’effetto di contaminazione non pubblicando le domande “contaminanti”.

Effetti analoghi si possono ottenere con l’uso di batterie di domande, che possono facilitare l’espressione di *pseudo-opinioni* (v. *supra*) e produrre un effetto distorsivo noto come *response set*¹⁵. Al problema, ben noto, si fa fronte alternando la polarità delle modalità di risposta (“controsalatura”).

Per verificare la presenza di possibili effetti di contaminazione e di *response set* è indispensabile che il testo completo del questionario sia pubblicato in appendice alla SQA.

Se la formulazione delle domande e la loro sequenza sono proposte dal Committente, l’Istituto si riserva di valutarne la conformità agli standard Assirm e di proporre modifiche o alternative praticabili. Le controproposte dell’Istituto sono riportate nella SQA.

¹² È questo il nocciolo della critica costruttivista all’idea oggettivista che l’intervista possa in modo asettico e neutrale rilevare opinioni e atteggiamenti dell’intervistato.

¹³ Nei sondaggi politici la desiderabilità sociale delle risposte si può manifestare sotto forma di “effetto *band wagon*”, come nel caso della cosiddetta “luna di miele” tra vincitori delle elezioni e opinione pubblica. Diverso è il caso delle domande sui precedenti comportamenti di voto. Alla domanda “Lei è andato/a a votare alle scorse elezioni politiche?” la percentuale di risposte “Sì” supera spesso e di molto il dato ufficiale dell’affluenza alle urne.

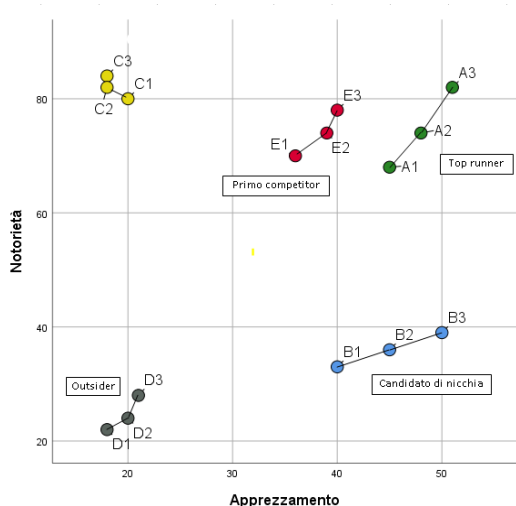
¹⁴ Ad esempio, iniziare l’intervista con domande “facili” o poco impegnative che mettono l’intervistato a suo agio, porre le domande più impegnative o delicate a metà del questionario, lasciare per ultime quelle sulle caratteristiche sociografiche, porre prima le domande più generali rispetto a quelle più specifiche, ecc.

¹⁵ “Una forma di distorsione delle risposte [...] è quella che va sotto il nome di *response set* (potremmo chiamarla «uniformità delle risposte», consistente nel fatto che di fronte a una batteria di domande tutte contemplanti lo stesso tipo di alternative di risposta (per esempio molto, abbastanza, poco, per niente d’accordo) ci possono essere intervistati che, per pigrizia (pur di finire in fretta l’intervista) o per mancanza di opinioni, rispondono sempre allo stesso modo (per esempio sempre «abbastanza d’accordo») indipendentemente dal contenuto delle domande”. P. Corbetta, *op. cit.*, p. 162.

3. DIGRESSIONE SUI CONCETTI DI NOTORIETÀ, POPOLARITÀ E CONSENSO

Molte delle variabili che un sondaggio politico ha l'obiettivo di misurare fanno riferimento ai concetti di notorietà, popolarità e consenso, riferiti sia a personalità politiche che a

Figura 1
Notorietà e apprezzamento dei candidati
Sindaco in tre momenti della campagna elettorale



partiti o istituzioni. I primi due concetti, notorietà e popolarità, riferiti a attori politici, dovrebbero essere tenuti separati per evitare che vengano impropriamente scambiati. Notorietà si riferisce alla percentuale di intervistati che conoscono, sanno chi è/che carica o ruolo ricopre/il soggetto della domanda. Il concetto di popolarità include quello di notorietà, ma vi aggiunge una connotazione generalmente positiva. La notorietà si misura, in diversi modi, con domande di tipo cognitivo, la popolarità con domande di tipo valutativo. Misurare la notorietà e proporla come popolarità, o addirittura consenso, è una forzatura semantica che comporta sempre un errore di validità, errore sistematico che può essere anche di notevole entità¹⁶.

Un sondaggio politico di qualità separa la rilevazione della notorietà da quella del giudizio, con domande diverse, anche opportunamente filtrate. Indipendentemente da come le due dimensioni vengono rilevate¹⁷. Misure ripetute dello stesso concetto rafforzano l'attendibilità della misurazione, mentre aumenta la sua validità la costruzione di indici (notorietà + giudizio = popolarità) di tipo additivo o moltiplicativo. L'adozione di questi criteri di rilevazione permette anche una efficace rappresentazione grafica del posizionamento dei soggetti nello "spazio" della popolarità, come esemplificato nella Figura 1¹⁸.

4. IL CAMPIONAMENTO

Il sondaggio politico non presenta problemi di campionamento diversi da quelli di qualsiasi altra ricerca sociale e di mercato. Un numero elevato di casi e una accurata stratificazione garantiscono in genere la bontà del campione, cioè la sua rappresentatività statistica, e margini ridotti delle stime campionarie, almeno della componente *random*.

La popolazione di riferimento per i sondaggi politici viene fatta coincidere convenzionalmente con quella degli elettori (aventi diritto al voto) che corrisponde in larga misura con la popolazione residente di 18 anni e oltre. L'utilizzo di quest'ultima per il piano

¹⁶ È il caso ad esempio dei personaggi politici molto noti (c.d. *top of mind*), sia a livello nazionale che locale, ma che dividono nel giudizio l'opinione pubblica per questioni politiche, morali o di personalità.

¹⁷ Ad esempio, la notorietà può essere rilevata in modalità *aided* o *unaided*, il giudizio con variabili ordinali, voti "scolastici", barometri di vario tipo, ecc.

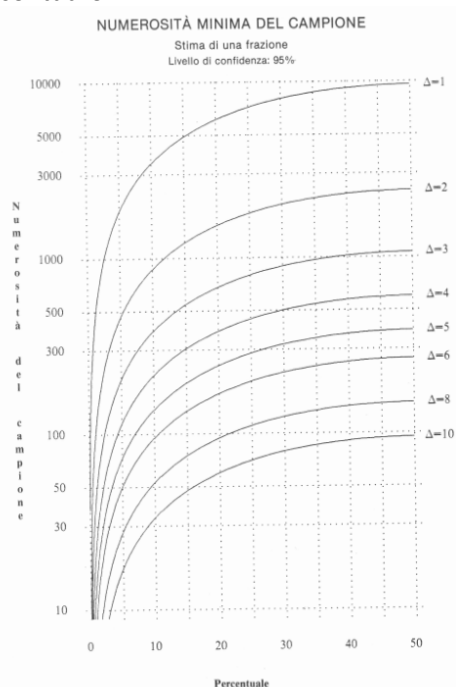
¹⁸ Grafico che dice molto sul possibile esito dell'elezione a complemento delle stime sulle intenzioni di voto.

di campionamento di un sondaggio politico comporta modesti svantaggi sul fronte della rappresentatività, mentre consente, grazie alla disponibilità dei dati istituzionali sulla popolazione, un'accurata stratificazione del campione.

4A. LA NUMEROSITÀ DEL CAMPIONE

Nella pratica della ricerca si ritiene che un sondaggio politico (nazionale, ma anche locale),

Figura 2
Numerosità del campione e errore campionario nella stima di una percentuale



per stimare ad esempio una percentuale, possa essere condotto su un campione di 1.000 casi con sufficiente accuratezza. Nel caso di un campione casuale semplice (a puro titolo di esempio) il margine di errore campionario si collocherebbe attorno al $\pm 3\%$. Cioè se la stima campionaria (ad esempio la percentuale di intervistati che apprezza il candidato X o il ministro Y) è pari al 40%, ciò significa che il valore "vero" nella popolazione è compreso tra 37% e 43% con almeno 95 probabilità su 100. Un margine forse sopportabile nell'*issue polling*, assai meno in un sondaggio preelettorale sulle intenzioni di voto ai partiti.

Aumentare la numerosità del campione per ridurre il margine di errore è una strategia costosissima e di fatto impraticabile. Per scendere dal 3 al 2% bisognerebbe quasi triplicare la numerosità del campione, come mostra il grafico della Figura 2. Soluzione improponibile nella maggior parte dei casi, e ancor più nel caso dei sondaggi politici che possono normalmente far affidamento su budget

molto limitati. L'alternativa è una precisa e appropriata stratificazione del campione e un accurato controllo sui meccanismi di autoselezione del campione.

In linea di principio campioni inferiori ai 1.000 casi, se non in casi limitati e particolari, sono quindi da ritenersi di qualità inferiore allo standard ASSIRM, sia a livello nazionale che locale.

La numerosità del campione è comunque una variabile decisiva per la qualità del sondaggio politico. Nei sondaggi preelettorali perché è pur vero che un campione più ampio non contribuisce molto alla riduzione dell'errore campionario, ma in genere – se associato a un questionario più articolato – consente di ridurre l'incertezza delle stime con opportune tecniche di analisi¹⁹ dell'area del non voto²⁰, ma soprattutto degli elettori indecisi o reticenti dai quali dipende in sostanza la bontà delle "previsioni"²¹.

¹⁹ Ad esempio l'analisi discriminante o simili.

²⁰ Astenuti, schede bianche e nulle.

²¹ Per quanto ci si sforzi di chiarire che un sondaggio elettorale non è una previsione di voto, perché può esserci molta differenza tra intenzioni dichiarate e comportamenti, perché ci sono ancora molti indecisi,

Ma anche nell'*issue polling* un campione numericamente robusto è condizione determinante della qualità del sondaggio, perché di norma i risultati vengono presentati in tavole bidimensionali²² e in questi casi l'errore delle stime campionarie relative alle sottocategorie aumenta perché l'aumento dell'errore derivante dalla diminuzione del numero di casi non è mai o quasi mai compensato da una diminuzione di pari entità indotta dalla riduzione della varianza interna. Ne deriva un'incertezza nell'interpretazione delle differenze tra i dati delle sottocategorie, in genere poi neppure sostenuta da calcoli di significatività statistica.

4B. CAMPIONI PROBABILISTICI E NON, MULTICANALITÀ

Il ricorso al campionamento probabilistico (legato soprattutto all'utilizzo del telefono fisso) è ormai in molti casi superato dalla pratica della ricerca, che fa largo uso di canali alternativi per l'esecuzione delle interviste. Tutti i "campioni" (o "sotto campioni") che si ottengono da popolazioni di cui è incerta la stima della numerosità e della composizione sono campioni per definizione non probabilistici e come tali non potrebbero dare alcuna garanzia scientifica sulla generalizzabilità dei risultati alla popolazione degli elettori.

L'uso di campioni probabilistici è quindi ancora preferibile quando l'esigenza di assicurare che i risultati rientrino entro un certo intervallo di confidenza (non calcolabile nei campioni non probabilistici) prevale su quella della più elevata precisione possibile delle stime. Così ad esempio in un sondaggio sulle intenzioni di voto non avrebbe molto senso garantire che la percentuale di voti stimata per un partito sia il 5% con un intervallo del $\pm 3\%$, mentre lo avrebbe in un *issue poll* che stima al 50% il giudizio positivo degli elettori su un determinato provvedimento.

Da molto tempo per la verità nella pratica della ricerca sociale e politica l'uso di campioni non probabilistici, o di disegni di ricerca misti, ha dato risultati comparabili, in termini di accuratezza delle stime, a quelli ottenuti con procedure di *random sampling*. La pratica non ha quindi controindicazioni²³ per un sondaggio di qualità, soprattutto se è finalizzata all'aumento della numerosità campionaria. L'uso anche parziale di campioni non probabilistici (auto-selezionati, con compenso, panel internet, ecc.) va quindi chiaramente indicato e documentato nella SQA, riportando per ciascuno numerosità, criteri di autoselezione (*opt-in*) e stratificazione.

8

perché tanti elettori decidono solo all'ultimo momento come votare, ecc. ecc., la pubblicazione di un sondaggio preelettorale su un quotidiano, un periodico o una trasmissione tv di rilevanza nazionale mantiene una forza profetica che vanifica spesso i pur sinceri tentativi di ridimensionarne il potere previsionale. Cfr. più avanti quanto detto in proposito sulla pubblicazione dei sondaggi preelettorali.

²² "L'approvazione dell'azione del Governo è maggiore nel Nord o nel Sud?", "tra i giovani o tra gli anziani?" ecc.

²³ Ad esempio l'AAPOR non ha ancora espresso un giudizio definitivo sull'uso di campioni non probabilistici, ma si limita a mettere in guardia che "questo tipo di campioni può non rappresentare tutta la popolazione oggetto della ricerca". Cfr. AAPOR, *Sampling Methods for Political Polling*, p. 3.

4C. LA STRATIFICAZIONE DEL CAMPIONE

Almeno nella situazione italiana le correlazioni tra orientamenti politico-elettorali e variabili sociodemografiche sono piuttosto deboli, se non quella della distribuzione territoriale del voto. Per questo nei sondaggi politici la stratificazione del campione in base a queste sole variabili porta in genere modesti margini di miglioramento delle stime campionarie.

L'alternativa di qualità è quella di aggiungere ai criteri di stratificazione tradizionali almeno una variabile politica, di solito un precedente comportamento di voto a un'elezione o a un referendum, anche se questa soluzione può aggravare tempi e costi per il completamento della campagna di interviste e la chiusura delle "quote". Infatti la domanda "Può dirmi come ha votato alle ultime elezioni?" implica sempre un errore sistematico derivante da quell'effetto noto come "spirale del silenzio", effetto mediatico che tende appunto a "silenziare" i comportamenti "perdenti" o le opinioni minoritarie²⁴.

L'effetto spirale del silenzio si manifesta non solo con una maggiore reticenza a dichiarare il proprio comportamento di voto passato, ma anche nell'autoselezione degli elettori che concedono l'intervista. Tenerlo sotto controllo con la stratificazione del campione è quindi condizione dirimente per valutare la qualità del sondaggio.

Un secondo elemento da tenere in considerazione perché può creare un insidioso *bias* nella composizione del campione è il livello di istruzione dell'intervistato. È noto, e ciò vale per tutte le ricerche sociali, ma non solo, che la propensione a rispondere è proporzionale al livello di istruzione e può essere molto difficoltoso ottenere un campione correttamente proporzionato in termini di titolo di studio dell'intervistato.

Poiché riempire tutte le quote può rilevarsi impraticabile a causa dei tempi e dei costi, si potrà chiudere la fase di rilevazione quando la somma degli scarti in valore assoluto tra valori ponderati e non, non superi i 10-15 punti percentuali, lasciando poi alla ponderazione il compito di ridurre al minimo questo valore.

9

4D. LA PONDERAZIONE DEI DATI

Per la ponderazione dei dati si utilizza la procedura di *sample balancing* su tutti i criteri di stratificazione utilizzati. È una procedura iterativa di ponderazione del campione che riequilibra contemporaneamente le distribuzioni di tutte le variabili di stratificazione e minimizza gli scarti dalle distribuzioni della popolazione con l'utilizzo dell'algoritmo di Stephan-Deming²⁵.

La SQA riporta la procedura di *sample balancing* utilizzata e per ciascuna modalità delle variabili di stratificazione il valore del parametro della popolazione, il valore campionario non ponderato e quello ponderato. Meno rilevante, e forse non consigliabile nel caso dei

²⁴ L'effetto può assumere dimensioni notevoli nei momenti in cui l'opinione pubblica è percorsa da eventi squassanti o di forte impatto emotivo. Emblematico il caso dei sondaggi condotti dopo le elezioni politiche del 1992 nel pieno dell'operazione "Mani pulite". Alle elezioni del 1992 i partiti dell'area governativa (DC, PSI, PSDI, PRI, PLI) avevano ottenuto oltre il 50% dei voti validi, ridotti a meno del 15-20% nei sondaggi successivi dopo l'esplosione del "ciclone" giudiziario.

²⁵ O metodi analoghi già presenti nei principali package di analisi dei dati, come il metodo Newton-Raphson utilizzato in SPSS rel. 25.

sondaggi politici, è la ponderazione in base al canale utilizzato per l'intervista (CATI, CAMI, CAWI). L'utilizzo di formule, come ad esempio quella riportata nella figura seguente, può eventualmente essere utilizzata prima della ponderazione in base alle variabili sociodemografiche e politiche.

Figura 3 Esempio di ponderazione per canale (telefono fisso e mobile)

The first-stage weight for the i^{th} case can be expressed as:

$$WT_i = \left[\left(\frac{S_{LL}}{F_{LL}} \times \frac{1}{AD_i} \times LL_i \right) + \left(\frac{S_{CP}}{F_{CP}} \times CP_i \right) - \left(\frac{S_{LL}}{F_{LL}} \times \frac{1}{AD_i} \times LL_i \times \frac{S_{CP}}{F_{CP}} \times CP_i \right) \right]^{-1}$$

Where S_{LL} = the size of the landline sample
 F_{LL} = the size of the landline sample frame
 S_{CP} = the size of the cell sample
 F_{CP} = the size of the cell sample frame
 AD_i = Number of adults in household i
 $LL_i=1$ if respondent has a landline phone, otherwise $LL=0$.
 $CP_i=1$ if respondent has a cell phone, otherwise $CP=0$.

5. I SONDAGGI SULLE INTENZIONI DI VOTO

I sondaggi preelettorali sulle intenzioni di voto, pubblicati con gran risalto da giornali e tv e inevitabilmente interpretati come previsioni (soprattutto in imminenza delle elezioni, ma non solo)²⁶, sono per i non esperti una prova della validità scientifica della teoria statistica del campionamento e dell'attendibilità delle rilevazioni tramite intervista. Sono quindi la principale (se non l'unica) fonte che rassicura sull'attendibilità dei risultati di tutte le ricerche sociali e di mercato. Anche per questo, pur essendo i sondaggi politici in genere una piccola parte del lavoro degli istituti, un insuccesso delle previsioni elettorali effettuate con sondaggi (e ancor più gli *exit polls*) può essere molto negativo per l'attendibilità presso i media e il grande pubblico delle tecniche di intervista e dei metodi di campionamento e per l'immagine degli istituti, con ripercussioni in tutta l'area della ricerca sociale e di mercato²⁷.

Sembra quindi ragionevole pretendere che un sondaggio sulle intenzioni di voto abbia standard di qualità anche superiori a quelli di un *issue poll*. Ma la definizione di tali standard non è facile né univoca.

Il primo problema per i sondaggi sulle intenzioni di voto è quello della definizione della popolazione di riferimento, che in teoria dovrebbe essere la popolazione residente da 18

²⁶ Per quanto *pollster* e giornalisti dichiarino (ma forse senza troppa convinzione) che un sondaggio pubblicato sulle intenzioni di voto "non è una previsione" del risultato delle elezioni, i dati, per il modo con cui sono raccolti, presentati e commentati ("se si votasse oggi...") hanno una componente "profetica" ineliminabile, che è poi quella che in larga misura fa la notizia e dà risalto alla testata.

²⁷ "Pre-election polls make up only a minority of published surveys. They are however, a very public test of sampling theory and survey research in action. Polls have a good track record for accuracy but the occasional poll which appears to be wrong gets extensive media coverage with a very negative impact on the image of opinion polls and surveys and opinion research in general". ESOMAR / WAPOR, Guideline on Opinion Polls and Published Surveys, 2014, p. 5.

anni in poi con diritto di voto (alla Camera). In realtà ottenere un campione proporzionale per tutti gli elettori è praticamente impossibile, perché, qualunque sia il canale di reclutamento, il processo di autoselezione del campione che deriva dalla volontarietà dell'intervista determina una sotto rappresentazione degli elettori marginali, meno interessati o coinvolti nella politica e che dunque tenderanno a rifiutare l'intervista o a rispondere in modo reticente. Sono questi poi in larga misura gli stessi elettori che non andranno a votare, che non daranno un voto valido o che non risponderanno alla domanda sulle intenzioni di voto.

È pur vero che è sempre possibile insistere fino a quando non viene riempita anche la quota dei non votanti (ad esempio alle precedenti elezioni per ponderare correttamente il campione), ma è un'operazione giudicata il più delle volte inutile e dispendiosa, perché comporta tempi più lunghi e costi più elevati per ottenere interviste in larga misura non utilizzabili per la stima delle intenzioni di voto²⁸. Anche per questo motivo è in genere necessario stratificare e ponderare per una variabile di precedente comportamento elettorale, pur sapendo della possibilità di introdurre una fonte di errore se è forte l'effetto "spirale del silenzio"²⁹.

Con il procedimento tradizionale le stime percentuali delle intenzioni di voto ai singoli partiti o candidati si ottengono utilizzando solo i voti validi, cioè solo le intenzioni di voto dichiarate. È questa una possibile fonte di errore perché tali stime sono basate su tre assunzioni spesso non esplicitate³⁰

- i non votanti sono gli stessi della precedente elezione
- è nullo l'effetto "spirale del silenzio"
- è uguale la distribuzione delle intenzioni di voto degli elettori reticenti e non

a cui si aggiunge il maggior margine di errore campionario derivante dalla notevole diminuzione del campione. Tra intervistati non votanti, reticenti e indecisi il campione originario di 1000 casi si può anche ridurre della metà, facendo così salire il solo margine di errore campionario a oltre il $\pm 4,5\%$.

La qualità di un sondaggio sulle intenzioni di voto si misura in primo luogo sul metodo utilizzato per far fronte a queste possibili fonti di distorsione, che se si aggiungono a un elevato errore campionario possono portare il *total survey error* a valori incompatibili con la finalità stessa della ricerca. Ciò è vero a maggior ragione se il sondaggio è condotto a ridosso di una tornata elettorale, quando il valore previsivo dei dati è più difficilmente negabile.

²⁸ Si tratta di una quantità di interviste notevole. Ad esempio, alle ultime elezioni politiche la percentuale di votanti è stata del 72,9% e in un campione non distorto di 1000 casi dovremmo quindi prevedere 271 interviste a non votanti, proprio quelle più difficili da ottenere per l'alto numero di rifiuti e l'alta probabilità che si tratti di elettori che non andranno a votare alle prossime elezioni o che risponderanno in modo reticente.

²⁹ Vedi su questo quanto detto al precedente par. 4C sulla stratificazione del campione.

³⁰ E sempre meno verosimili per il continuo aumento dell'astensionismo e della fluidità dei comportamenti di voto.

Rispetto a quello tradizionale prima richiamato, i metodi utilizzati per una più accurata stima delle intenzioni di voto possono essere diversi³¹ e anche egualmente apprezzabili. Piuttosto che l'uso dell'una o dell'altra tecnica, a rivelare la qualità del sondaggio sarà la presenza nel questionario di domande formulate appositamente per la determinazione della base di calcolo e la profilazione degli elettori reticenti e indecisi.

In queste operazioni ciascun istituto di ricerca dovrebbe avere la libertà di adottare le soluzioni che ritiene più appropriate con l'unico vincolo, se tiene al giudizio di qualità, di presentarle nella SQA nel modo più chiaro e completo possibile.

³¹ Uno, ad esempio, è quello di utilizzare come base per il calcolo delle percentuali di voto non i voti validi dichiarati, ma i “probabili votanti” selezionati in base a una apposita domanda del questionario e stimando con opportune tecniche statistiche le probabilità di voto a ciascuna lista o candidato.