

# Metodologia CATI w badaniach sondażowych

SONDAŻ POLSKI. METODA, ETYKA, MEDIA



Stanisław Mirowski, Research Services Director, Millward Brown  
Warszawa, 8 listopada 2013 r.

# CATI – warunki realizacji

---

# CATI to:

- Telefon
  - Dobra jakość połączeń
  - Wysoka penetracja (dobre pokrycie populacji badanej)
  - Automatyczne centrale
  - Jednolity plan numeracyjny
- Komputer
  - Tani i powszechnie dostępny
  - Połączony w sieć z serwerem (sieci LAN)
  - Coraz częściej – Internet – podstawa CATI @ HOME , w tym komunikacji VOIP

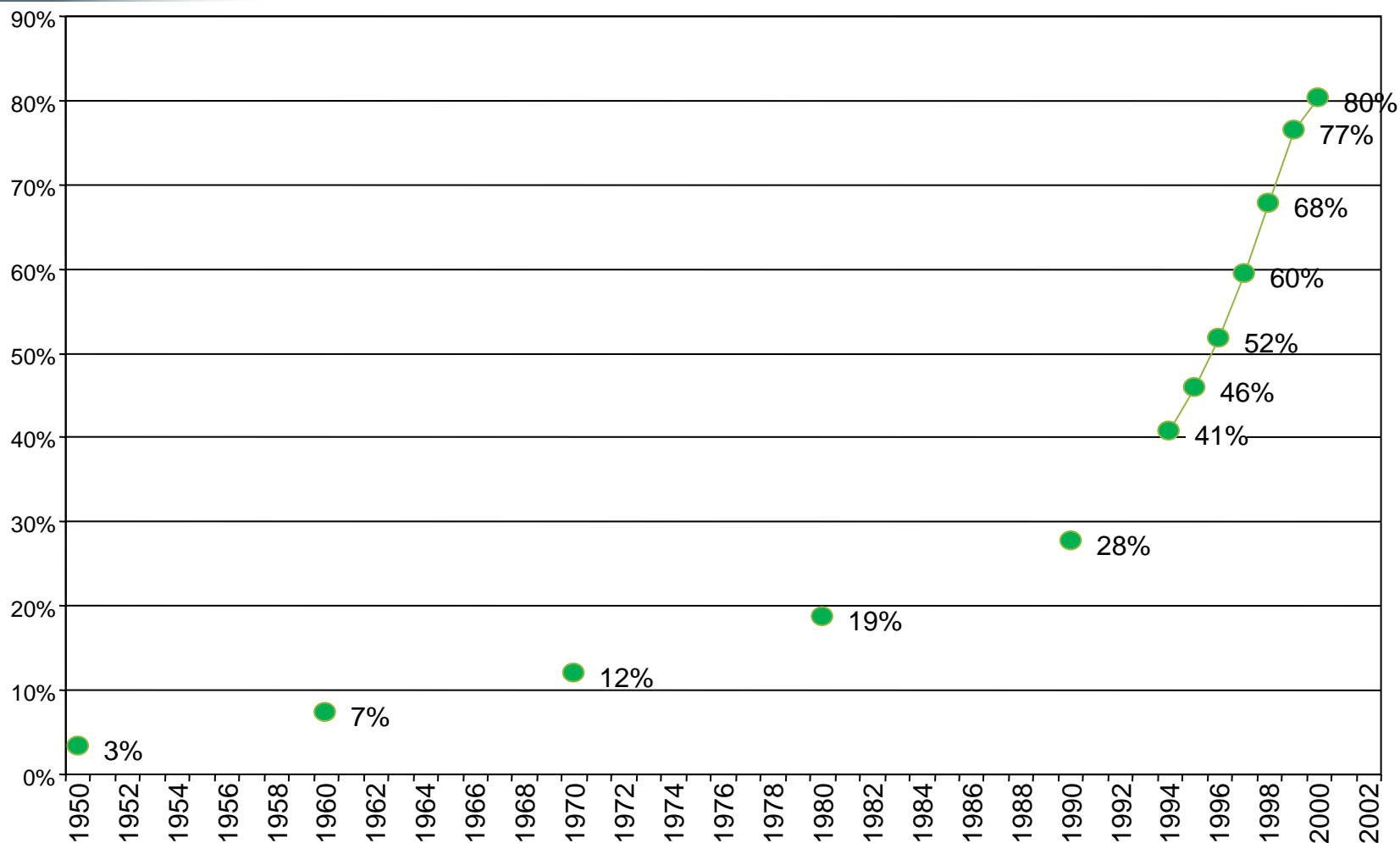


*W całym cywilizowanym świecie telefon poprzedzał komputer, w Polsce i innych byłych krajach bloku wschodniego, to komputer poprzedzał telefon*

# Rozwój badań telefonicznych – CATI (świat)


- 1966 pierwsza próba badań telefonicznych
- 1970 początek szerszych zastosowań
- 1980 konkurencja dla F2F
  - Metodologia sondaży
  - RDD
  - CATI
- 1990 maksymalny rozwój
  - Bardziej ekonomiczne
  - Szybsze
  - Lepsza kontrola ankietera
  - Spadek response rate w badaniach F2F
- 2000 spadek efektywności
  - Konieczność wprowadzenia tel. komórkowych

# Odsetek gospodarstw w Polsce posiadających telefon stacjonarny (1950-2000)



Źródło: GUS

# Rozwój badań telefonicznych – CATI (Polska)

- 1995 pierwsze instalacje CATI (PENTOR, SMG/KRC)
- 1996 badania B2B i B2C w dużych miastach
- 1998-2000 PTBRiO raportuje osobno badania CATI i „inne badania telefoniczne”
- 1999 pierwsze ogólnopolskie próby CATI 
- 2005 likwidacja ostatnich centrali ręcznych, wprowadzenie jednolitego planu numeracji
- 2005 maksymalna penetracja telefonii stacjonarnej ok. 35 łączy na 100 mieszkańców
- 2006-2007 zrównanie dostępności populacji 15-75 pod telefonem komórkowym i stacjonarnym
- 2009 S. Mirowski i W. Hołdakowski prezentują referat na IX Kongresie Badaczy Rynku i Opinii „Próba mix telefonów stacjonarnych i komórkowych”
- 2010 komisja powołana przez OFBOR (prof. Domański, prof. Markowski, dr Sawiński i prof. Sztabiński) sporządza raport „Ocena metodologii i rezultatów badań poprzedzających 1 i 2 turę wyborów prezydenckich” i konstatuje, że „zastosowanie techniki CATI umożliwiło sformułowanie trafniejszych prognoz niż zastosowanie wywiadu osobistego” (str.41 raportu)

*Raport Komisji nie przesądzał o wyższości CATI nad F2F...*



# Przykład przygotowania próby CATI – SMG/KRC 1999

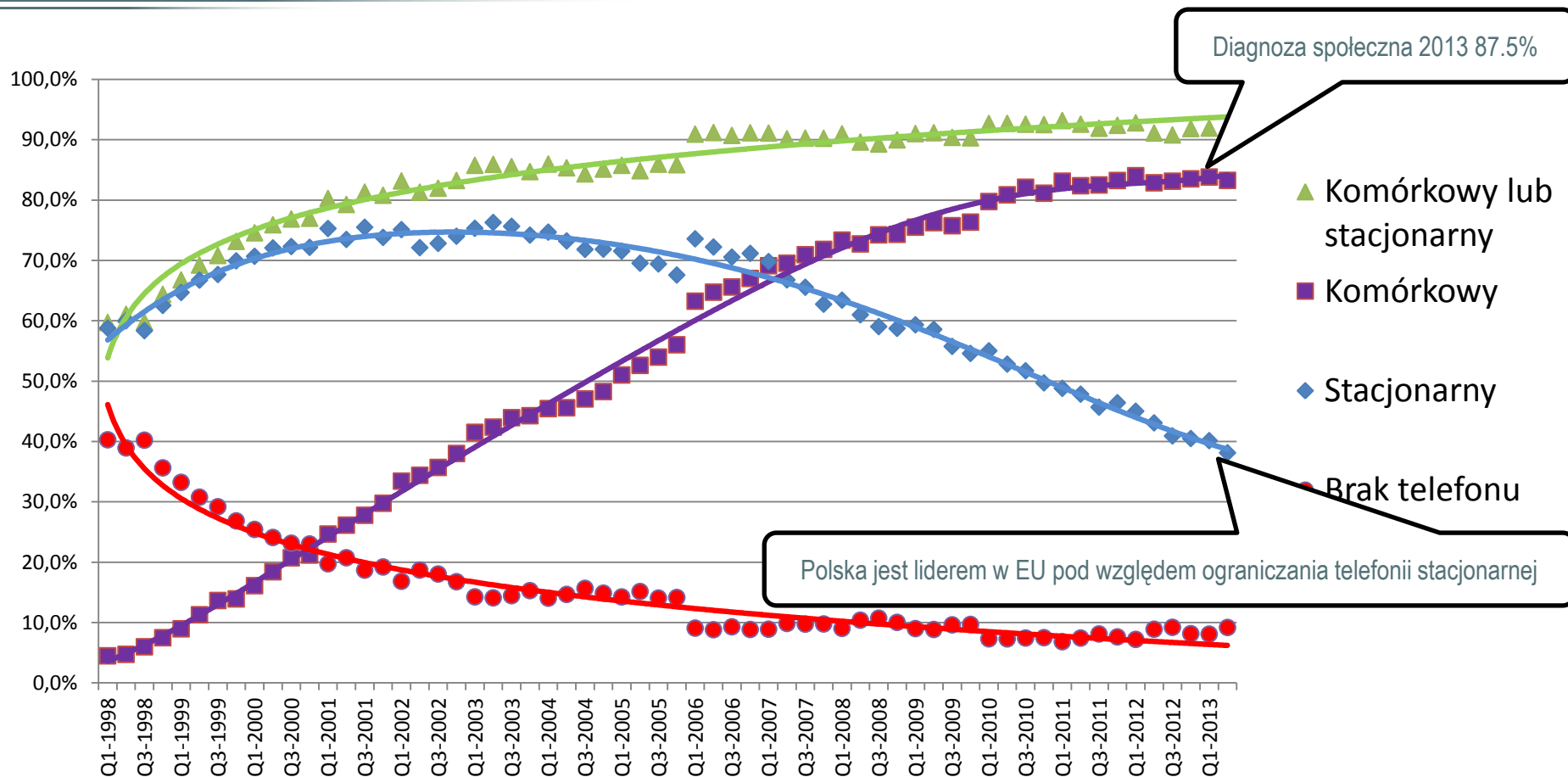
20.12.1999

MIASTA				MIASTA							
Kasa	Makro	KodGUS	Nazwa	Województwo-49	Kierunek	Telefon1	Telefon2				
2	1	13133	Gliniów	Ciechanowskie	0-23	6740029		3	5	77261	Ustka
2	1	73233	Warta	Sieradzkie	0-43	8294027		3	5	81241	Nowogard
2	2	31213	Pyzdry	Koniejskie	0-63	768139		3	6	5269	Siemiatycze
2	2	83173	Klasko	Poznańskie	0-61	4270019		3	6	5273	Sokołka
2	3	27309	Ogródzieniec	Katowickie	0-32	2946447		3	6	45157	Kolno
2	3	53301	Niemodlin	Opolskie	0-77	606128		3	6	51205	Lidzbark Warmiński
2	4	21301	Witnica	Gorzowskie	0-95	7515087		3	6	51241	Nidzica
2	4	39189	Przemków	Legnickie	0-76	8310438		3	6	55173	Meków Mazowiecki
2	4	89105	Bardo Śląskie	Wałbrzyskie	0-74	171207		3	6	79153	Goldap
2	5	19129	Gniew	Gdańskie	0-69	1352048		3	6	79205	Pisz
2	5	81321	Trzcianko-Zdrój	Szczecińskie	0-91	146807		3	7	3165	Międzyrzec Podlaski
2	6	5305	Wasilków	Białostockie	0-85	7185012		3	7	3193	Radzyń Podlaski
2	6	51269	Rzeszel	Olsztyńskie	0-89	7550005		3	7	11185	Włodawa
2	8	37117	Brzozów	Krośnieńskie	0-13	4341052		3	7	43133	Dęblin
2	8	65201	Radymno	Przemyskie	0-16	6281037		3	7	43297	Ryki
2	8	69177	Kolbuszowa	Rzeszowskie	0-17	2271090		3	7	71133	Garwolin
3	1	1107	Brwinów	Warszawskie	0-22	7295307	dobry nr.	3	7	71297	Sokołów Podlaski
3	1	1151	Konstancja Jeziorna	Warszawskie	0-22	7175399	dobry nr.	3	7	71337	Węgrów
3	1	1175	Marki	Warszawskie	0-22	7811056	dobry nr.	3	8	29137	Miechów
3	1	1183	Milanówek	Warszawskie	0-22	7242290	dobry nr.	3	8	29441	Włoszczowa
3	1	47119	Konstantynów Łódzki	Łódzkie	0-42	2111036	dobry nr.	3	8	35267	Wieliczka
3	1	59193	Koluszki	Piotrkowskie	0-44	7140056	dobry nr.	3	8	49193	Limanowa
3	1	67173	Grójec	Radomskie	0-48	6641030	dobry nr.	3	8	49193	Łęka
3	1	75213	Rawa Mazowiecka	Świętokrzyskie	0-49	8142152	dobry nr.	3	8	89201	Łęka
3	2	9301	Solec Kujawski	Bydgoskie	0-52	3871060	dobry nr.	3	8	83217	Nisko
3	2	9333	Tuchola	Bydgoskie	0-52	3342010	dobry nr.	3	8	85121	Brzesko
3	2	25185	Kapno	Kaliskie	0-62	7822258	dobry nr.	4	1	1123	Grodzisk Mazowiecki
3	2	25249	Ostrzeszów	Kaliskie	0-62	7302180	dobry nr.	4	1	1195	Nowy Dwór Mazowiecki
3	2	25257	Pleszew	Kaliskie	0-62	7421178	dobry nr.	4	1	1207	Piaseczno
3	2	25297	Syców	Kaliskie	0-62	7852240	dobry nr.	4	1	1211	Piastów
3	2	57117	Czarnków	Piłskie	0-67	2552043	dobry nr.	4	1	13101	Ciechanów
3	2	57229	Trzcianka	Piłskie	0-67	2161101	dobry nr.	4	1	13185	Mława
3	2	57277	Zielonka	Piłskie	0-67	2632189	dobry nr.	4	1	47103	Aleksandrów Łódzki
3	2	63313	Szamotuły	Poznańskie	0-61	2920080	dobry nr.	4	1	47131	Ozorków
3	3	7177	Kęty	Bielskie	0-33	450156	dobry nr.	4	1	61153	Kosylin
3	3	7297	Wadowice	Bielskie	0-33	731061	dobry nr.	4	1	67213	Kozienice
3	3	7321	Wisła	Bielskie	0-33	8551419	dobry nr.	4	1	67245	Pionki
3	3	15105	Błachowina	Częstochowskie	0-34	3204017	dobry nr.	4	1	73101	Sieradz
3	3	27247	Lędziny	Katowickie	0-32	2169000	dobry nr.	4	1	73269	Zduńska Wola
3	3	27355	Pszów	Katowickie	0-32	4540049	dobry nr.	4	1	75101	Skiermiewice
3	3	27433	Trzebinia	Katowickie	0-32	6110095	dobry nr.	4	1	75257	Zyrardów
3	3	53297	Namysłów	Opolskie	0-77	100068	dobry nr.				

Pierwsze podejścia do losowania prób ogólnopolskich wiązały się z koniecznością sprawdzania operatu losowania (aktualności numerów telefonów); warstwowanie próby i losowanie jednostek administracyjnych, jak w próbie terenowej...



# Penetracja telefonii stacjonarnej i komórkowej w populacji 15+



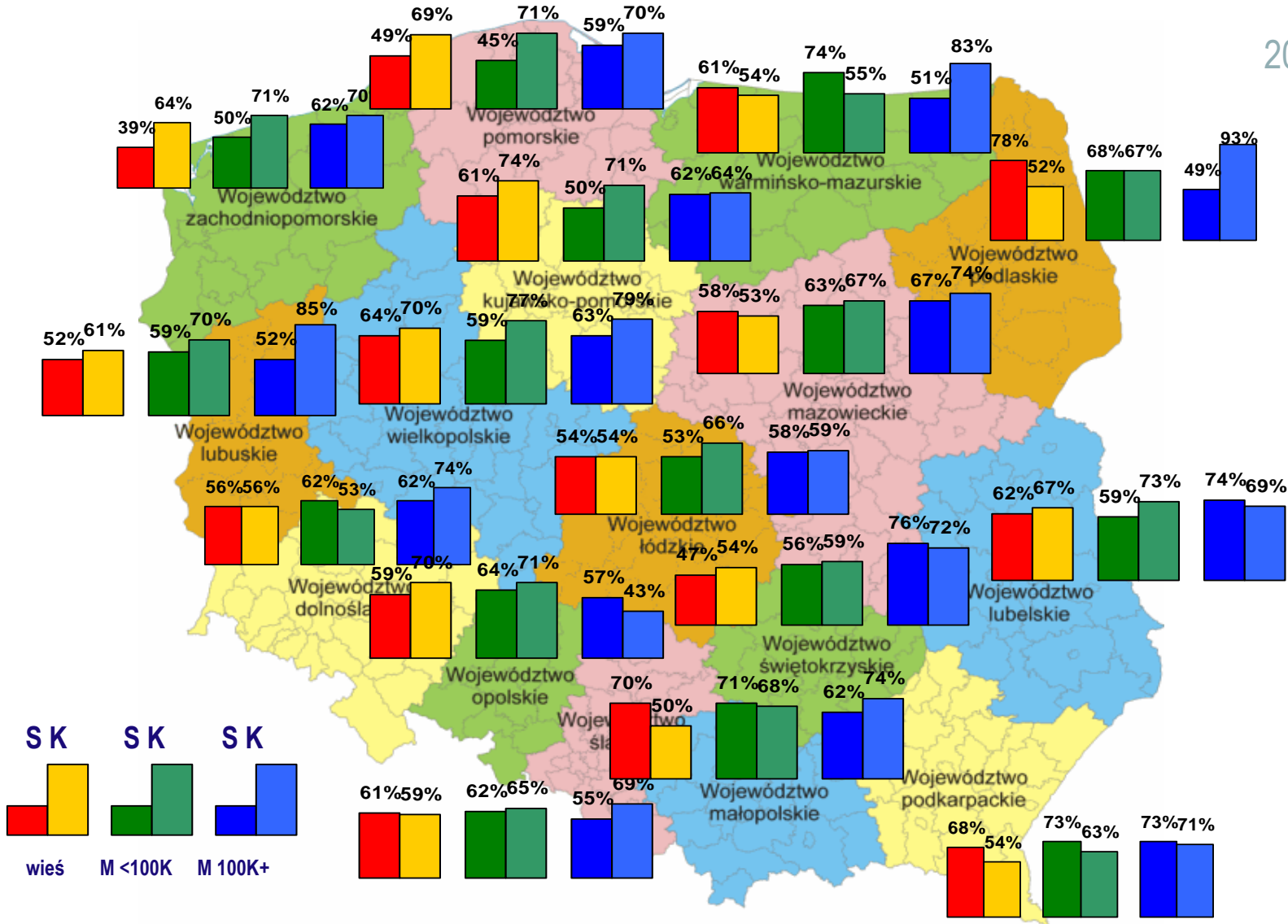
źródło: TGI, MB, N=9000/kwartał

*Na przełomie 2006-2007 dostępność osób 15+ stała się większa pod telefonem komórkowym niż pod telefonem stacjonarnym w domu*



# Posiadanie telefonów stacjonarnych i komórkowych

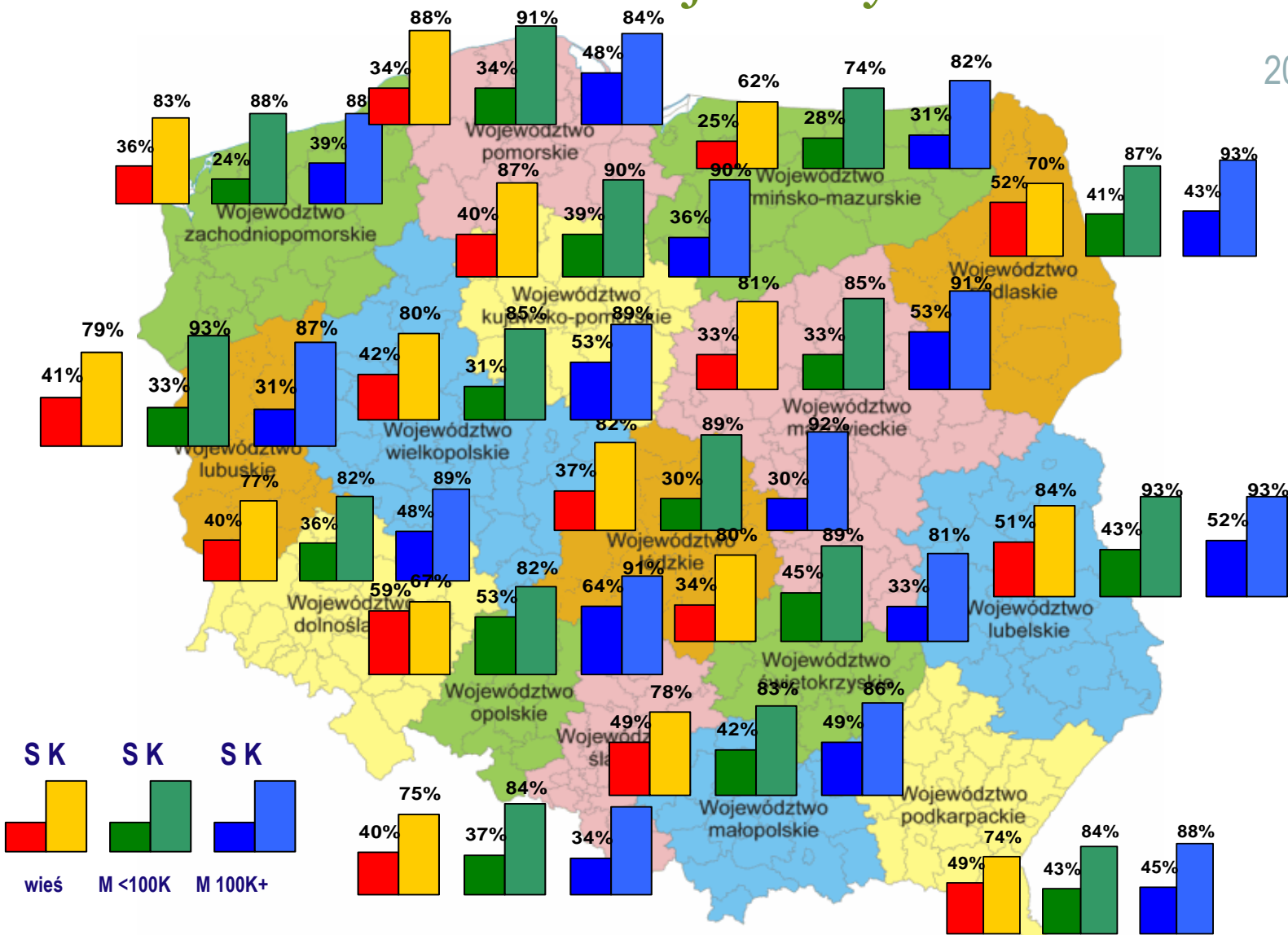
2007/2008



źródło: AGB Establishment Survey, 2007 / 2008, N = 8000 (realizacja MB SMG/KC)

# Posiadanie telefonów stacjonarnych i komórkowych

2012/2013



źródło: Millward Brown, TGI VII/2012-VI/2013 N=31968



# Próby CATI – Dual frame

---

# Błędy pokrycia populacji, telefony stacjonarne

Pewne kategorie badanych osób są niedoreprezentowane w dobieranej próbie (bądź w ogóle nie występują)  
(np.: osoby w wieku 25-34 lat; osoby nieposiadające telefonu stacjonarnego w domu)

Pewne kategorie osób są nadreprezentowane  
(np.: osoby w wieku 60-75 lat)

2013

Osoby należące do populacji niedostępne pod telefonem stacjonarnym w miejscu zamieszkania nie wchodzi do próby badawczej

Osoby dostępne pod telefonem stacjonarnym wyłącznie w miejscu pracy nie podlegają badaniu

Mimo, że operat doboru próby pokrywa się z populacją badaną, niektóre jednostki występują nieproporcjonalnie często, a inne rzadko

Osoby dostępne pod telefonem stacjonarnym

Populacja badana

# Błędy pokrycia populacji, telefony komórkowe

2013

Pewne kategorie badanych osób są niedoreprezentowane w dobieranej próbie (bądź w ogóle nie występują)

(np.: osoby starsze; osoby nieposiadające telefonu stacjonarnego w domu)

Pewne kategorie osób są nadreprezentowane

(np.: osoby młodsze)

Osoby należące do populacji niedostępne pod telefonem komórkowym

W praktyce trudno odróżnić telefony komórkowe prywatne od służbowych, badania są prowadzone na obu typach, więc nie ma osób poza populacją, do których się można dodzwonić

Osoby dostępne pod telefonem komórkowym

Mimo, że operat doboru próby pokrywa się z populacją badaną, niektóre jednostki występują nieproporcjonalnie często, a inne rzadko

Populacja badana

# Błędy pokrycia populacji, mix telefonów stacjonarnych i komórkowych

Pewne kategorie badanych osób są niedoreprezentowane w dobieranej próbie (bądź w ogóle nie występują)  
(np.: osoby nieposiadające telefonu stacjonarnego w domu i niedysponujące komórkami)

Pewne kategorie osób są nadreprezentowane  
(np.: osoby dysponujące wieloma komórkami lub/i telefonem stacjonarnym w domu)

Osoby należące do populacji niedostępne ani pod telefonem stacjonarnym w miejscu zamieszkania ani nie dysponujące telefonem komórkowym, obszar zdecydowanie ograniczony

Osoby dostępne pod telefonem stacjonarnym wyłącznie w miejscu pracy i nie posiadające telefonu komórkowego nie podlegają badaniu

Mimo, że operat doboru próby pokrywa się z populacją badaną, niektóre jednostki występują nieproporcjonalnie często, a inne rzadko, efekt ten częściowo jest kompensowany przez odmienne profile populacji stacjonarnej i komórkowej

Osoby dostępne pod telefonem komórkowym

Osoby dostępne pod telefonem stacjonarnym

Populacja badana

# Proporcje populacyjne 2013



Źródło: Millward Brown, TGI 2Q2013

- Na początku badań z próbą mieszaną (stacjonarne – komórki) proporcje były odwrotne, a koszt połączenia na telefon komórkowy powodował, że do próby telefonów stacjonarnych dołączano niewielką próbę telefonów komórkowych – i to jedynie osób, które nie miały telefonów stacjonarnych
- Obecnie proporcje i ceny nie preferują połączeń stacjonarnych
- Dążenie do odwzorowania w próbie populacyjnej struktury dostępu do telefonii stacjonarnej i komórkowej wymaga:
  - Dokładnego poznania tej proporcji dla badanej populacji (trudne przy szybko zmiennym rynku telefonii stacjonarnej)
  - Dokładnego określenia statusu respondenta jako użytkownika telefonii

*Przy krótkich, kilku pytaniowych sondażach wypytywanie respondenta o jego status telefoniczny może okazać się czynnikiem zakłócającym wywiad*

# Łączenie próby telefonów stacjonarnych i komórkowych

## Telefony stacjonarne

- jako dobro przynależne gospodarstwu domowemu – są podstawą próby osób dobieranej poprzez gospodarstwa
- mają zwykle określoną lokalizację – możliwe warstwowanie terytorialne

## Telefony komórkowe

- stanowią próbę osób – odbierający jest respondentem
- Nie mają określonej lokalizacji – trzeba ją wyznaczyć w trakcie wywiadu i uwzględnić w warstwowaniu

*Charakter obu prób wskazuje na konieczność „domykania” stratyfikacji terytorialnej poprzez próby telefonów stacjonarnych*

*W obu przypadkach jako operat losowania próby można przyjąć bazy losowych numerów, generowanych z uwzględnieniem formatu 9-cyfrowego i zakresów numeracji przydzielanych poszczególnym operatorom*



# Łączenie próby telefonów stacjonarnych i komórkowych

- Poprawne określenie prawdopodobieństw doboru poprzez oba rodzaje telefonii, nastrocza – jak wspomniano – dużych trudności; konieczne jest określenie w trakcie wywiadu:
  - Liczby wszystkich aktywnych telefonów komórkowych w dyspozycji respondenta
  - Liczby wszystkich linii telefonii stacjonarnej w gospodarstwie domowym
  - Liczby członków gospodarstwa domowego korzystających z telefonu stacjonarnego w domu i spełniających kryterium przynależności do badanej populacji
- Niezbędne są też dane populacyjne dotyczące liczby łącz telefonii stacjonarnej w gospodarstwach i liczby telefonów komórkowych (te liczby powinny dotyczyć jedynie urządzeń służących do prowadzenia rozmów)
  - Nie jest łatwo o takie dane; „aktywna karta SIM” może służyć do transmisji danych, a do łącza stacjonarnego może być podłączony fax...

*Brak precyzyjnych danych populacyjnych i błędne określenie statusu telefonicznego respondenta może być przyczyną pomyłek przy szacowaniu prawdopodobieństw doboru respondenta*

# Ważenie próby CATI

- Najbezpieczniej jest ważyć bez uwzględniania sposobu doboru respondenta
- Jako cechy ważeniowe należy przyjąć rozkład płci i wieku respondenta oraz rozkład terytorialny (klasa wielkości miejscowości i region)
- W badaniach telefonicznych obserwuje się niedomiar respondentów z wykształceniem podstawowym, poziom wykształcenia powinien być również uwzględniany w procesie ważenia

# CATI – zalety metody

---

# Podstawowe zalety metody CATI

- Nie ma problemu wiązkania próby – większe rozproszenie próby
- Koszt dotarcia do respondenta nie zależy od lokalizacji respondenta, a realizacja próby ogólnopolska kosztuje tyle, co lokalnej
- Ankieterzy zapisują umówienia na dogodną dla respondenta godzinę i dzień; powroty do respondenta realizowane są przez dowolnego ankietera – nie stanowi to problemu organizacyjnego ani dodatkowego kosztu
- Ankieter realizuje próbę w całej populacji – ograniczenie efektu ankieterskiego
- Płace ankieterów opierają się na stawkach godzinnych, a nie osiąganym sukcesie – mniejsza pokusa oszustw ankieterskich
- Bieżący nadzór supervisorski zwalnia z konieczności prowadzenia kontroli po badaniu
- Rozmowa telefoniczna zapewnia większą anonimowość respondentowi
- Pokrycie populacji próbą mix telefonów komórkowych i stacjonarnych (92-96%) nie znajduje konkurencji w żadnym operacie stosowanym do losowania prób terenowych

*Sondaż CATI ma przewagę wszędzie tam, gdy trzeba zbadać szybkozmiennie zjawisko i wystarczy do tego celu zadać kilka prostych pytań*

# Pozycjonowanie badań telefonicznych

- Wywiady CATI to obecnie najpopularniejsza metoda badań ilościowych - ok. 30% wszystkich wywiadów realizowanych przez firmy poddające się audytowi OFBOR
- Niestety po dziś dzień pokutują nawet w środowisku badawczym stwierdzenia bazujące na wiedzy sprzed dwóch dekad:
  - Że badania CATI bazują na telefonii stacjonarnej
  - Że telefony są dostępne tylko w dużych miastach
  - Że penetracja telefonii jest znikomo niska
- Tymczasem CATI to mix telefonii stacjonarnej i komórkowej, pokrywający ok. 95% populacji dorosłej w Polsce

*Przyszłość – patrz Andy Peytchev, Benjamin Neely, „RDD Telephone Surveys. Toward a Single Frame Cell-Phone Design”, Public Opinion Quarterly, Vol. 77, No.1 Spring 2013, pp. 283-304*

# Dziękuję za uwagę

---