

RDD サンプルングにおける稼働局番法の再評価

Re-evaluating Working Exchange Design of RDD Sampling

島田喜郎

Yoshiro Shimada

株式会社日経リサーチ (Nikkei Research Inc.)

客員

連絡先

〒101-0048 東京都千代田区神田司町 2-2-7

E-mail: yshimada@mbd.nifty.com

Summary:

Working exchange design of sampling for RDD (random digit dialing) telephone survey uses a frame consisting of all possible telephone numbers in all the working exchanges and draws a telephone number sample in a single stage random sampling. This type of sample was believed to be practically useless, as the residential hit rate of such samples are very low.

In RDD sampling of recent years, however, it has become standard practice to apply computer assisted screening of non-working numbers to telephone number samples. Unfortunately, we can hardly find any papers which take into consideration the effect of screening of non-working numbers in evaluating sampling designs. If samples drawn with working exchange design are executed with computer assisted screening, their residential hit rate can reach a practical executable level. The design should be re-evaluated since it has become suitable for practical use.

The list-assisted design, which is recently used frequently because of its high household hit rate, has a problem of truncating telephone households in banks which contains only unlisted residential numbers. This paper tries to demonstrate that the working exchange design can attain a residential hit rate which is almost comparable with that of list-assisted design, as well as the full coverage of telephone household population.

キーワード: RDD 電話調査 サンプルング 稼働局番法 電話帳準拠法

Key words: RDD, telephone survey, sampling, working exchange design, list-assisted design

1. はじめに

1.1. RDDサンプリングの原理

電話調査に使われるRDD(random digit dialing)サンプリングの母集団は固定電話の加入世帯(以後「電話世帯」という)である。NTTの統計によると、「加入電話」と「ISDN」を併せた固定電話の契約件数は、2002 年末で「住宅用契約」が 4,265 万件、「事務用契約」が 1,819 万件^注であった。世帯数に対する住宅用契約件数の比率はほぼ 90%になる。契約件数の中には、一世帯で2以上の契約を持つものも数えられている。一方、「事務用契約」で住宅用として併用されているものも存在する。これらを勘案すると、全国の全世帯に対する固定電話加入世帯の比率は 90%前後に達しているものと考えられる。また固定電話加入率が 96%であるという調査結果(佐藤、2002)もある。このように高い普及率を持つ電話システムを、標本抽出と対象者へのアプローチの両方の手段として活用しようとするのがRDD電話調査である。

世帯用固定電話の電話帳掲載率が 100%であれば、電話帳を電話世帯の母集団リストとすることができる。しかし「個人名電話帳」の掲載件数は 2,800 万程度で、世帯用電話の契約件数の 65%程度にしかならない。RDD サンプリングの目的は、電話帳掲載番号にも非掲載番号にも等しい抽出確率を与えて世帯用番号の標本を作ることである。

電話帳掲載の有無に関わらず世帯用番号の等確率標本をつくる目的のために、世帯用番号全

^注 事務用契約には1契約に複数の番号が割り当てられるビル電話という契約があるため、事務用電話番号の件数はこの契約数を相当上回ると考えられる。

体を、あるいはほぼ全体を包含し、かつリスト可能な抽出フレームを定義する。このフレームには使われていない電話番号や事業所用の電話番号が混入するのを避けることはできない。この抽出フレームの電話番号から無作為抽出を行って電話番号の標本をつくる。得られた標本が RDD 電話番号標本であり、この電話番号標本の中で世帯用番号と判明したものを電話世帯の標本とする。

1.2. 二つの抽出フレーム

1.2.1. 稼働局番フレーム

上記の抽出フレームを単純に具体化したものが、「稼働局番フレーム」である。固定電話システムで稼働中の局番、すなわち現に存在している市外局番と市内局番の組み合わせ(たとえば 012-345-xxxx の数値部分)をすべて集め、その中の 0000 から 9999 の番号を抽出単位とする。以下では、稼働局番フレームから単純な一段階の無作為抽出で RDD 標本を得る方法を「稼働局番法」とし、抽出フレームそのものと区別する。(稼働局番フレームは Waksberg 法や非比例層化抽出法にも使われる。これらの手法については後に説明する。)

稼働局番フレームは現在の固定電話システム内に存在し得るすべての電話番号を含む。当然、すべての世帯用電話番号もその中に包含される。稼働局番フレームに含まれる局番の数は全国で約 1 万 9 千、この中の電話番号の数は 1 億 9 千万に上る。全国の世帯用番号の件数は約 4 千 3 百万であるので、稼働局番フレーム中の世帯用番号の比率は 20%強程度である。このフレームから単純無作為抽出した電話番号が世帯に当たるヒット率も平均で 20%強程度に過ぎない。

1.2.2. 電話帳準拠フレーム

稼働局番フレームの世帯ヒット率の低さを改善するために考えられたのが、「電話帳準拠(list-assisted)フレーム」である。電話帳準拠フレームは、電話帳に掲載されている世帯用番号が1以上あるいは一定件数以上含まれるバンクを有効バンクとして、有効バンクだけを集めて抽出フレームを作る。

バンクは局番の中を一定の大きさのブロックに分けたもので、012-345-67xxのように局番の先頭から百位までの数字が同じ100個の番号のブロックを百位バンクと呼ぶ。十位や、千位のバンクも考えられるし、局番は万位のバンクである。実際のRDD調査で最もよく使われるのは百位バンクである。

局番の中の電話番号は、世帯用、事業所用、非使用のものが入り混じっているが、それぞれが特定のバンクに比較的集中して存在する。その結果、電話帳掲載の世帯用番号が1以上または数件以上ある「有効バンク(residential bank)」には世帯用番号がある程度固まって存在すると考えられるし、「無効バンク(non-residential bank)」には世帯用番号がまったく無いか、わずかしき無いと考えられる。無効バンクを排除し、有効バンクだけを集めて抽出フレームを作ることにより、抽出フレーム内の世帯用番号の密度が大幅に高まり、標本の世帯ヒット率を高めることができる。この抽出フレームでは有効バンクのリストが作られるが、抽出単位は当然その中の個々の電話番号である。百位バンクならバンク内に末尾が00から99の電話番号が列挙されていると想定して処理する。

千位バンクを使うより百位バンクを使うほうが、電話帳掲載番号を含まないバンクをきめ細かく排除できるので世帯ヒット率は大幅に高まる。しかしバンクの大きさを小さくすると、後に述べるよう

に、総ての世帯用番号が非掲載であるバンクが排除される「脚きり」の問題がより深刻になる。また、バンクを有効とする掲載件数の基準値によっても世帯用番号の密度は違ってくる。有効判定基準が1の場合よりも3の場合の方が世帯ヒット率は高まるが、「脚きり」の問題は大きくなる

1.3. 本論文の目的

稼働局番法はRDDの原理をそのまま具体化したともいえるものであり、RDDを考える人々の思考の中に最初から存在していた。しかし、この方法は世帯ヒット率が非常に低く、実用的な手法として利用できないと考えられていたため、この手法の実施例の報告はほとんどない。

最近のRDD調査では、抽出後の電話番号標本に対して、コンピュータを使って「使われていない番号」のスクリーニングが行なわれるのが普通である。従来のRDD手法の議論はこのスクリーニングの効果を手法の評価に結びつけることをしなかった。スクリーニングによって、稼働局番法の実査段階における世帯ヒット率は大幅に改善される。稼働局番法が現実的に実施不能な手法ではなくなったのであり、実用的な有力な手法として再評価されるべきである。この論文では、日経リサーチが日経新聞社のために実施している「日経電話世論調査」のデータを使って、稼働局番法を再評価するための実証的なデータを提示する。

日本でも、米国でも現在主流と考えられているRDD手法は電話帳準拠法(list-assisted design)である。電話帳準拠法で作った電話番号標本は、スクリーニング処理をする前でも、非常に高い世帯ヒット率を示す。しかし、この手法には電話世帯全体がカバーされないという「脚きり」問題をはじめ、電話帳情報を使うことに関係するさまざまな問題

がある。この論文のもう一つの目的は、稼働局番号が電話世帯全体をカバーする最も単純な方法でありながら、スクリーニング処理をした後の世帯ヒット率では電話帳準拠法に近い水準に達することを実証することである。

2. 非使用番号のスクリーニング

2.1. スクリーニングとは

非使用番号のスクリーニングとは、抽出した RDD 電話番号標本に対して機械的な処理を行って「使われていない番号」を事前に判定することである。これにより RDD サンプリングの最大の問題点である実査の非効率性を大幅に改善することができる。これは、標本抽出後に実査を開始する前に行う重要なプロセスである (Battaglia, M., Starer, A., Oberkofler, J., Zell, E.(1995), 松田(2002), 鈴木(2003))。

電話システムは、ダイヤルされた電話番号を受け取ってから、実際に交換作業を開始する前に、その番号の性状を示す電子信号を発信側に返してくる。その電子信号によって使われていない番号を判定できる。コンピュータを使って RDD 標本の番号をダイヤルしてこの処理を行うと、迅速に、多数の電話番号を非使用番号としてふるい分けできる。スクリーニングの結果、非使用と判定された番号には当然電話をかけることはしない。しかし、非使用の番号を別の番号で入れ替えることもしない。その番号のディスポジション (disposition=どのように処理されたかという分類) を「非使用」と確定し、その番号に対する処理が終わったものとする。結果として、それらの番号は電話をかける実査業務からは排除される。

非使用番号のスクリーニングについて、もう一つ重要なポイントはそれが非常に安価なことであ

る。電話番号一件につき数円の費用しかかからない。

電話システムが発信側に返してくる信号はさまざまである。あるものは、単純明快に使われている、いないを判定出来るが、あるものは状況に依存していて判定が困難なものもある。スクリーニングの目的からすると、明確に「使われていない」と判定出来るもの以外は「使われている」として分類の方が安全である。そのためこの論文で「使われていない番号」「非使用番号」というときは、使われていないことが確かな番号だが、スクリーニング段階で「使われている番号」と称するものの中には曖昧なものも含まれていて、実査の結果、最終的には「非使用」と処理されるケースも少なくない。スクリーニングで「使われている番号」と判定された番号が実査の対象になるので、以下では、状況によって「実査対象番号」と呼ぶこともある。

機械的スクリーニングの対象として、事業所用の電話番号の判定も実務的に活用されていることが多い。この論文では後に述べるように、実証のために使ったデータにおいて、事業所用電話番号のスクリーニングを行っていないので、これについては詳述しなかった。

2.2. スクリーニングの効果

機械的なスクリーニングで非使用と判定される番号の比率は、使用した抽出フレームにより大幅な違いがある。全国の稼働局番フレームを使った日経リサーチの経験では、非使用番号の比率は64%(表3参照)に達する。前田・土屋(2001) は一都三県で千位バンクの電話帳準拠フレームを使って、自動判定された非使用番号が26%(表3参照)であると報告している。両者の間にさまざまな条件の違いがあるとはいえ、非使用番号の比率

の差の主因は抽出フレームの違いにあると考えられる。電話帳準拠フレームでは、電話帳掲載番号を含まない非有効バンクを排除することによって、非使用番号の混入率の低い抽出フレームが得られている。しかし、事前における非使用番号の混入率の大小にかかわらず、いずれの方法においても、スクリーニングを経た後には、使用されていると判定された電話番号だけのリストが入手できる。

3. 電話世帯カバレッジと脚きり問題

日本でも、米国でも現在主流の RDD 手法は、電話帳準拠法(list-assisted design)である。電話帳準拠フレームには脚きりの問題がある。すなわち、世帯用番号がすべて非掲載であるバンクや、電話帳編集完了後に使われるようになったバンクは、世帯用番号があっても抽出フレームから排除される。そのためこの抽出フレームは、別名脚きりフレーム(truncated frame)と言われることがある(Casady and Lepkowski 2003)。脚きりに対する対処の仕方としては1)脚きりはやむを得ないとして容認する、2)脚きりによるバイアスを検証した上で脚きりを容認する、3)脚きりが発生しない手法を採用する、などのタイプ考えられる。日本において脚きりについてどのような対応がとられているか、そして、RDD の長い歴史を持つ米国においてはどうかであったのかを概観してみよう。

3.1. 脚きりについての日本の現状

日本では最近の文献に電話帳準拠法を扱った論文が相次いでいる(前田・土屋 2001、松田 2002、佐藤 2002、鈴木 2003、土屋・前田 2003)。鈴木(2003)は、電話帳準拠法について「現行の

実用的 RDD 法は…電話帳に依存する RDD 法である」と述べている。

松田(2002)は、「Bank もれの問題がつきまとうが」、運用面を考えれば百位バンクを使うのが最善と考えている。残念ながら松田(2002)はバンクを有効とする基準の件数がいくつであるかを記述していない。鈴木(2003)は、電話帳掲載件数 1 以上の百位バンクを前提に議論しているが、脚きりされるバンクに含まれる世帯用電話が全体の中で「どのくらいの率になるかを確かめた例をみたことはない」としながら、「実査の効率に重点を置くだけでは問題が生じる可能性がある」という危惧を表明している。前田・土屋(2001)と土屋・前田(2003)は同じ一つの実施例についての報告であり、1都3県で電話帳掲載番号を10個以上含む有効千位バンクを抽出フレームとして使っている。掲載番号10件以上の千位バンクと掲載件数1件以上の百位バンクでは、どちらが脚きりが発生しやすいだろうか。著者らはバンクの大きさが異なる場合のヒット率の検討にとどまっている。

3.2. 米国の初期の電話帳準拠法

Frankel and Frankel (1977)によると、1976年の A. C. Nielsen 社の冊子に、電話帳掲載件数3以上の百位バンクで構成した抽出フレームを使う方法が紹介されていたという。また、1980年代になると、SSI(Survey Sampling Inc.)が、掲載件数3以上の百位バンクを使って作成した RDD 標本を調査会社などに販売するサンプリング専門のビジネスを行っていた(Lepkowski 1988)。「掲載件数3以上の百位バンク」という有効バンクの決め方から、この当時の電話帳準拠法が、ヒット率を高めて実査の効率を高めることに重点を置き、ある程度の脚きりが発生することは容認するというものであったことがうかがえる。

3.3. Waksberg法の登場

脚きりが発生しない手法として成功をおさめたのが、Waksberg (1978)の方法である。Waksberg法の実施例の報告は日本では非常に少ない。城川(2003)がその希有の例であり、米国のBRFSS調査(Behavioral Risk Factor Surveillance System = 行動における危険因子の調査)に倣って医療関係の分野にWaksberg変法を適用する試みを行っている。

Waksbergの方法を要約すると、確率比例抽出法を用いた二段抽出に相当する。稼働局番フレームから、一段目で百位バンクを抽出し、二段目で百位バンクの中の個別の番号を抽出する。一段目の抽出で、世帯用番号を含まない百位バンクが選ばれないように工夫されているので、二段目の抽出では世帯ヒット率が大幅に向上する。この方法が強く支持された理由は、稼働局番フレームを使って電話世帯全体をカバーしたうえで、高い世帯ヒット率を達成したことである。その結果、それまで正統な調査手法として評価されていなかった電話調査が、政府機関の重要な調査にも使われるようになり、Waksberg法は電話調査の標準的なサンプリング手法としての地位を確立した。(Waksberg法の詳しい説明は城川(2003)または島田(2004)参照のこと。)

しかし、Waksberg法にも大きな問題があった。サンプリングと実査が複雑に絡み合っているため管理が難しく、人手を要し時間がかかるという問題だ。時間やコストを重視する最近の調査利用者からこの問題は非常に重大視されるようになった。

3.4. 新しい電話帳準拠法

Brick, Waksberg, Kulp and Starer(1995)の研究は、

脚きりによるバイアスを検証した上で電話帳準拠フレームを容認することにより、Waksberg法の使用を回避しようとする試みである。1以上の掲載番号を持つ百位バンクで抽出フレームを作ると、1)脚切りによるカバレッジの損失は電話加入世帯の3-4%に過ぎないこと、また2)脚切りされたバンクの中の世帯と標本中の世帯の間には、さまざまな特性値でみて有意な差が見られないことを確認した。そして、脚きりによるバイアスは重大ではないと結論した。Giesbrecht, Kulp and Starer (1996)も別の方法を使ってほぼ同様の結論を出した。これらの論文は非常に大きな影響力を持った。多くの重要な調査がWaksberg法を捨てて電話帳準拠フレームの採用に踏みきり、調査報告書の中でこれらの論文を引用して手法の正当性を主張するようになった(島田2003)。

電話帳掲載の状況は時代とともに変化する。1995年以前の電話帳掲載率がほぼ70%であった米国の状況をベースに行われた研究の結果を現在の米国の状況に当てはめることが果たして妥当といえるだろうか疑問がある。もちろん、電話帳掲載状況は国によっても大きな違いがある。米国の研究結果を、そのまま現在の日本に当てはめることが無理なことは自明といえよう。(ちなみに、現在の日本の電話帳掲載率は65%程度である)。

3.5. 非比例層化抽出法

城川(2003)がWaksberg法の参考にしたBRFSSはどうなっただろうか。BRFSSは、米国保健省の指導の下に各州が主体になって実施する調査であり、サンプリング手法の選択も各州が自主的に行う。BRFSSは標本設計のルールとして電話世帯全体をカバーすることを要求している。電話帳準拠法はこのルールのために対象外とな

る。島田(2003)によると、1995年には50州中35州でWaksberg法が使われたが、2000年にWaksberg法を使ったのは53州中1州だけだった。2000年に53州中50州で採用された手法は、非比例層化抽出法(disproportionate stratified design)であった。これはCasady and Lepkowski(1993)が提案した方法であり、彼らもまたWaksberg法の煩雑さを回避するためにこの方法を考案したのである。

非比例層化抽出法では、稼働局番フレームの中を電話帳掲載番号の件数で層化し、層により抽出率を変えることで精度または費用を最適化しようとする。たとえば、掲載1件以上の百位バンク層(1+バンク層)と掲載0件の百位バンク層(0バンク層)に層化して、一段階の抽出で電話番号の標本を作る。重要なパラメータは、「 γ =適格な番号で調査を完了する費用 / 不適格な番号を確定する費用」である。 γ が10倍程度のとき、0バンク層での抽出率を1+バンク層の抽出率の3分の1程度にするのが最適とされる。この手法の特徴は、0バンク層の抽出率を低くして費用を抑制しながら、この層も抽出の対象とすることにより電話世帯全体をカバーすることである。すなわち、Waksberg法に代わる新しい「脚きりが発生しない手法」が生まれたのである。

3.6. 非比例層化抽出法の先にあるもの

Casady and Lepkowski(1993)は0バンク層から抽出しないという選択もあり得ると考えたが、脚きりによるバイアスの程度が不明であることが問題だと指摘した。これをうけて電話帳準拠フレーム(すなわち0バンク層から抽出しない方法)のバイアスの大きさを検証したのが先のBrick et al.(1995)の研究である。0バンク層の抽出率を0にするのは不適格な番号を確定する費用が巨大

で、 γ の値が非常に小さい場合に相当する。一方不適格な番号を確定する費用が非常に安価で、 γ の値が50とか100のような大きい値になると、0バンク層でも1+バンク層と同じ抽出率を適用するのが効率的ということになる。結果としてこの抽出法は2層に共通の抽出率を適用する単純な層化抽出法になる。先に述べたように現在では、コンピュータを使って、自動的に、安価に、非使用番号のスクリーニングができるので、 γ の値は非常に大きいと見なすことができる。それを前提とすると層により抽出率を変える意味はほとんどない。さらに、何らかの理由で(たとえば層化のためのデータとして、バンク別に電話帳掲載の世帯用番号の件数を確認することがコスト的に見合わないとして)電話帳情報による層化を断念すれば、結果としてこの手法は、この論文の主題である稼働局番法に帰着する。

4. 稼働局番法の実際

4.1. 稼働局番法の実施方法概要

この節では、脚きりが発生しない手法として、稼働局番フレームから一段階の無作為抽出でRDD標本を作る稼働局番法の実施例を紹介する。この方法は、RDDの原理そのものであり、原理的には1.2.1節で述べた稼働局番法に、2節で述べた非使用番号のスクリーニングを加えたもので、それ以外に特筆すべきことはあまりない。

日経リサーチでは2002年の9月から、日経新聞社の定期世論調査をこの方法で実施している。この調査の実施方法の概要は以下の通りである。

- 1) 本来の母集団は日本全国の有権者だが、実質の母集団は日本全国の電話世帯の有権者とする。

- 2) 稼働中の約 18,900 の局番の 1 億 8900 万件の電話番号から 18,000 の電話番号を系統抽出する。
- 3) 抽出した電話番号標本に対してコンピュータを使って非使用番号のスクリーニングを実施する。非使用と判定した番号は処理済みとし、残りの使用されている番号だけを実査の対象とする。事業所の電話番号のスクリーニングはしない。標本の追加は一切行わない。
- 4) 実査は木曜日から日曜日の午前まで(以前は日曜日の全日を含む金土日の 3 日で実施していたことがある)。また臨時にもっと短期間に実施したケースもある。日経リサーチのウェブサイト過去のすべての調査の結果が掲示されている。
- 5) 世帯内の個人は、世帯の有権者数以内の乱数を発生させ年齢順に乱数番目に当たる人を選ぶ。選ばれた人が電話に出ている人でないならその人を電話口と呼んでもらう。その人が不在なら、再度電話する。
- 6) 回収率の計算は有権者がいることが判明した世帯を分母とする。
- 7) 集計では、属性構成を調整するための重み付けを行わない。(ただし、世帯内有権者数と世帯の電話番号数による抽出率の違いの調整は行う。)

この調査について、表 1 に最近の二回分の調査概要を示した。二回分の調査概要は、各回の調査の変動が比較的小さく、調査が安定した枠組みの元を実施されていることを示している。また、表 2 にはその二回の調査の性別・年齢別の属性構成を示した。属性構成は母集団の構成とよく一致している。ただし、20 歳代の構成比が低めにでる傾向が見られる。20 歳代の構成比が低くなるのは、面接調査などでも見られる現象であり、この調査やRDD調査の固有の問題とはいえない。たとえば、内閣府の「国民生活に関する世論調査」は、住民基本台帳からの無作為標本を使った訪問面接調査だが、平成 6 年以来、10 回の調査で、20 歳代の構成比が 12%2 回、11%4 回、10%4 回となっている。

表1 日経電話調査の実施概要

調査実施日		年	2003	2004
		日	12/18-12/21	2/12-2/15
		曜日	木金土日	木金土日
		掲載	12月22日	2月16日
抽出総数			18,000	18,000
スクリーニング		非使用番号	11,541	11,590
		使用番号(実査対象番号)	6,459	6,410
実査結果概要	世帯	有権者数	3,694	3,712
		有権者いる	3,116	3,033
		回収数	1,916	1,913
		回収率(%)*1	61.5	63.1
	非世帯	有権者無し又は有無確認不能	1,200	1,120
		事業所など*2	578	679
		非使用番号	1,617	1,576
		世帯か事業所か不明	22	24
判定不能	世帯か事業所か不明	189	207	
	不対話 *3	937	891	

- *1 分母は「有権者がいる」と確認できた世帯
 *2 事業所、公衆電話、FAX、外国人世帯など
 *3 呼出音、話中音、留守番電話を含む

表2 日経電話世論調査の標本構成

	2003年12月	2004年2月	母集団構成
全体	1,916	1,913	100%
性別			
男性	47%	46%	48%
女性	53%	54%	52%
年代別			
20歳代	9%	10%	18%
30歳代	17%	20%	17%
40歳代	17%	18%	16%
50歳代	21%	19%	19%
60歳代	18%	18%	15%
70歳代	16%	13%	15%
拒否	2%	1%	

- ※ 世帯内有権者数と電話本数で抽出確率の調整を
 ※ 上記以外の重み付けは行っていない
 ※ 母集団構成は2003年1月の住民基本台帳人口で

4.2. 電話帳準拠法との比較

以下では、前田・土屋(2001)の中での RDD 調査の部分(以下では「前田土屋調査」とする)と比較しつつ、日経電話世論調査(以下では「日経調査」とする)の特徴を論じる。前田土屋調査は 1 都 3 県の電話帳掲載件数 10 以上の千位バンクを使った電話帳準拠法である。地域が全国でなく、比較の対象として最適とはいえないが、比較可能な形でデータが提供されている数少ない例である。

表 3 は、両調査のスクリーニングの結果を比較している。前田土屋調査の実査対象番号(すなわち、非使用でない電話番号)の比率は 74%で、日経調査の 36%に比べて非常に高い。この比率の高さが電話帳情報を使った効果の一つである。両者の比率の差を埋めるために、人間が一つ一つ電話をかけて確認するのだとしたらこの差は非常に大きな意味を持つ。しかし、両調査ともコンピュータで非使用電話番号を判定するスクリーニングを行っており、スクリーニングが終わると両者の差は埋め合わされてしまう。

前田土屋調査は事業所の電話帳データベースを使って事業所用番号を判定するスクリーニングを行っている。一方、日経調査は事業所用番号のスクリーニングを行っていない。スクリーニングを行わない理由は、1) 事業所用番号のスクリーニングのコストが比較的高いこと、2) 事業所用番号としてスクリーニングできるのは、電話帳に掲載

されている番号、その多くが代表番号、だけであり、事業所用番号がすべてスクリーニングできるわけではない、3) 飲食業や床屋などで事務用と契約した電話が世帯用に併用されている可能性が考えられ、それらが自動的に排除されてしまう危険がある、などである。表 3 では、両調査の比較を可能にするため、前田土屋調査のスクリーニングで事業所用と判定された 838 件の電話番号を実査対象の電話番号に加えている。

スクリーニングして「使用されている電話番号」と判定されたものが実査の対象である。表 4 によると、実査の対象となった電話番号のうち、世帯に当たる比率(世帯ヒット率)は日経調査で 58%、前田土屋調査で 61%である。電話帳準拠フレームでは、使われている番号があっても、それが事業所を中心とする非世帯用電話番号だけのバンクが排除される結果として、世帯ヒット率が高まるのが期待されているのだが、前田土屋調査と日経調査の世帯ヒット率の差は意外に小さい。千位バンクでは非世帯用番号の排除の効果をあまり期待できないということかもしれない。百位バンクであれば、一つのバンクを一つの企業が占有するケースが多いため、そうしたバンクを排除した後の世帯ヒット率がもう少し高まることは当然期待できる。しかし、それを判断できる材料は今手元にない。一方、逆の視点でみると、電話帳情報を使わない日経調査でも相当高い世帯ヒット率を達成できているといえるのではないだろうか。

表3 非使用番号のスクリーニングの効果

	日経調査(a)		前田土屋調査(b)	
標本抽出数	18,000	100.0%	10,769	100.0%
非使用番号	11,590	64.4%	2,780	25.8%
使用番号 (実査対象番号)	6,410	35.6%	7,989	74.2% 注1

出典 (a) 日経電話世論調査2004年2月
(b) 前田・土屋(2001)

注1 前田土屋調査では、839件が事業所用番号としてスクリーニングされたが、ここでは比較のため実査対象番号に含めた。

表4 実査対象電話番号に対する実査結果の比較

実査対象番号	日経調査 (a)		前田土屋調査 (b)	
	数	割合	数	割合
世帯	6,410	100.0%	7,989	100.0% 注4
回収	3,712	57.9%	4,901	61.3%
非回収	1,913	29.8%	1,651	20.7%
非世帯	1,799	28.1% 注1	3,250	40.7% 注5
判定不能	1,600	25.0% 注2	1,892	23.7% 注4、注6
判定不能	1,098	17.1% 注3	1,196	15.0%

(a), (b)とも出典は表3と同じ

注1 有権者なし又は有権者確認不能 (679) を含む

注2 事業所, 公衆電話, FAX, 外国人世帯など (1576) と非使用番号 (24)

注3 不対話 (891) を含む

注4 スクリーニングした事業所用電話 (838) を含む

注5 該当者なし (84) を含む

注6 事業用番号 (1441) に加え非使用番号 (253)、Fax (189) を含む

(ちなみに、Waksberg 法の第 2 段階の世帯ヒット率は 60%を越える程度(Lepkowski 1988)であるので、日経調査の世帯ヒット率は、Waksberg 法と比べても大差ないといえる。)

稼働局番法の特徴は 1)原理的に電話世帯全体をカバーできること、2)1 万 9 千件程度の局番データベースでサンプリングができることである。一方、電話帳準拠法を実施するには、バンクの世帯用番号の件数を知るために、電話帳に掲載されている 3 千万件に近い世帯用電話番号のデータベースを構築し、電話帳が更新されるたびにデータベースも更新すること、あるいはそのようなデータベースを利用するサービスを購入することが必要である。電話帳準拠法は、1)電話世帯の脚きりと、2)巨大なデータベースの構築と維持、という対価を払って、実査段階での世帯ヒット率の高さを買っているといえる。しかし、スクリーニングを介することによって、両者の世帯ヒット率の差はほとんど無くなってしまふ。電話帳準拠法が払った大きな対価に対して得たものは上記の例ではたった 3 ポイント(61%-58%)の世帯ヒット率の高さであった。百位バンクを使えばもう少し得るものは大きいという期待は持てるが、はたして十分満足できるレベルになるだろうか。

5. まとめ

稼働局番法は世帯ヒット率が低く、RDD の原理を考える思考過程においてのみ存在し得ると考えられていた。しかし、スクリーニングを介することにより、稼働局番法の世帯ヒット率は他の手法に比肩するレベルに高めることができる。その結果、稼働局番法は電話世帯全体をカバーする最も単純で、十分に効率的な手法として再評価されるべきである。

現在、米国でも、日本でも主流を占めるといわれる電話帳準拠法は、脚きりという問題を抱えている。Brick et al.(1995)が、脚きりの問題は重大ではないと実証したとはいえ、その根拠は世帯用電話番号の電話帳掲載率という変化の激しい事象をベースにしたものであり、Brick 等のお墨付きがいつまでも意味をもつとは言い難い。電話帳準拠法は、電話帳データベースを維持しなければならない。そのデータベースの規模は、稼働局番法で使われる局番データベースに比べて遙かに巨大なものである。さらには、個人情報保護法との関連で、調査の目的の為に電話帳データベースを維持することが問題とされる状況が発生しないとも限らない。こうしたことを全て容認した上で、得られる世帯ヒット率の優位が、稼働局番法に比

べて数ポイントにすぎないとしたら電話帳準拠法がより現実的と考える理由はない。

5.1. 今後の課題

RDD サンプルングでは、ある電話番号の世帯がどの地域に属するかを確定することが困難である。局番がどの市区町村をカバーしているかという正確な情報は購入することができる。しかし、多くの局番が、市区町村の境界をまたいで、場合によっては都府県の境界を越えたサービスエリアを持っている。これは、全国レベルの世論調査などではほとんど問題にならないが、選挙区別を実施する調査のように、正確な地域区分が必要な調査では大きな障害になる。電話の相手に質問して適格地域内の世帯かどうかを確認しなければならない。

もし、百位バンクや千位バンクのレベルで見ると、その中の電話番号の世帯の住所が非常に狭い範囲に限定されるのが普通であるとすれば、正確な地域区分が必要な調査において電話帳に掲載されている電話番号の住所情報を活用する価値が大きいだらう。しかし、稼動局番法の中で、このような住所情報をどのように活用することができるかは未検討の課題である。

5.2. 謝辞

草稿にたいし貴重な助言をいただいた鈴木督久氏(日経リサーチ・マーケティング局長)に深く感謝します。当然のことながら、稼動局番法の実現は、同氏をはじめ、日経リサーチの多くの人々の協同の成果であること付け加えておきたいと思えます。また、拙論に丁寧なコメントをいただいた審査員の先生方に心から感謝します。

6. 参考文献

- Battaglia, M., Starer, A., Oberkofler, J., Zell, E. (1995), Pre-identification of nonworking and business telephone numbers in list-assisted random-digit-dialing samples, *Proceedings of the Survey Research Methods Section*, American Statistical Association, pp 957-962.
- Brick, J., Waksberg, J., Kulp, D. and Starer, A. (1995), Bias in list-assisted telephone samples, *Public Opinion Quarterly*, Vol. 59, No. 2, pp. 218-235.
- Casady, R. and Lepkowski, J. (1993), Stratified telephone survey designs, *Survey Methodology*, Vol 19, No 1: pp. 103-113.
- Frankel, M. R., and Frankel, L. R. (1977), Some recent developments in sample survey design. *Journal of Marketing Research*, Vol. 14, pp. 280-293.
- Giesbrecht, L., Kulp, D., Starer, A. (1996) Estimating coverage bias in RDD samples with Current Population Survey data, in *Proceedings of the Survey Research Method Section*, American Statistical Association, pp. 503-508.
- 城川 美佳 (2003) 医療・保険領域の電話調査 Waksberg 変法、行動計量学、58号、pp. 111-119.
- Lepkowski, J. M. (1988), Telephone sampling methods in the United State, in Groves, R., Biemer, P., Lyberg, L., Massey, J., Nicholas, W., and Wakesberg, J. (eds.) (1988), *Telephone Survey Methodology*, John Wiley & Sons, pp. 73-98..
- 前田忠彦,土屋隆裕 (2001), 「日本人の国民性

- 2000 年度吟味調査報告 -電話・郵送・面接調査の比較-, 統計数理研究所
- 松田 英二 (2002) 朝日新聞社の RDD 調査について、行動計量学、56 号、pp. 81-89.
- 佐藤武嗣(2002), RDD 電話調査, 林知己夫(編)「社会調査ハンドブック」,朝倉書店, pp. 191-200.
- 島田喜郎 (2003), 米国における RDD サンプリングの最近の動向, 「日本行動計量学会第 35 回大会発表論文抄録集」, 日本行動計量学会, pp. 126-129
- 島田喜郎 (2004), RDD サンプリング手法の比較研究, 「よろん」, 93 号, 2004 年 3 月, 日本世論調査協会, pp. 45-55
- 鈴木 達三 (2003) 電話調査についての一考察、行動計量学、58 号、pp. 73-91.
- 土屋 隆裕・前田 忠彦 (2003) 二種類の電話調査法の比較実験調査、行動計量学、58 号、pp. 93-107.
- Waksberg, J. (1978), Sampling methods for random digit dialing, *Journal of the American Statistical Association*, Vol. 73, No. 367, pp. 40-46.