

Comparación de unidades de muestreo en encuestas de telefonía móvil en Costa Rica. Muestreo de conglomerados *versus* elementos

Comparing Sampling Units in Cellphone Surveys in Costa Rica. Cluster Versus Element Sampling

Fernanda Alvarado-Leiton
Johnny Madrigal Pana

Recibido el 4 de marzo de 2025
Aceptado el 30 de junio de 2025

Resumen

Dos procedimientos de selección de unidades de muestreo fueron comparados para obtener muestras de números telefónicos celulares. El primero selecciona conglomerados (siguiendo a Waksberg y Mitofsky) y el segundo, números individuales. Se seleccionó al azar una muestra de 13,000 números celulares en conglomerados y otra de 12,000 celulares individuales. 25 entrevistadores fueron aleatorizados para realizar las llamadas con los dos procedimientos. Los resultados coinciden en las características de las muestras totales: activos y modalidad prepago/pospago. En el trabajo de campo los porcentajes de inelegibles fueron similares, pero en los conglomerados fueron mayores los rechazos que en los números individuales (35% *versus* 15%, respectivamente) y no contactados (31% *versus* 24%). En la muestra final las distribuciones por edad y educación fueron iguales, pero se reclutaron más mujeres usando conglomerados (59.3%) que teléfonos independientes (48.9%). La recomendación es que se opte por la selección de teléfonos individuales, lo que incrementará la eficiencia del trabajo de campo.

Palabras clave

Muestreo aleatorio, encuestas telefónicas, Mitofsky y Waksberg, experimento, metodología de encuestas

* Doctora en Metodología de Encuestas y Ciencia de Datos por la University of Michigan (UM, EE.UU.). Docente e investigadora de la Escuela de Estadística de la Universidad de Costa Rica (UCR). ORCID: 0009-0002-8327-9320. Correo electrónico: maria.alvaradoleiton@ucr.ac.cr

** Maestro en Estadística con énfasis en Población por la Universidad de Costa Rica (UCR). Es investigador catedrático jubilado de la UCR. ORCID: 0009-0000-3966-1170.

Abstract

In this study, two sampling approaches for cellphone numbers were compared. The first selects number clusters (following Waksberg and Mitofsky) and the second selects individual numbers. A random sample of 13,000 clustered and 12,000 individual cellphone numbers were selected. 25 interviewers were randomly assigned to one of the two sampling approaches. The results show that the two approaches yield a similar number of active numbers and a prepaid/postpaid distribution of numbers. During fieldwork the two approaches resulted in similar percentages of ineligible respondents. Nonetheless, cluster sampling elicited higher rates than individual sampling of rejection to the survey invitation (35% vs 15%) and of non-contacts (31% vs 24%). The recruited respondents were similar by age and education for the two approaches, although a disproportionate number of females was recruited for cluster sampling (59.3%) when compared to individual numbers (48.9%). We recommend sampling individual numbers for cellphone surveys as it can potentially increase the efficiency of fieldwork.

Keywords

Random sampling, telephone surveys, Mitofsky and Waksberg, Experiment, Survey Methodology

Las encuestas telefónicas, como en muchos otros países, son uno de los métodos de recolección de información más utilizados para realizar encuestas de opinión pública en Costa Rica. Su uso en el país es relativamente reciente, en comparación con los países desarrollados, con coberturas de telefonía fija altas, como por ejemplo Estados Unidos, Finlandia o Reino Unido (Groves, 1990), donde las encuestas telefónicas se convirtieron en una práctica estándar desde la década de los 80 (Tucker y Lepkowski, 2008). En esa época, la tenencia de telefonía fija era baja en el país y alcanzó su máximo en 2005, cuando 67.4% de los hogares contaba con este servicio (INEC, 2005), lo que impedía implementar encuestas telefónicas con una cobertura nacional efectiva.

La Encuesta de Confianza del Consumidor (ENCON), iniciada en Costa Rica en el año 2002, es ejemplo de uno de los primeros esfuerzos que brindó cifras a nivel nacional basándose en una muestra de teléfonos fijos (Madrigal-Pana, 2002). Aunque la cobertura de la telefonía fija excluía una parte importante de los hogares del país (solamente 61.4% de los hogares tenían servicio de telefonía fija en ese momento; (Ministerio de Planificación Nacional y Política Económica [MIDEPLAN], 2009), en el primer informe de esta encuesta se indica que "... la encuesta debe contemplar consumidores que posean la capacidad de realizar gastos discrecionales para comprar o reemplazar sus bienes de consumo" (Madrigal-Pana, 2002, p. 4), por lo que los hogares con teléfono fijo eran la población objetivo ideal para la encuesta, al ser los estrados de ingresos medios y altos los que podían costear la conexión telefónica. Realizada en sus inicios con una frecuencia semestral, y posteriormente de forma trimestral, la ENCON utilizó la telefonía fija como método de recolección por cerca de 16 años más desde su inicio.

Desde 2007 la cobertura de la telefonía fija cayó paulatinamente, hasta alcanzar 18.9% de los hogares en 2023 (INEC, 2023). Simultáneamente, la difusión de la telefonía móvil se incrementó considerablemente en las primeras dos décadas del siglo XXI, alcanzando una cobertura casi universal dentro del territorio nacional. Cifras de tres encuestas realizadas por el Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT) en 2015, 2016 y 2017, revelan que el porcentaje de personas adultas con un teléfono celular propio era de 90.6%, 95% y 94%, respectivamente (MICITT, 2015, 2016, 2017).

Como consecuencia de lo anterior, la migración a la telefonía celular era inminente para las encuestas que se realizaban por teléfono en el país, por lo que entre 2018 y 2019 se estudió, entre otros, el impacto que tendría el cambio de la telefonía fija por la móvil como medio para entrevistar a la muestra de personas en la ENCON (Madrigal-Pana et al., 2020), nuevamente siendo la encuesta telefónica de referencia en el país. El objetivo del estudio de Madrigal-Pana et al. (2020), fue conocer qué efecto tendrían los cambios metodológicos en la tendencia de los principales indicadores de la encuesta. Como conclusión se obtuvo que, usando ambos tipos de telefonía, la magnitud de las estimaciones del Índice de Confianza

del Consumidor (ICC), y de otros indicadores de interés, no mostraron diferencias significativas en tres de las cuatro encuestas realizadas y que todos esos indicadores reprodujeron la misma tendencia. Por este motivo, desde mayo de 2019, la telefonía móvil se convirtió en el medio de reclutamiento de participantes en la muestra de la ENCON. Este ejercicio metodológico que respaldó la migración hacia la telefonía móvil en la ENCON fue vital, pues sirvió como base para que otros estudios de opinión pública acogieran el muestreo de teléfonos móviles como su principal método de reclutamiento de participantes (e.g., INEP, 2020; LAPOP, 2021a; CIEP, 2022).

Del cambio metodológico de la ENCON destaca que la transición de telefonía fija a la móvil no contempló cambios en el procedimiento de selección de las unidades de muestreo para obtener los números telefónicos de la muestra, el que se mantiene invariante hasta la fecha (Alvarado-Leiton, 2024). Por este motivo, el objetivo principal del presente estudio es comparar dos unidades de muestreo para seleccionar muestras de teléfonos móviles. Como consecuencia, este trabajo constituye una etapa adicional de investigación metodológica, que plantea una aproximación más simple a la generación aleatoria de los números telefónicos y que, además, tiene el potencial de incrementar la eficiencia del recurso humano durante la etapa de la recolección de datos.

En el siguiente apartado se detalla la metodología del estudio, seguido por la descripción de los resultados y finalizando con la discusión de las implicaciones de los resultados en las encuestas telefónicas realizadas en Costa Rica.

Metodología

El marco muestral

Desde una perspectiva estadística, para generar una muestra probabilística se requiere de un marco muestral. Este instrumento, por lo general, se trata de un listado físico o digital que contiene todas las unidades a muestrear de una población específica, para seleccionar una o varias muestras. Este concepto, de acuerdo con Kish (1979), va más allá de un listado de unidades de muestreo, pues también abarca los procedimientos que pueden tomar en cuenta a todos los elementos de una población sin el esfuerzo físico de listarlos efectivamente (Kish, 1979, p. 78).

La selección de la muestra de la ENCON es un ejemplo de este concepto ampliado del uso de un procedimiento como marco muestral. Esto es así porque en el país no está disponible, para el público general, un listado que contenga todos los números de teléfono celular activos. De ser así, la selección de los teléfonos sería mucho más simple y expedita. Debido a esto, se han utilizado procedimientos que toman en cuenta a todos los números telefónicos existentes, sin tenerlos listados efectivamente, para seleccionar las muestras que se requieran. En la ENCON, y otras encuestas realizadas bajo diseños muestrales similares, se

ha implementado el procedimiento desarrollado por Joseph Waksberg y Warren Mitofsky (Waksberg, 1978), que utiliza información básica de los códigos telefónicos activos para construir números telefónicos con algún procedimiento aleatorio y así obtener las muestras requeridas.

Actualmente, para aplicar la estrategia de Waksberg y Mitofsky se posee un algoritmo computarizado que facilita la selección de teléfonos. Bajo la perspectiva de Kish (1979), este programa puede concebirse como el marco muestral y a partir de este se seleccionarán las unidades de muestreo que van a contener los números telefónicos de la muestra de interés. La información básica para alimentar el marco muestral proviene de la Superintendencia de Telecomunicaciones (SUTEL), entidad que ejerce la rectoría sobre los servicios de telecomunicaciones en el país y se explica a continuación.

Procedimientos para generar los números telefónicos

Para comprender los procedimientos de generación de la muestra de números telefónicos, es necesario familiarizarse con la estructura de los números telefónicos móviles en Costa Rica y su funcionamiento administrativo. En el país, los números de telefonía celular se conforman de ocho dígitos (e.g. 8888-9999). En este trabajo se definen como prefijos los primeros cuatro dígitos del número telefónico (e.g., 8888) y como sufijos los últimos cuatro dígitos (e.g. 9999). Los prefijos son administrados por la SUTEL, que designa los prefijos que pueden utilizar los diferentes operadores de telecomunicaciones para que estos a su vez asignen los números telefónicos a los clientes finales.

Como se mencionó anteriormente, el muestreo de telefonía móvil que utiliza la ENCON sigue el procedimiento desarrollado originalmente por Waksberg y Mitofsky (Waksberg, 1978). Este consiste en seleccionar aleatoriamente los prefijos telefónicos activos y a cada prefijo en la muestra se le agregan sufijos al azar para que cada prefijo en la muestra genere una cantidad específica de teléfonos celulares. Más concretamente, en el caso de la ENCON, cada prefijo seleccionado se repite 10 veces y a cada uno le agregan cuatro sufijos, para completar un total de 10 números telefónicos de celulares. Al grupo de 10 números de teléfonos, así construido, se le conoce como un banco telefónico, lo que desde un punto de vista del muestreo estadístico constituye un conglomerado de teléfonos. En otras palabras, el procedimiento selecciona conglomerados de números de telefonía celular.

Posteriormente, una vez generados todos los bancos telefónicos, la selección de las personas en la muestra se realiza en dos etapas. En la primera, para cada banco telefónico, se llama al primer número telefónico hasta un máximo de cuatro intentos, en diferentes días y horas para intentar verificar que el número está asignado a un cliente final. Si el teléfono marcado provee un indicador de que está en funcionamiento (contesta una persona o hay una contestadora personalizada) entonces se determina que el banco telefónico se encuentra activo. En resumen, identificar que un banco está activo consiste en saber, de alguna manera, que el

número celular está en uso, por lo que una persona, o una contestadora, debe responder, independientemente del resultado de la entrevista (realizada, pendiente, rechazo u otras situaciones de encuesta). Si se determina que el banco está activo se procede con la segunda etapa, que es llamar a todos los teléfonos celulares de ese banco telefónico para concretar las entrevistas correspondientes. Por otra parte, si se concluye que el primer número telefónico marcado está inactivo, se desecha todo el banco telefónico de la muestra. Por lo general, la inactividad del banco se sabe porque la operadora telefónica emite un mensaje mencionando que el teléfono no existe o porque no se obtiene respuesta.

El uso de este procedimiento para seleccionar los números telefónicos de una muestra requiere que se cumpla un supuesto básico, descrito por sus proponentes. Los primeros pioneros del muestreo telefónico identificaron que los operadores telefónicos en Estados Unidos asignaban los números telefónicos de forma sistemática en prefijos telefónicos que ya estuviesen en uso (Brick y Tucker, 2007). El aporte de Mitofsky fue identificar que las combinaciones de códigos de áreas y de prefijos podían ser utilizadas como unidades primarias de muestreo (Brick y Tucker, 2007). De esta forma, si se identificaba un número telefónico válido era probable que existieran más números válidos con ese mismo prefijo. Diversos trabajos mostraron las ganancias en eficiencia del método de Waksberg-Mitofsky, en comparación con el *RDD*, en países donde se implementó, aumentando el número de teléfonos válidos contactados de 20% a 60% (Brick y Tucker, 2007; Kalsbeek y Agans, 2008). En síntesis, el método fue diseñado para facilitar aspectos operativos al realizar las llamadas a una muestra aleatoria de teléfonos, minimizando el número de llamadas a números inexistentes o improductivos, por lo que aumentó la eficacia en la recolección de la información y se disminuyeron sus costos.

Desde su inicio en el 2002, la ENCON utilizó el procedimiento de Waksberg-Mitofsky para la selección de los números de telefonía fija y la transición a telefonía móvil no varió esta característica del diseño de la muestra. Aunque puede resultar extraño, y a pesar de que en otros países los números móviles se seleccionan sin el uso de los bancos telefónicos (Kennedy et al., 2018), el uso del método de Waksberg y Mitofsky en Costa Rica no es necesariamente injustificado. Los servicios de telecomunicaciones en Costa Rica fueron provistos por una única empresa estatal hasta el año 2011, cuando se abrió por primera vez el mercado de las telecomunicaciones al sector privado (Trejos-Lalli, 2012; OCDE, 2020). Además, inicialmente el servicio de telefonía móvil solamente existía bajo la figura de telefonía pospago (Canales, 2007). En este modelo de servicio, para cada persona que adquiría una línea, registrada además con sus datos personales se asignaba un número telefónico con una metodología similar a la de las líneas fijas (Rojas-Suárez, 2008). La apertura del mercado trajo consigo, también, que no se dispusiera de un listado con los números de celulares activos para mantener la privacidad de ciertos datos que pueden utilizarse para desarrollar estrategias de mercado.

Este procedimiento experimentó cambios en la ENCON y en los últimos años la determinación del banco telefónico como activo o inactivo incluyó también al

segundo número celular de cada banco. En otras palabras, si con el primer número telefónico del banco no podía determinarse si era activo o no, entonces se recurrió al segundo número en lista de ese banco. En este caso, si el primero o segundo número marcado muestra evidencia de que es activo, entonces el banco queda en la muestra y, en caso contrario, se desecha. Esto se hizo con el afán de reducir el tiempo para catalogar un banco telefónico como activo o no.

La apertura del mercado de telecomunicaciones al sector privado, así como la implementación de la modalidad de telefonía prepago, condujeron a la alta cobertura de telefonía celular en el país (Rodríguez-Céspedes, 2015), que también es característica de la región latinoamericana (LAPOP, 2021b). No obstante, ante esta explosión de oferta y demanda de la telefonía móvil en el país no es claro si se mantienen los procedimientos de asignación de los números telefónicos a los usuarios finales que justificaba la utilización del procedimiento de Waksberg-Mitofsky.

Objetivo e hipótesis

El objetivo de este trabajo es determinar si el uso de los bancos telefónicos, o conglomerados de números telefónicos celulares, continúa siendo una estrategia operativa adecuada frente a los cambios que ha experimentado el servicio de telefonía celular en los últimos años.

La hipótesis principal del estudio es que la selección de números telefónicos móviles individuales produce los mismos resultados que una selección usando el procedimiento de bancos de números de teléfonos celulares, utilizado en la ENCON.

Es importante realizar dos aclaraciones. La primera es que los dos procedimientos de selección que serán analizados utilizan los prefijos activos que provee la SUTEL y que la construcción de los números de telefonía celular se realiza con técnicas aleatorias. La segunda es que, desde el punto de vista del muestreo, se realizará una comparación para analizar la conveniencia de utilizar entre dos procedimientos de selección de unidades de muestreo a partir de un marco muestral. Esto es así, pues se asume que la escogencia de estas unidades de muestreo (números telefónicos) se hará seleccionando grupos de números telefónicos (bancos telefónicos o conglomerados) o números telefónicos individuales (elementos). Se aclara que las unidades de muestreo son aquellas unidades que conforman el marco muestral y que contienen a toda la población de estudio. Estas unidades son las que se seleccionan del marco muestral para obtener una muestra y cada una de ellas puede contener un elemento único (un número de teléfono celular) o varios elementos (un grupo o conglomerado de números telefónicos celulares). Para este trabajo, el procedimiento de muestreo se desarrolló con el lenguaje de programación R, el que contiene todos los bancos telefónicos activos, con lo que la persona especialista en muestreo puede generar las unidades de la muestra en forma de conglomerados o de elementos para seleccionarlas.

Para evaluar la hipótesis del estudio se implementó un experimento en una encuesta telefónica con cobertura nacional. Se contrastaron dos formas de seleccionar números de telefonía móvil en términos de las características de la muestra de teléfonos seleccionada, los resultados del trabajo de campo y las características de la muestra final de personas. Los detalles del experimento se explican a continuación.

Diseño experimental

El experimento se ejecutó mientras se realizó el trabajo de campo de la Encuesta Actualidades 2023 (<https://www.estadistica.ucr.ac.cr/index.php/vida-academica/encuestas/encuesta-actualidades>). Esta encuesta es planeada y ejecutada anualmente por estudiantes de tercer año del programa del Bachillerato en Estadística de la Universidad de Costa Rica. La encuesta se realiza dentro del marco del curso Diseño de Encuestas por Muestreo y tiene como objetivo formar al estudiantado en las diferentes etapas que requiere el desarrollo de una encuesta. La encuesta tiene una cobertura nacional y se realiza por medio de telefonía móvil.

El trabajo de campo se realizó en dos etapas. La primera consistió en una prueba piloto realizada entre el 26 y 29 de septiembre de 2023 y la segunda comprendió el trabajo de campo principal de la encuesta, ejecutado entre 1 y 31 de octubre de 2023. Para desarrollar el diseño experimental se trabajó en ambas etapas.

El experimento consistió en seleccionar dos muestras independientes, para un total de 25,000 números telefónicos (13,000 en conglomerados y 12,000 en elementos). Para el primer grupo experimental se seleccionaron los números de telefonía móvil aplicando el procedimiento de Waksberg-Mitofsky, esto es un muestreo de conglomerados de teléfonos, tal y como se realiza actualmente en la ENCON. Específicamente, se seleccionaron 1,300 prefijos de telefonía celular activos (130 en la piloto y 1,170 en la encuesta) y cada uno se repitió 10 veces, con un agregado de cuatro dígitos aleatorios para completar los bancos telefónicos con los números de teléfonos celulares para la muestra. Para el segundo grupo experimental, la muestra se generó seleccionando 12,000 prefijos de telefonía celular activos (1,200 en la piloto y 10,800 en la encuesta) y cada prefijo se completó con cuatro dígitos aleatorios para formar los teléfonos celulares requeridos para la muestra.

La metodología utilizada para la selección de los prefijos y sufijos consistió en lo siguiente.

La selección de los prefijos se realizó con un diseño muestral de Probabilidades Proporcionales al Tamaño (PPT), donde la medida de tamaño fue la cantidad de prefijos asignada a cada operadora telefónica por la SUTEL dentro del Registro Nacional de Numeración (SUTEL, 2024). Una ilustración de la asignación que realiza la SUTEL se muestra en la Figura 1. Nótese que esta asignación a las operadoras no implica que todos los números telefónicos estén activos. Por otra parte, los sufijos se generaron de la misma manera para los dos grupos experimentales:

cuatro dígitos entre cero y nueve se seleccionaron con igual probabilidad y de forma independiente.

Figura 1. Extracto del Registro de Numeración (SUTEL, 2024)

Número (IND + NA)	Número (IND + NA) 2	Cantidad	Operador
60002000	61001999	1000000	LIBERTY
70000000	70009999	10000	CLARO
83000000	83999999	1000000	ICE

Fuente: Registro Nacional de Numeración. Superintendencia de Telecomunicaciones, 2024.

Trabajo de campo

Las 25 personas inscritas en el curso cumplieron el rol de entrevistadores y se asignaron de forma aleatoria a uno de los dos grupos experimentales en cada etapa del estudio. A cada estudiante se le entregó un archivo con los números telefónicos a los que debía de llamar y el registro que debía llenar para cada número. Un ejemplo de este registro se muestra en la Figura 2. A las personas entrevistadoras de cada grupo experimental se les solicitó realizar los contactos correspondientes, con el procedimiento de selección respectivo, llamando hasta un máximo de cuatro veces, en diferentes horas y días, hasta completar 25 entrevistas. Además, se les pidió buscar cada número telefónico en el registro de teléfonos móviles de la SUTEL¹ previo a iniciar las llamadas y registrar si el número estaba activo, el operador asociado al número y el tipo de telefonía (prepago o pospago; Figura 2).

Indicadores de interés

Para evaluar la hipótesis del estudio se utilizan cinco indicadores de interés que describen las características de la muestra de teléfonos seleccionada, el trabajo de campo y las características de la muestra de personas finalmente reclutada. Estos indicadores se detallan a continuación.

- *Porcentaje de números activos.* Como número activo se definió aquel que se encontró como activo en un operador en el registro de consulta de la SUTEL. El porcentaje de números activos en cada grupo experimental se calculó como el total de números activos entre el total de números seleccionados en la muestra.

¹ Este registro se actualiza cada seis meses y está disponible en: <https://registroprepago.sutel.go.cr/busquedaConsultaOperadorERPN.action>

Figura 2. Ejemplo de registro de llamadas utilizado

Bco	No.	Teléfono	1a llamada			Cod. final	Tipo telefonía	Operador Sutel
			Día	Hora	Minuto			
1	1	7016-5XXX	27	17	40	3-no contacto	3-no contacto	Claro
2	1	7037-4XXX	27	17	42	5-inelegible	5-inelegible	Claro
3	1	7066-0XXX	27	17	43	3-no contacto	3-no contacto	Liberty
4	1	7008-8XXX	27	17	44	3-no contacto	3-no contacto	Claro
5	1	7078-6XXX	27	17	44	5-inelegible	5-inelegible	Claro
6	1	7051-0XXX	27	17	45	5-inelegible	5-inelegible	Claro
7	1	7067-8XXX	27	17	46	5-inelegible	5-inelegible	Claro
8	1	7037-0XXX	27	17	47	3-no contacto	3-no contacto	Claro
9	1	7069-7XXX	27	17	48	5-inelegible	5-inelegible	Claro
10	1	7106-0XXX	27	17	49	3-no contacto	6-No realizada por otras razones	

Fuente: Registro de operaciones, Encuesta Nacional del Consumidor. Escuela de Estadística, Universidad de Costa Rica.

- *Porcentaje de números prepago y pospago.* Se asignó esta clasificación a todos los números activos de acuerdo con la información disponible en el registro de teléfonos móviles de la SUTEL. El porcentaje de números prepago y pospago para cada grupo se calculó como el total de números en cada categoría entre el total de números activos.
- *Porcentaje de rechazos.* Se definió como rechazo cuando una persona elegible rehúsa explícitamente la entrevista durante la llamada o cuando la persona accede a participar, reprograma la entrevista, pero no responde las llamadas posteriores. El porcentaje de rechazos en cada grupo se calculó como el número total de rechazos entre el total de números activos con personas elegibles.
- *Porcentaje de inelegibles.* Los números inelegibles se definieron como aquellos para los cuales se identificaba con certeza que el número no pertenecía a una persona elegible (persona mayor de 18 años residiendo en Costa Rica por al menos 1.5 años). El porcentaje de inelegibles en cada grupo se calculó como el número total de números inelegibles entre el total de números seleccionados.
- *Porcentaje de elegibilidad desconocida.* Un número telefónico se clasificó con elegibilidad desconocida cuando, después de cuatro llamadas, no fue posible determinar si estaban asociados a una persona elegible para realizar la encuesta. El porcentaje de elegibilidad desconocida en cada grupo se definió como el total de números clasificados como no contacto entre el total de números activos.
- *Características demográficas de las personas entrevistadas.* Se examinaron las frecuencias relativas de la muestra sin ponderar de las variables sexo al nacer (hombre, mujer), educación (primaria o menos, secundaria completa e incompleta, técnica o universitaria completa e incompleta) y edad (menos de 35 años, 35 a 50 años, más de 50 años).

Resultados

En el Cuadro 1 se detallan las características de las muestras totales de teléfonos seleccionadas. En total se seleccionaron 25,000 números telefónicos, de los cuales 13,000 se seleccionaron usando conglomerados (1,300 en la piloto más 11,700 en la encuesta) y 12,000 de manera individual (1,200 en la piloto más 10,800 en la encuesta).

Cuadro 1. Resultado total y por etapas del número de teléfonos celulares seleccionados y número de entrevistadores para los dos grupos experimentales (conglomerados y elementos de números celulares)

Unidad de muestreo	Números telefónicos	Número de entrevistadores
Total	25,000	25
Conglomerados	13,000	26
Elementos	12,000	24
Piloto Actualidades 2023	2,500	25
Conglomerados	1,300	13
Elementos	1,200	12
Actualidades 2023	22,500	25
Conglomerados	11,700	13
Elementos	10,800	12

El Cuadro 2 muestra dos de los indicadores de interés en el estudio. En primer lugar, se examinó el porcentaje de números activos para cada grupo experimental y se encontró que, del total de números telefónicos generados en cada grupo, el porcentaje entre ellos resultó similar, cercano a 58%. El cuadro también muestra la distribución porcentual de los números válidos de acuerdo con la modalidad de contrato (prepago y pospago). Este cálculo tiene como base de cada grupo el número total de celulares activos y el resultado muestra que no se encontraron diferencias entre los grupos.

Cuadro 2. Resultado total y por etapa de la situación de actividad y de pago para los dos grupos experimentales (conglomerados y elementos de número celulares)

Unidades de muestreo	% de números activos ^a	% números prepago ^b	% números pospago ^b
Total	57.7	42.9	57.1
Conglomerados	57.5	42.8	57.2
Elementos	58.0	43.2	56.8
Piloto Actualidades 2023	58.0	43.6	56.4
Conglomerados	56.9	44.1	55.9
Elementos	59.2	43.1	56.9
Actualidades 2023	57.7	42.9	57.1
Conglomerados	57.5	42.6	57.4
Elementos	57.9	43.2	56.8

^a Del total de números generados ^b Del total de números activos

El Cuadro 3 resume los resultados de las métricas de interés del trabajo de campo. En este caso se trabaja tomando como base el total de teléfonos activos en cada grupo. Estos resultados revelan que no existen diferencias significativas en los porcentajes de números inelegibles entre los grupos experimentales. Para los porcentajes de rechazos y no contactos se observó una diferencia en la etapa del trabajo de campo de la encuesta (Actualidades 2023). Cuando los teléfonos se seleccionaron con conglomerados 34.9% de las llamadas terminaron en un rechazo explícito, en comparación con 14.8% cuando los teléfonos se seleccionaron de forma independiente, es decir, una diferencia de 20.1 p.p. Asimismo, cuando los números se seleccionaron mediante conglomerados, el 31.1% de los números se clasificó con elegibilidad desconocida frente a un 23.7% cuando se seleccionaron de forma independiente.

Cuadro 3
Resultado total y por etapa del trabajo de campo para los dos grupos experimentales (conglomerados y elementos de números celulares)

Unidades de muestreo	% inelegibles	% de rechazos	% de elegibilidad desconocida
Total	20.8	28.4	15.6
Conglomerados	20.9	33.9	16.9
Elementos	20.7	23.6	13.5
Piloto actualidades 2023	25.4	30.4	12.4
Conglomerados	24.5	30.4	12.9
Elementos	26.3	30.3	11.8
Actualidades 2023	20.3	28.1	27.6
Conglomerados	20.5	34.9 ^a	31.1 ^a
Elementos	20.1	14.8 ^a	23.7 ^a

^a Diferencia estadísticamente significativa con $p < 0,05$ entre los dos grupos experimentales.

El cuadro 4 muestra la distribución porcentual de algunas variables sociodemográficas de la muestra final en cada grupo experimental. No se encontraron diferencias significativas en las distribuciones por edad y nivel educativo entre los grupos experimentales. Por sexo, se reclutaron más mujeres cuando se seleccionaron los números mediante conglomerados de teléfonos (59.3%) que cuando se seleccionaron de forma independiente (48.9%).

Cuadro 4
Distribución porcentual de características sociodemográficas
de la muestra final para los dos grupos experimentales
(conglomerados y elementos de números celulares)

Características sociodemográficas	Unidades de muestreo		
	Total (n=701)	Conglomerados (n= 339)	Elementos (n=362)
Total	100	100	100
Sexo			
Hombre	45,9	40,7a	50,8a
Mujer	53,9	59,3a	48,9a
Nivel educativo			
Primaria o menos	20,4	22,7	18,2
Secundaria	41,4	40,4	42,3
Técnica y universitaria	38,2	36,9	39,5
Edad			
Menos de 35	34,0	33,0	34,8
35 a 49	29,1	28,3	29,8
50+	36,8	38,3	35,4

^a Diferencia estadísticamente significativa con p < 0,05 entre los dos grupos experimentales.

Se exploró como posible razón para la diferencia de reclutamiento de hombres y mujeres el sexo de la persona entrevistadora. Esto debido a que, la diferencia en el reclutamiento de hombres y mujeres se observó principalmente en la prueba piloto (Anexo 1), en la cual, aunque la asignación de las personas entrevistadoras al tipo de muestreo fue aleatoria, las mujeres entrevistadoras resultaron desproporcionalmente asignadas al grupo que realizó trabajo de campo usando conglomerados. En la prueba piloto, de las 11 entrevistadoras, 7 fueron asignadas al tratamiento de muestreo por conglomerados y 4 al de elementos en la prueba piloto. En el trabajo de campo, en cambio, 6 entrevistadoras fueron asignadas al muestreo de elementos y 5 al de conglomerados.

En el Cuadro 5 se muestra que el balance en el reclutamiento entre hombres y mujeres se dio para el muestreo de elementos (el que se caracterizó por tener pocas mujeres entrevistadoras). Esto, aunado a la existente literatura sobre las tasas de respuesta por sexo de la persona entrevistada y entrevistadora (e.g., Nealon, 1983; Groves y Fultz, 1985; Korkeila et al., 2001), explica la diferencia observada en el porcentaje de hombres y mujeres reclutados en la muestra final.

Cuadro5

Distribución porcentual por sexo y grupo experimental (conglomerados y elementos de números celulares) de las personas reclutadas en la prueba piloto

Sexo de la persona entrevistada	Elementos a (n= 32)	Conglomerados b (n= 38)
Total	100.00	100.00
Hombre	31,2	44,7
Mujer	68,8	55,3

^a Equipo entrevistador: 4 mujeres y 8 hombres

^b Equipo entrevistador: 7 mujeres y 6 hombres

Discusión

En esta investigación se compararon dos procedimientos para seleccionar unidades de muestreo y obtener muestras de números telefónicos celulares. El primero corresponde al utilizado en la ENCON desde el año 2002 y para el cual se seleccionan los números en dos etapas. En la primera se seleccionan prefijos activos de cuatro dígitos y en la segunda se seleccionan 10 combinaciones de cuatro dígitos para cada prefijo seleccionado. Es decir, se seleccionan conglomerados de números de teléfono. Este procedimiento supone que los números se asignan a los usuarios finales en bloques definidos por los primeros cuatro dígitos y que la probabilidad de encontrar un número válido es más alta cuando ya se encontró un número válido con los mismos primeros cuatro dígitos. El segundo procedimiento abandona la noción de conglomerados por elementos y selecciona una combinación de cuatro dígitos para cada prefijo seleccionado. Esta segunda estrategia es una respuesta a la apertura del mercado de telefonía móvil costarricense que produjo la incursión en el mercado de diversos productos y empresas y, por ende, se alejó de la asignación de números a los usuarios finales por medio de bloques con algún significado práctico, desde la perspectiva del muestreo. En este estudio hipotetizamos que la selección de números telefónicos celulares individuales es equivalente a la que se realiza por conglomerados en tres áreas de interés: 1) características de los números telefónicos seleccionados, 2) métricas del trabajo de campo y 3) la muestra final de personas reclutada.

Los resultados muestran una clara coincidencia entre los dos procedimientos estudiados en lo que respecta a números válidos, lo que implica que el supuesto de que los números telefónicos se asignan a los usuarios en bloque no refleja lo que sucede en el mercado de telecomunicaciones del país. Es decir, no existe una superioridad operacional en la identificación de los números telefónicos válidos

como la que se plantea en el trabajo de Waksberg (1978). En consecuencia, se pierde la principal ventaja del muestreo de teléfonos por conglomerados y más bien el método se transforma en una restricción innecesaria en el trabajo de campo al requerir la validez de los primeros dos números del banco para llamar al resto.

Los demás resultados muestran que ambos métodos de selección producen un número similar de números prepago y pospago y porcentajes similares de números inelegibles. No obstante, en el trabajo de campo de la encuesta se encontraron diferencias en los porcentajes de no contactos y rechazos (ocurrieron con más frecuencia cuando los números se seleccionaron con conglomerados) y se reclutaron más mujeres en la muestra seleccionada con conglomerados. Las personas entrevistadoras se asignaron aleatoriamente a los grupos experimentales, por lo que errores en la clasificación de casos como no contactos y en el reclutamiento de los participantes deberían estar distribuidos de forma uniforme en los dos grupos. Queda para una futura exploración, que se escapa de los objetivos del presente trabajo, dilucidar las posibles razones de estas diferencias, aunque posiblemente se deban a diferencias en la ejecución del trabajo de campo, adherencia a protocolos por parte de las personas entrevistadoras y sus características particulares. Por ejemplo, es posible que aquellos entrevistadores asignados al muestreo de conglomerados clasificaran una llamada como un rechazo con mayor facilidad, ya que, su protocolo de llamada era más demandante que el del muestreo de elementos. Por otro lado, las diferencias por sexo en la muestra reclutada en los dos tipos de muestreo parecen estar explicadas por la asignación desproporcional (aunque aleatoria) de mujeres entrevistadoras a uno de los dos grupos experimentales. Lo anterior debido a que los resultados obtenidos en este estudio son consistentes con la evidencia existente sobre el efecto del género de la persona entrevistadora, que sugiere que las mujeres tienden a lograr tasas de respuesta más altas que las de los hombres entrevistadores, y estos últimos tienden a reclutar menos mujeres en las muestras. También, los hallazgos de este estudio son consistentes con la evidencia empírica que demuestra ser más sencillo el reclutar mujeres participantes en una encuesta que hombres.

Por otra parte, vale destacar que los resultados que podrían considerarse desfavorables (más rechazos, más no contactos y desbalance en el reclutamiento de hombres y mujeres), ocurren en la muestra seleccionada con conglomerados, por lo que no se evidencia alguna razón que impida adoptar la selección independiente de números telefónicos.

En conclusión, se recomienda abandonar la selección de números telefónicos móviles por medio de conglomerados, debido a que su supuesto principal no corresponde con el mercado costarricense de telecomunicaciones. Las implicaciones de esta transición radican en un incremento de la eficiencia del trabajo de campo, puesto que, al eliminar los conglomerados de números telefónicos, los entrevistadores evitarían el trabajo de verificar que exista un número válido en los primeros dos de cada 10 números para poder llamar a los ocho siguientes, permitiendo un avance más rápido en el intento de contacto a la muestra seleccionada.

Referencias

- Alvarado-Leiton, F. (2024). *La confianza de los consumidores. Encuesta No 82*. Universidad de Costa Rica. https://estadistica.ucr.ac.cr/images/EEs/Documentos/Encuestas/Consumidor/Informes/UCR_ESTADISTICA_ICC_MAYO_2024.pdf
- Brick, M. y Tucker, C. (2007). Mitofsky-Waksberg: Learning from the past. *Public Opinion Quarterly*, 71(5), 703-716. <https://doi.org/10.1093/poq/nfm049>
- Canales, D. (29 de noviembre, 2007). Líneas TDMA se convertirán en servicio prepago. *La República*. https://www.larepublica.net/noticia/lineas_tdma_se_convertiran_en_servicio_prepago
- CIEP. (12 de marzo, 2022). ¿Cómo se generan las bases de datos de números telefónicos para las encuestas?. Centro de Investigación y Estudios Políticos-Universidad de Costa Rica. <https://ciep.ucr.ac.cr/como-se-generan-las-bases-de-datos-de-numeros-telefonicos-para-las-encuestas/>
- Groves, M. (1990). Theories and Methods of Telephone Surveys. *Annual Review of Sociology*, 16, 221-240. <https://doi.org/10.1146/annurev.so.16.080190.001253>
- Groves, R. M. y Fultz, N. H. (1985). Gender effects among telephone interviewers in a survey of economic attitudes. *Sociological Methods & Research*, 14(1), 31-52. <https://doi.org/10.1177/0049124185014001002>
- IDESPO. (2020). *Encuesta: Cultura política e intención de voto en las elecciones municipales 2020*. Instituto de Estudios Sociales en Población-Universidad Nacional de Costa Rica. https://repositorio.una.ac.cr/bitstream/handle/11056/17241/informe_encuesta_cultura_politica_e_intencion_voto_elecciones_municipales_2020.pdf?sequence=1&isAllowed=y
- INEC. (2005). *Cantidad y porcentaje de servicios de telefonía residencial y telefonía celular*. Instituto Nacional de Estadística y Censos de Costa Rica. https://admin.inec.cr/sites/default/files/media/recientecehpm2005-02_2.xls
- INEC. (2023). *Cantidad y porcentaje de viviendas que poseen diferentes TIC por zona según tenencia de TIC*. Instituto Nacional de Estadística y Censos de Costa Rica. https://admin.inec.cr/sites/default/files/2024-05/recientecenaho2023-02_0.xlsx
- Kalsbeek, W. y Agans, R. (2008). Sampling and Weighting in Household Telephone Surveys. En Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, P. J. Lavrakas, M. W. Link, R. L. Sangster (Eds.), *Advances in Telephone Survey Methodology* (pp. 27-55). Wiley-Interscience. <https://doi.org/10.1002/9780470173404>
- Kennedy, C., McGeeney, K., Keeter, S., Patten, E., Perrin, A., Lee, A. y Best, J. (2018). Implications of moving public opinion surveys to a single-frame cell-phone random-digit-dial design. *Public Opinion Quarterly*, 82(2), 279-299. <https://doi.org/10.1093/poq/nfy016>
- Kish, L. (1979). *Muestreo de Encuestas*. Editorial Trillas.
- Korkeila, K., Suominen, S., Ahvenainen, J., Ojanlatva, A., Rautava, P., Helenius, H. y Koskenvuo, M. (2001). Non-response and related factors in a nation-wide health survey. *European Journal of Epidemiology*, 17, 991-999. <https://doi.org/10.1023/a:1020016922473>
- LAPOP. (2021a). *AmericasBarometer 2021: Costa Rica Technical Information*. Vanderbilt University. <https://www.vanderbilt.edu/lapop/cr/ABCRI2021-Technical-Report-v1.0-FINAL-eng-110921.pdf>

- LAPOP. (2021b). *AmericasBarometer 2021: Technical Information*. Vanderbilt University. <https://www.vanderbilt.edu/lapop/ab2021/AB2021-Technical-Report-v1.0-FINAL-eng-030722.pdf>
- Madrigal-Pana, J. (2002). *El sentimiento del consumidor sobre la economía nacional*. Centro de Investigaciones en Estadística-Universidad de Costa Rica. https://estadistica.ucr.ac.cr/images/EEs/Documentos/Encuestas/Consumidor/Informes/UCR_01ESTADISTICA_ICC_SET_2002.pdf
- Madrigal-Pana, J., Flores-Alvarado, S., Benavides-Prado, M. F. y Fiallos-González, M. (2020). Cambio de Metodología en la Encuesta de Confianza de los Consumidores en Costa Rica. *Revista de Ciencias Económicas*, 38(1), 71-97. <https://doi.org/10.15517/rce.v38i1.38671>
- MICITT. (2015). *Acceso y uso de los servicios de telecomunicaciones en Costa Rica 2015*. Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones de Costa Rica. https://www.micitt.go.cr/sites/default/files/publicaciones/estadisticas/acceso_y_uso_de_los_servicios_de_telecomunicaciones_en_costa_rica_2015-1.pdf
- MICITT. (2016). *Encuesta de acceso y uso de los servicios de telecomunicaciones en Costa Rica 2016*. Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones. <https://www.micitt.go.cr/sites/default/files/publicaciones/estadisticas/encuesta-de-acceso-y-uso-de-los-servicios-de-telecomunicaciones-en-costa-rica-2016.pdf>
- MICITT. (2017). *Acceso y uso de los servicios de telecomunicaciones en Costa Rica 2017*. Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones. <https://www.micitt.go.cr/sites/default/files/publicaciones/estadisticas/acceso-y-uso-de-los-servicios-de-telecomunicaciones-en-costa-rica-2017.pdf>
- MIDEPLAN. (2009). Costa Rica: estadísticas regionales 2001-2008. Ministerio de Planificación Nacional y Política Económica. <https://share.google/ZT7dxfNIIx9kLp4u1>
- Nealon, J. (1983). The effects of male vs. female telephone interviewers. En *Proceedings of the Survey Research Methods Section of the American Statistical Association* (pp. 139-41). American Statistical Association.
- OCDE. (2020). *Estudios Económicos de la OCDE: Costa Rica 2020*. Organización para la Cooperación y el Desarrollo Económicos-OECD Publishing. <https://doi.org/10.1787/84ccb575-es>
- Rodríguez-Céspedes, E. (2015). *Evaluación del impacto de la apertura del mercado de telecomunicaciones en Costa Rica*. Universidad de Costa Rica. <http://www.iice.ucr.ac.cr/serie2.html>
- Rojas-Suárez, E. (12 de febrero, 2008). 80.323 GSM están disponibles. *Al Día*. http://www.aldia.cr/ad_ee/2008/febrero/12/nacionales1420970.html
- SUTEL. (2024). *Registro de Nacional de Numeración*. Superintendencia de Telecomunicaciones. <https://sutel.go.cr/sites/default/files/Registro%20Numeración%20Marzo%202024.pdf>
- Trejos-Lalli, E. (2012). Apertura de las telecomunicaciones: 5 años de avances y desafíos. En M. Guzmán-Hidalgo (Ed.), *Hacia la sociedad de la información y el conocimiento*. Universidad de Costa Rica.
- Tucker, C. y Lepkowski, J. M. (2008). Telephone Survey Methods: Adapting to Change. En J. M. Lepkowski, C. Tucker, J. M. Michael Brick, E. D. de Leeuw, L. Japec, P. J. Lavrakas, M. W. Link y R. L. Sangster (Eds.), *Advances in Telephone Survey Methodology*. Wiley-Interscience. <https://doi.org/10.1002/9780470173404.ch1>
- Waksberg, J. (1978). Methods for Random Digit Dialing. *Journal of the American Statistical Association*, 73(161), 40-46. <https://doi.org/10.1080/01621459.1978.10479995>

Distribución porcentual de las personas reclutadas en la encuesta por sexo y etapa del trabajo de campo

Sexo de la persona entrevistada	Piloto Actualidades 2023 (n=70)	Actualidades 2023 (n= 630)
Total	100,0	100,0
Hombre	38,6	46,8
Mujer	61,4	53,2

^a Equipo entrevistador: 4 mujeres y 8 hombres

^b Equipo entrevistador: 7 mujeres y 6 hombres

