

Panel vs Snowball: Un análisis de contraste

Víctor Fernández Araiz, Víctor Sanz Suarez-Lledó y Javier Lorenzo Rodríguez

Junio 2017

1.Introducción

Este trabajo está realizado en el marco del proyecto de investigación Horizon2020 MOVE (Grant Agreement N^o 649263), en el que se estudia la movilidad de los jóvenes en la Unión Europea y la reducción de los impactos negativos a través de la identificación de buenas prácticas que fomentan el desarrollo sostenible y el bienestar. Este proyecto está formado por nueve socios de seis países: Alemania, Hungría, Luxemburgo, Noruega, Rumanía y España.

De estos nueve socios, para la consecución de los objetivos que marca el proyecto MOVE, el Ilustre Colegio Nacional de Doctores y Licenciados en Ciencias Políticas y Sociología ha sido el encargado de llevar a cabo la obtención de dos bases de datos mediante la aplicación de encuesta *online*. La primera de ellas se obtuvo a través de una encuesta panel *online* (n=5.499) subcontratada en base a la mejor relación calidad-precio a GfK (Art. 13 AMGA) después de dos convocatorias abiertas (publicadas el 4 de enero y el 11 de marzo de 2016). La segunda se obtuvo, también, mediante una encuesta *online* pero a través de un muestreo “snowball.” “bola de nieve” (n=3.207). Ambas, dirigidas sólo a jóvenes de entre 18 y 29 años que hayan tenido o no experiencias de movilidad (en el caso de la encuesta panel) y que solamente hayan tenido experiencias de movilidad (en el caso de la encuesta “snowball”).

En este proyecto, el concepto de “movilidad” se ha definido como “haber estado alguna vez en el extranjero por una duración de más de dos semanas y por alguna razón distinta al turismo o a la visita de familiares”, con el fin de incluir todos los tipos de movilidad estudiadas en este proyecto: movilidad por motivos laborales, académicos, intercambio de estudiantes, servicio voluntario, etc. Las encuestas realizadas fueron dirigidas a los jóvenes cuya nacionalidad se correspondiese con alguno de los países del consorcio del proyecto (Alemania, Hungría, Luxemburgo, Noruega, Rumanía y España), o que hubiesen cursado sus estudios secundarios (o el último año de sus estudios en caso de no haber completado los estudios secundarios) en alguno de los seis países que forman el consorcio del proyecto MOVE.

Los principales objetivos de las dos encuestas realizadas son:

- Conocer la función y el valor de la información y de los servicios de apoyo a los jóvenes en su proceso de toma de decisiones para trasladarse al extranjero.
- Explorar el papel de las redes transnacionales de apoyo y como potencial “factor de atracción” para la movilidad.
- Examinar la agencia de los jóvenes con o sin experiencia de movilidad.
- El estudio de la formación del capital social y de las dimensiones de la desigualdad social de los jóvenes móviles y sus efectos en sus perspectivas de futuro, así como la reproducción de las desigualdades sociales.
- El estudio comparado sobre la formación de la identidad entre los jóvenes con y sin experiencia de movilidad.
- Examinar los planes de carrera de los jóvenes y sus apegos personales relacionados con sus compromisos en su país de origen (por ejemplo, el envío de dinero a sus hogares, apoyo a sus familias, etc.).
- Obtener información sobre la (re)producción de la desigualdad social en relación con la movilidad y la no movilidad.

A partir de los resultados de ambas encuestas proponemos en este artículo **un estudio comparado de ambas técnicas**. Para ello, nos centraremos en **el análisis de la composición de la muestra de ambos casos, así como en su utilidad para alcanzar los objetivos propuestos en el proyecto, entre ellos: los patrones de la movilidad juvenil, la participación política, electoral y asociativa. Además, profundizaremos en los pros y los contras de la aplicación de esta doble metodología y los desafíos que implica el análisis conjunto de ambas encuestas.**

2. Metodología

El estudio las migraciones ha supuesto, tradicionalmente, un desafío en términos metodológicos. Varios académicos han señalado una brecha entre los métodos cualitativos y los métodos cuantitativos (Bakewell 2010) Con el propósito de llenar este vacío, se ha desarrollado en este proyecto MOVE un estudio metodológico complejo, que ha abarcado tanto métodos cualitativos como cuantitativos, diseñando una triangulación metodológica en tres etapas: un primer paso en el que a partir de un conjunto de estudios de caso se obtuvieron las claves para desentrañar los patrones de comportamiento que desencadenan los individuos como agentes para tomar la decisión de emigrar; en paralelo, se llevó a cabo un estudio comparativo basado en el análisis a nivel macro; en la tercera etapa, de la que trata este documento, se realizó una encuesta online

para preguntar a las poblaciones jóvenes móviles y no móviles sobre sus opiniones, experiencias y motivaciones, así como sobre las barreras, temores o razones que tienen los jóvenes no móviles en sus respectivos países.

Para alcanzar esta tercera etapa, se ha diseñado un modelo original con el propósito de maximizar el resultado de los análisis cuantitativos. Siendo conscientes de la ausencia de un censo real de la juventud móvil, sobre todo al considerar la amplia variedad tipológica de movilidades y las diferencias en el uso de Internet y redes sociales, ha sido aplicado un proceso de recolección de datos de método mixto a través de una estrategia dual: por un lado, una encuesta panel *online* (n=5.499) que permite el estudio de los jóvenes móviles y no móviles en sus respectivos países –tratándose de una muestra proporcional basada en las distribuciones poblacionales por sexo y edad–, y complementando este enfoque en conexión directa con los jóvenes de los países que participan en el proyecto que están o han estado inmersos en un proceso de movilidad, se ha llevado a cabo otra encuesta *online* (aplicando el mismo cuestionario), en este caso, mediante la metodología de muestreo "snowball".

2.1. La encuesta panel online

El uso de encuestas *online* ha sido ampliamente discutido en el campo de las ciencias sociales en los últimos veinte años (Couper, Traugott and Lamias 2001, Gosling et al. 2004). Aunque las particularidades en la aplicación de este tipo de encuestas son nuevas, la controversia en torno a la metodología de investigación no lo es. De hecho, gran parte del actual debate en torno a las encuestas *online* es una reminiscencia de una era pasada cuando las encuestas telefónicas y por correo postal eran vistas con sospecha (Konstan et al. 2005, Sue and Ritter 2012).

Con independencia de las claras ventajas que presuponen las encuestas por Internet (bajo costo, respuesta rápida y menores dificultades relativas al trabajo de campo como principales beneficios), también enfrenta algunos inconvenientes a tener en cuenta: falta de representatividad muestral debida al sesgo de auto-selección, además de los relacionados con la anonimidad y confidencialidad del proceso (Dillman 1978, Couper 2000, Gosling et al. 2004, Landau and Jacobsen 2005).

Tradicionalmente, ha resultado poco frecuente encontrar análisis cuantitativos en estudios de migración o movilidad, y aún más raro con tamaños muestrales grandes (Fawcett and Arnold 1987, McKenzie and Mistiaen 2009). Una de las principales razones para evitar este tipo de enfoque metodológico es la dificultad de reconocer el universo, es decir, la población a analizar. Por tanto, los principales argumentos en contra del uso de encuestas a la hora de analizar poblaciones migrantes o móviles giran en torno a: la representatividad, el sesgo, la inferencia

causal, la escasez de datos estadísticamente analizables y la falta de transparencia metodológica (Landau and Jacobsen 2005). Más allá de esto, las cuestiones prácticas y logísticas tales como las limitaciones presupuestarias y de tiempo también han desalentado su desarrollo ulterior.

No obstante, en los últimos años, varios proyectos han desarrollado estudios con aplicación de metodologías mixtas que incluyen análisis cuantitativos y cualitativos (Hussein, Manthorpe and Stevens 2011, Lados 2013, Bilsborrow 1997, Zander, Blümel and Busse 2013, Rérat 2014). La creciente atención que este tema ha recibido en la última década y la necesidad de contar con datos más precisos y comparables entre los países han alentado al mundo académico a superar tales obstáculos anteriormente mencionados, llevando a cabo estudios especializados que les ha permitido realizar un marco muestral y, en consecuencia, desarrollar un muestreo probabilístico o, por otra parte, diseñar diferentes métodos con muestras no probabilísticas (Bloch 2007, McKenzie and Mistiaen 2009).

En nuestro caso, la población objeto de estudio es bastante grande, además de estar ampliamente distribuida geográficamente, por lo que la opción de haber llevado a cabo una encuesta presencial (cara a cara) habría resultado económicamente inalcanzable. Por otro lado, dada la diversidad lingüística de la población objeto y su dispersión geográfica, también se descartó la opción de una encuesta telefónica. Por estas razones, la encuesta *online* supuso la mejor opción (Kwak, Radler et al. 2002, Duffy et al. 2005). Además, este tipo de encuesta permite acceder rápidamente a los resultados incluso en tiempo real (Bloch 2007). Especialmente, en la medida en que el equipo de investigación del proyecto MOVE está formado por siete equipos de seis países, todos ellos distribuidos a lo largo de Europa, por lo que la posibilidad de monitorear y rastrear todo el proceso de campo ahorra mucho tiempo y recursos (Dillman 2000, Paolo et al. 2000, Hamilton 2009). Académicos como Landau and Jacobsen (2005) y Bloch (2007) han argumentado contra la falta de anonimato y confidencialidad en las respuestas de las encuestas *online*. Sin embargo, cuando el cuestionario incorpora conjuntos de preguntas sensibles, o las respuestas requieren a los participantes más tiempo de lo esperado, como ocurre con el cuestionario del proyecto MOVE, las encuestas *online* parecen ser la mejor opción. Además, algunos autores han considerado vulnerable a la población inmigrante, pudiendo ser identificada como una barrera para la participación en investigación (Atkinson and Flint 2001, Tyldum and Brunovskis 2005). En estos casos, las encuestas *online* se comportan como cuestionarios autoadministrados por correo ordinario, evitando la “intimidación” debida a la presencia de los entrevistadores telefónicos o presenciales, por lo que los encuestados tienden a sentirse más seguros ofreciendo respuestas más honestas (Davis et al. 2004, Wright 2005).

Además, la principal razón por la que se lanzó una encuesta *online* es debida a la existencia de una perfecta correspondencia entre nuestra población objetivo y su conexión con las TIC,

especialmente con Internet y el uso de las redes sociales. Una de las mayores preocupaciones en la disciplina con las encuestas *online* es la falta de representatividad, argumentando que no toda la población tiene acceso a Internet (Wright 2005). Este no es el caso de las cohortes de los jóvenes europeos, donde el acceso a Internet entre los jóvenes de 16 a 29 años en los países de la UE-28 es del 94 %. 85 % de ellos también navegan por la red todos los días y 80 % de ellos han creado un perfil en alguna red social. Estos porcentajes presentan un potencial escenario con mayor accesibilidad que las entrevistas presenciales (Duffy et al. 2005) o las telefónicas a teléfono fijo, cuyos ratios han caído en la última década en los países europeos, especialmente entre los jóvenes, que son más propensos a utilizar sus teléfonos móviles cuando viven solos (70 % de ellos) (Häder 2012). Por lo tanto, el desarrollo de una encuesta *online* permite superar el obstáculo de la representatividad. Además, el papel de la información y los servicios de apoyo, aunque también el impacto de las redes sociales en la población migrante, tanto reales como virtuales (TIC), constituyen temas tratados en la encuesta, en tanto que este tipo de redes pueden contribuir a nuevas actividades e identidades transnacionales (Binder and Sutcliffe 2014, Baltar and Brunet 2012, Guarnizo, Portes and Haller 2003, Portes 2001, Waldinger 2008).

2.2. La encuesta snowball

Después de haber discutido los beneficios y los inconvenientes de llevar a cabo una encuesta *online*, trataremos la conveniencia de complementar la muestra antes mencionada llevando a cabo una encuesta mediante la aplicación de muestreo “bola de nieve.” “snowball”.

Sin lugar a dudas, el muestreo es uno de los principales desafíos al que se enfrentan los investigadores que desean realizar una investigación mediante encuesta a la población migrante. Al preguntar a toda la población si se mudaron o no a sus países de nacimiento, aseguramos la representatividad y la capacidad de generalizar los resultados utilizando técnicas basadas en probabilidades. Sin embargo, las encuestas frecuentemente sub-representan a las poblaciones ocultas, en tanto que un gran número de sujetos es necesario para revelar fenómenos de “baja incidencia” en la población (Watters and Biernacki 1989,419). Por lo tanto, los estudios en migración y movilidad no han pasado por alto el desafío de enfrentarse a poblaciones que podrían ser definidas como “elementos ocultos” o “raros”(Kendall et al. 2008, Muhib et al. 2001, Platt et al. 2006, Heckathorn 1997), en cuyo caso resulta bastante infrecuente la probabilidad de lograr marcos muestrales completos.

Algunos estudios en ciencias de la salud (Pequegnat et al. 2007), ciencias sociales (Platt et al. 2006), y más específicamente en estudios de migración y refugiados (Perez et al. 2013, Baltar and Brunet 2012), han superado este obstáculo mediante el diseño de estrategias mixtas, incluyendo muestras no probabilísticas. Una muestra no probabilística es aquella en la que algunos miem-

bros de la población encuestada tienen una probabilidad desconocida (o ninguna) de selección, ya que no se utiliza o no se dispone de un marco de muestreo completo y preciso (Bloch 1999). En este sentido, se han utilizado en la práctica tres enfoques para muestrear elementos “raros” (Kalton and Anderson 1986) para estudiar el comportamiento de los migrantes: muestreo estratificado utilizando fracciones de muestreo desproporcionadas con muestreo en dos fases; muestreo “snowbal” y muestreo de punto de intersección, también conocido como muestreo de tiempo y espacio, muestreo de ubicación o muestreo de puntos de agregación (McKenzie and Mistiaen 2009).

Sin embargo, al analizar poblaciones ocultas, el muestreo “snowball” (Goodman 1961) es la forma más eficiente de acceder a los participantes a través de personas cuyos contactos compartan características similares (Biernacki and Waldorf 1981, Buchanan and Smith 1999). De acuerdo con esto, algunas instituciones y gobiernos también lo han llevado a la práctica con el fin de acceder a poblaciones migrantes; es el caso del Proyecto de Migración Mexicana (Massey and Singer 1995), del Netherlands Interdisciplinary Demographic Institute junto con Eurostat para encuestar a inmigrantes en España (Groenewold and Bilsborrow 2004), entre otros. Además, el uso de un amplio abanico de plataformas de redes sociales para el mantenimiento de la red está bien documentado en anteriores investigaciones (Haythornthwaite 2005, Ledbetter 2009). Las tecnologías de la información, generalmente, han ayudado a los migrantes a mantenerse conectados en relación a sus orígenes geográficos (Baltar and Brunet 2012, Illenberger and Flötteröd 2012, Rérat 2014), mejorando su calidad de vida social, aunque también puede plantear desafíos adicionales. El uso de las Redes Sociales (Facebook, Twitter, etc.) entre los migrantes, por tanto, puede indicar la necesidad de mantener los lazos culturales y las prácticas cotidianas entre pares (Binder and Sutcliffe 2014, Dekker and Engbersen 2014).

En tanto que no hay disponibles cifras reales sobre el número de jóvenes inmersos en un proyecto migratorio o en una experiencia de movilidad, especialmente teniendo en cuenta toda la variedad tipológica de movilidades (por motivos de estudios, voluntariado y profesional, etc.), se ha seguido la literatura previa combinando y complementando la encuesta panel *online* ($n=5.499$), con una encuesta mediante muestreo *snowball* que involucra, en este caso, a jóvenes inmersos en un proceso de movilidad en el momento de realización de la encuesta.

La metodología *snowball* consiste en una muestra inicial de individuos, a la que se le pide que proporcione a otros individuos de la población de interés (Goodman 1961). A cada individuo de la muestra se le pide que nombre k diferentes individuos que forman parte de la población rara” mencionada; y a cada uno de estos se les pide lo mismo, y así sucesivamente. Una condición necesaria para la aplicación exitosa del *snowball* es que los miembros de una población rara” se conozcan entre si (Kalton and Anderson 1986). Heckathorn (1997) ha demostrado que es posible obtener una muestra representativa a través de estas metodologías, basado en la idea de los

“seis grados de separación”, según la cual cada persona de una población está ligada entre sí a través de seis intermediarios en promedio. La mayor preocupación, entonces, cuando se ejecuta el proceso *snowball*, es obtener un número suficiente de puntos de partida, más probable es que la muestra comparta características y sea más homogénea en naturaleza que la población de la que se extrae. Maximizar el número de grupos iniciales aumenta la probabilidad de obtener una muestra más representativa, al igual que utilizar más puntos de partida en este proceso. El uso de muestras grandes, de cuotas por variables sociodemográficas claves para asegurar la representación, y el uso de múltiples puntos de partida para iniciar el proceso de muestreo *snowball* son algunas de las técnicas que han sido utilizadas (Atkinson and Flint 2001). Entonces, una manera de ampliar la inclusión y, por tanto, la representatividad de las encuestas con migrantes, es el uso de múltiples enfoques de redes.

2.2.1. Proceso de difusión de la encuesta “snowball

Dada la composición de la muestra de la encuesta panel (jóvenes de 18 a 29 años residentes en alguno de los países del consorcio del MOVE al momento de realización de la encuesta) nos sugiere que la proporción de encuestados no móviles prevalecería sobre el número de móviles. El requisito de la ubicación geográfica apuntaba que el muestreo de la población joven móvil era en sí mismo un factor crítico en el proceso de recolección de datos. Para abordar este desafío y captar un mayor número de jóvenes con experiencia de movilidad, tiene sentido la encuesta *online* mediante muestreo *snowball* (n=3.207) para contrabalancear los datos obtenidos del panel sobremuestreando el subconjunto de la población joven móvil.

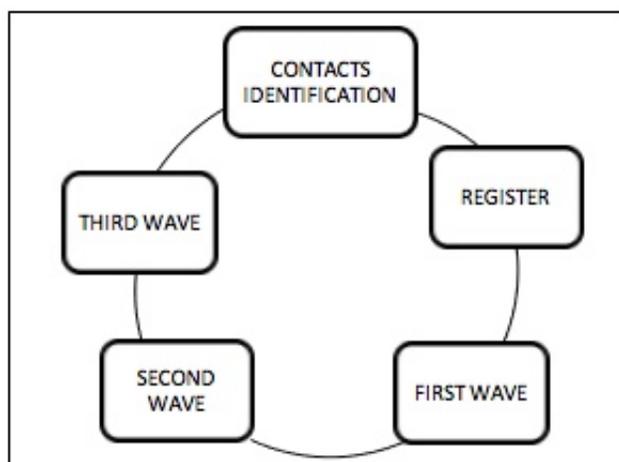
Basado en análisis previos de datos macro, los esquemas de movilidad juvenil que representan las principales razones de movilidad fueron divididas en seis grupos: empleo, formación profesional, voluntariado social, movilidad por motivos de estudios e intercambio de alumnos. Por tanto, se centraron los esfuerzos en estos seis subgrupos para la difusión de la encuesta mediante el muestreo *snowball*.

En términos prácticos, la encuesta se alojó en servidores de GfK, previamente subcontratados para la encuesta panel *online*, y difundidos por los miembros del equipo del consorcio siguiendo una estrategia coordinada. Esta segunda encuesta fue lanzada el 16 de noviembre de 2016, alcanzando su mayor pico de actividad del 19 de diciembre al 31 de enero de 2017, finalizando el 5 de febrero.

Con el fin de asegurar una tasa de respuesta positiva, cada punto de contacto se trató de la siguiente manera:

1. Investigar los contactos clave pertenecientes a cada tipo de movilidad.
2. Registro de los datos de contacto en una plantilla.
3. Una primera oleada de difusión, presentando los objetivos generales del proyecto MOVE, solicitando su colaboración para difundir la encuesta a través de su grupo objetivo de jóvenes móviles relacionados, registrando la respuesta.
4. Una segunda oleada de difusión en cuyo fin de asegurar una colaboración continua en todo el proceso los miembros del consorcio lanzaron dos recordatorios más.
5. Tercera y última oleada de difusión.

Gráfico 1: Proceso de difusión de la encuesta snowball



Los puntos de contacto para la difusión abarcaron desde instituciones hasta empresas, contactos personales, jóvenes móviles o grupos de Facebook¹. Con el propósito de implementar la estrategia de difusión a través de redes sociales, cada país del consorcio creó perfiles oficiales y perfiles *ad hoc*. Los primeros se utilizaron para dirigirse a organizaciones formales y los segundos para participar en grupos informales, en el que sólo se permite que los perfiles individuales se unan a los grupos.

El número total de puntos de contacto alcanzados durante el proceso de difusión fue de 5.173, siendo la movilidad por razones de estudios y la movilidad general los grupos más grandes.

¹Ver tabla en anexo

Tabla 1: Puntos de contacto registrados durante el proceso de difusión del snowball

Tipo de movilidad	Nº inicial de contactos	Total nº
General	651	1,856
Intercambio de alumnos	29	134
Movilidad por estudios	665	2,384
Voluntariado social	240	131
Empleo	78	103
Formación profesional	140	322
Emprendimiento	202	243

Teniendo en cuenta la duración de la encuesta (30 minutos) y la progresiva caída de las tasas de respuesta en las encuestas de investigación (Baruch and Holtom 2008, Amundsen and Lie 2013), los miembros del consorcio acordaron el uso de incentivos entre los jóvenes móviles con el fin de aumentar la tasa de respuesta (Deutskens et al. 2004).

Los mensajes de difusión de la encuesta Snowball también proporcionaron información clara sobre la posibilidad de participar en un sorteo de 20 tarjetas de regalo, 50 por cada una, para una suscripción de seis meses a una plataforma *online* de música, películas y series o una tarjeta de regalo para gastar en libros, música, productos tecnológicos, productos éticos hechos a mano o una ONG a elección. Una vez completada la encuesta, se permitió a los participantes la posibilidad de recibir el informe final de la investigación y también la de participar en el sorteo.

2.3. Fusión del Panel y el Snowball

¿Por qué utilizar dos conjuntos de datos? Por un lado, el uso de la encuesta Panel se justifica por razones de representatividad de la muestra, pero también por presencia de dos subpoblaciones de nuestro interés (móviles y no móviles).

La muestra solo tiene en cuenta la población en general. Por lo tanto, cuando segmentamos por subconjuntos de la población tales como los móviles y los no móviles, o incluso, por el género, la edad, o la ocupación, las muestras pierden representatividad. Lo que debilita la capacidad de generalización de los resultados para esas submuestras.

Debido a los propósitos generales del proyecto MOVE, y con el fin de responder a los principales objetivos marcados en el mismo, sobrerrepresentar a la población móvil supone una necesidad. En consecuencia, poseemos dos conjuntos de datos con dos conjuntos diferentes de

población, móviles y no móviles, y dentro de los primeros dos etapas diferentes de movilidad.

¿Por qué fusionar dos conjunto de datos, es decir, por qué unir el conjunto de datos de la encuesta Panel y el de la encuesta *snowball*? Con el propósito de obtener resultados más robustos en nuestra población de interés e identificar los patrones de la movilidad de los jóvenes en Europa, necesitamos combinar ambos conjuntos de datos, dejando fuera las observaciones de la población no móvil del panel proporcionándonos un n total de 5.096.

¿Es este proceso teórica y metodológicamente válido? Teóricamente es válido porque estamos analizando una misma población móvil, a la vez que la realización de ambas encuestas se realizó al mismo tiempo, aunque con una diferencia: el Panel se refiere a la experiencia de movilidad retrospectiva, mientras que el *snowball* se refiere a la experiencia de movilidad actual. En ambos casos se realizaron las mismas preguntas, siendo fácilmente identificables. Con el fin de combinar ambos conjuntos de datos y dar cuenta del potencial *bias* en relación a la variación temporal en sus respuestas, hemos codificado la fecha de inicio de la experiencia de movilidad por año, correspondiéndose en el caso del *snowball* con 2017. Por otro lado, la duración de la experiencia de movilidad ha sido codificada como inferior a tres meses y de tres meses o más. Por fortuna, la distribución entre ellos es normal en el panel.

¿Cómo fusionar los dos conjuntos de datos? Con el fin de utilizar los datos del panel como encuesta de referencia, los hemos ponderado a partir de datos oficiales provenientes de Eurostat mediante la técnica de **post-estratificación**. La post-estratificación se calcula en el conjunto de datos final para variables sociodemográficas tales como el país, la edad, el género y el nivel educativo. No obstante, en cuanto al acceso a internet no se ha tenido en cuenta el porcentaje general de acceso y se asume que la población joven de nuestros países de interés alcanza casi la universalización. Hemos descartado también utilizar en este proceso el desempleo, en tanto que nuestra población de interés (18-29) puede figurar como desempleada pero a la vez estar estudiando.

¿Cómo ponderar la muestra? La post-estratificación es un método bastante extendido a la hora de ponderar muestras. En base a variables de clasificación se diseñan estratos utilizados *a posteriori* para comparar la composición de cada uno de ellos con estadísticas oficiales. Las variables cualitativas son necesarias en este método para el correcto diseño de los estratos (Cobben and Bethlehem 2013).

Siguiendo la notación propuesta por Cobben and Bethlehem (2013), suponiendo que tenemos una variable X con L categorías, la población es dividida en L estratos. Esta variable dividirá a la población en U_1, U_2, \dots, U_L estratos. El número de target population objeto de estudio en

cada uno de los estratos U_h es denominada como N_h , siendo $h \in \{1, 2, \dots, h\}$. El tamaño de la población N es igual a $N = N_1 + N_2 + \dots + N_L$.

Por otro lado, suponiendo una muestra aleatoria de esa población con un n total de n_R . $n_{R,h}$ denota el número total de encuestados en el estrato h , entonces $n_R = n_{R1} + n_{R2} + \dots + n_{RL}$. El proceso de pos-estratificación otorga los mismos factores de ajuste en cada uno de los elementos de los mismos estratos. El factor de corrección w_k para un elemento k en el estrato h es igual a:

$$w_k = \frac{N_h/N}{n_{R,h}/n_R} \quad (1)$$

El bias del estimador basado en la tasa de respuesta media es según Cobben and Bethlehem (2013) y basado en Bethlehem (2009):

$$B(\bar{y}_{R,PS}) = \sum_{h=1}^L W_h B(\bar{y}_R^{(h)}) = \sum_{h=1}^L W_h \frac{R_{\rho Y}^{(h)} S_{\rho}^{(h)} S_Y^{(h)}}{\bar{\rho}^{(h)}} \quad (2)$$

Siendo $R_{\rho Y}^{(h)}$ la correlación entre Y y ρ en el estrato h , $S_{\rho}^{(h)}$ es la desviación estandar de las probabilidades de respuesta en el estrato h , and $S_Y^{(h)}$ es la desviación estandar de la variable target en el estrato h .

Partiendo del hecho de que nuestra encuesta panel fue hecha vía internet y no a través de un panel probabilístico tenemos en cuenta los sesgos derivados de este método. Como los objetivos del proyecto incluyen extrapolar conclusiones para la movilidad general de los países del consorcio los primeros factores de ponderación fueron calculados en base a las siguientes variables:

- País de residencia del encuestado, con un total de 6 categorías: Alemania, España, Hungría, Luxemburgo, Noruega y Rumanía.
- Edad del encuestado con dos categorías de respuesta 18-24 y de 25-29 años.
- Género, con dos categorías de respuesta: Hombre y Mujer.
- Nivel educativo del encuestado, con tres categorías de respuesta: Primaria, secundaria y superior.

En un principio fueron diseñados 73 estratos en total. Las frecuencias de cada uno de los estratos están disponibles en EUROSTAT. Al mismo tiempo estos estratos fueron elaborados en la base de datos del panel. En la tabla 2 están ordenados los primeros veinte estratos con el factor de corrección más alto.

Tabla 2: Valores más altos para el factor de corrección

	STRATUMS	COUNTRY	AGE	GENDER	EDUCATION	wf
1	5211	Romania	25-29	Male	Primary	42.40
2	5221	Romania	25-29	Female	Primary	35.94
3	6221	Spain	25-29	Female	Primary	11.10
4	6211	Spain	25-29	Male	Primary	10.18
5	1222	Germany	25-29	Female	Secondary	5.77
6	5111	Romania	18-24	Male	Primary	5.59
7	5121	Romania	18-24	Female	Primary	5.24
8	1212	Germany	25-29	Male	Secondary	5.13
9	1111	Germany	18-24	Male	Primary	5.02
10	6111	Spain	18-24	Male	Primary	4.66
11	6121	Spain	18-24	Female	Primary	4.41
12	1121	Germany	18-24	Female	Primary	4.03
13	1112	Germany	18-24	Male	Secondary	3.52
14	5222	Romania	25-29	Female	Secondary	2.75
15	1122	Germany	18-24	Female	Secondary	2.25
16	1211	Germany	25-29	Male	Primary	2.20
17	6122	Spain	18-24	Female	Secondary	2.17
18	1223	Germany	25-29	Female	Tertiary	2.07
19	6212	Spain	25-29	Male	Secondary	2.02
20	1213	Germany	25-29	Male	Tertiary	1.86

Los estratos menos representados en el panel corresponden a Rumania y España, con edades comprendidas entre los 25-29 años y con nivel "Primary" de educación. Estos resultado están en consonancia con los resultados obtenido por Isaksson and Forsman (2003). Donde la muestra obtenida a través de internet resultó tener un nivel educativo más alto. A fin de evitar factores de ponderación tan elevados fueron diseñados unos nuevos estratos recodificando la variable edad en dos estratos "Secondary or less" y "Tertiary", obteniéndose finalmente un total de 48 estratos. La Tabla 3 contiene un resumen descriptivo de los valores para las dos ponderaciones. En esta ocasión el valor máximo de los estratos es de 4.12.

Tabla 3: Estadísticos de resumen para los dos procesos de estratificación

	Min.	Median	Mean	Max.
1	0.01	0.52	2.51	42.40
2	0.01	0.36	1.01	4.12

Después de este primer paso, tomamos los resultados ponderados del panel como estadística de referencia y realizamos el mismo proceso de post-estratificación con el conjunto de datos del *snowball*

Además, de acuerdo con nuestro compromiso de realizar un análisis en perspectiva compara-

da, así como a nivel de país, calculamos el análisis de post-estratificación tanto a nivel agregado como a nivel de país.

Resultados esperados:

Llevando a cabo este proceso, es decir, al aumentar la muestra de población móvil, podríamos esperar obtener resultados más robustos y significativos en nuestros análisis posteriores. La razón es que, dada una población oculta, no hay disponibles estadísticas oficiales de nuestro universo, por lo que utilizamos el panel para obtener una muestra representativa de esta población.

3.Resultados

En este apartado procedemos a comparar la composición de la muestra tanto para el Panel como para el Snowball, y en su versión final conjunta; teniendo en cuenta los factores de ponderación calculados mediante el proceso de post-estratificación anteriormente descrito y atendiendo a las principales variables sociodemográficas: país, edad (en dos grandes grupos –18-24 y 25-29), género y nivel educativo, que además constituyen las variables auxiliares para la construcción de los estratos. También procedemos en el análisis de algunas variables de interés que forman parte de los objetivos del proyecto MOVE, dados los propósitos de este Grupo de Trabajo y del XIII Congreso Nacional de AECPA, tales como la participación política y electoral de los jóvenes móviles y no móviles.

3.1.Composición de la muestra

En primer lugar, procedemos a señalar la composición de la muestra por país comparando según las estadísticas oficiales.

Tabla 4: Composición muestral del Panel y Snowball por país

Country	Eurostat		Panel		Snowball	
	Frec. (jóvenes 18-29)	%	Frec.	%	Frec.	%
 Alemania	11.720.972	51,7 %	961	17.5 %	1.124	35.0 %
 Hungría	1.444.897	6,4 %	980	17.8 %	157	4.9 %
 Luxemburgo	90.824	0,4 %	739	13.4 %	231	7.2 %
 Noruega	837.198	3,7 %	877	15.9 %	176	5.5 %
 Rumanía	2.862.824	12,6 %	976	17.7 %	354	11.0 %
 España	5.708.794	25,2 %	966	17.6 %	1.165	36.3 %
 Total	22.665.509	100 %	5.499	100 %	3.207	100 %

Comparando la composición por edad entre el panel y el snowball observamos en ambas una proporción mayor en el grupo de 18-24 años, sobre todo en el snowball en el que se agudiza la diferencia respecto a los de 25-29 años. De acuerdo a las estadísticas oficiales, al menos, la muestra es proporcional en el caso del panel.

Tabla 5: Composición de la muestra por edad

	Eurostat		Panel		Snowball	
	Frec. jóvenes de 18 a 29	%	Frec.	%	Frec.	%
18-24	12.336.405	54,43 %	3.068	55,79 %	2070	64,5 %
25-29	10.329.104	45,57 %	2.431	44,21 %	1.137	35,5 %
Total	22.665.509	100 %	5.499	100 %	3.207	100 %

Atendiendo a las diferencias por género (Tabla 6) se observa un claro desajuste entre el panel y el snowball. Las mujeres representan el 71,7 % de la muestra del snowball, mientras que en el panel constituyen el 53,3 %. Los datos de Eurostat, para el conjunto de mujeres jóvenes de los seis países, indican un 48,4 %. No obstante, el hecho de una mayor proporción de mujeres en ambas muestras consistente con la literatura, en tanto que éstas tienden a participar más en las encuestas (Curtin, Presser and Singer 2000, Moore and Tarnai 2002, Singer, Hoewyk and Maher 2000).

Tabla 6: Distribución de la muestra por género

	Eurostat		Panel		Snowball	
	Frec. jóvenes de 18 a 29	%	Frec.	%	Frec.	%
Hombres	11.705.258	51,6 %	2.567	46,7 %	907	28,3 %
Mujeres	10.960.251	48,4 %	2.932	53,3 %	2.300	71,7 %
Total	22.665.509	100 %	5.499	100 %	3.207	100 %

4. Bibliografía

- Amundsen, Bard and Else Lie. 2013. “Do surveys still give an accurate picture?”.
URL: https://www.forskningradet.no/en/Newsarticle/Fewer_willing_to_participate_in_surveys/1253986812629
- Atkinson, Rowland and John Flint. 2001. “Accessing hidden and hard-to-reach populations: Snowball research strategies.” *Social research update* 33(1):1–4.
- Bakewell, Oliver. 2010. “Some reflections on structure and agency in migration theory.” *Journal of Ethnic and Migration Studies* 36(10):1689–1708.
- Baltar, Fabiola and Ignasi Brunet. 2012. “Social research 2.0: virtual snowball sampling method using Facebook.” *Internet Research* 22(1):57–74.
- Baruch, Yehuda and Brooks C. Holtom. 2008. “Survey response rate levels and trends in organizational research.” *Human Relations* 6(8):1139–1160.
- Bethlehem, Jelke. 2009. *Sample Surveys: Design, Methods and Applications*. Vol. 29A Elsevier chapter Computer software for sample surveys, pp. 317–350.
- Biernacki, Patrick and Dan Waldorf. 1981. “Snowball sampling: Problems and techniques of chain referral sampling.” *Sociological methods & research* 10(2):141–163.
- Bilsborrow, Richard E. 1997. *International migration statistics: Guidelines for improving data collection systems*. International Labour Organization.
- Binder, Jens F and Alistair G Sutcliffe. 2014. “The Best of Both Worlds? Online Ties and the Alternating Use of Social Network Sites in the Context of Migration.” *Societies* 4(4):753–769.
- Bloch, Alice. 1999. “Carrying out a survey of refugees: some methodological considerations and guidelines.” *J. Refugee Stud.* 12:367.
- Bloch, Alice. 2007. “Methodological challenges for national and multi-sited comparative survey research.” *Journal of Refugee Studies* 20(2):230–247.
- Buchanan, Tom and John L Smith. 1999. “Using the Internet for psychological research: Personality testing on the World Wide Web.” *British journal of Psychology* 90(1):125–144.
- Cobben, Fannie and Jelke Bethlehem. 2013. Web panels for official statistics. Technical report Statistics Netherlands.
- Couper, Mick P. 2000. “Review: Web surveys: A review of issues and approaches.” *Public opinion quarterly* pp. 464–494.

- Couper, Mick P, Michael W Traugott and Mark J Lamias. 2001. "Web survey design and administration." *Public opinion quarterly* 65(2):230–253.
- Curtin, R, S Presser and E Singer. 2000. "The effects of response rate changes on the index of consumer sentiment." *Public Opinion Quarterly* .
URL: <http://poq.oxfordjournals.org/content/64/4/413.short>
- Davis, Mark, Graham Bolding, Graham Hart, Lorraine Sherr and Jonathan Elford. 2004. "Reflecting on the experience of interviewing online: perspectives from the Internet and HIV study in London." *AIDS care* 16(8):944–952.
- Dekker, Rianne and Godfried Engbersen. 2014. "How social media transform migrant networks and facilitate migration." *Global Networks* 14(4):401–418.
- Deutskens, Elisabeth, Ko De Ruyter, Martin Wetzels and Paul Oosterveld. 2004. "Response rate and response quality of Internet-based surveys." *Marketing Letters* 15(1):21–36.
- Dillman, Don A. 1978. *Mail and telephone surveys*. Vol. 3 Wiley Interscience.
- Dillman, Don A. 2000. *Mail and internet surveys: The tailored design method*. Vol. 2 Wiley New York.
- Duffy, Bobby, Kate Smith, George Terhanian and John Bremer. 2005. "Comparing data from online and face-to-face surveys." *International Journal of Market Research* 47(6):615.
- Fawcett, James T and Fred Arnold. 1987. "The role of surveys in the study of international migration: an appraisal." *International Migration Review* pp. 1523–1540.
- Goodman, Leo A. 1961. "Snowball sampling." *The annals of mathematical statistics* pp. 148–170.
- Gosling, Samuel D, Simine Vazire, Sanjay Srivastava and Oliver P John. 2004. "Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires." *American Psychologist* 59(2):93.
- Groenewold, George and Richard Bilsborrow. 2004. Design of samples for international migration surveys: methodological considerations, practical constraints and lessons learned from a multi-country study in Africa and Europe. In *Population Association of America 2004 General Conference*.
- Guarnizo, Luis Eduardo, Alejandro Portes and William Haller. 2003. "Assimilation and Transnationalism: Determinants of Transnational Political Action among Contemporary Migrants1." *American journal of sociology* 108(6):1211–1248.
- Häder, Michael. 2012. *Data Quality in Telephone Surveys via Mobile and Landline Phone*. Springer.

- Hamilton, MB. 2009. "On line survey response rates and times." *Bethesda, MD: ltpath* 250.
- Haythornthwaite, Caroline. 2005. "Social networks and Internet connectivity effects." *Information, Community & Society* 8(2):125–147.
- Heckathorn, Douglas D. 1997. "Respondent-driven sampling: a new approach to the study of hidden populations." *SOCIAL PROBLEMS-NEW YORK-* 44:174–199.
- Hussein, Shereen, Jill Manthorpe and Martin Stevens. 2011. "The experiences of migrant social work and social care practitioners in the UK: findings from an online survey." *European Journal of Social Work* 14(4):479–496.
- Illenberger, Johannes and Gunnar Flötteröd. 2012. "Estimating network properties from snowball sampled data." *Social Networks* 34(4):701–711.
- Isaksson, Annica and Gösta Forsman. 2003. "A comparison between using the web and using telephone to survey political opinions." *American Association for Public Opinion Research* pp. 100–106.
- Kalton, Graham and Dallas W Anderson. 1986. "Sampling rare populations." *Journal of the royal statistical society. Series A (general)* pp. 65–82.
- Kendall, Carl, Ligia RFS Kerr, Rogerio C Gondim, Guilherme L Werneck, Raimunda Hermelinda Maia Macena, Marta Kerr Pontes, Lisa G Johnston, Keith Sabin and Willi McFarland. 2008. "An empirical comparison of respondent-driven sampling, time location sampling, and snowball sampling for behavioral surveillance in men who have sex with men, Fortaleza, Brazil." *AIDS and Behavior* 12(1):97–104.
- Konstan, Joseph A, BR Simon Rosser, Michael W Ross, Jeffrey Stanton and Weston M Edwards. 2005. "The story of subject naught: A cautionary but optimistic tale of Internet survey research." *Journal of Computer-Mediated Communication* 10(2):00–00.
- Kwak, Nojin, Barry Radler et al. 2002. "A comparison between mail and web surveys: Response pattern, respondent profile, and data quality." *JOURNAL OF OFFICIAL STATISTICS-STOCKHOLM-* 18(2):257–274.
- Lados, Gábor. 2013. The impact and importance of return migration in East Central Europe. In *Forum geografic*. Vol. 12 pp. 132–137.
- Landau, Loren B and Karen Jacobsen. 2005. "The value of transparency, replicability and representativeness." *Forced Migration Review* 22:46.
- Ledbetter, Andrew M. 2009. "Patterns of media use and multiplexity: Associations with sex, geographic distance and friendship interdependence." *New Media & Society* 11(7):1187–1208.

Massey, Douglas S and Audrey Singer. 1995. "New Estimates of Undocumented Mexican Migration and the Probability of Apprehension*." *Demography* 32(2):203–213.

McKenzie, David J and Johan Mistiaen. 2009. "Surveying migrant households: a comparison of census-based, snowball and intercept point surveys." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 172(2):339–360.

Moore, DL and J Tarnai. 2002. "Evaluating nonresponse error in mail surveys." *Survey nonresponse* .

URL: https://scholar.google.es/scholar?q=%3B+Moore+%26+Tarnai%2C+2002&btnG=&hl=es&as_sdt=0

Muhib, Farzana B, Lillian S Lin, Ann Stueve, Robin L Miller, Wesley L Ford, Wayne D Johnson, Philip J Smith, Community Intervention Trial for Youth Study Team et al. 2001. "A venue-based method for sampling hard-to-reach populations." *Public health reports* 116(Suppl 1):216.

Paolo, Anthony M, Giulia A Bonaminio, Cheryl Gibson, Ty Partridge and Ken Kallail. 2000. "Response rate comparisons of e-mail-and mail-distributed student evaluations." *Teaching and learning in medicine* 12(2):81–84.

Pequegnat, Willo, BR Simon Rosser, Anne M Bowen, Sheana S Bull, Ralph J DiClemente, Walter O Bockting, Jonathan Elford, Martin Fishbein, Laura Gurak, Keith Horvath et al. 2007. "Conducting Internet-based HIV/STD prevention survey research: considerations in design and evaluation." *AIDS and Behavior* 11(4):505–521.

Perez, Daniel F, Jason X Nie, Chris I Ardern, Natasha Radhu and Paul Ritvo. 2013. "Impact of participant incentives and direct and snowball sampling on survey response rate in an ethnically diverse community: Results from a pilot study of physical activity and the built environment." *Journal of Immigrant and Minority Health* 15(1):207–214.

Platt, Lucy, Martin Wall, Tim Rhodes, Ali Judd, Matthew Hickman, Lisa G Johnston, Adrian Renton, Natalia Bobrova and Anya Sarang. 2006. "Methods to recruit hard-to-reach groups: comparing two chain referral sampling methods of recruiting injecting drug users across nine studies in Russia and Estonia." *Journal of Urban Health* 83(1):39–53.

Portes, Alejandro. 2001. "Introduction: the debates and significance of immigrant transnationalism." *Global networks* 1(3):181–194.

Rérat, Patrick. 2014. "The selective migration of young graduates: Which of them return to their rural home region and which do not?" *Journal of Rural Studies* 35:123–132.

Singer, E, J Van Hoewyk and MP Maher. 2000. "Experiments with incentives in telephone surveys." *Public Opinion Quarterly* .

URL: <http://poq.oxfordjournals.org/content/64/2/171.short>

- Sue, Valerie M and Lois A Ritter. 2012. *Conducting online surveys*. Sage.
- Tyldum, Guri and Anette Brunovskis. 2005. "Describing the unobserved: Methodological challenges in empirical studies on human trafficking." *International Migration* 43(1/2):17–34.
- Waldinger, Roger. 2008. "Between "Here" and "There": Immigrant Cross-Border Activities and Loyalties." *International Migration Review* 42(1):3–29.
- Watters, John K and Patrick Biernacki. 1989. "Targeted sampling: options for the study of hidden populations." *Social problems* 36(4):416–430.
- Wright, Kevin B. 2005. "Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services." *Journal of Computer-Mediated Communication* 10(3):00–00.
- Zander, Britta, Miriam Blümel and Reinhard Busse. 2013. "Nurse migration in Europe—Can expectations really be met? Combining qualitative and quantitative data from Germany and eight of its destination and source countries." *International journal of nursing studies* 50(2):210–218.