

# Las dificultades de vincular datos de encuestas y administrativos

Combinar datos de encuestas y administrativos está cada vez más en boga, aunque el acceso a los datos está todavía muy restringido

## DISCURSO DE ASCENSOR

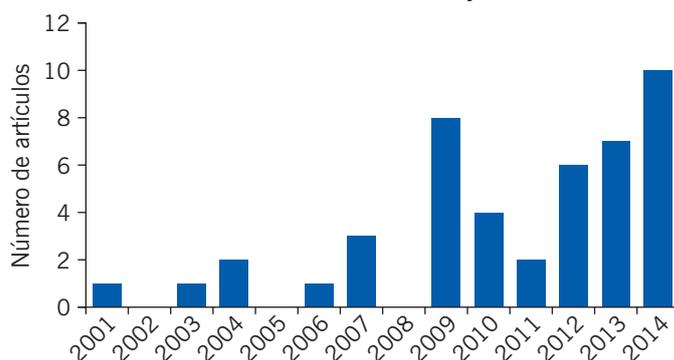
El uso de datos de registros administrativos y datos de encuestas para optimizar ambos ofrece un gran potencial a la investigación científica y relacionada con las políticas. Dos cambios recientes han ampliado el potencial para crear esta vinculación de datos: la mayor disponibilidad de fuentes de datos y los avances en la tecnología de correspondencia de datos. Estos desarrollos se reflejan en el mayor número de artículos académicos en economía del trabajo que utilizan datos administrativos y de encuestas vinculados. Aunque el número de estudios que utilizan datos vinculados es aún pequeño, la tendencia va claramente en aumento. Sin embargo, el crecimiento se ralentiza por preocupaciones sobre seguridad y privacidad de los datos, lo que dificulta el acceso a los mismos.

## HALLAZGOS CLAVE

### Pros

- ⊕ Vincular datos supera algunas de las deficiencias de las dos fuentes de datos por separado.
- ⊕ Vincular datos abre nuevas oportunidades de investigación al combinar datos administrativos muy fiables con datos de encuestas detallados.
- ⊕ Los registros administrativos, que ya se recogen de manera habitual, son una fuente de datos barata y fidedigna para enriquecer los datos de encuestas.
- ⊕ La vinculación de datos puede reducir los costes de las encuestas al requerir menos preguntas.
- ⊕ La vinculación de datos permite que los datos sensibles, como los salarios, se extraigan de los registros administrativos, lo que reduce la carga para los encuestados y evita que se abandone la encuesta o que se dejen preguntas sin contestar.

### El uso de datos administrativos y de encuestas vinculados en los artículos de economía del trabajo ha aumentado



Fuente: cálculos basados en el análisis del autor de la serie de artículos de debate de IZA.

IZA  
World of Labor

### Contras

- ⊖ Vincular datos puede ser muy caro y llevar mucho tiempo, sobre todo debido a las negociaciones de extracción con los proveedores de datos.
- ⊖ Las preocupaciones sobre privacidad y las limitaciones legales resultantes, así como la necesidad de la anonimización de los datos restringen el acceso a los datos y a su contenido.
- ⊖ Solicitar el consentimiento para usar los datos vinculados puede introducir sesgos (los otorgantes difieren de los no otorgantes) o puede reducir las tasas de respuesta, lo que presenta otro sesgo de selección.
- ⊖ Una buena vinculación requiere un identificador único para cada individuo; sin estos identificadores, la vinculación se complica y la calidad de la correspondencia puede verse afectada.

## MENSAJE PRINCIPAL DE LOS AUTORES

Vincular datos abre nuevas oportunidades de investigación al combinar registros administrativos muy fiables con datos de encuestas detallados. Los investigadores que desean vincular las dos fuentes de datos deben incluir identificadores personales únicos y hacer que la encuesta incluya una solicitud de consentimiento bien redactada para los encuestados. Más importante es que cualquier preocupación sobre la seguridad de los datos del proveedor de datos (normalmente, una institución del gobierno) debe resolverse con antelación para evitar que se produzcan restricciones de acceso o demandas de una anonimización estricta de los datos, lo que reduce el potencial de investigación.