

Calling from the outside: The role of networks in residential mobility

Büchel, K., Ehrlich, M. Puga, D., & Viladecans-Marsal, E. (2020)

Samantha Cervantes

Motivación

- Relevancia por uso de nuestro tiempo y dinero, así como oportunidades laborales y sociales
- Análisis de equilibrio espacial tradicional analiza factores de decisión comunes a muchas personas
 - Mercado laboral, precios de vivienda, amenidades, impuestos
 - *Moving costs* y características individuo-lugar
- Bajo un contexto de equilibrio espacial los pros y contras de cada ubicación se "compensan"
- Pero entonces, **¿por qué personas similares se mudan a lugares diferentes, y por qué lugares similares tienen efectos diferentes?**

Preguntas de investigación

¿Cómo afecta la cercanía de las redes sociales a la probabilidad de que un individuo emigre?

¿Cuánto cambia preferencia por un lugar con amigos y familia sobre otro sin ellos?

¿Cuál es el rol de los contactos sobre estas elecciones?

¿Cuál es el moving cost percibido de dejar a amigos y familia atrás?

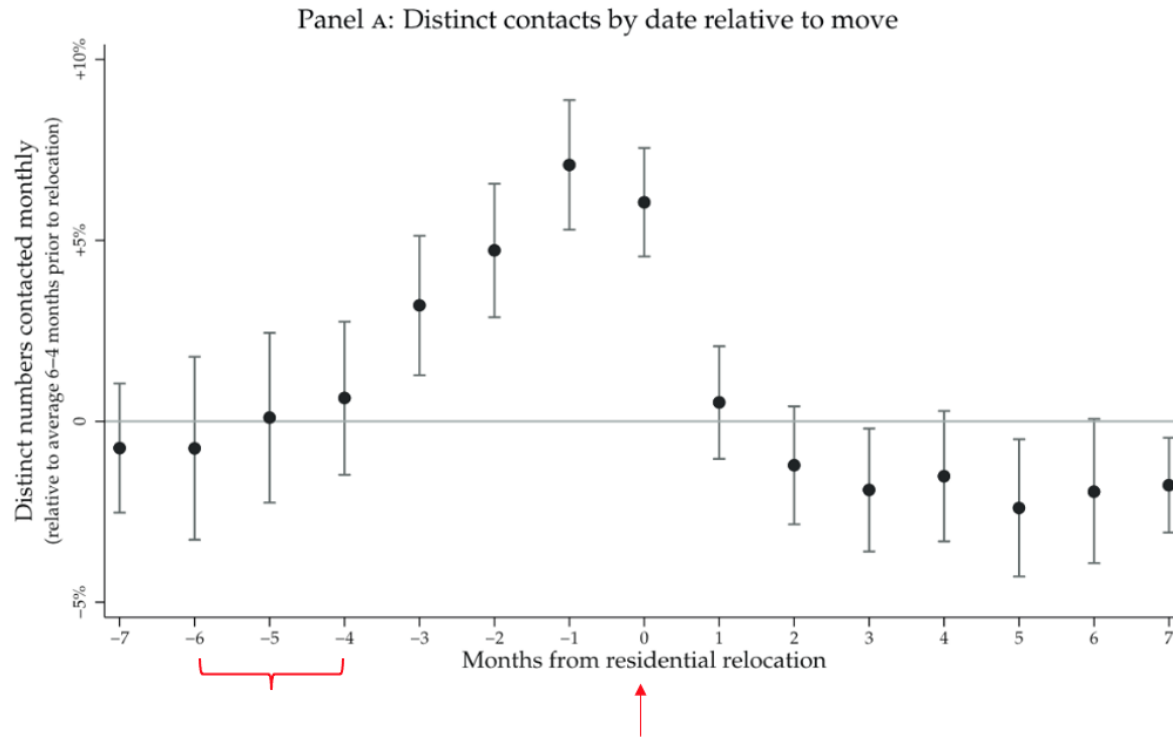
¿Cómo abordaron la pregunta?

- Usando datos de quién llamó a quién, desde dónde y a dónde, filtrando a los posibles amigos y familiares, y midiendo qué tan concentrados están alrededor de ellos.
 - 10% de las interacciones más frecuentes componen el 90% de las llamadas
 - 93% de las personas que se llaman se han visto en el último año
 - Tiempo de viaje mejor predictor de cercanía que *social media* y distancia geográfica
- Asimismo, usan con características individuales/demográficas, y espaciales.

Datos

- Operadora telefónica suiza proveyó datos de:
 - **Llamadas:** Muestra final consistió de 2.1 millones de personas y 410 millones de llamadas, realizadas de junio 2015 - mayo 2016.
 - **Demográficos:** Sexo, grupo de edad, lenguaje de correspondencia
 - **Residencia:** Código postal mensual de usuarios y contactos, 3,152 disponibles. Identificar a *movers* (5% de la muestra) y *long-term residents*
 - Distancia de individuo a sus contactos
- Datos de ubicaciones:
 - **Housing market tightness.** Meta-Sys & Swiss Federal Statistics Office
 - Guarderías. Federal Social Insurance Office
 - Eventos culturales. Guide database.
 - Crimen. Swiss Federal Statistics Office

Matriz de redes sociales



Matriz de redes sociales

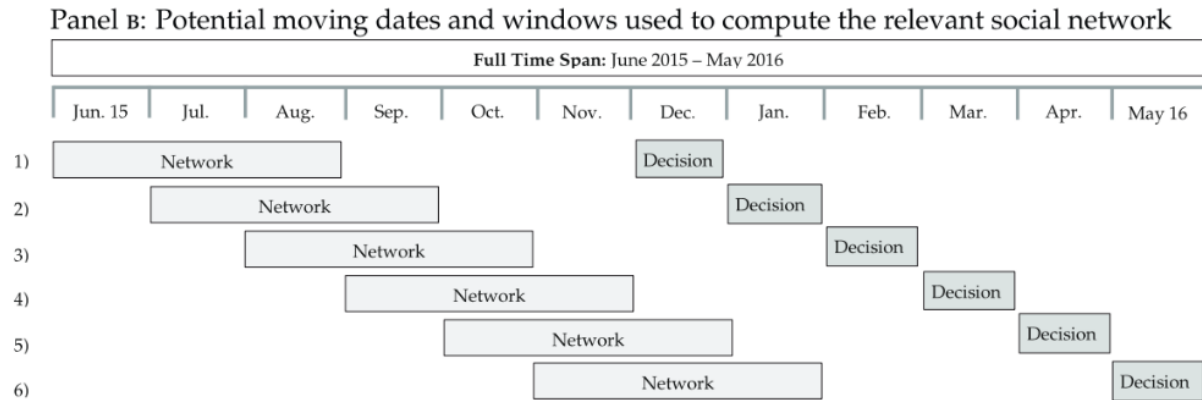


Fig. 1. Timing of relocation decisions.

Matriz de redes sociales

A partir de esto, construyen:

1. Matriz de adyacencia indicando cuántos contactos i han hablado con alguien en código postal j , en los meses 4 a 6 antes a la mudanza
2. Matriz de tiempo de viaje (coche) de individuos - contactos calculando tiempo de distancia entre todas las parejas de códigos postales

Al combinarlas, obtienen la variable de interés de las redes sociales de un individuo i en el CP j en mes t : la proporción de las personas con las que i habló por teléfono en los meses $t - 6$ y $t - 4$ que vivían a menos de 10 minutos de distancia del CP j .

Modelo

- La utilidad indirecta que el individuo i obtiene de la ubicación j viene de:

$$V_{ij} = f(X_i) + \lambda_j + \pi_{ij} + I_{ijt-1}(\bar{\lambda}_j + \bar{\pi}_{ij}) + \epsilon_{ij}$$

- donde:
 - $f(X_i)$ = set de características individuales. Incluye total de contactos de i , de primer o segundo orden, y si son del mismo origen o de importancia central, así como demográficos.
 - λ_j = set características de la ubicación, capturadas con efectos fijos
 - π_{ij} = componente de características individuo - ubicación
 - I_{ijt-1} = dummy que toma valor de 1 si el individuo i ya había vivido en j , (**stayer**), y 0 eoc (**mover**).
 - $I_{ijt-1}(\bar{\lambda}_j + \bar{\pi}_{ij})$ = heterogeneidad en los componentes de ubicación e individuo-ubicación
 - ϵ_{ij} = componente no observado de preferencias

Modelo

- Dicho componente individuo-ubicación consiste en:

$$\pi_{ij} = \kappa(N_{ij}, N_{ij}Z_j, N_{ij}X_iZ_j) + h(X_iZ_j)$$

donde:

- N_{ij} = vector con la proporción de contactos del individuo i que viven a diferentes distancias de j .
- $N_{ij}Z_j$ = vector que captura grado en que ciertas características de las ubicaciones Z_j adquieren relevancia según las redes sociales de los individuos.
- $N_{ij}X_iZ_j$ = vector que permite que características de las ubicaciones adquieran relevancia según las redes sociales del individuo y sus características demográficas.
- $h(X_iZ_j)$ = interacciones que capturan características individuales no sociales y de la ubicación

Estimación

- Primera etapa: decisión de migración. Esta ocurrirá si hay otra ubicación que le provea de más utilidad que su residencia r :

$$Prob[\max_{j \neq r}(\lambda_j + \pi_{ij} + \epsilon_{ij}) > \lambda_j + \pi_{ij} + \bar{\lambda}_j + \bar{\pi}_{ij} + \epsilon_{ij}]$$

- Segunda etapa: condicional en decisión de mudarse, estimar la probabilidad de que la alternativa k dé la utilidad máxima entre todas las j , dado que $j \neq r$:

$$Prob[\max_{j \neq r}(\lambda_j + \pi_{ij} + \epsilon_{ij}) > \lambda_j + \pi_{ij} + \epsilon_{ij}]$$

- Asumiendo que ϵ_{ij} viene de una distribución de valor extremo, calculan la probabilidad de que el individuo i escoja la ubicación k :

$$P_{ik} = \frac{\exp(\lambda_k + \pi_{ik})}{\sum_j \exp(\lambda_k + \pi_{ik})}$$

Resultados: probabilidad de cambiar de ubicación

La probabilidad media de mudarse es de 4.8%. Al...

- Tener contactos muy cerca, se *reduce* en 0.9%
- Tener muchos contactos, *aumenta* en 0.1%
- Compartir lenguaje y ser long-term resident, *disminuye*
- Tener "amigos de amigos", se *reduce*

Table 1
Probability of changing residential location.

	Linear probability model			Logit		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of contacts						
0-10 min	-3.008*** (0.039)	-3.596*** (0.050)	-3.199*** (0.061)	-1.426*** (0.019)	-1.606*** (0.022)	-1.518*** (0.029)
10-20 min		-1.450*** (0.059)	-1.327*** (0.060)		-0.478*** (0.023)	-0.458*** (0.024)
20-30 min		-0.717*** (0.072)	-0.671*** (0.072)		-0.211*** (0.028)	-0.204*** (0.028)
30-40 min		-0.365*** (0.086)	-0.343*** (0.086)		-0.151** (0.034)	-0.148*** (0.034)
Share of 2nd-order contacts 0-10 min			-1.025*** (0.094)			-0.217*** (0.049)
Total number of contacts	0.010*** (0.001)	0.011*** (0.001)	0.011*** (0.001)	0.005*** (0.000)	0.005*** (0.000)	0.005*** (0.000)
Long-term resident	-0.815*** (0.022)	-0.798*** (0.022)	-0.795*** (0.022)	-0.323*** (0.009)	-0.321*** (0.009)	-0.321*** (0.009)
Speaks same language as majority	-0.508*** (0.076)	-0.388*** (0.077)	-0.379*** (0.077)	-0.123*** (0.027)	-0.083*** (0.027)	-0.081*** (0.027)
Language, age, gender	Yes	Yes	Yes	Yes	Yes	Yes
Location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.016	0.016	0.017	-	-	-
Pseudo R ²	-	-	-	0.180	0.180	0.181
N	2,136,093	2,136,093	2,136,093	2,136,093	2,136,093	2,136,093

Notes: Dependent variable is expressed as a percentage in the linear probability model. Location fixed effects defined at the postcode level in columns (1)-(3) and at the employment region level in columns (4)-(6). The pseudo R² in columns (4)-(6) is calculated following [McKelvey and Zavoina \(1975\)](#) and reflects the proportion of the variance of the dependent variable that is explained by the covariates. ***, **, and * indicate significance at the 0.1, 1, and 5 percent levels.

Resultados: probabilidad de elegir cierta ubicación dada decisión de migrar

- Se quitan las características individuales
 - Entran efectos fijos de ubicación
 - Fuente principal de variación es la individuo-ubicación
-
- Tener 10% más de contactos muy cerca la *aumenta* 1 pp
 - Considerando +3000 opciones, y contactos en 540
 - Redes más relevantes para dónde, que para irse o no
 - Mismo C.P. que trabajo y mismo idioma, la aumenta

Table 2
Residential location choice.

	Linear probability model					Cond. logit (6)
	(1)	(2)	(3)	(4)	(5)	
Dep. var.: Probability of choosing a location conditional on moving						
Share of contacts						
0-10 min	9.259*** (0.015)	8.855*** (0.015)	9.085*** (0.015)	9.084*** (0.014)	9.044*** (0.015)	6.263*** (0.027)
10-20 min	0.949*** (0.007)	0.909*** (0.007)				
20-30 min	0.049*** (0.005)	0.045*** (0.005)				
30-40 min	-0.038*** (0.004)	-0.038*** (0.004)				
Return migration		15.318*** (0.036)	15.386*** (0.036)	15.383*** (0.036)		2.765*** (0.026)
Individual × location controls	Yes	Yes	Yes	No	Yes	Yes
Individual controls × location f. e.	No	No	No	Yes	No	No
Location-specific return migration	No	No	No	No	Yes	No
Location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.024	0.030	0.030	0.030	0.030	-
Pseudo R ²	-	-	-	-	-	0.207
N	25,555,189	25,555,189	25,555,189	25,555,189	25,555,189	25,342,595

Notes: Dependent variable is expressed as a percentage in the linear probability model. Location fixed effects defined at the postcode level in columns (1)-(5) and at the employment region level in column (6). Return migration indicates the individual was a resident at the same location at an earlier time in 2012-2015. Individual × location controls are an indicator for the potential new location being in the same employment region as the current residence, an indicator for the individual's preferred language being the local majority language, an interaction between an indicator for the individual's preferred language being English and the local share of foreign immigrants, an interaction between an indicator for the individual having multiple cellphone numbers on the same bill and the local average household size, an interaction between six age-group indicators and local population density, an interaction between six age-group indicators and the local share of homeowners, and an interaction between six age-group indicators and the local tax burden. Individual controls × location fixed effects uses as individual controls age group and gender indicators. The pseudo R² in column (6) is calculated following McFadden (1973). ***, **, and * indicate significance at the 0.1, 1, and 5 percent levels.

Problemas de identificación

- Posibles confusores: *chain mobility* y *sorting*
 - Personas parecidas toman elecciones parecidas, sin conocerse
- Solución: EF a nivel origen-destino y control por *movers* del mismo origen sin conocerse
- Coeficiente refleja *solo* cercanía a contactos
- Efecto de sorting significativo
- Amigos que vivían recientemente en el mismo lugar, aumentan probabilidad x2

Table 3
Chain mobility.

	Linear probability model				
	(1)	(2)	(3)	(4)	(5)
Share of contacts 0-10 min	6.931*** (0.181)	8.518*** (0.015)	8.166*** (0.017)	8.702*** (0.015)	
& recent movers			-0.459** (0.051)		
& from same origin			9.107*** (0.094)		
Share non-contact movers from same origin		88.881*** (0.470)	88.109*** (0.470)		
Return migration	13.525*** (0.039)	15.275*** (0.036)	15.261*** (0.036)	16.640*** (0.036)	15.336*** (0.036)
Moving distance				-0.208*** (0.002)	-0.438*** (0.002)
Individual × location controls	Yes	Yes	Yes	Yes	Yes
Location fixed effects	No	Yes	Yes	Yes	Yes
Origin-destination pair fixed effects	Yes	No	No	No	No
R ²	0.147	0.031	0.031	0.030	0.017
N	25,555,189	25,555,189	25,555,189	25,555,189	25,555,189

Notes: Dependent variable is expressed as a percentage. Location fixed effects defined at the postcode level. Recent movers captures the additional effect of contacts who, in addition to residing in any postcodes that can be reached by car within 10 min, moved there between January 2013 and three months prior. From same origin further restricts these to those who moved from within 10 min driving distance of where the individual is also moving. Share non-contact movers from same origin considers those individuals who moved between January 2013 and three months prior from within 10 min driving distance of where the individual is also moving and who are not one of their contacts and then calculates what share of these chose a postcode that can be reached by car within 10 min. Individual × location controls as in Table 2. ***, **, and * indicate significance at the 0.1, 1, and 5 percent levels.

Mediciones alternativas de relevancia de redes

1. Capacidad de predicción del modelo

- Incorporar información de la red de *cada* individuo **mejora predicciones x2**

2. Proporción de costos de migración que implica estar lejos de familia y amigos

- El **50% de los costos de migrar** se puede atribuir a dejar a amigos y familia

3. WTP por cercanía a amigos/familia, usando *commuting costs*

- Vivir en un lugar donde + 30pp de amigos viven cerca **vale tanto como vivir en un lugar donde no hacen commuting**

Rol de la información

- Hay información relevante difícil de conseguir si no es por contactos.
 - Contactos pre-existentes en una ubicación atractiva "aumenta probabilidad de ir ahí "jalan", y si el lugar no es bueno los "mandan" para otros lugares
- Amigos cercanos: información + compañía
- Amigos de amigos: solo información

Table 5

The role of information.

	Linear probability model					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of contacts 0-10 min	6.620*** (0.016)	6.394*** (0.017)	7.198*** (0.019)	6.987*** (0.019)	5.083*** (0.009)	9.038*** (0.019)
Share of contacts 0-10 min × location fixed effect	13.601*** (0.032)	12.239*** (0.034)	14.577*** (0.039)	14.265*** (0.039)	6.949*** (0.091)	6.724*** (0.243)
× childcare slots						-1.603*** (0.014)
× recent crimes						0.042*** (0.001)
× cultural events						-238.758*** (1.623)
× housing turnover						
Share of contacts 0-10 min & central		2.992*** (0.052)				
& central × loc. fixed effect		6.655*** (0.115)				
& strong (duration)			-1.610*** (0.038)			
& strong (duration) × loc. f. e.			-5.967*** (0.081)			
& strong (frequency)				-0.570*** (0.039)		
& strong (frequency) × loc. f. e.				-4.836*** (0.081)		
Share of 2nd-order contacts 0-10 min					4.420*** (0.037)	
Share of 2nd-order contacts 0-10 min × location fixed effect					8.991*** (0.066)	
Share non-contact movers from same origin	80.333*** (0.475)	88.199*** (0.469)	88.142*** (0.469)	89.057*** (0.469)	74.755*** (0.478)	78.734*** (0.478)
Return migration	14.603*** (0.036)	14.668*** (0.036)	14.626*** (0.036)	14.656*** (0.036)	14.540*** (0.036)	15.091*** (0.036)
Individual × location controls	Yes	Yes	Yes	Yes	Yes	Yes
Location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.038	0.038	0.038	0.037	0.039	0.033
N	25,555,189	25,555,189	25,555,189	25,555,189	25,555,189	25,538,167

Notes: Dependent variable is expressed as a percentage. Location fixed effects defined at the postcode level. Central contacts are those in the top 10% in terms of eigenvector centrality in the overall Swiss network. Strong (duration/frequency) contacts are those in the top 10% in terms of total call duration/frequency in the individual's contact network. Location fixed effects defined at the postcode level. All local characteristics are centred at zero. Share non-contact movers from same origin considers those individuals who moved between January 2013 and three months prior from within 10 min driving distance of where the individual is also moving and who are not one of their contacts and then calculates what share of these those a postcode that can be reached by car within 10 min. Individual × location controls as in Table 2. ***, **, * and * indicate significance at the 0.1, 1, and 5 percent

Rol de la información

- Heterogeneidad de tipo de información + demográficos captura canales de información
- Mecanismos:
 1. **Ranking**
 - *Childcare*
 - Crimen
 2. **Amenidades**
 - Eventos culturales
 3. **Reducción fricción asimetrías información**
 - Rotación de vivienda

Importancia de familias y amigos, por edad

- Autores infieren qué contactos son amigos o familia
- En promedio, decisión pesa más por concentración de amigos que de familia
- Efecto pierde fuerza a mayor edad

Table 6
Relative importance of friends and family by age.

	All		Age 25–34		Age 35–54		Age ≥ 55	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share family contacts	2.694***	2.394***	2.378***	2.176***	3.171***	2.485***	3.791***	2.854***
within 0–10 min	(0.019)	(0.019)	(0.027)	(0.028)	(0.045)	(0.045)	(0.075)	(0.074)
Share friend contacts	7.803***	5.695***	8.101***	5.735***	7.565***	5.222***	6.727***	4.734***
within 0–10 min	(0.034)	(0.035)	(0.051)	(0.0522)	(0.078)	(0.079)	(0.129)	(0.126)
Share family contacts 0–10 min		5.560***		4.190***		6.456***		6.077***
× location fixed effects		(0.046)		(0.107)		(0.086)		(0.087)
Share friend contacts 0–10 min		15.540***		16.054***		17.893***		12.420***
× location fixed effects		(0.077)		(0.580)		(0.144)		(0.148)
Share non-contact movers	70.514***	68.555***	74.100***	73.369***	66.314***	58.932***	65.190***	48.915***
from same origin	(0.856)	(0.849)	(1.322)	(1.311)	(1.959)	(1.930)	(3.131)	(3.028)
Return migration	15.684***	14.411***	13.993***	12.877***	13.628***	12.645***	13.608***	12.522***
	(0.069)	(0.068)	(0.093)	(0.092)	(0.194)	(0.191)	(0.350)	(0.339)
Individual × location controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.041	0.056	0.040	0.055	0.037	0.067	0.041	0.104
N	6,235,721	6,235,721	2,793,125	2,793,125	1,067,625	1,067,625	383,510	383,510

Notes: All columns estimated using a linear probability model, with the dependent variable expressed as a percentage. Location fixed effects defined at the postcode level. Individual × location controls as in Table 2. ***, **, and * indicate significance at the 0.1, 1, and 5 percent levels.

Conclusiones

- **Primera etapa:** los que tienen a sus **amigos** concentrados muy cerca de ellos tienen menos probabilidad de moverse.
- **Segunda etapa:** Condicional a haber decidido mudarse, la presencia de amigos en una ubicación dada aumenta la probabilidad de ir ahí.
 - Efectos muy locales (hasta a 10 minutos de distancia)
- **Mecanismos.** Los amigos dan:
 - Información difícil de conseguir, lo cual permite **rankear** ubicaciones
 - Compañía, con la que se pueden disfrutar las amenidades locales
 - Información de mercados sujetos a fricciones, como el de bienes inmuebles
- Los amigos de los amigos están en un canal meramente informativo, y la familia gana importancia con la edad

Estar cerca de amigos y familia da utilidad, lo cual crea un trade-off entre proximidad a ellos a otras ventajas locales.