

Análisis y Proyecciones del RAIS y el RPM a partir de las Transiciones

Eduardo Lora y Andrés Gómez-Liévano

Marzo de 2017

Seminario de Colpensiones

Contenido

- Base de datos y definiciones
- Caracterización de los afiliados (resumen)
- Metodología de las matrices de transición
- Análisis de la estabilidad de las probabilidades de transición
- Método de construcción de cadenas de Markov
- Método de proyección de afiliados y semanas cotizadas

Base de datos y definiciones

PILA 2008 a 2014

Datos mes a mes por individuo:

Cotizó a pensiones?

Sí:

dependiente o independiente (tomamos ambos)

en qué entidad? (RAIS, RPM, otro)

cuántos días?

qué ingreso base de cotización?

qué sector de actividad económica CIIU (sólo dependientes)?

qué municipio?

No: no aparece ese mes

Sólo en 2013 y 2014: edad y sexo.

Base de datos y definiciones

Cotizante: quien haya hecho al menos una cotización en el año en cuestión, bien sea en RAIS o el RPM.

Cobertura: número de cotizantes en el año en cuestión / población nacional del grupo de edad.

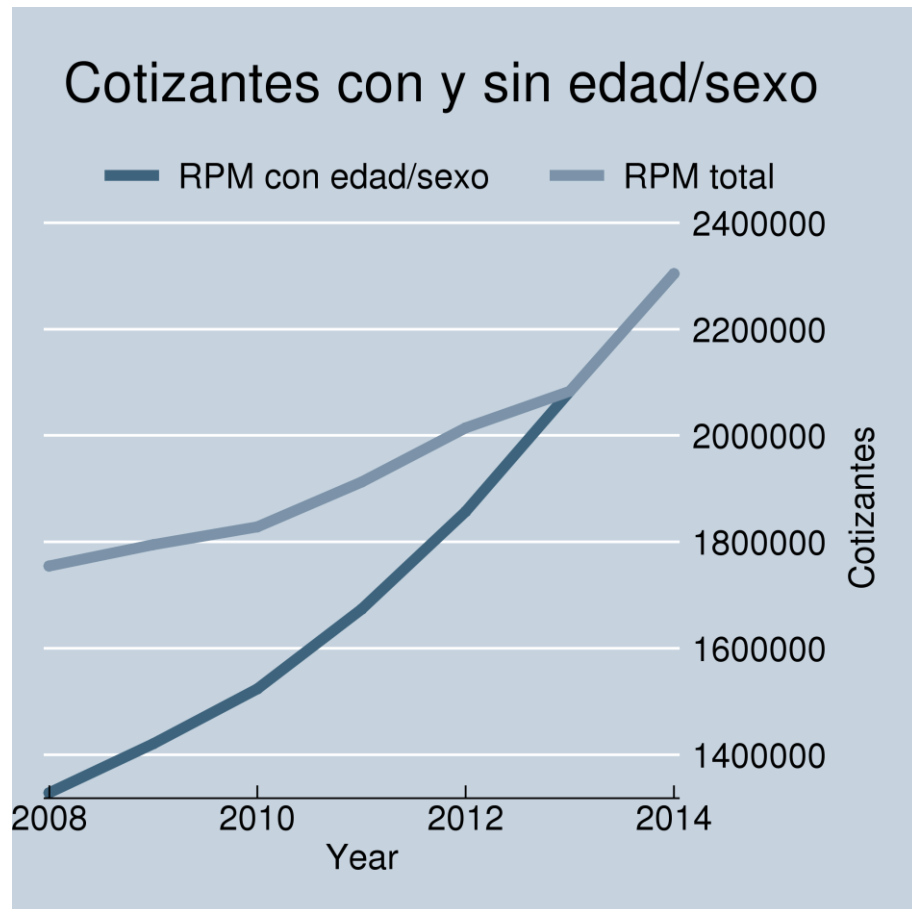
Fidelidad (o densidad de cotizaciones): número de semanas de cotización

Salario base de cotización: equivalente mensual (en los días cotizados) del salario base de cotización reportado.

Traslados entre regímenes: en la última cotización en el año $t+1$ el individuo aparece en un régimen distinto al de la última cotización en el año t .

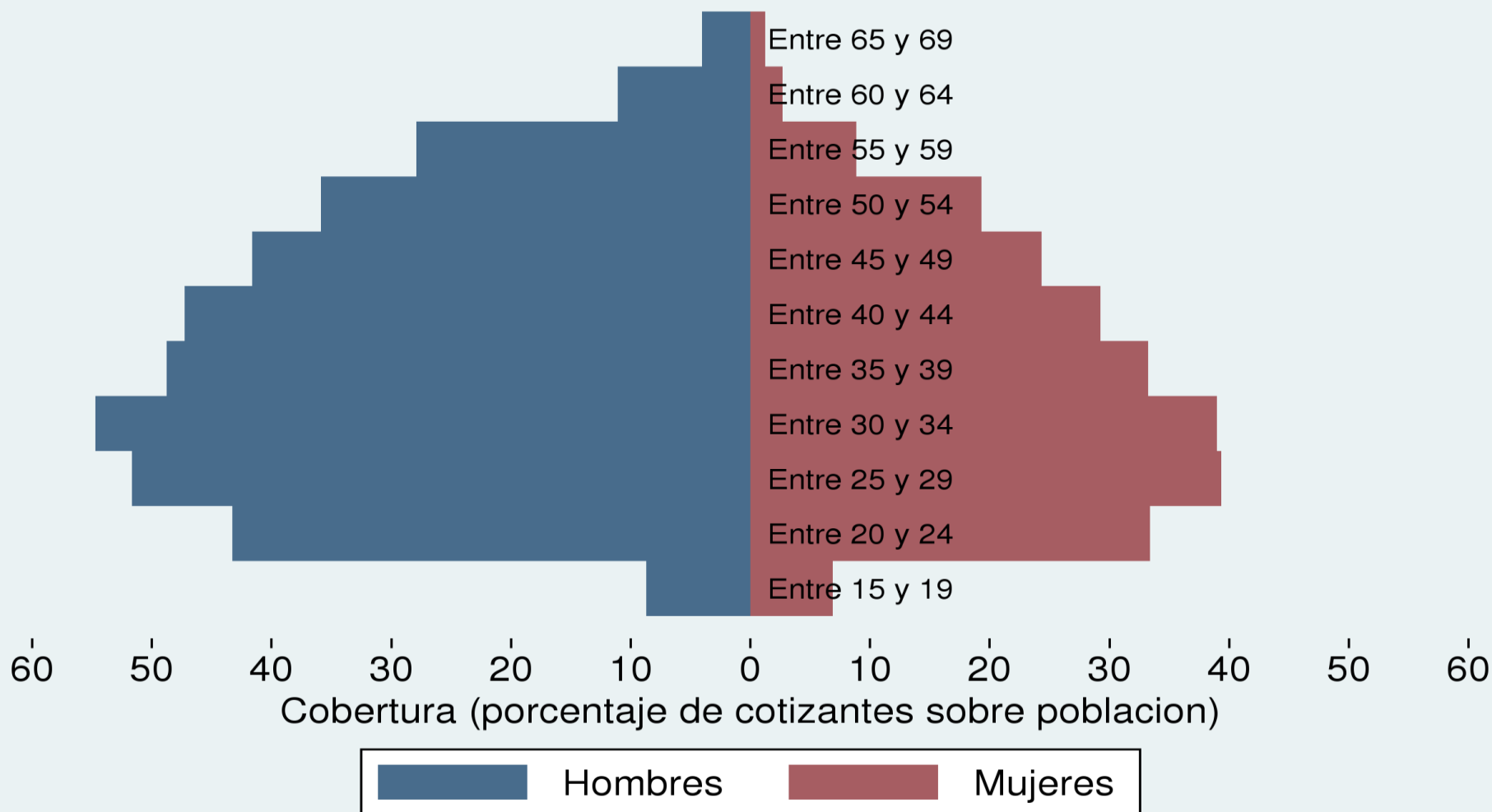
Tasa de traslados son los traslados en $t+1$ dividido por los cotizantes en t .

Fuerte aumento de cotizantes en RAIS RPM

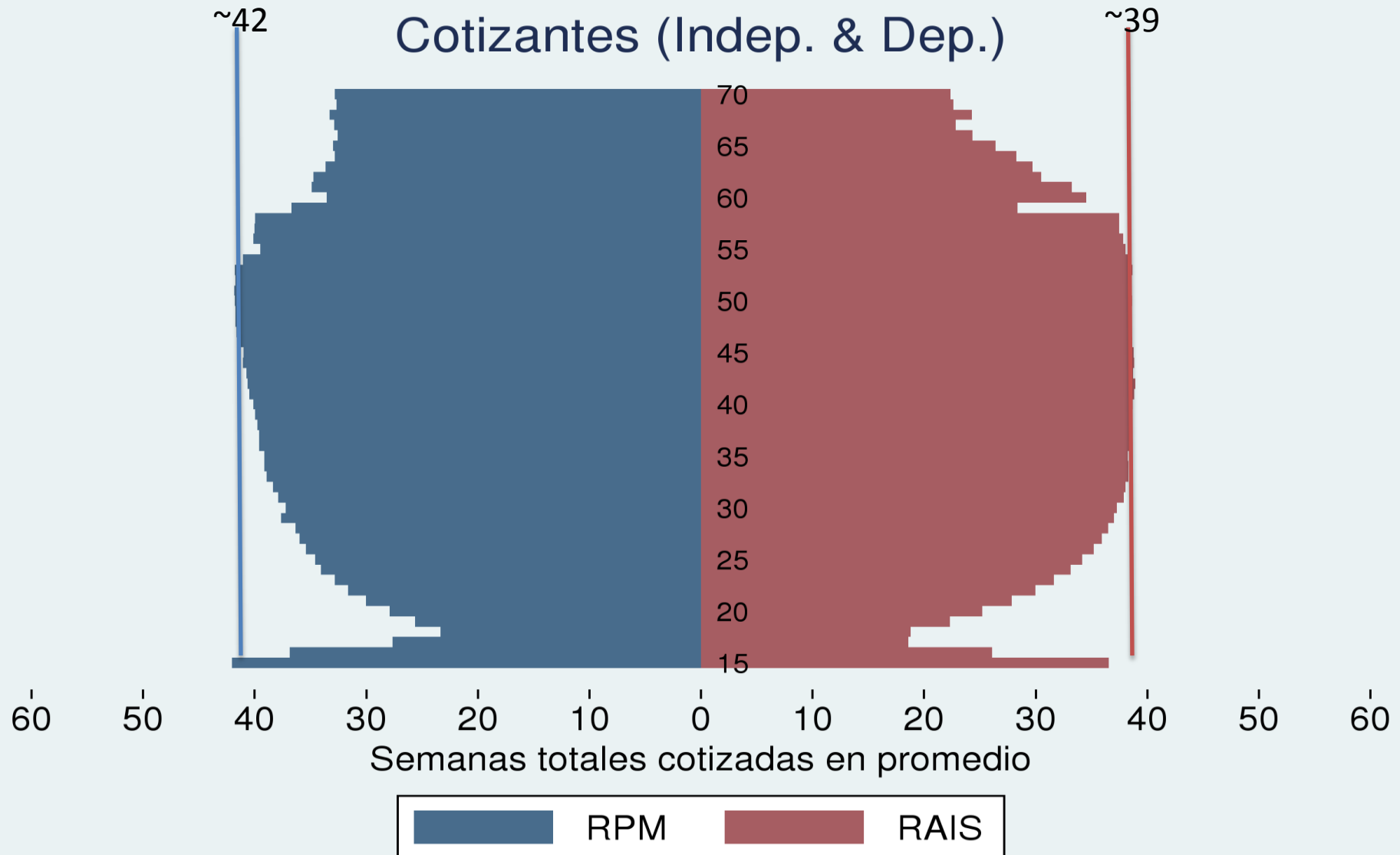


Mayor cobertura masculina y picos a los 25-34 años

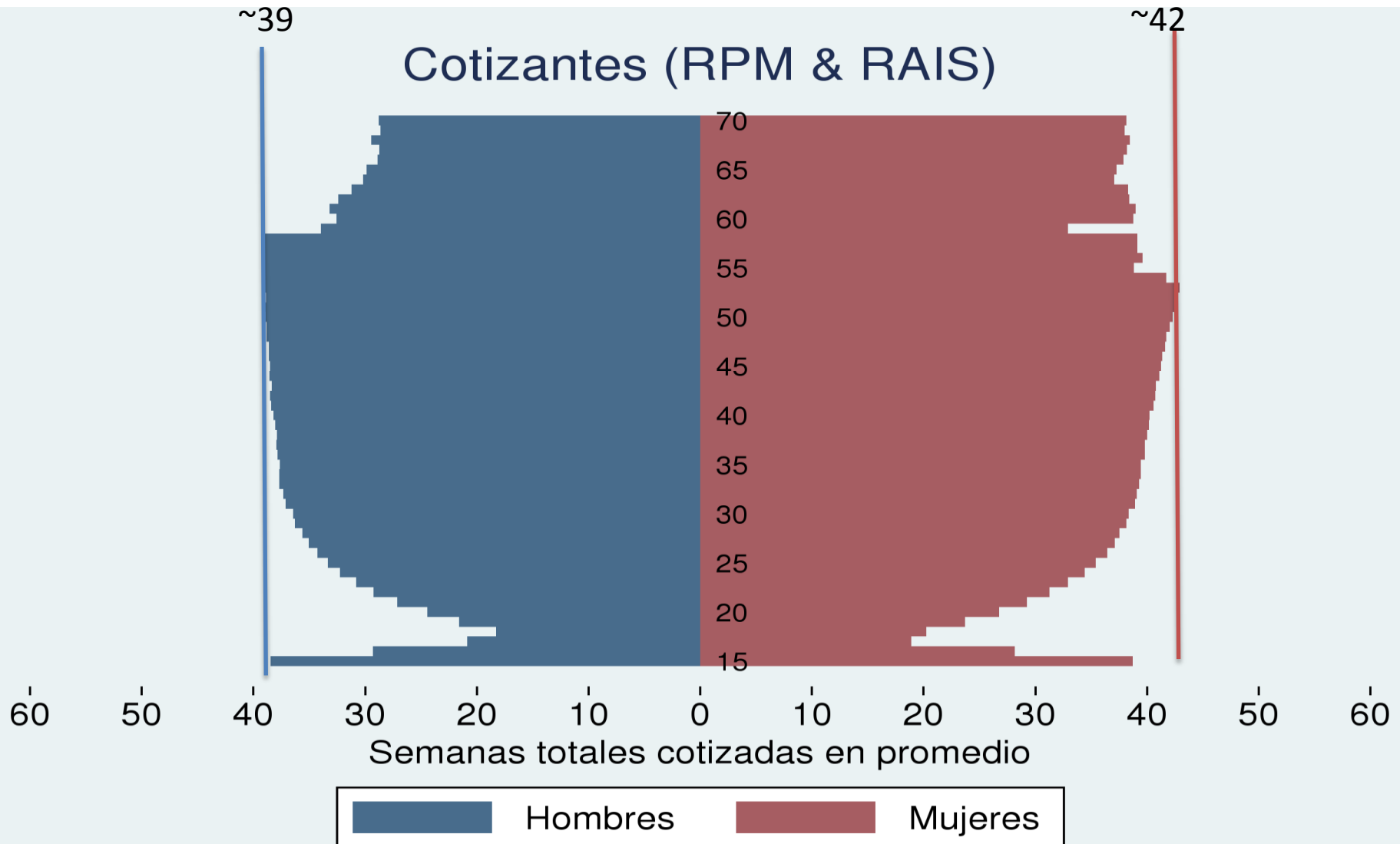
Tasas de Cobertura por Quinquenios



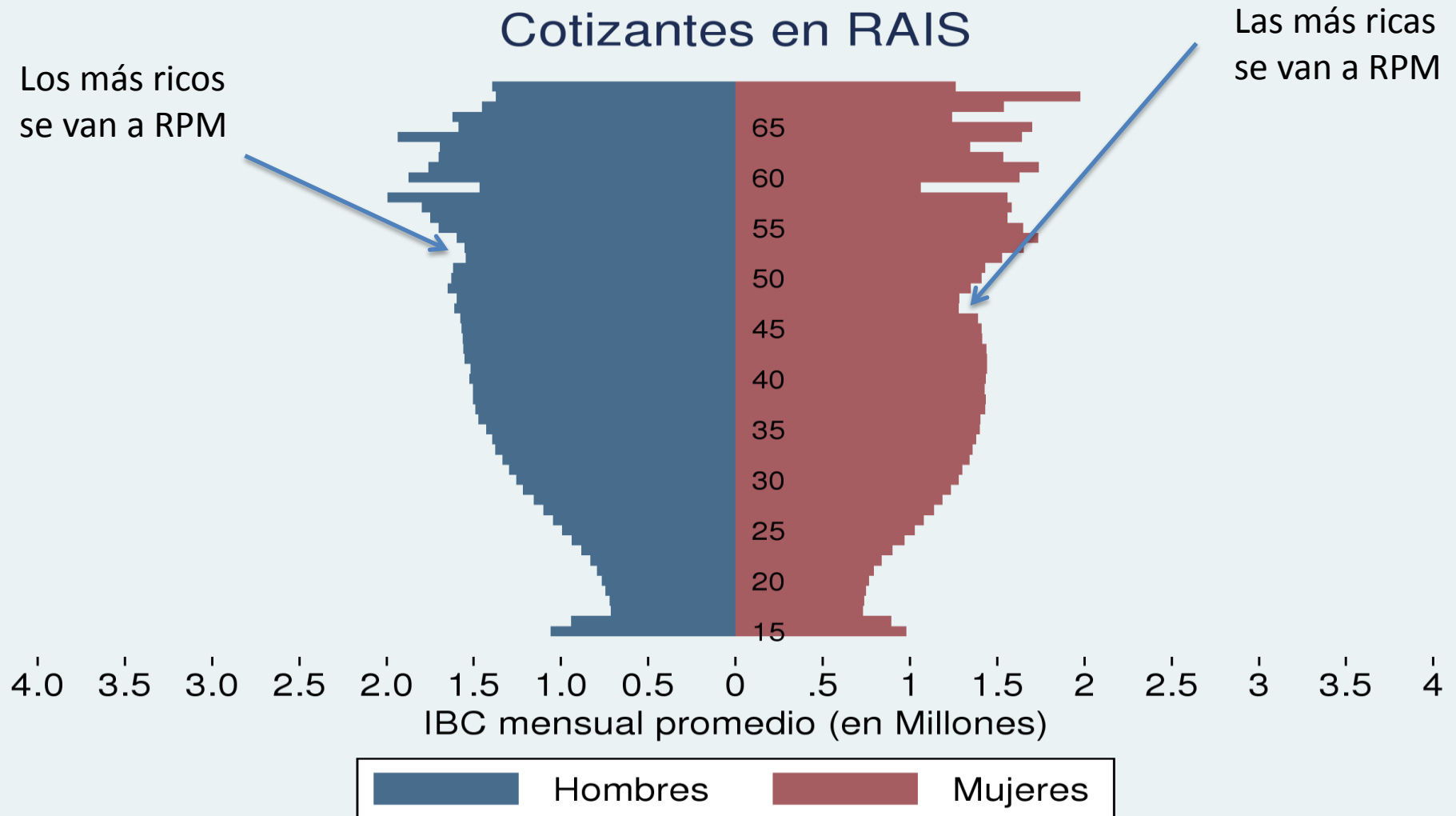
Fidelidad creciente con la edad en ambos sistemas



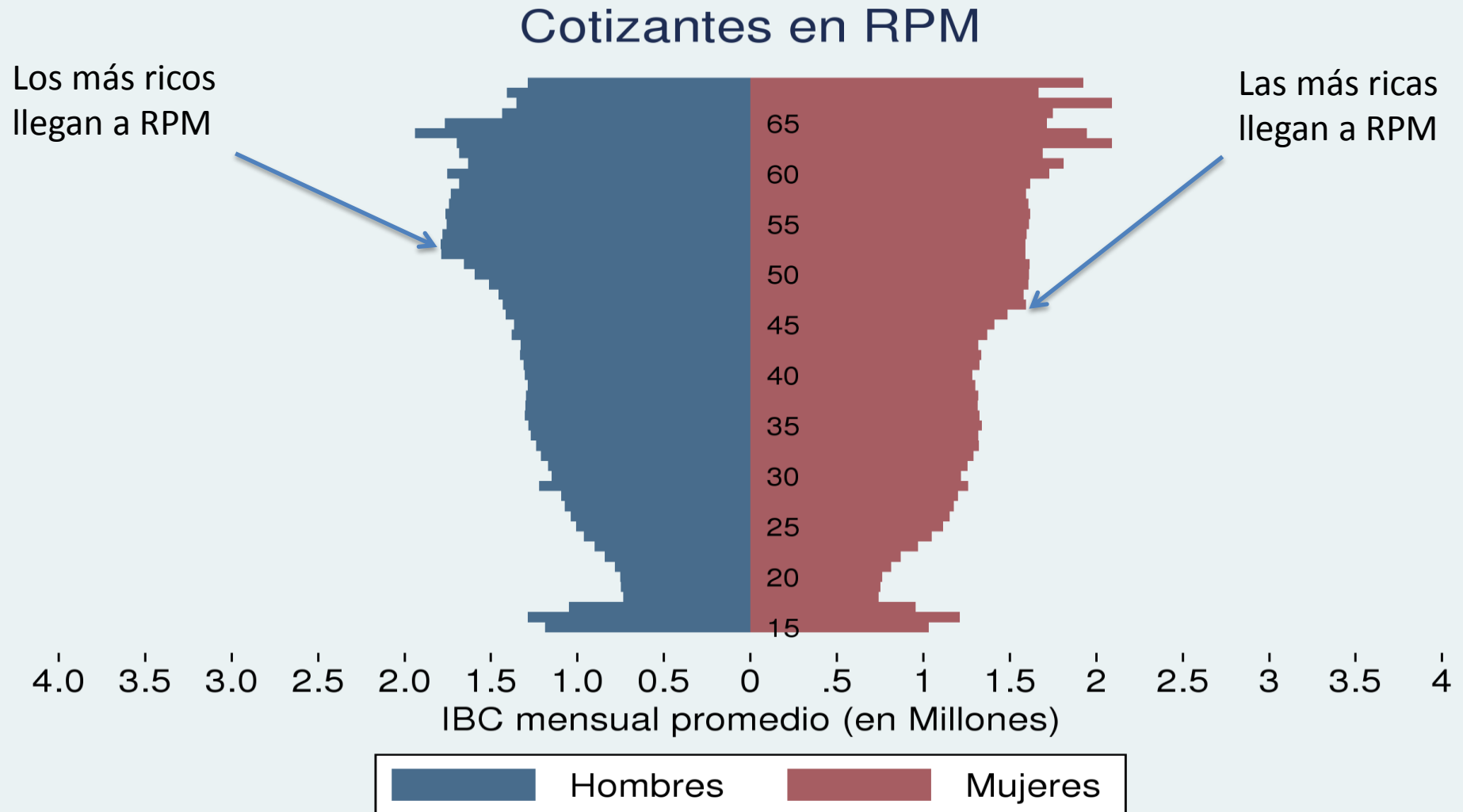
Fidelidad ligeramente mayor en mujeres



Ingresos base de cotización en forma de copa, pero....

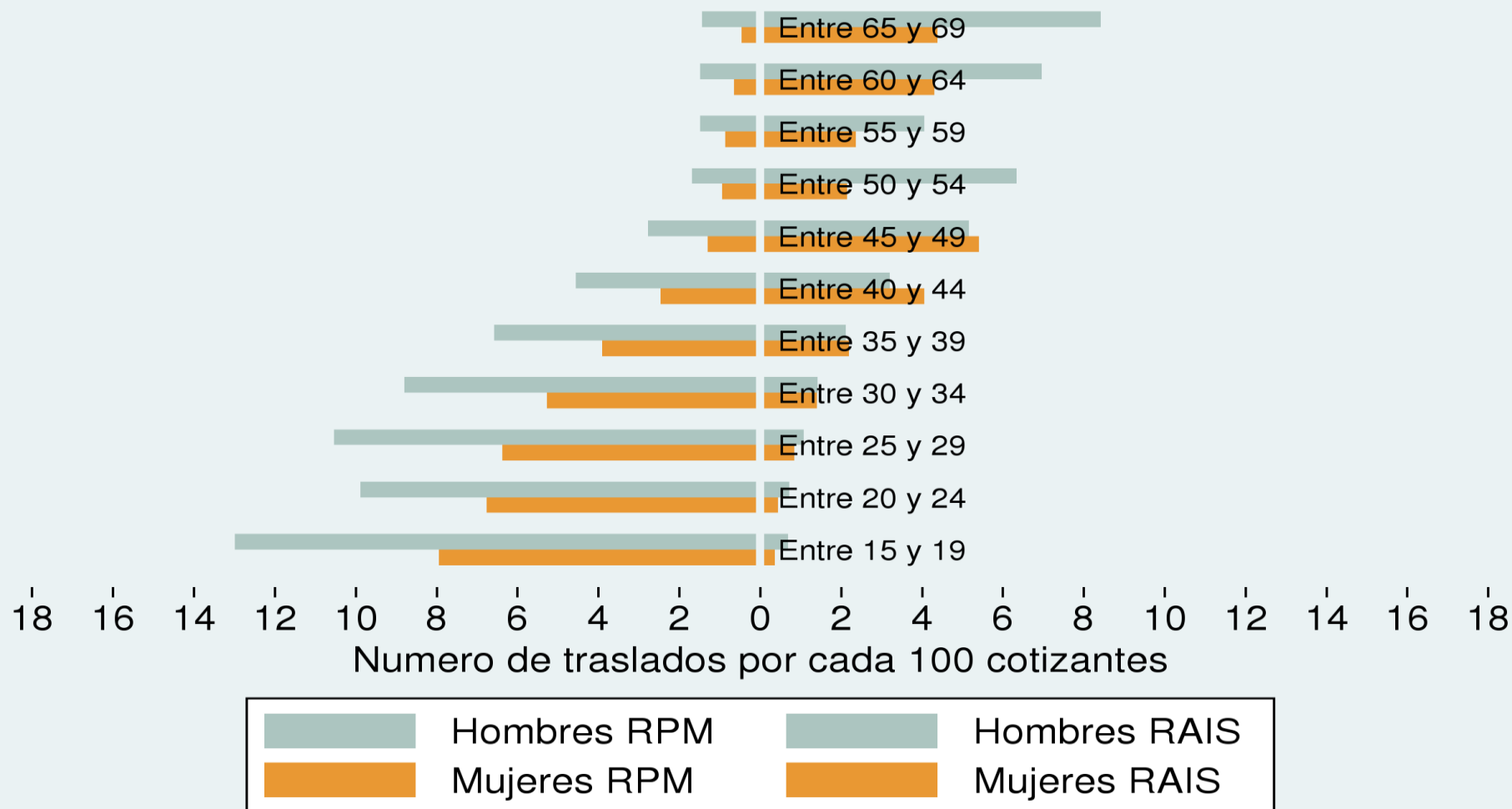


Ingresos base de cotización en forma de copa, pero....



Los jóvenes se trasladan de RPM a RAIS, y los viejos de RAIS a RPM

Tasas de Traslado por Quinquenios (izq=RPM; der=RAIS)



Dimensiones de las matrices de transición

Para cada grupo de edad simple y sexo se tiene el estado inicial y final según:

- Régimen de afiliación:
 - 0. No afiliado
 - 1. Afiliado a RPM
 - 2. Afiliado a RAIS
- Grupos de ingreso básico de cotización:
 - 1. Menos de 1 salario mínimo legal
 - 2. De 1 a 1.2 salarios mínimos legales
 - 3. De 1.2 a 2.5 salarios mínimos legales
 - 4. De 2.5 a 5 salarios mínimos legales
 - 5. Más de 5 salarios mínimos legales
- Grupos de fidelidad:
 - 1. Entre 1 y 13 semanas
 - 2. Entre 14 y 26 semanas
 - 3. Entre 27 y 39 semanas
 - 4. Entre 40 y 46 semanas
 - 5. Entre 47 y 52 semanas

Número de matrices:
 $(62 - 16) + (57 - 16) = 87$

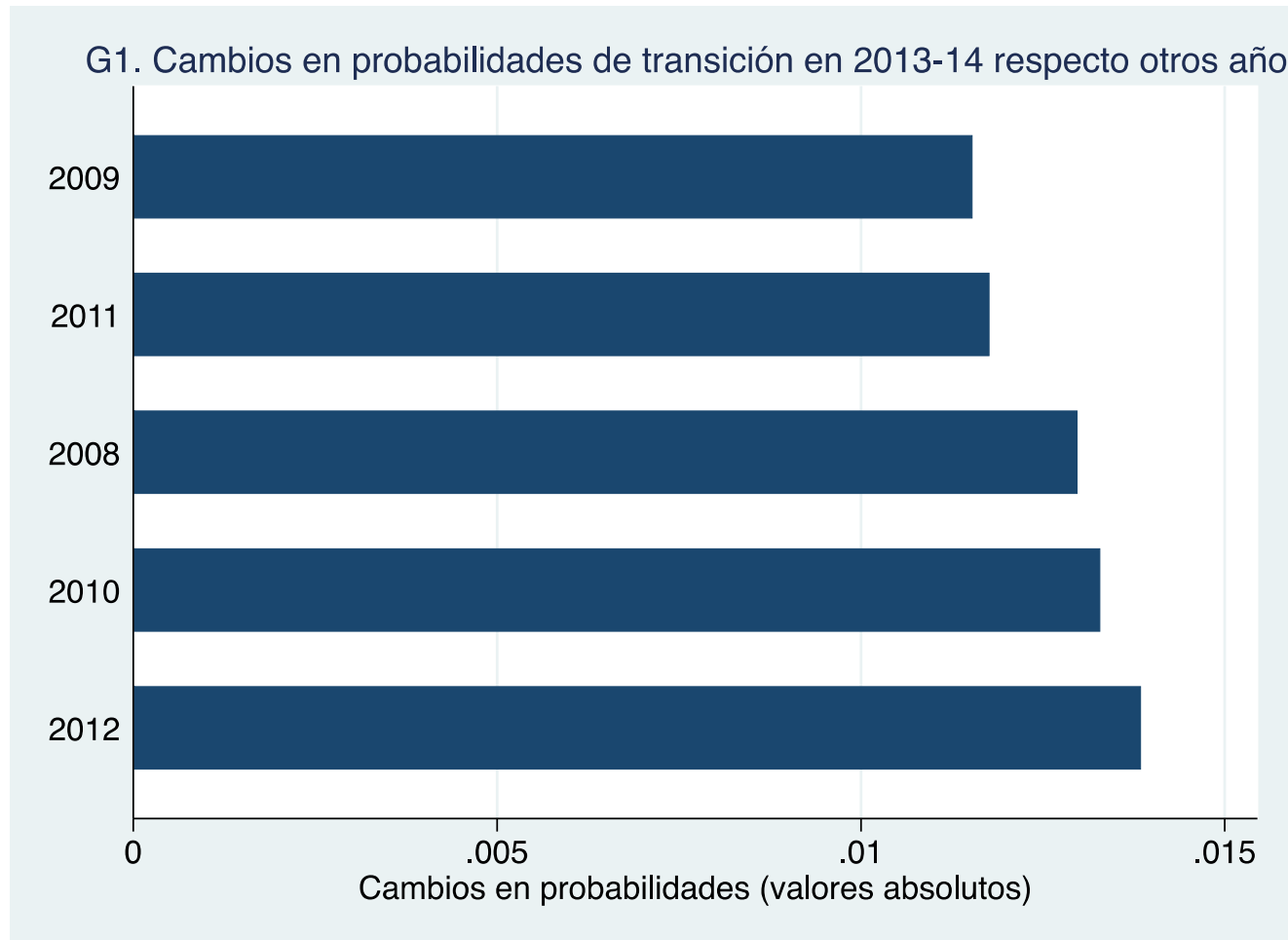
Celdas:
 $((2 \times 5 \times 5) + 1)^2 = 2551$

Mujeres de 44 años en 2013		RPM en 2014					RAIS en 2014					No afiliados en 2014	Totales provenientes del anterior grupo de edad
		< 1 sml	1 - 1.2 sml	1.2 - 2.5 sml	2.5 - 5 sml	> 5 sml	< 1 sml	1 - 1.2 sml	1.2 - 2.5 sml	2.5 - 5 sml	> 5 sml		
RPM en 2013	< 1 sml	1743	659	177	13	1	28	29	12	0	0	412	3074
	1 - 1.2 sml	2146	3340	1073	45	1	50	108	35	1	0	961	7760
	1.2 - 2.5 sml	265	1077	6248	485	30	12	43	122	4	0	593	8879
	2.5 - 5 sml	19	44	434	2887	320	0	3	5	28	6	132	3878
	> 5 sml	3	8	21	251	2128	1	1	2	3	17	69	2504
RAIS en 2013	< 1 sml	138	84	34	5	0	2504	1124	375	27	11	885	5187
	1 - 1.2 sml	164	408	150	8	2	3244	6138	2034	90	20	1953	14211
	1.2 - 2.5 sml	31	183	885	81	11	451	2189	9690	734	59	1207	15521
	2.5 - 5 sml	0	12	85	467	54	19	83	681	3650	444	315	5810
	> 5 sml	2	4	7	62	545	6	19	74	412	3546	142	4819
Total afiliados 2013		4511	5819	9114	4304	3092	6315	9737	#####	4949	4103	6669	71643
No afiliados 2013		302	3690	1043	148	61	355	4509	1629	338	128	207404	219607

Probabilidades de transición de las mujeres de 44 años en 2013		RPM en 2014					RAIS en 2014					No afiliados en 2014	Totales provenientes del anterior grupo de edad
		< 1 sml	1 - 1	1.2 -	2.5 -	> 5 s	< 1 s	1 - 1	1.2 -	2.5 -	> 5 s		
RPM en 2013	< 1 sml	0.57	0.21	0.06	0	0	0.01	0.01	0	0	0	0.134	1
	1 - 1.2 sml	0.28	0.43	0.14	0.01	0	0.01	0.01	0	0	0	0.124	1
	1.2 - 2.5 sml	0.03	0.12	0.7	0.05	0	0	0	0.01	0	0	0.067	1
	2.5 - 5 sml	0	0.01	0.11	0.74	0.08	0	0	0	0.01	0	0.034	1
	> 5 sml	0	0	0.01	0.1	0.85	0	0	0	0	0.01	0.028	1
RAIS en 2013	< 1 sml	0.03	0.02	0.01	0	0	0.48	0.22	0.07	0.01	0	0.171	1
	1 - 1.2 sml	0.01	0.03	0.01	0	0	0.23	0.43	0.14	0.01	0	0.137	1
	1.2 - 2.5 sml	0	0.01	0.06	0.01	0	0.03	0.14	0.62	0.05	0	0.078	1
	2.5 - 5 sml	0	0	0.01	0.08	0.01	0	0.01	0.12	0.63	0.08	0.054	1
	> 5 sml	0	0	0	0.01	0.11	0	0	0.02	0.09	0.74	0.029	1
Total afiliados 2013		0.06	0.08	0.13	0.06	0.04	0.09	0.14	0.18	0.07	0.06	0.093	1
No afiliados en 2013		0	0.02	0	0	0	0	0.02	0.01	0	0	0.944	1

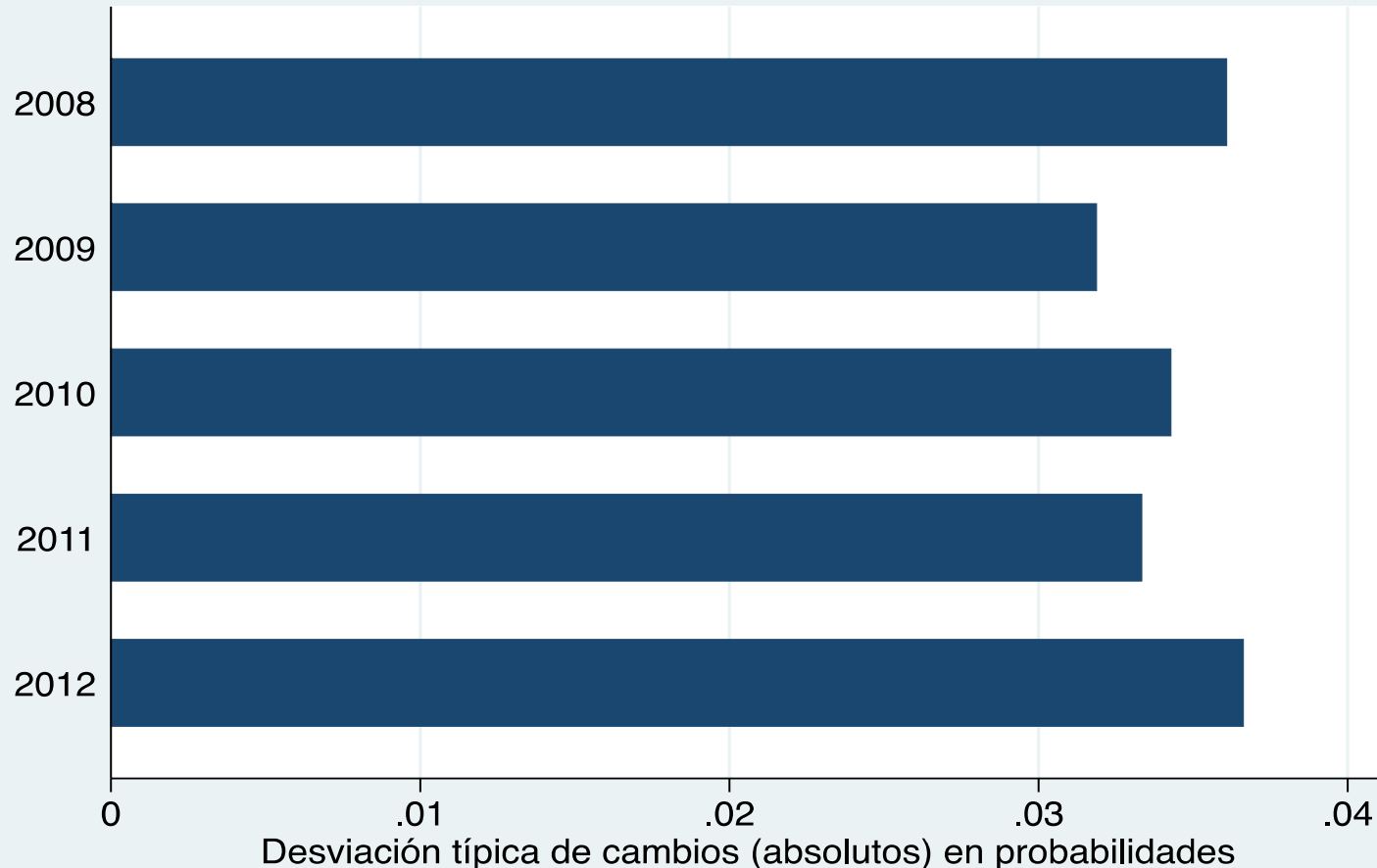
Mujeres de 44 años en 2013	RPM en 2014	RAIS en 2014	Total afiliados en 2014	Total no afiliados en 2014	Totales
RPM en 2013	0.90	0.02	0.92	0.08	1.00
RAIS en 2013	0.08	0.83	0.90	0.10	1.00
Total afiliados 2013	0.37	0.53	0.91	0.09	1.00
No afiliados 2013	0.02	0.03	0.05	0.95	1.00
Totales	0.11	0.15	0.26	0.74	1.00

Las transiciones cambian bastante



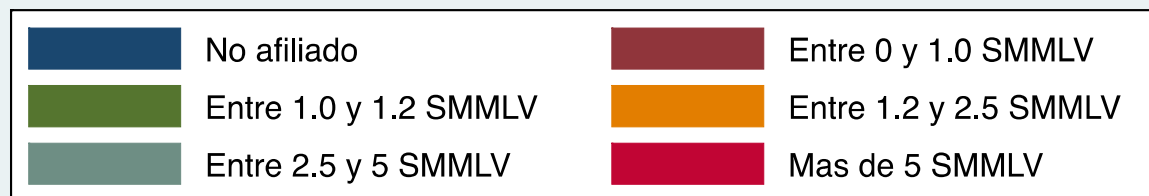
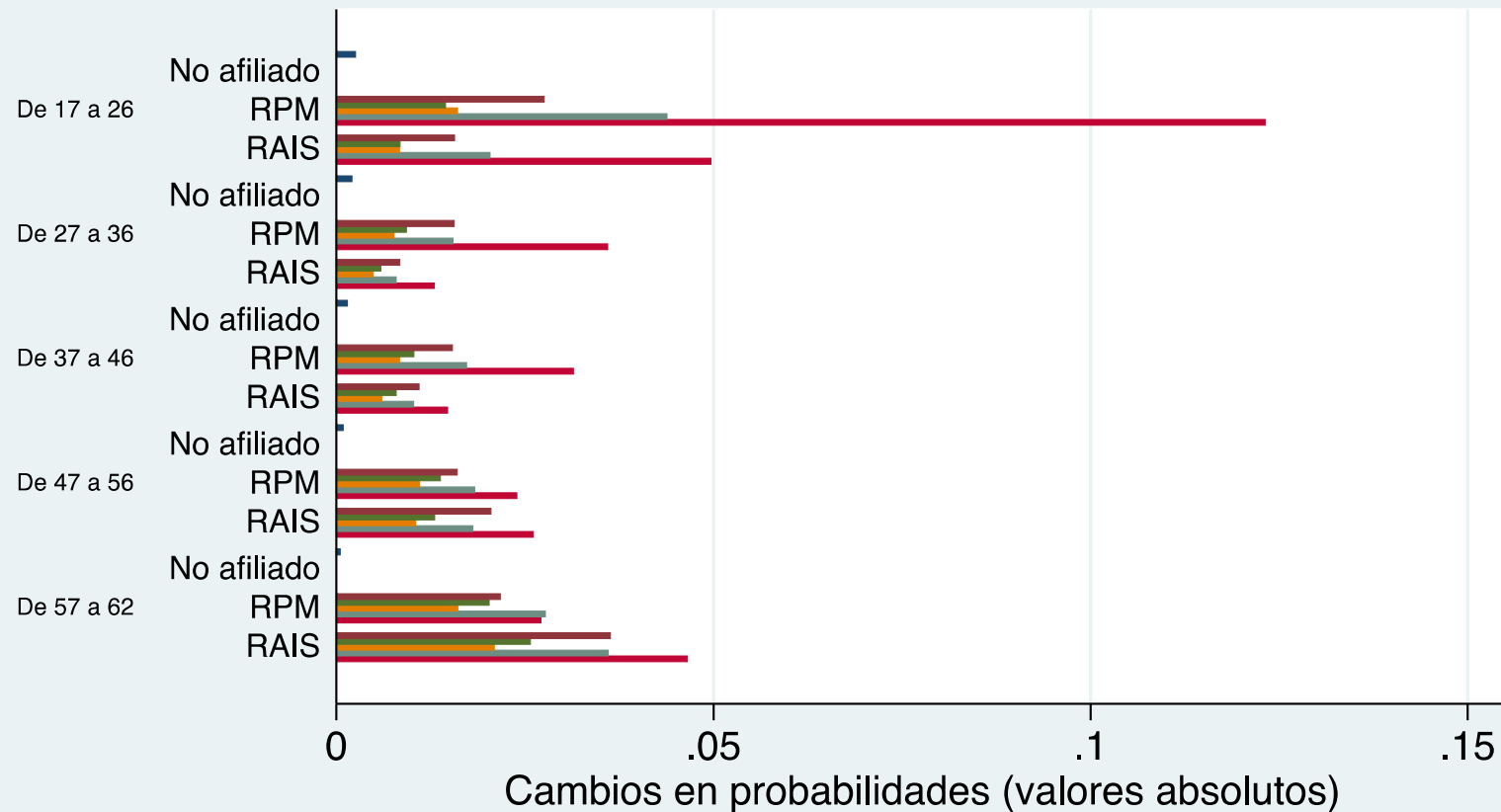
...y en magnitudes variables

G2. Desviación típica de los cambios en probabilidades 2013 vs año



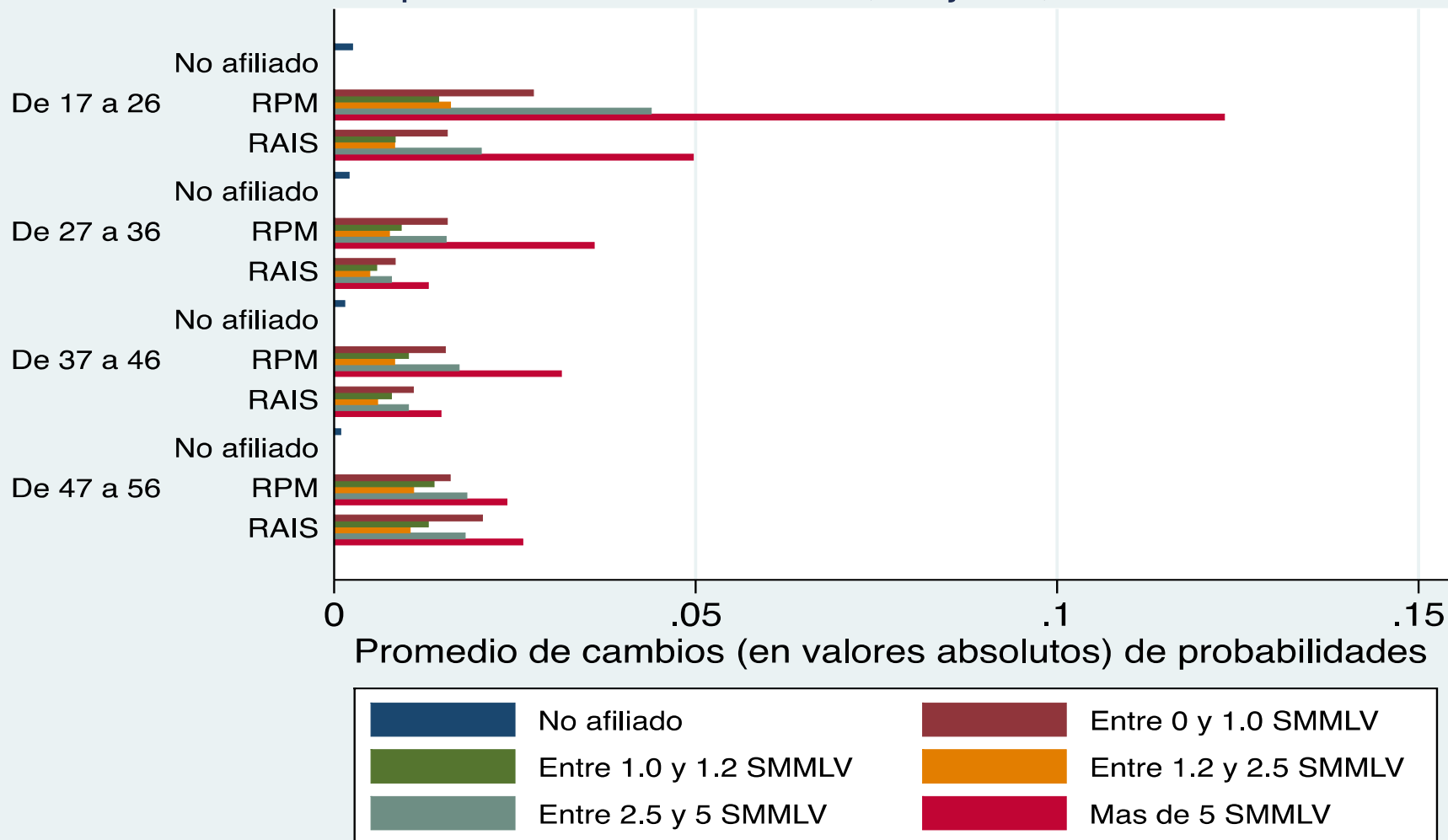
Los jóvenes son inestables (sorpresa?)

G3. Cambios en transiciones hombres 2013 vs 2008-2012



Las jóvenes también

G5. Cambios probabilidades transición, mujeres, 2013 vs 2008-12



...pero ni jóvenes ni viejos son racionales

Cuadro 6. Resultados de regresiones del efecto del rendimiento de las AFPs sobre las probabilidades de traslados (hombres y mujeres)

Traslados de RPM a RAIS				
Grupo de edad	Coficiente	Estadístico T	Número de observaciones	R cuadrado ajustado
17-26	0.0139544	0.42	4920	-0.0002
27-36	0.0797649	4.65	4920	0.0042
37-46	0.0323568	2.38	4920	0.001
47-56	0.0095412	1.28	4920	0.0001
57-62	0.0091551	0.8	1722	-0.0002
Todos	0.032023	3.4	21156	0.0005

Traslados de RAIS a RPM				
Grupo de edad	Coficiente	Estadístico T	Número de observaciones	R cuadrado
17-26	0.0330894	2.03	4920	0.0006
27-36	0.0391731	10.11	4920	0.0202
37-46	0.0621727	8.38	4920	0.0139
47-56	0.1062116	8.25	4920	0.0135
57-62	0.3189489	7.19	1722	0.0287
Todos	0.0789562	12.49	21156	0.0073

Nota: la variable explicativa es el rendimiento promedio de los últimos cinco años. La variable dependiente es la probabilidad de pasar de un régimen a otro por parte de los individuos de un grupo de edad simple y sexo, que comparten el mismo rango de ingreso básico de cotización y el mismo grupo de semanas de cotización. Tanto a la variable explicativa como a la dependiente se les ha restado su valor de 2014.

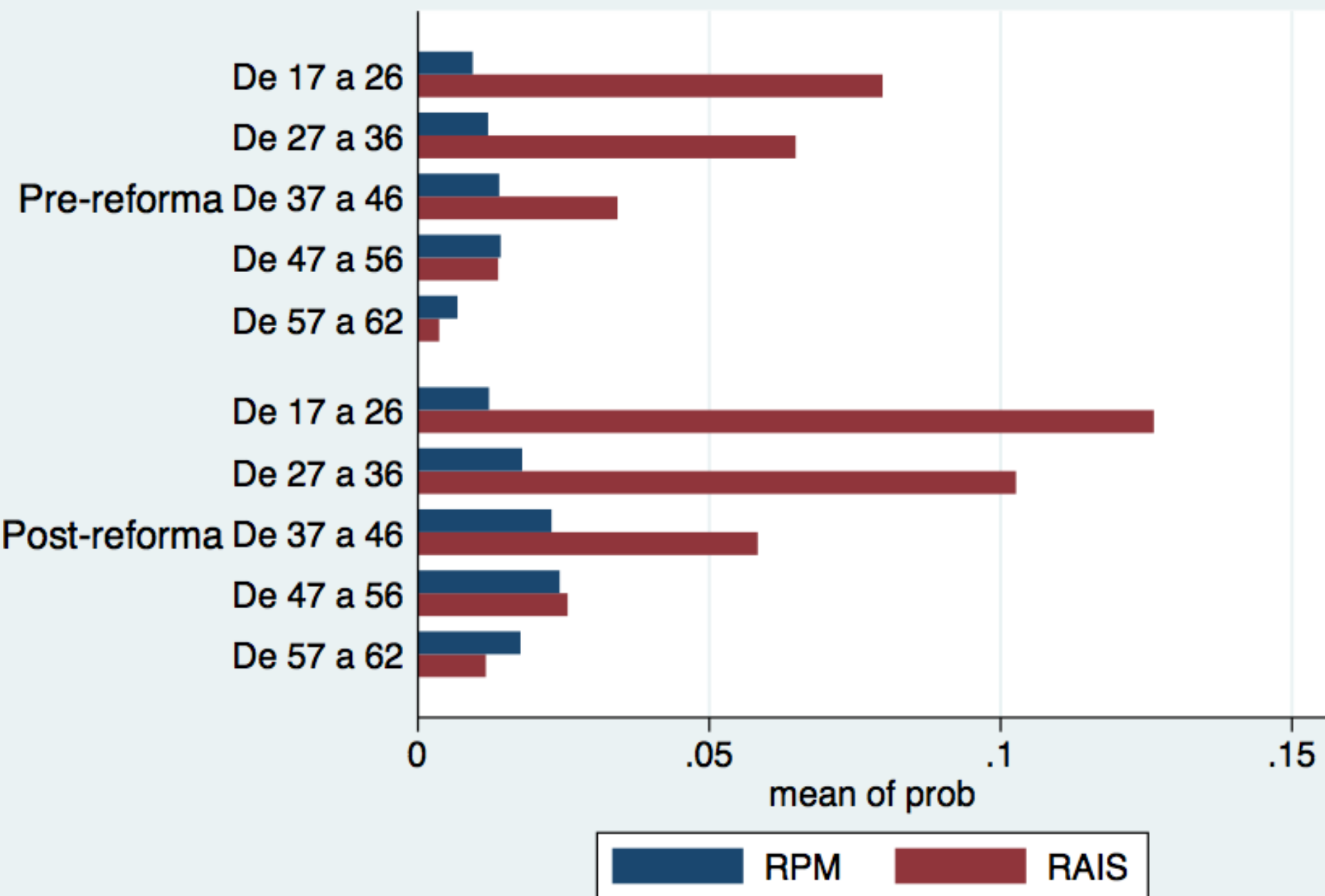
El único cambio consistente ha sido de no afiliado a afiliado

Probabilidades de transición de los no afiliados

	No afiliado	RPM	RAIS
TOTAL			
Pre-reforma	95.5%	1.2%	3.2%
Post-reforma	90.7%	2.0%	7.3%
HOMBRES			
Pre-reforma	94.6%	1.7%	3.8%
Post-reforma	88.7%	2.5%	8.8%
MUJERES			
Pre-reforma	96.6%	0.7%	2.6%
Post-reforma	93.0%	1.4%	5.7%

Nota: promedio no ponderado de todas las edades simples. Pre-reforma es 2008 a 2012, post-reforma es 2013 y 2014.

Probabilidades de transición de no afiliados a RPM o RAIS



Algunas implicaciones para hacer proyecciones

- No se pueden ignorar las transiciones (proyecciones con grupos predefinidos)
- Pero tampoco se puede suponer que resultan de decisiones racionales (modelos de optimización)
- Hay que reconocer las limitaciones de las probabilidades de transición que tenemos:
 - Información incompleta de sexo y edad pre 2013
 - Las proyecciones deben entenderse como “expectativas dadas las probabilidades actuales”
 - Especialmente relevante para 2014 vs años anteriores
 - Las probabilidades de transición dependen de las historias laborales de los individuos.

Cómo se proyecta en cadenas de Markov simples

Para todos los individuos de un grupo de edad simple y sexo:

- Se clasifican en los 51 estados iniciales (columna inicial)
- Se aplican las probabilidades por filas para saber en qué estado final quedan
- Sumando se obtienen los 51 estados finales
- Se corrige por mortalidad la celda de no afiliados (con información del DANE).
- Así pasan a ser el grupo de edad siguiente, y así sucesivamente...

Mujeres de 44 años en 2013		Estados de origen (2013)	Estados de destino (2014)										No afiliados	Suma de probabilid ades
			RPM					RAIS						
			< 1 sml	1 - 1.2 sml	1.2 - 2.5 sml	2.5 - 5 sml	> 5 sml	< 1 sml	1 - 1.2 sml	1.2 - 2.5 sml	2.5 - 5 sml	> 5 sml		
RPM	< 1 sml	3074	0.57	0.21	0.06	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.13	1
	1 - 1.2 sml	7760	0.28	0.43	0.14	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.12	1
	1.2 - 2.5 sml	8879	0.03	0.12	0.70	0.05	0.00	0.00	0.00	0.01	0.00	0.00	0.07	1
	2.5 - 5 sml	3878	0.00	0.01	0.11	0.74	0.08	0.00	0.00	0.00	0.01	0.00	0.03	1
	> 5 sml	2504	0.00	0.00	0.01	0.10	0.85	0.00	0.00	0.00	0.00	0.01	0.03	1
RAIS	< 1 sml	5187	0.03	0.02	0.01	0.00	0.00	0.48	0.22	0.07	0.01	0.00	0.17	1
	1 - 1.2 sml	14211	0.01	0.03	0.01	0.00	0.00	0.23	0.43	0.14	0.01	0.00	0.14	1
	1.2 - 2.5 sml	15521	0.00	0.01	0.06	0.01	0.00	0.03	0.14	0.62	0.05	0.00	0.08	1
	2.5 - 5 sml	5810	0.00	0.00	0.01	0.08	0.01	0.00	0.01	0.12	0.63	0.08	0.05	1
	> 5 sml	4819	0.00	0.00	0.00	0.01	0.11	0.00	0.00	0.02	0.09	0.74	0.03	1
No afiliados en 2013		207404	0.00	0.02	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.94	1

Mujeres de 44 años en 2013		Estados de origen (2013)	Estados de destino (2014)										No afiliados	Total por origen
			RPM					RAIS						
			< 1 sml	1 - 1.2 sml	1.2 - 2.5 sml	2.5 - 5 sml	> 5 sml	< 1 s	1 - 1.2 s	1.2 - 2.5 s	2.5 - 5 s	> 5 s		
RPM	< 1 sml	3074	1743	659	177	13	1	28	29	12	0	0	412	2662
	1 - 1.2 sml	7760	2146	3340	1073	45	1	50	108	35	1	0	961	6799
	1.2 - 2.5 sml	8879	265	1077	6248	485	30	12	43	122	4	0	593	8286
	2.5 - 5 sml	3878	19	44	434	2887	320	0	3	5	28	6	132	3746
	> 5 sml	2504	3	8	21	251	2128	1	1	2	3	17	69	2435
RAIS	< 1 sml	5187	138	84	34	5	0	2504	1124	375	27	11	885	4302
	1 - 1.2 sml	14211	164	408	150	8	2	3244	6138	2034	90	20	1953	12258
	1.2 - 2.5 sml	15521	31	183	885	81	11	451	2189	9690	734	59	1207	14314
	2.5 - 5 sml	5810	0	12	85	467	54	19	83	681	3650	444	315	5495
	> 5 sml	4819	2	4	7	62	545	6	19	74	412	3546	142	4677
No afiliados en 2013		207404	285	3485	985	140	58	335	4258	1538	319	121	195879	11525
Total por destino			4796	9304	#####	4444	3150	6650	####	#####	5268	4224	202548	

Matricialmente,

$$N(t+1; y+1; s) = P(t; y; s) N(t; y; s);$$

Donde N es un vector columna de estados, y P es la matriz de transición. Cuando se encadenan:

$$N(t+n; y+n; s) = P(t+n; y+n; s) * \dots * P(t; y; s) N(t; y; s).$$

Cuadro 10. Errores de predicción a 2013 de las matrices de transición de 2008 (% de valores reales en 2013)

	Hombres 30 años	Hombres 40 años	Hombres 50 años	Hombres 60 años
Totales	0%	0%	0%	0%
Por régimen				
No afiliados	11%	8%	5%	4%
RAIS	-13%	-10%	-10%	-18%
RPM	4%	-5%	-7%	-15%
Por ingreso				
De 0 a 1 SML	53%	50%	47%	32%
De 1.2 a 2.5 SML	-12%	-6%	-5%	-12%
De 1 a 1.2 SML	-20%	-27%	-33%	-43%
De 2.5 a 5 SML	-20%	-15%	-9%	-19%
Mas de 5 SML	-28%	-18%	-11%	-11%
Sin información	11%	8%	5%	4%
Por fidelidad				
De 1 a 13 semanas	-32%	-42%	-51%	-67%
De 14 a 26 semanas	-20%	-26%	-33%	-50%
De 27 a 39 semanas	-26%	-33%	-37%	-44%
De 40 a 46 semanas	-36%	-39%	-39%	-37%
De 47 a 52 semanas	17%	25%	29%	48%
Sin información	11%	8%	5%	4%

	Mujeres 25 años	Mujeres 35 años	Mujeres 45 años	Mujeres 55 años
Totales	0%	0%	0%	0%
Por régimen				
No afiliados	13%	8%	3%	1%
RAIS	-26%	-14%	-9%	-4%
RPM	-30%	-22%	-13%	-5%
Por ingreso				
De 0 a 1 SML	32%	40%	39%	44%
De 1.2 a 2.5 SML	-28%	-16%	-9%	-8%
De 1 a 1.2 SML	-32%	-30%	-29%	-25%
De 2.5 a 5 SML	-32%	-12%	-8%	-8%
Mas de 5 SML	-41%	-17%	-11%	0%
Sin información	13%	8%	3%	1%
Por fidelidad				
De 1 a 13 semanas	-40%	-47%	-57%	-63%
De 14 a 26 semanas	-34%	-38%	-43%	-51%
De 27 a 39 semanas	-39%	-43%	-46%	-44%
De 40 a 46 semanas	-44%	-44%	-40%	-29%
De 47 a 52 semanas	0%	20%	27%	48%
Sin información	13%	8%	3%	1%

Los errores de proyección pueden ser muy grandes:

- Demasiados no afiliados y muy pocos afiliados (jóvenes, sobre todo)
- Demasiados IBC < 1 SM y muy pocos (sobre todo entre 1 y 1.2 SM)
- Demasiada alta fidelidad (47-52 semanas).

Estamos en el proceso de corregir estos sesgos:

- Asignamos edad y sexo a quienes no tienen esa información pre 2013
- Tomamos la historia laboral de dos períodos, en vez de uno...

El tamaño del problema estadístico

- 14,295,851 individuos únicos que vemos al menos una vez afiliados entre 2008 y 2014.
- Hay $51 \times 51 \times \dots \times 51 = 51^7 = 9 \times 10^{11}$ posibles secuencias de estados en esos siete años.
- Sólo vemos la instanciación de 3,925,477 secuencias únicas de longitud $L=7$.
- 213,429 de esas secuencias no tienen edad/sexo.
- En promedio, 3.6 individuos por cada secuencia observada.

Los “nodos” (nxxx) son las dimensiones de las matrices de transición

Para cada grupo de edad simple y sexo se tiene el estado inicial y final según:

- Régimen de afiliación:
 - 0. No afiliado
 - 1. Afiliado a RPM
 - 2. Afiliado a RAIS
- Grupos de ingreso básico de cotización:
 - 1. Menos de 1 salario mínimo legal
 - 2. De 1 a 1.2 salarios mínimos legales
 - 3. De 1.2 a 2.5 salarios mínimos legales
 - 4. De 2.5 a 5 salarios mínimos legales
 - 5. Más de 5 salarios mínimos legales
- Grupos de fidelidad:
 - 1. Entre 1 y 13 semanas
 - 2. Entre 14 y 26 semanas
 - 3. Entre 27 y 39 semanas
 - 4. Entre 40 y 46 semanas
 - 5. Entre 47 y 52 semanas

Ejemplo:
n135

Asignación de edad y sexo

- Se compara
 1. la secuencia de estados durante **2008-2012** de un individuo sin edad/sexo,
e.g., “**n000-n121-n121-n133-n145**”,
 2. con las secuencias de estados durante **2010-2014** de los individuos que sí tienen edad/sexo
 - i. Con la misma secuencia,
“**n000-n121-n121-n133-n145**”;
 - ii. Si no los hay, con secuencias que difieren en el 1er estado,
“**nxyz-n121-n121-n133-n145**”;
 - iii. Si no los hay, con secuencias que difieren en el 2do estado,
“**n000-nxyz-n121-n133-n145**”;
 - iv. Etc, ...
 3. Se toma el vector de edades/sexos que resulta del punto 2, y se muestrea aleatoriamente de ahí.

Cuadro 10. Errores de predicción a 2013 de las matrices de transición de 2008 (% de valores reales en 2013)

	Hombres 30 años	Hombres 40 años	Hombres 50 años	Hombres 60 años
Totales	0%	0%	0%	0%
Por régimen				
No afiliados	11%	8%	5%	4%
RAIS	-13%	-10%	-10%	-18%
RPM	4%	-5%	-7%	-15%
Por ingreso				
De 0 a 1 SML	53%	50%	47%	32%
De 1.2 a 2.5 SML	-12%	-6%	-5%	-12%
De 1 a 1.2 SML	-20%	-27%	-33%	-43%
De 2.5 a 5 SML	-20%	-15%	-9%	-19%
Mas de 5 SML	-28%	-18%	-11%	-11%
Sin información	11%	8%	5%	4%
Por fidelidad				
De 1 a 13 semanas	-32%	-42%	-51%	-67%
De 14 a 26 semanas	-20%	-26%	-33%	-50%
De 27 a 39 semanas	-26%	-33%	-37%	-44%
De 40 a 46 semanas	-36%	-39%	-39%	-37%
De 47 a 52 semanas	17%	25%	29%	48%
Sin información	11%	8%	5%	4%

	Mujeres 25 años	Mujeres 35 años	Mujeres 45 años	Mujeres 55 años
Totales	0%	0%	0%	0%
Por régimen				
No afiliados	13%	8%	3%	1%
RAIS	-26%	-14%	-9%	-4%
RPM	-30%	-22%	-13%	-5%
Por ingreso				
De 0 a 1 SML	32%	40%	39%	44%
De 1.2 a 2.5 SML	-28%	-16%	-9%	-8%
De 1 a 1.2 SML	-32%	-30%	-29%	-25%
De 2.5 a 5 SML	-32%	-12%	-8%	-8%
Mas de 5 SML	-41%	-17%	-11%	0%
Sin información	13%	8%	3%	1%
Por fidelidad				
De 1 a 13 semanas	-40%	-47%	-57%	-63%
De 14 a 26 semanas	-34%	-38%	-43%	-51%
De 27 a 39 semanas	-39%	-43%	-46%	-44%
De 40 a 46 semanas	-44%	-44%	-40%	-29%
De 47 a 52 semanas	0%	20%	27%	48%
Sin información	13%	8%	3%	1%

Cuadro 10A. Errores de predicción a 2013 de las matrices de transición de 2008 asignando edades y sexo (% de valores observados en 2013)

	Hombres 30 años	Hombres 40 años	Hombres 50 años	Hombres 60 años
Por ingreso				
De 0 a 1 SML	71%	61%	46%	28%
De 1.2 a 2.5 SML	-8%	-4%	-4%	-15%
De 1 a 1.2 SML	-10%	-18%	-30%	-41%
De 2.5 a 5 SML	-17%	-13%	-8%	-15%
Mas de 5 SML	-25%	-15%	-9%	-10%
Sin información	5%	3%	4%	4%
Por fidelidad				
De 14 a 26 semanas	-8%	-13%	-21%	-38%
De 1 a 13 semanas	-10%	-20%	-34%	-40%
De 27 a 39 semanas	-21%	-28%	-33%	-42%
De 40 a 46 semanas	-33%	-37%	-40%	-42%
De 47 a 52 semanas	19%	24%	24%	25%
Sin información	5%	3%	4%	4%

Cuadro 10A. Errores de predicción a 2013 de las matrices de transición de 2008 asignando edades y sexo (% de valores observados en 2013)

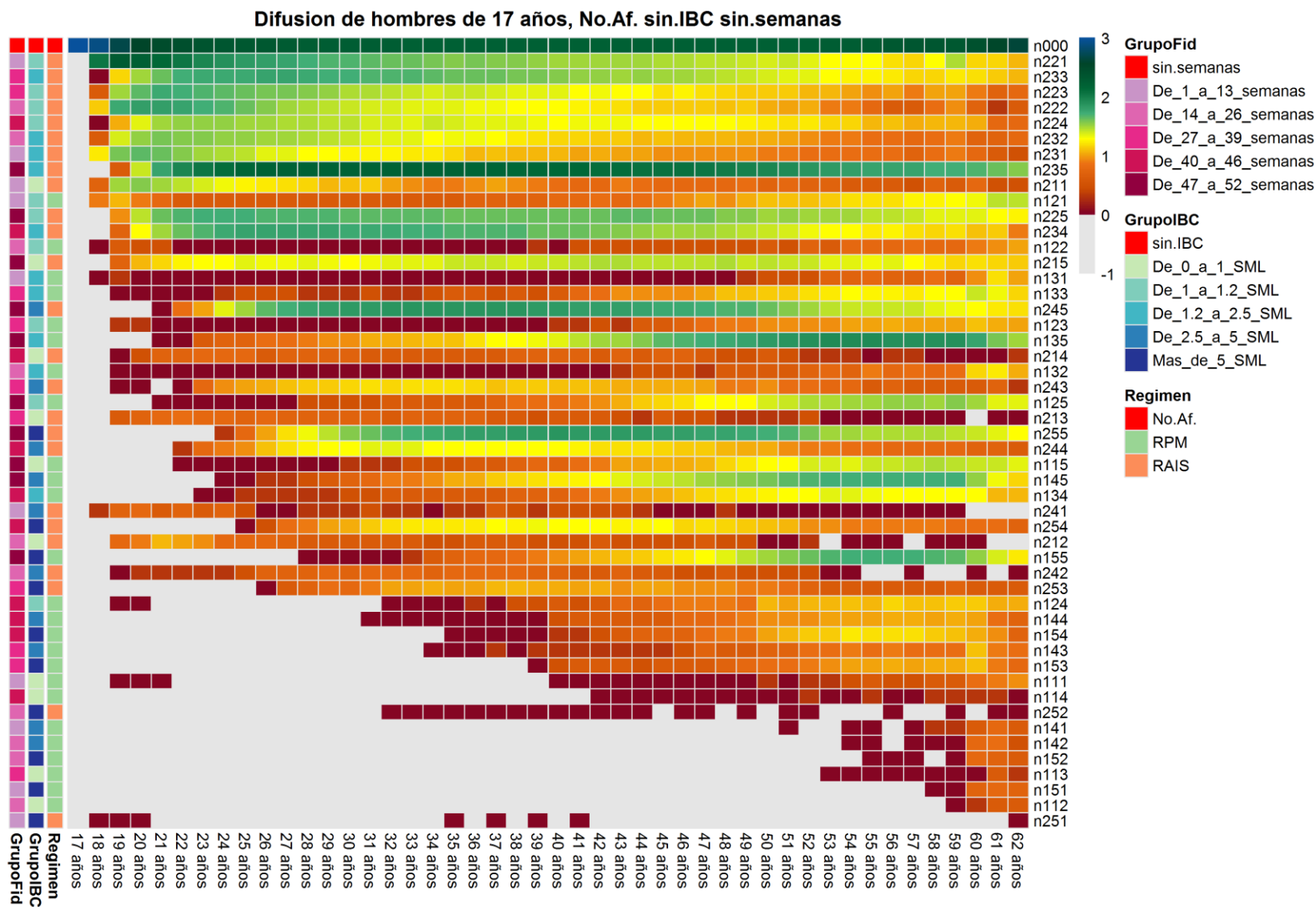
	Mujeres 25 años	Mujeres 35 años	Mujeres 45 años	Mujeres 55 años
Por ingreso				
De 0 a 1 SML	48%	48%	41%	35%
De 1.2 a 2.5 SML	-25%	-15%	-9%	-9%
De 1 a 1.2 SML	-25%	-25%	-26%	-24%
De 2.5 a 5 SML	-27%	-12%	-7%	-16%
Mas de 5 SML	-25%	-17%	-10%	1%
Sin información	11%	6%	3%	1%
Por fidelidad				
De 14 a 26 semanas	-24%	-26%	-27%	-18%
De 1 a 13 semanas	-23%	-28%	-33%	1%
De 27 a 39 semanas	-34%	-38%	-39%	-38%
De 40 a 46 semanas	-41%	-44%	-40%	-40%
De 47 a 52 semanas	1%	16%	21%	19%
Sin información	11%	6%	3%	1%

Cómo tener en cuenta el pasado de los individuos (al menos un año atrás)

- Cadenas de Markov de Longitud Variable (VLMC)
 - Generalización: “Arboles Probabilísticos de Sufijos” (PST)
 - Ref: Gabadinho, A. and Ritschard, G., 2016. *Analyzing state sequences with probabilistic suffix trees: the PST R package*. Journal of Statistical Software, 72(3), pp.1-39.
- Markov: $\Pr(\text{estado } i \mid \text{estado } j)$
- PST: $\Pr(\text{estado } i \mid \text{contexto})$,
donde “contexto” = secuencia de estados que precede el estado a pronosticar.

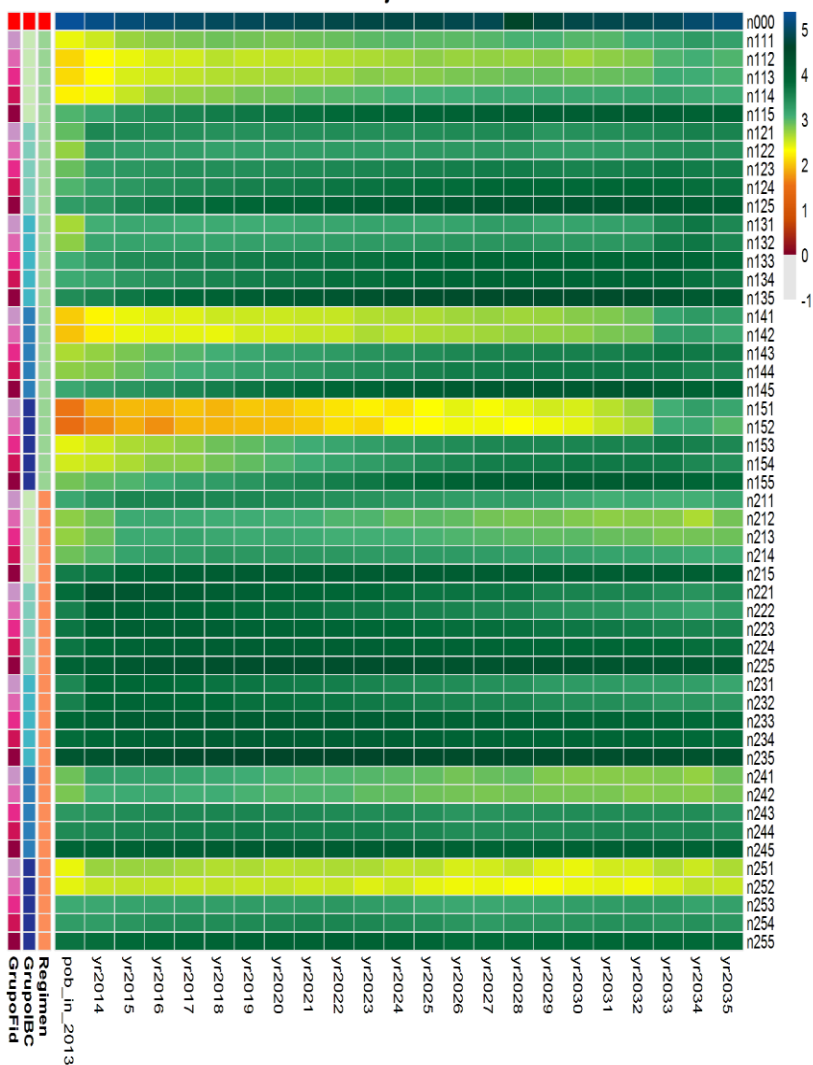
$$\hat{P}(\sigma|c) \simeq \hat{P}(\sigma|\text{suf}(c)), \sigma \in A,$$

Ejemplo de matriz de difusión

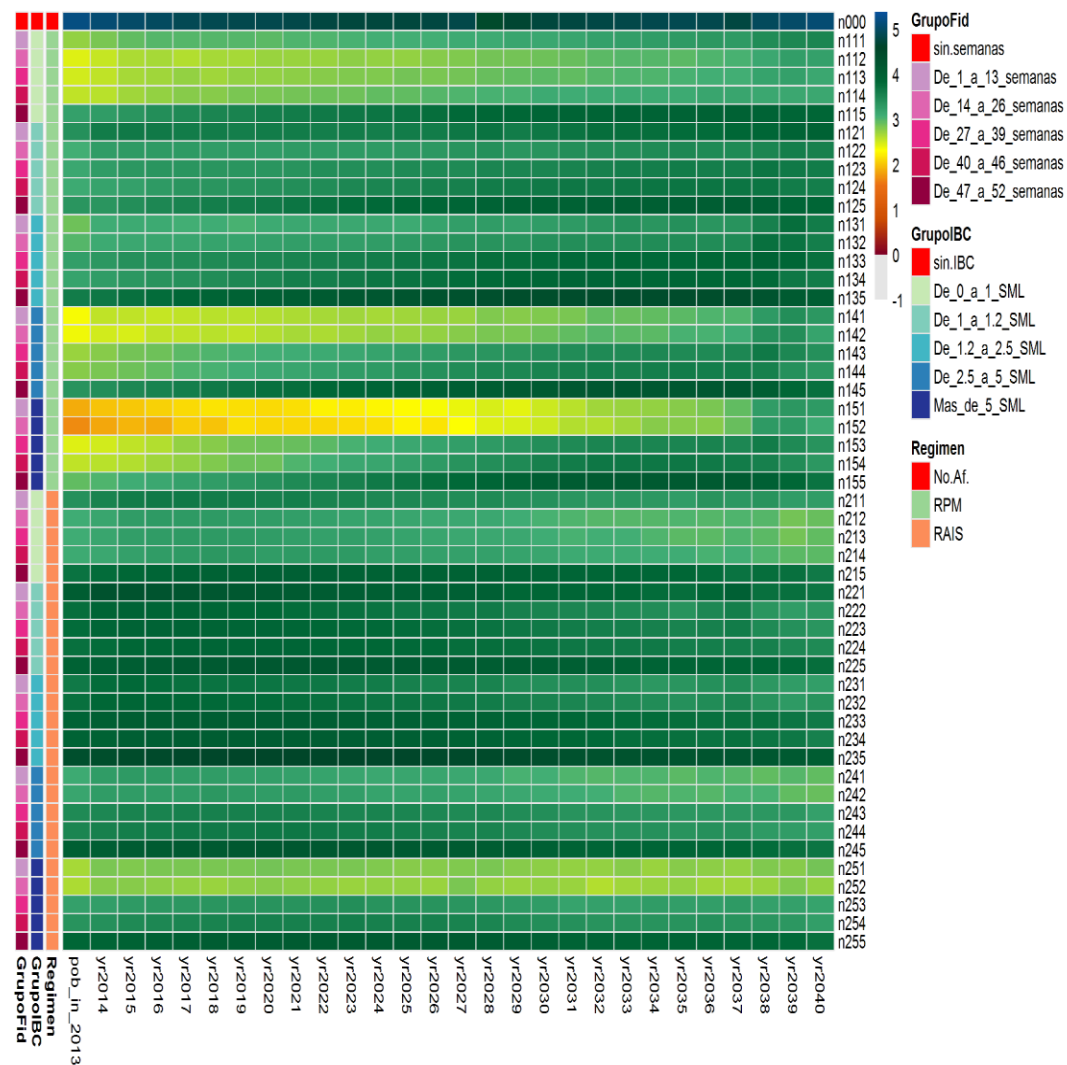


Difusión de una población real

Diffusion de mujeres de 35 en 2013



Diffusion de hombres de 35 en 2013



Cuadro 13. Proyecciones de afiliación a la edad de jubilación

	Datos observados en 2014			Proyecciones a 2050		
	Personas	Porcentajes totales	Porcentajes de afiliados	Personas	Porcentajes totales	Porcentajes de afiliados
Hombres						
Total	156184	100%		329423	100%	
Por régimen						
No afiliados	134791	86%		135494	41%	
RAIS	6137	4%	29%	80848	25%	42%
RPM	15256	10%	71%	113081	34%	58%
Por IBC						
De 0 a 1 SML	4523	3%	21%	26838	8%	14%
De 1.2 a 2.5 SML	5680	4%	27%	59108	18%	30%
De 1 a 1.2 SML	6392	4%	30%	52481	16%	27%
De 2.5 a 5 SML	2452	2%	11%	27403	8%	14%
Mas de 5 SML	2346	2%	11%	28098	9%	14%
Sin información	134791	86%		135494	41%	
Por fidelidad						
De_1_a_13_semana:	4928	3%	23%	35716	11%	18%
De_14_a_26_semana:	2395	2%	11%	21561	7%	11%
De_27_a_39_semana:	2649	2%	12%	30203	9%	16%
De_40_a_46_semana:	1934	1%	9%	27956	8%	14%
De_47_a_52_semana:	9487	6%	44%	78493	24%	40%
Sin información	134791	86%		135494	41%	

Cuadro 13. Proyecciones de afiliación a la edad de jubilación

	Datos observados en 2014			Proyecciones a 2050		
	Personas	Porcentajes totales	Porcentajes de afiliados	Personas	Porcentajes totales	Porcentajes de afiliados
Mujeres						
Total	224886	100%		372411	100%	
Por régimen						
No afiliados	200779	89%		144197	39%	
RAIS	7840	3%	33%	109677	29%	48%
RPM	16267	7%	67%	118536	32%	52%
Por IBC						
De 0 a 1 SML	4850	2%	20%	28056	8%	12%
De 1.2 a 2.5 SML	6430	3%	27%	68946	19%	30%
De 1 a 1.2 SML	6796	3%	28%	56220	15%	25%
De 2.5 a 5 SML	3202	1%	13%	40942	11%	18%
Mas de 5 SML	2829	1%	12%	34050	9%	15%
Sin información	200779	89%		144197	39%	
Por fidelidad						
De_1_a_13_semanas	2808	1%	12%	21101	6%	9%
De_14_a_26_semanas	1883	1%	8%	17088	5%	7%
De_27_a_39_semanas	3152	1%	13%	37514	10%	16%
De_40_a_46_semanas	3069	1%	13%	39331	11%	17%
De_47_a_52_semanas	13195	6%	55%	113180	30%	50%
Sin información	200779	89%		144197	39%	

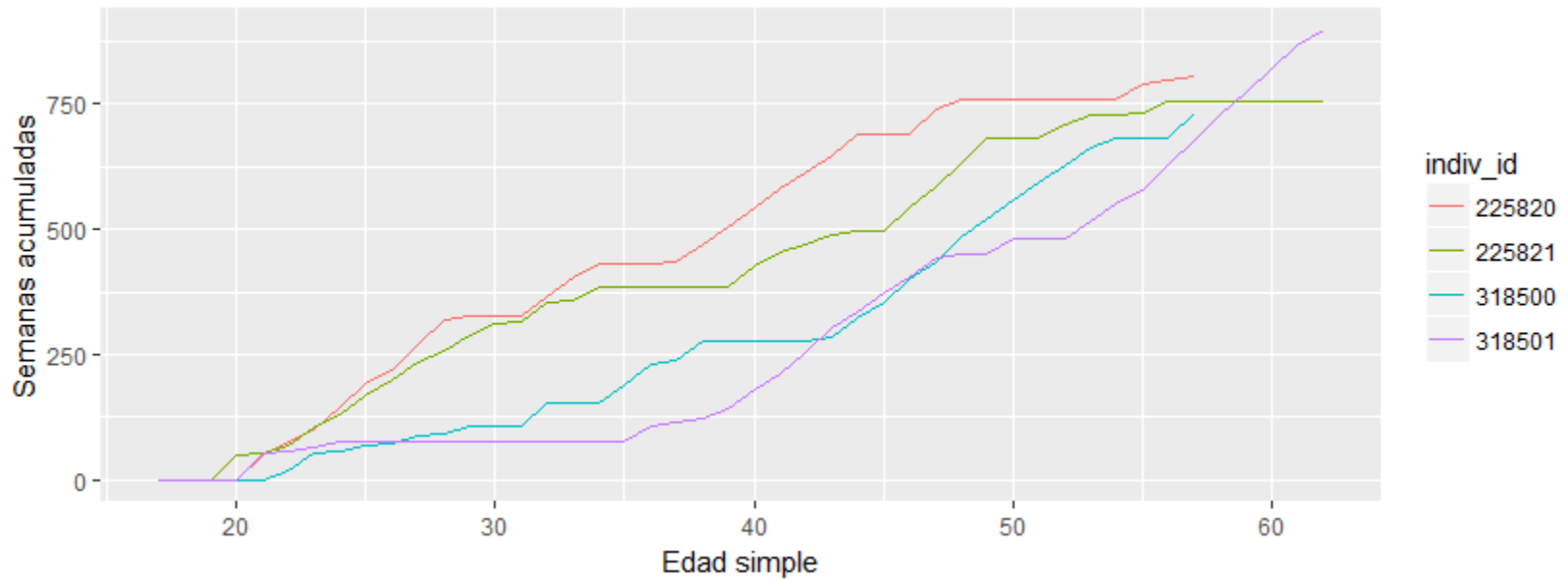
Cómo calcular las semanas

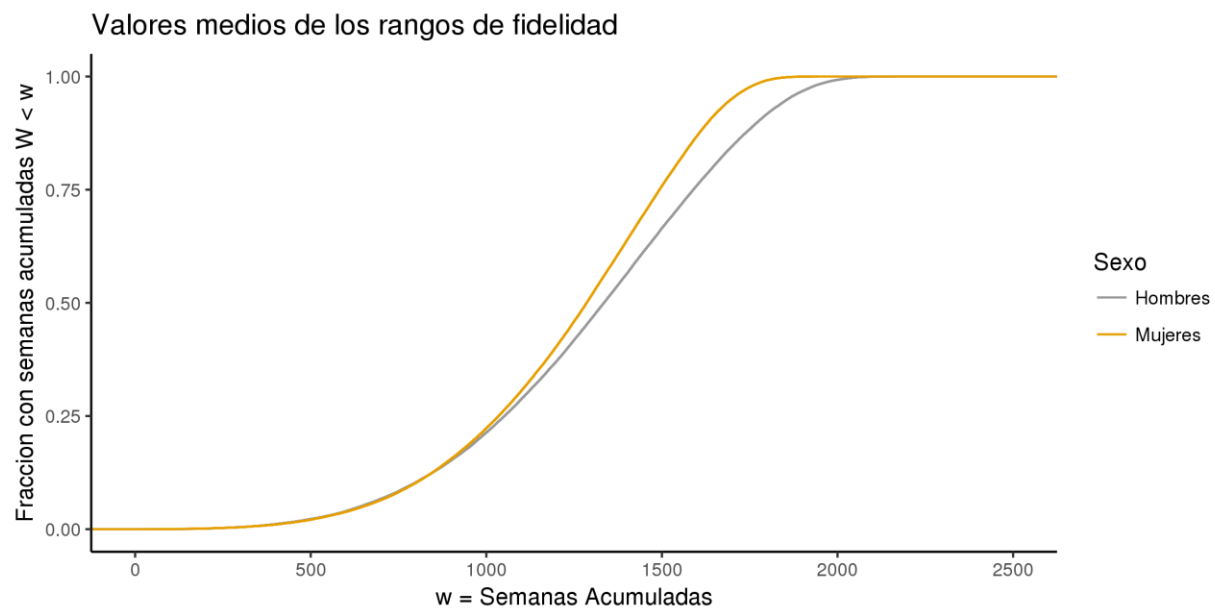
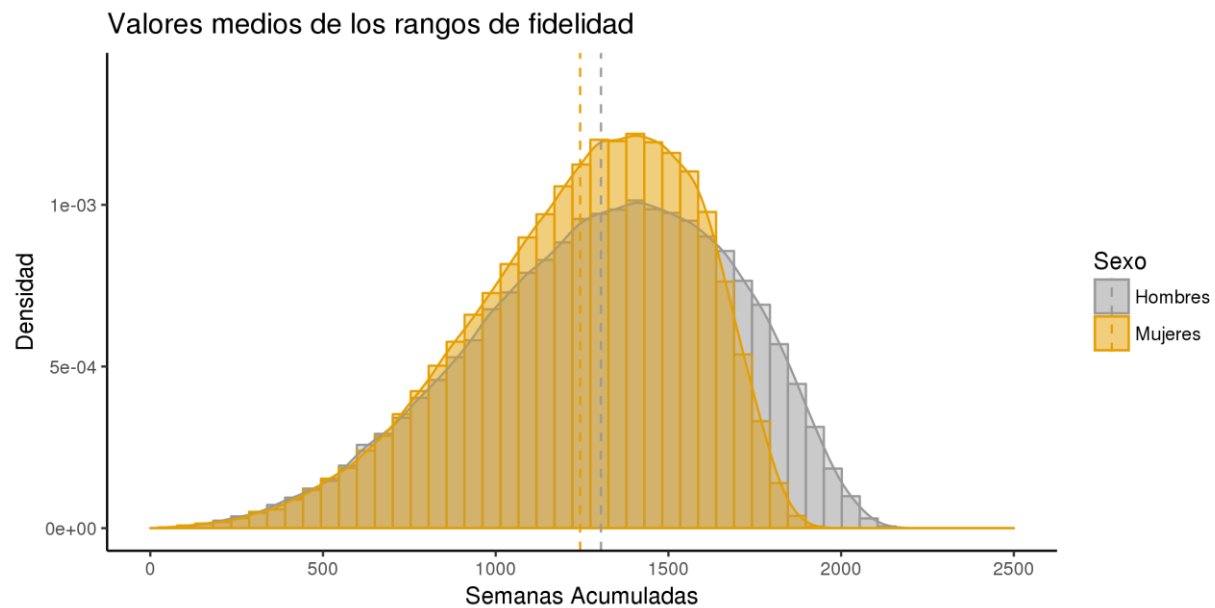
- Es necesario simular las historias laborales, no basta saber cuántas personas hay en cada estado porque hay muchas rutas posibles...
- Esto se ve con las matrices de difusión
- Con las matrices de difusión, para cada estado y grupo simple de edad-sexo, se calcula la *distribución* de las semanas acumuladas.
- De esta manera se tienen las semanas acumuladas hasta hoy de cada individuo (simulado)
- Y se sigue el mismo procedimiento a partir de los estados de la gente hoy hasta que lleguen el 2050.
- Pero además hay que saber cuántas semanas tendría cada individuo en el momento de arranque (2014).

Cómo calcular las semanas

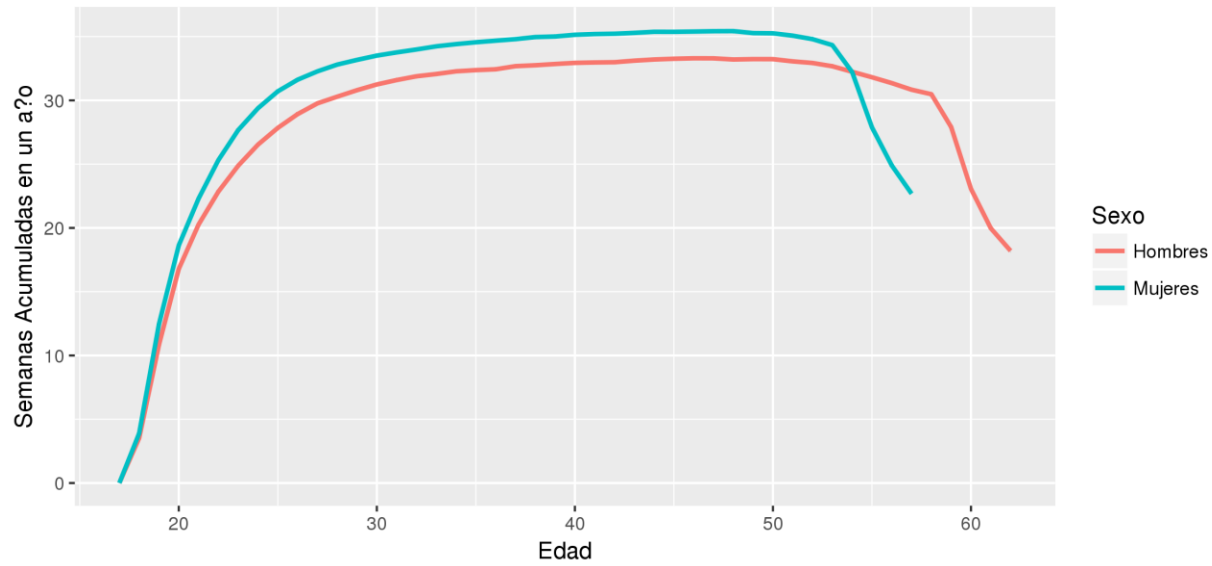
- Es necesario simular las historias laborales, no basta saber cuántas personas hay en cada estado porque hay muchas rutas posibles...
- Esto se ve con una “matriz de difusión”
- Con las matrices de difusión, para cada estado y grupo simple de edad-sexo, se calcula la *distribución* de las semanas acumuladas.
- De esta manera se tienen las semanas acumuladas hasta hoy de cada individuo (simulado)
- Y se sigue el mismo procedimiento a partir de los estados de la gente hoy hasta que lleguen el 2050.
- Pero además hay que saber cuántas semanas tendría cada individuo en el momento de arranque (2014).

Ejemplos de simulaciones dos mujeres y dos hombres

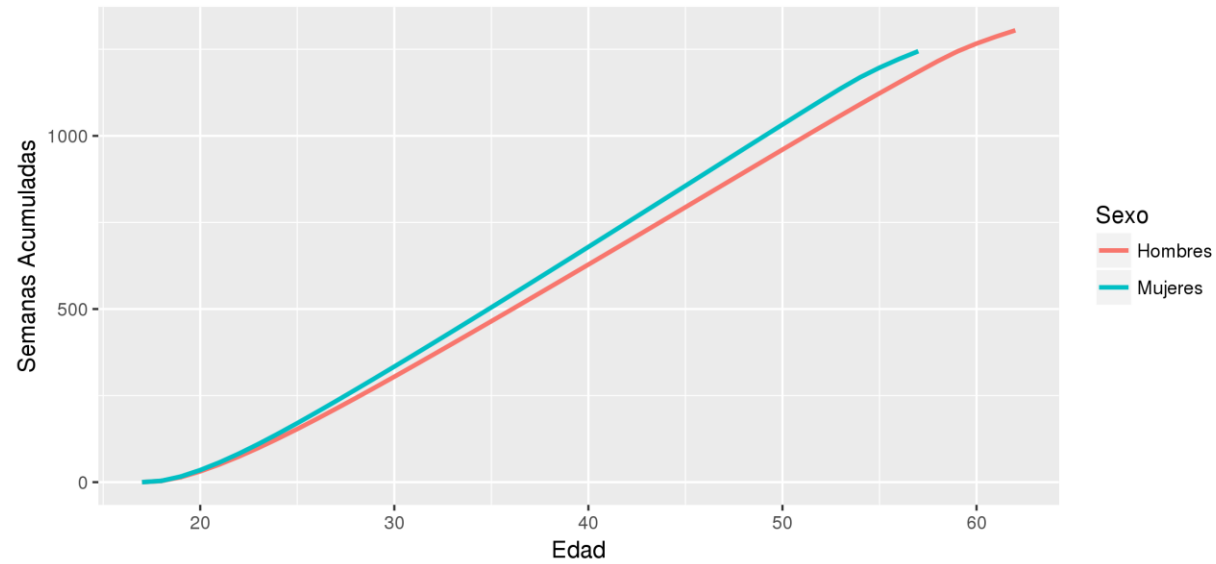




Promedio por edad de semanas acumuladas por año



Promedio por edad de semanas acumuladas desde los 17 años



Cuadro 12. Cómo se distribuye una cohorte simulada de 1000 personas a la edad de jubilación por semanas acumuladas de cotización

Hombres					
Semanas acumuladas	Total	No afiliados	RAIS	RPM	
Menos de 550	2121	1852	114	155	
550 a 700	725	566	83	76	
700 a 850	812	625	93	94	
850 a 1000	866	597	127	142	
1000 a 1150	848	570	119	159	
1150 a 1300	909	571	145	193	
Más de 1300	3719	1689	909	1121	
Totales	10000	6470	1590	1940	

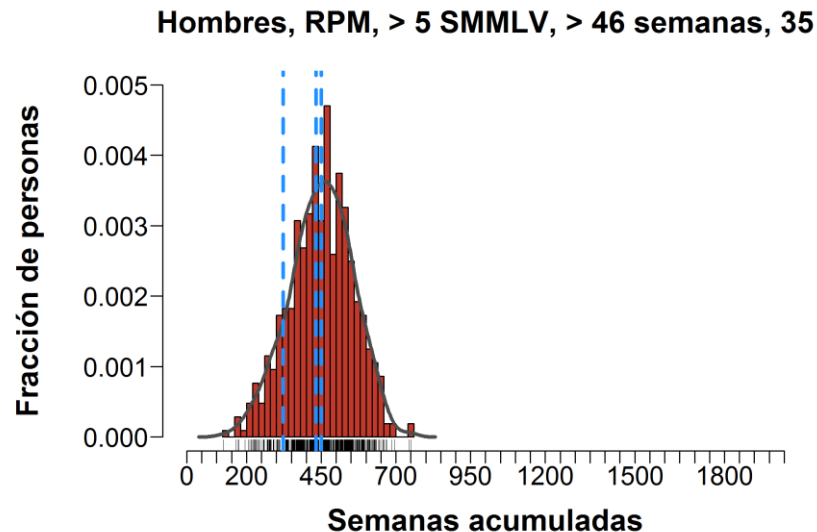
Mujeres					
Semanas acumuladas	Total	No afiliadas	RAIS	RPM	
Menos de 550	3678	3227	191	260	
550 a 700	785	585	103	97	
700 a 850	775	529	137	109	
850 a 1000	789	472	170	147	
1000 a 1150	795	439	167	189	
1150 a 1300	825	409	199	217	
Más de 1300	2353	717	773	863	
Totales	10000	6378	1740	1882	

Cómo calcular las semanas

- Es necesario simular las historias laborales, no basta saber cuántas personas hay en cada estado porque hay muchas rutas posibles...
- Esto se ve con una “matriz de difusión”
- Con las matrices de difusión, para cada estado y grupo simple de edad-sexo, se calcula la *distribución* de las semanas acumuladas.
- De esta manera se tienen las semanas acumuladas hasta hoy de cada individuo (simulado)
- Y se sigue el mismo procedimiento a partir de los estados de la gente hoy hasta que lleguen el 2050.
- Pero además hay que saber cuántas semanas tendría cada individuo en el momento de arranque (2014).

Estimación de las semanas actuales ya acumuladas

- No tenemos manera de saber cuántas semanas la gente hoy en día ha acumulado.
- Pero,
 - 1) Sabemos que más del 97% de los individuos de 17 años (en 2013) son “no afiliados” (estado n000).
 - 2) Sabemos las transiciones actuales para cada cohorte de edades simples.
 - 3) Podemos simular la “trayectoria laboral” de n individuos que comienzan como “no afiliados” a los 17 años,
 - 4) Y así, obtener una posible distribución de semanas acumuladas, por edad y sexo,
 - 5) de la cual se puede muestrear.



Posibilidades analíticas de los resultados que esperamos tener

- Proyecciones anuales hasta 2050 de los 51 estados para cada grupo simple de edad y sexo
 - Número de afiliados a RAIS y RPM por IBC y fidelidad
 - Coberturas
 - Distribución de los trabajadores por salario
- Proyecciones anuales de la distribución de semanas acumuladas de cada grupo simple de edad y sexo:
 - Cuántas personas completarán las semanas en cada régimen al llegar a la edad de jubilación
- Trayectorias posibles desde cualquier estado inicial:
 - Probabilidades de jubilarse en RPM (y de completar semanas en RAIS)
 - Probabilidades de permanecer en RPM o RAIS
- Revisiones utilizando un nuevo año de transiciones de PILA:
 - Tienden a mejorar las coberturas de afiliación y jubilación?

Gracias!

Bienvenidas las sugerencias

Representación de los estados

Cada trabajador en la PILA lo asignamos a un “estado” o “nodo”, y los nodos los representamos con la siguiente notación:

$$n\{\textit{regimen pensional}\}\{\textit{grupo ibc}\}\{\textit{grupo fidelidad}\}$$

Hay **2 regímenes pensionales**, **5 grupos de IBC**, y **5 grupos de fidelidad**, y los numeramos con los dígitos correspondientes:

- Régimen:

- “RPM”=1;
- “RAIS”=2;

- Grupo IBC:

- “> 0 y \leq 1 SMMLV” = 1;
- “> 1 y \leq 1.2 SMMLV” = 2;
- “> 1.2 y \leq 2.5 SMMLV” = 3;
- “> 2.5 y \leq 5 SMMLV” = 4;
- “> 5 SMMLV” = 5;

- Grupo Fid.:

- “> 1 y \leq 13 semanas” = 1;
- “> 13 y \leq 26 semanas” = 2;
- “> 26 y \leq 39 semanas” = 3;
- “> 39 y \leq 46 semanas” = 4;
- “> 46 y \leq 52 semanas” = 5.

Vemos que hay 50 diferentes estados en los que un individuo puede estar dentro de la PILA ($n111, n112, \dots, n255$). A estos 50 estados hay que adicionarle el estado de “no afiliados”, que representamos con los dígitos todos igual a 0,

$n000$.

Resumen del trabajo técnico

MANIPULACION DE DATOS

1. Limpieza microdatos PILA 2008-2014.
 - Eliminar observaciones irregulares.
 - Retener cotizantes a pensión, dependientes e independientes.
 - Corrección salarios y edades.
 - Creación de los estados = (régimen, grupo IBC, grupo Fidelidad)
2. Conversión a formato de *individuos x secuencia de estados* 2008-2014.
3. Asignación edades/sexos para todos los individuos.
4. Separar los individuos en diferentes archivos por sexo y edad simple (edad que tenían en 2008).

ESTIMACION DE MODELO

5. Para cada edad simple y sexo, para los años 2011-2012-2013, estimar el modelo de VLMC (*Variable Length Markov Chain*) con memoria $L = 2$ (donde Markov simple sería $L=1$). Es decir tomamos dos periodos (2011-2012) para estimar la probabilidad de transición a cada uno de los 51 estados (en 2013).
6. Guardamos en archivos CSV todas las matrices de transición, por sexo y edad.

POSIBILIDADES Y EJEMPLOS DE TRABAJO ESTADÍSTICO

7. **Generar la difusión de poblaciones** específicas de individuos (e.g., 1,000 mujeres de 35 años de edad de alto salario y fidelidad).
8. **Proyectar las poblaciones** de individuos en los diferentes estados en 2013 hacia adelante en el futuro, e.g., los años 2014 al 2050.
9. **Simular trayectorias laborales** individuales.
 - Puesto que en los datos reales más del 97% de individuos con 17 años son “no afiliados”, es decir el estado ‘n000’, al comenzar todas nuestras simulaciones con individuos de 17 años que comienzan como “no afiliados” obtenemos una muestra que nos sirve para estimar el número de semanas de las poblaciones de individuos reales actuales.
10. **Asignar estimación de semanas acumuladas** a individuos en 2013/2014.
11. **Analizar distribuciones** (no sólo promedios!) de semanas acumuladas por edad, sexo, año, régimen.

Ejemplo de la base final

personabasicaid	edadsimple_in_2008	sex	year2008	year2009	year2010	year2011	year2012	year2013	year2014
5896054	44	1	n245	n245	n244	n233	n000	n233	n234
5896057	20	1	n222	n214	n215	n223	n213	n232	n222
5896061	30	1	n000	n000	n000	n000	n000	n221	n221
5896065	40	0	n235	n235	n235	n235	n235	n235	n225
5896066	21	1	n222	n222	n213	n223	n222	n232	n233
5896073	40	1	n122	n115	n115	n115	n115	n125	n115
5896095	37	1	n225	n223	n234	n235	n233	n223	n224
5896101	26	1	n211	n231	n215	n215	n212	n000	n000
5896102	22	1	n000	n000	n000	n000	n000	n133	n125
5896104	19	1	n000	n000	n000	n222	n221	n215	n215
5896113	32	1	n241	n000	n221	n233	n213	n000	n223
5896118	32	0	n000	n000	n000	n000	n000	n243	n223
5896125	27	0	n223	n235	n234	n235	n234	n234	n235
5896128	38	1	n235	n235	n235	n234	n221	n222	n223
5896132	25	1	n214	n212	n222	n234	n233	n223	n224
5896152	33	0	n245	n245	n245	n245	n245	n245	n245
5896153	43	0	n255	n255	n255	n255	n255	n255	n255
5896171	22	0	n000	n000	n000	n000	n000	n000	n224
5896175	34	0	n221	n000	n000	n000	n000	n222	n215