

Movilidad Social y Transición Demográfica en Colombia-1870

Hermes Tovar¹ y B. Piedad Urdinola²

Este resumen ampliado recoge los resultados preliminares de dos temas disímiles que corresponden a poblaciones geográfica y demográficamente opuestas que harán parte de la colección de libros resultado de la investigación en curso: “La población y desarrollo regional a partir del censo de 1870”. Este proyecto, que lleva cinco años continuos de investigación, consiste en la digitalización sistemática del micro-dato del censo de 1870 que reposa disperso en los archivos históricos del país. Este proyecto es completamente innovador, pues durante del siglo XIX se realizaron al menos seis censos nacionales (1825, 1835, 1843, 1851, 1859 y 1870), pero no existe en el país un estudio detallado de los mismos. Conocemos sus resúmenes generales por provincias, cantones y parroquias pero no se ha llevado a cabo una investigación sistemática de la información detallada por familias e individuos, que es la que permite los análisis del tipo que aquí se presentan.

La digitalización del censo de 1870 requiere un trabajo extensivo y altamente demandante de recursos humanos, tal como hemos aprendido en el proceso, por tres razones principales. Primero, los archivos históricos no se encuentran todos en un único archivo en Colombia. Aunque la mayoría de los registros pueden estar en el Archivo General de la Nación (Bogotá), existen más registros en los Archivos regionales, como Boyacá, Antioquia y Popayán. Segundo, aunque la mayoría de los registros se encuentran en el Archivo General de la Nación, estos registros no están propiamente organizados y por tanto requiere de una búsqueda extensiva en los diversos fondos para lograr ubicarlos, identificarlos y digitalizarlos. Tercero, una vez identificados los folios que contienen los datos individuales de los censos, su lectura no es fácil para un digitador corriente tanto por la escritura como por el estado en que se encuentren. Sobre decir que el papel en el que están impresos tiene más de siglo y medio, lo que hace que muchos de los folios a pesar de ser encontrados en los Archivos están altamente deteriorados e incluso rotos. Asimismo, cualquier digitador contemporáneo no está entrenado en paleografía tanto para leer los nombres como para las profesiones y edades que se encuentran registrados en el censo de 1870.

El libro, actualmente en curso, será el resultado del análisis de dichos datos digitalizados y constituirá el primer volumen de una colección que busca presentar los resultados por provincia y al final un volumen de análisis comparativo entre regiones. El análisis utiliza técnicas estadísticas y demográficas modernas que permiten tener una visión más cercana de la población colombiana de 1870. Los datos utilizados pasaron por un proceso de transcripción, verificación, validación e imputación. Estos datos nos permiten reconstruir la forma de vida de sociedades para confirmar si son agrícolas y con patrones estáticos o si, siguiendo el espíritu de Peter Laslett (1965), encontramos que las sociedades del pasado se comportan muy diferentes a lo que imaginamos.

I. Movilidad Social

El primer tema hace referencia a los patrones de movilidad social en Colombia de 1870. Los datos del censo nos permite estudiar no sólo las diferencias entre generaciones en la ocupación, sino que también incluir un perfil por edad, sexo y estado civil, ya que se ha encontrado que las variables demográficas son determinantes de la movilidad social de otras sociedades históricas (Hajnal, 1965; 1982; Ohlin 1961). El censo colombiano de 1870 capturó nombre y apellidos, estado civil, edad y ocupación. La última variable ha sido recogida por todos los miembros del hogar mayores de 8 años de edad, lo que constituye una rareza en los censos históricos que usualmente no incluyen

¹ Profesor Titular, Facultad de Economía. Universidad de los Andes-Bogotá

² Profesora Asociada, Departamento de Estadística. Universidad Nacional de Colombia-Bogotá

información de ocupación y cuando lo hacen, solo incluye varones jóvenes (en su mayoría blancos). De manera que la información en nuestras manos nos permite utilizar modernas técnicas estadísticas y demográficas para estudiar la movilidad social, y aquí proponemos el análisis de correspondencias simples (Lebart, 1995).

El análisis de correspondencias es la representación simultánea de dos variables representadas en dos espacios en un mismo plano. La tabla se construye a partir de perfiles fila columna, así como una tabla de frecuencias relativas. Al tener la tabla de perfiles fila se tienen n-puntos en el espacio de las p-columnas y similarmente, con la tabla de perfiles columna se tienen p-puntos en el espacio de las n-filas. Luego se calculan las distancias de las filas entre sí y de las columnas entre sí, por aparte, y se usa la distancia χ^2 , ya que tiene en cuenta los pesos de las categorías fila así como la influencia de las categorías de las columnas. Lo mismo se hace para las columnas. Las distancias miden la similitud entre categorías de la misma variable, que se puede representar en cada uno de los espacios (fila y columna), pero es por medio de las relaciones cuasi-baricéntricas que se puede justificar la representación simultánea de filas y columnas. Los métodos factoriales se fundamentan en el cálculo de distancias entre puntos fila y puntos columna, sin embargo, la distancia entre un punto fila y un punto columna no tiene sentido debido a que los puntos están en espacios diferentes. Para este caso, el análisis de correspondencias ofrece la posibilidad de posicionar e interpretar un punto de un conjunto relativo a un espacio con respecto al conjunto de los otros puntos definidos en el otro espacio (Lebart 1995).

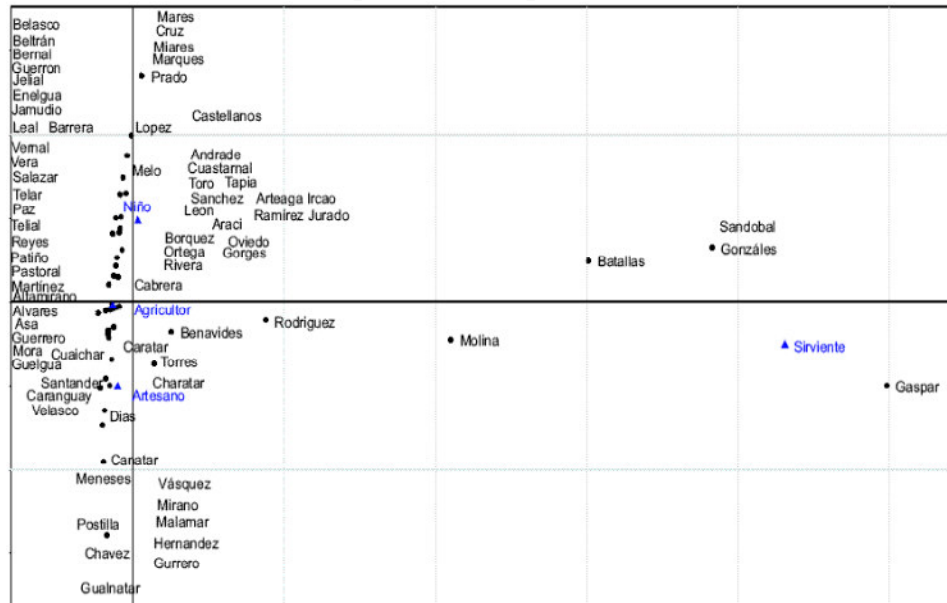
El análisis se presenta para una población pequeña que consta de 770 registros censales, después del proceso de verificación y de eliminar 4 registros que tenían dos o más variables faltantes, del Distrito de Guachávez (en el Estado soberano del Cauca, Círculo de Túquerres). Los resultados muestran que la aparente perpetuación de las profesiones entre familias (Figuras 1 y 2), encubren un efecto de ciclo vital y roles de género en esta pequeña comunidad agraria de 1870. Es decir, si bien puede existir una predominancia en las profesiones entre generaciones, es mucho más evidente la relación entre las variables demográficas, género y edad, (Figuras 3 y 4) atada al ciclo de vida de sus habitantes. La aplicación de este tipo de metodologías es ideal para las denominadas “poblaciones pequeñas”, que en términos demográficos y estadísticos no logran satisfacer los números suficientes como para aplicar las metodologías tradicionales.

La interpretación de los mapas factoriales es simple, en general, entre más cercanas visualmente estén las categorías de una variable a una categoría de la otra, hay una fuerte relación. Para tener un punto de referencia la gráfica presenta los dos primeros ejes principales, que son los que explican la mayor parte de los datos y genera un vértice diferente para cada mapa. Tradicionalmente el primer eje se presenta como una línea horizontal y el segundo eje como una línea vertical. Para ejemplificar su lectura veamos la Figura 1 que contrasta apellidos y ocupaciones. El primer eje (eje horizontal) es el contraste de “sirviente”, a la derecha, y “agricultor”, a la izquierda. Este eje explica casi en su totalidad la categoría “sirviente”. El segundo eje (eje vertical) compara las categorías “niños” y “artesanos” y el tercer eje contrasta “agricultor” con “artesano”. La tabla asociada a la Fig. 1, muestra que la categoría “sirviente”, del primer eje, está prácticamente explicada por los datos (96,83%), dejando el resto de las ocupaciones mal representadas; mientras que en el segundo plano, los ejes 2 y 3, las otras ocupaciones están mejor representados. De este mapa factorial, podemos decir que las familias Benavides, Rodríguez, Gaspar, Molina, Batallas, Gonzáles y Sandobal siempre se ocupan como sirvientes.

Pero como el efecto de sirviente en el primer plano es tan grande, se limpia el efecto presentando un segundo mapa con los planos 2 y 3, sin la categoría “sirviente” (Figura 2). Allí se observa que las familias Cabrera, Vera, Ramírez, Ircao, León y Gaspar están asociadas a ocupación de la agricultura, mientras las familias Meneses, Charalá, Rodríguez, Guerrero, Dias y Molina están más

asociadas a la de artesanos. También se observa que las familias Arteaga, Prado, Cuastarnal, Castellanos y Andrade son las que tienen más hijos menores de 8 años. Si no se hiciera un análisis posterior se pensaría que las familias por linaje tienden a preservar el mismo tipo de ocupación en el tiempo. Sin embargo, los siguientes planos factoriales muestran una fuerte relación con otras variables demográficas, como ya se explicó, y por ende son las determinantes un ciclo vital del tipo de ocupación en esta pequeña población, lo que nos lleva a pensar que mientras son niños-jóvenes y solteros trabajan en la agricultura, una vez se casan y hasta la adultez media a ser artesanos y en las últimas edades, cuando ya muchos son viudos, terminan su ciclo vital trabajando como sirvientes.

Figura 1. Primer plano factorial del análisis de correspondencias de la tabla de contingencia Apellidos × Ocupaciones



	Contribución del eje		
	Comp1	Comp2	Comp3
Niño	0.11	51.75	19.95
Agricultor	2.34	0.21	51.74
Artesano	0.73	47.21	27.66
Sirviente	96.83	0.83	0.66

Figura 2. Segundo plano factorial del análisis de correspondencias de la tabla de contingencia Apellidos × Ocupaciones

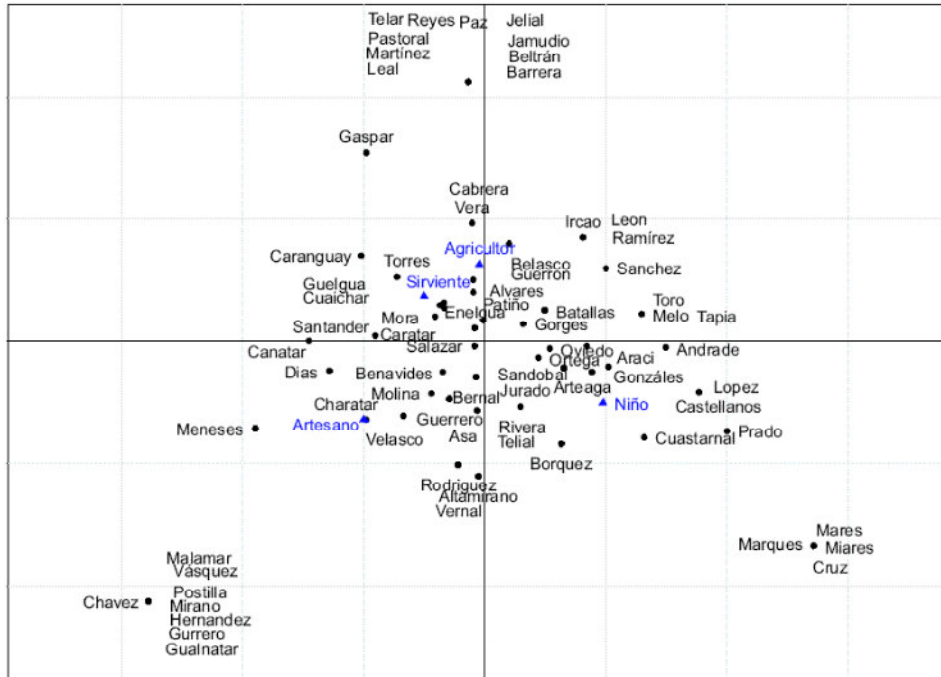


Figura 3. Primer plano factorial del análisis de correspondencias de la tabla de contingencia Edad × Ocupaciones

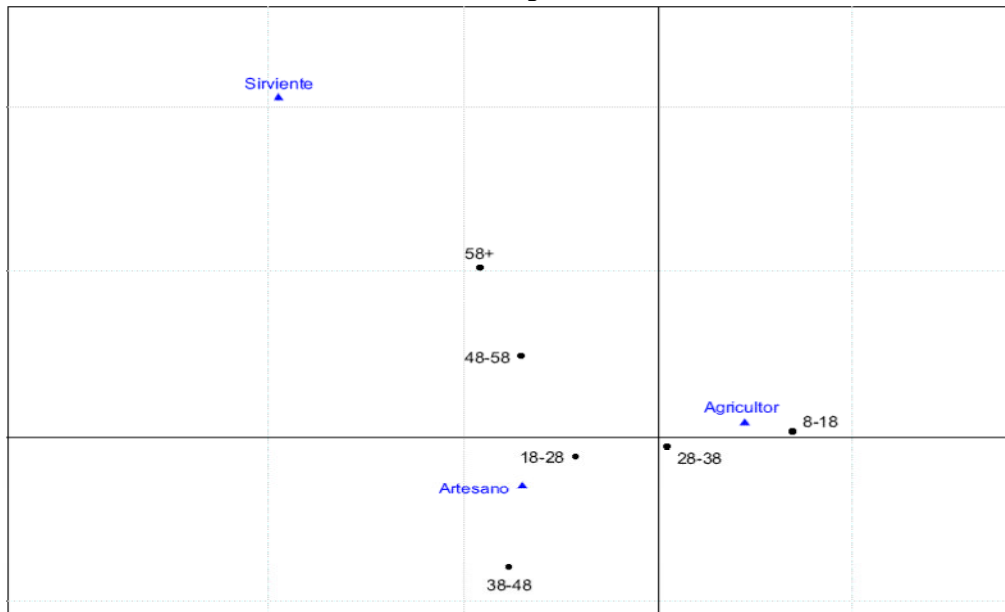


Figura 4. Primer plano factorial del análisis de correspondencias de la tabla de contingencia Estado Civil x Ocupaciones



II. Transición Demográfica y Control a la Natalidad

La segunda parte busca resolver la pregunta de si las mujeres colombianas de la época comenzaban a controlar la fecundidad como un comienzo de transición demográfica ya presente en varios países europeos, incluyendo aquellos del sur de Europa, mediante la medición de los modelos tradicionales de control a la fecundidad y proponemos una modificación estadística a tal modelo.

Louis Henry (1961) propone, por primera vez, que si queremos entender el control a la fecundidad es necesario describir primero los patrones de nacimientos *antes* que se impusieran las prácticas de lo que hoy llamamos control a la natalidad. De manera que se pudieran distinguir las poblaciones en “ausencia” de prácticas que llevan a la reducción de los nacimientos, de las que tienen “presencia” de dichas prácticas, que acercan a las parejas a un número deseado de hijos objetivo. Recordemos que el proceso de caída de la fecundidad, como paso determinante en la transición demográfica, comenzó desde el siglo XIX en Europa, lo que ya consistía en información histórica para cuando Henry propone sus ideas. Esta idea que Henry nos hereda de “fecundidad natural” (sin control a la natalidad), la recoge años más tarde la Oficina de Población de Princeton (OPR) de donde nacen exitosamente varios indicadores que cuantifican estos hechos demográficos, que revelan si una población está o no controlando la natalidad.

En particular, aquí seguimos el modelo matemático propuesto originalmente por Coale&Trussell (1974) y la modificación de Xie&Pimentel (1992). El modelo original permite una estimación indirecta a partir de las agendas de fecundidad de las mujeres casadas, y por su simplicidad en la aplicación es el método más popular en Demografía Histórica. El método propone la medición de las tasas específicas de fecundidad a cualquier edad como:

$${}_nF_x = Mn(x)e^{-mv(x)}$$

M= parámetro que denota el nivel de fecundidad natural

m= parámetro que denota el alcance de la limitación a la fecundidad

$n(x)$ = es una agenda de fecundidad dada para la fecundidad natural, para cada edad x

$v(x)$ = es un conjunto de pesos por el impacto del control a la fecundidad en las tasas para la edad x

Los parámetros a estimar son m y M que denotan el nivel del control a la fecundidad. Los valores de $n(x)$ originalmente corresponden al promedio de 10 poblaciones identificadas por Henry como con fecundidad natural. Mientras que $v(x)$ fue creado a partir de 48 agendas de fecundidad de las encuestas mundiales de fecundidad que ya seguían un control a la natalidad, identificadas por Coale&Trussell en 1974. Los autores estandarizaron estos valores y fijaron el valor de la primera edad a cero, de manera que fuera posible estimar los demás parámetros de la ecuación para la primera edad y poder resolver los parámetros para los siguientes grupos de edades.

Posteriormente, las variaciones de Trussell (1985) y Broström (1985) hacen que el modelo tenga una solución estadística, y no matemática, al cuantificar el valor esperado de los nacimientos, en lugar de tomar las tasas esperadas de fecundidad. Más recientemente Xie&Pimentel (1992) proponen una estimación log-lineal siguiendo una distribución Poisson bajo el modelo estadístico, vía máxima verosimilitud.

Para los datos del censo colombiano de 1870, nosotros proponemos trabajar con esta distribución Poisson, que es la distribución que mejor describe los datos que representan conteos, como son el número de hijos por mujer, pero re-estimando las agendas de fecundidad subyacentes. En lugar de tomarlas de muestras, trabajamos con los datos de registros vitales (universales) de excelente calidad, como los publicados en la base de fecundidad humana (www.humanfertility.org). De allí se obtuvo el número de nacimientos y la exposición al riesgo de la población femenina con las que se calcularon los valores esperados de los nacimientos a partir de la información de: Austria, República Checa, Estonia, Finlandia, Holanda, Rusia, Eslovaquia, Eslovenia, Suiza, Suecia, y Estados Unidos; para los años de 1951, 1970, y 2000. Cuando la información de estos años no estaba disponible se utilizaron los años de 1983, para Estonia y Rusia, y 1990 para Eslovenia. La aplicación se hace sobre las familias reconstruidas a partir del micro-dato censal de 1870 para datos parciales de la ciudad de Santa Marta y San Juan del Cesar (del Estado soberano del Magdalena), donde se reconstruyeron 32 familias para Santa Marta y 99 para San Juan del Cesar.

Los resultados sugieren que el uso de datos de estadísticas vitales modernos puede llevar a resultados encontrados. De hecho, los parámetros estimados sugieren que para 1870 existía un incipiente control a la fecundidad medido desde el método tradicional (Coale&Trussell, 1974) y la variación Xie&Pimentel (1992), mas no por el ajuste a metodología aquí propuesta (ver Tablas 1, 2 y 3). La metodología aquí propuesta puede estar generando resultados encontrados porque los niveles estimados [de control a la fecundidad] a partir de las agendas modernas, pueden ser demasiado estrictas para los datos históricos que estamos trabajando, a pesar de ser mucho más completas que las extraídas de muestras. Esto nos da un argumento más para seguir adelante en proyectos como este y motivar a otros investigadores a medir estos fenómenos demográficos a partir de otras fuentes, como archivos parroquiales o más datos censales históricos, para tener mayor certeza sobre las estimaciones encontradas.

Tabla 1. Estimaciones del Modelo Coale-Trussell para Santa Marta (SM) y San Juan del Cesar (SJC), 1870

Edad	${}_5F_x$ (S.M)	${}_5F_x$ (S.J.C)	$n(x)$	X	Y	X*X	X*Y
15	0.1875	0.3535	0.360	0	-0.236	0	0
20	0.5625	0.6666	0.460	0	0.008	0	0
25	0.53125	0.7070	0.431	0.316	-0.056	0.099	0.0179
30	0.25	0.4747	0.396	0.814	-0.141	0.662	0.115
35	0.21875	0.3131	0.321	1.048	-0.351	1.09	0.368
40	0.0625	0.1212	0.167	1.424	-1.005	1.971	1.411

Tabla 2. Estimaciones del Modelo Xie-Pimentel para Santa Marta (SM) y San Juan del Cesar (SJC), 1870

Edad	${}_5F_x(S.M)$	${}_5F_x(S.J.C)$	n(x)	X	Y	X*X	X*Y
15	0.1875	0.3535	0.360	0	-0.652	0	0
20	0.5625	0.6666	0.460	0	0.201	0	0
25	0.53125	0.7070	0.436	0.329	0.197	0.108	0.065
30	0.25	0.4747	0.392	0.713	-0.449	0.508	-0.320
35	0.21875	0.3131	0.333	1.194	-0.420	1.425	-0.501
40	0.0625	0.1212	0.199	1.671	-1.158	2.792	-1.935
Av	0.302	0.4561	0.363	0.651	-0.380	0.805	-0.448

Tabla 3. Estimaciones del Modelo Propuesto para Santa Marta (SM) y San Juan del Cesar (SJC), 1870

Edad	${}_5F_x(S.M)$	${}_5F_x(S.J.C)$	n(x)	X	Y	X*X	X*Y
15	0.1875	0.3535	0.360	0	-0.652	0	0
20	0.5625	0.6666	0.364	0	0.433	0	0
25	0.53125	0.7070	0.301	0.316	0.566	0.108	0.186
30	0.25	0.4747	0.240	0.814	0.037	0.508	0.026
35	0.21875	0.3131	0.172	1.048	0.235	1.425	0.281
40	0.0625	0.1212	0.093	1.424	-0.401	2.79	-0.671

Teniendo esto en mente, procedemos a analizar los resultados del modelo tradicional y la variante Xie-Pimentel. Los parámetros M y m por el modelo Coale-Trussell son 0.963 y 0.526 para Santa Marta, mientras que para San Juan del Cesar M=1.531 y m=0.47. Niveles cercanos a 1 en M y superiores a 0.2 denotan el comienzo del control a la fecundidad, mucho más evidente para el caso de Santa Marta que para el de San Juan del Cesar. En cuanto a la utilización de la variante Xie-Pimentel, se corrió la prueba Kolmogorov-Smirnov, que nos muestra que con un nivel de significancia de 10% no hay suficiente evidencia para rechazar la hipótesis que los nacimientos tienen una distribución Poisson, para ambas poblaciones y podemos estimar la variación aquí propuesta. En este caso los parámetros son para Santa Marta: M=0.957 y m=0.505. Mientras que para San Juan del Cesar son: M=1.511 y m=0.431.

La consistencia de los resultados bajo ambos modelos nos llevan a concluir que para la muestra de Santa Marta y San Juan del Cesar, a partir de los datos censales de 1870, ya comenzaba a gestarse un proceso de limitación o control a la fecundidad natural en el sentido de Henry. Seguramente, si esta tendencia no se hubiera visto interrumpida por los conflictos de fin de siglo, el proceso de transición demográfica habría sido lento y paulatino para Santa Marta y San Juan del Cesar, con procesos similares a los observados para los países Europeos. Este novedoso encuentro nos motiva a indagar estos resultados para niveles regionales más grandes (como los Estados Soberanos) y diferenciando entre ciudades y sitios rurales.

Referencias

- Broström, G. (1985), 'Practical aspects on the estimation of the parameters in Coale's model for marital fertility.', *Demography* (4), 625–631.
- Coale, A. & Trussell, T. J. (1974), 'Model Fertility Schedules: Variations in the age structure of childbearing in human populations.', *Population Index* (40), 185–258.
- Hajnal, J. 1965. 'European marriage patterns in perspective', in: D.V. Glass and D.E.C. Eversley (eds), *Population in History: Essays in historical demography*, pp. 101-143.
- Hajnal, J. 1982. 'Two kinds of preindustrial household formation system', *Population and Development Review*, 8:449-494.
- Henry, Louis (1961). "Some data on natural fertility", *Eugenics Quarterly* 8: 81–91.
- Lebart, M. (1995), '*Statistique exploratoire multidimensionnelle*', Dunod.
- Laslett, P. (1965), 'The World We Have Lost'
- Ohlin, G., 1961 'Mortality, marriage, and growth in Pre-Industrial Populations', *Population Studies*. 14:190-197.
- Xie, Y. & Pimentel, E. (1992), 'Age patterns of marital fertility: Revising the Coale-Trussell Method.', *Journal of the American Statistical Association* 87(420), 977–984.