

II Congreso de la Asociación Latinoamericana de Población

Guadalajara, México, 3 – 5 de Septiembre de 2006

**La demografía latinoamericana del siglo XXI
Desafíos, oportunidades y prioridades**

El estudio de migraciones en Latinoamérica a través del análisis geográfico de apellidos

Pablo Mateos

University College London
p.mateos@ucl.ac.uk

Paul Longley

University College London

Richard Webber

University College London

El estudio de migraciones en Latinoamérica a través del análisis geográfico de apellidos

Pablo Mateos¹, Paul Longley², y Richard Webber³

^{1,2,3} Department of Geography /Centre for Advanced Spatial Analysis (CASA)

University College London, Reino Unido

¹ p.mateos@ucl.ac.uk, ² plongley@geog.ucl.ac.uk; ³ richardwebber@blueyonder.co.uk

Palabras Clave: Análisis de Apellidos; Geografía de las Migraciones; Métodos Cuantitativos; Demografía Histórica; Poblamiento de América

1. Introducción

El análisis cuantitativo de los episodios históricos en el poblamiento de América, de las migraciones históricas y actuales, y del impacto de ambos en la estructura y distribución geográfica de la población actual, ha estado muy limitado debido a una escasez de fuentes de datos estadísticos apropiadas para el estudio de poblaciones completas.

Esta comunicación presenta una técnica de investigación geodemográfica; el análisis del origen de nombres y apellidos para el estudio de las migraciones y estructuras de la población, y demostrar su utilidad en su aplicación a la población de América Latina. Esta técnica ha experimentando un auge en los últimos años a raíz del incremento en la disponibilidad de información digitalizada sobre frecuencias de nombres y apellidos georreferenciadas. las mencionadas tendencias;

El análisis de la estructura de la población a través de los nombres y apellidos tiene una relativamente larga historia en estudios de genética y de salud pública (Jobling, 2001), y más recientemente el origen de los apellidos se ha utilizado para otorgar una probabilidad de origen étnico a la población en estudios demográficos en ausencia de otros datos (Lauderdale and Kestenbaum, 2000, Nanchahal et al., 2001). Esta técnica puede ser también utilizada para el estudio del origen (étnico-geográfico-lingüístico) de poblaciones completas, bien para comprender procesos actuales de inmigración o también para analizar procesos históricos de poblamiento, que como en el caso de los países de América Latina han dejado un registro en la distribución actual de nombres y apellidos. En este estudio se analizan los apellidos de los siguientes países; México, Argentina, Venezuela, España y Estados Unidos (principalmente poblaciones hispanas).

A través de esta investigación se pretende dar respuesta a las siguientes preguntas; ¿Es el estudio del origen de los nombres y apellidos una técnica adecuada para el estudio cuantitativo de migraciones históricas en América Latina? ¿Qué patrones diferenciales existen entre la distribución de frecuencias de apellidos en los países estudiados? ¿Qué tipologías de apellidos se pueden distinguir y qué procesos migratorios históricos puede haber detrás de ellas? ¿Cuáles son las principales limitaciones y ventajas de este método?

2. Fuentes de Información

En España se utilizó un directorio telefónico comercializado por la compañía Infobel en un CD-ROM, con 11.8 millones de abonados particulares en 2004, lo que representa el 94% de las líneas de telefonía fija existentes en ese año. En México se utilizó un listado con los 100 apellidos más frecuentes a nivel nacional y para cada Estado de la federación (31 Estados), proveniente del Registro Federal de Electores de 2006 y proporcionado por el Instituto Federal Electoral. En Argentina se obtuvo un listado de los 100 apellidos más frecuentes a nivel nacional proveniente del Padrón Electoral 2001 de la Cámara Nacional Electoral a través de un resumen facilitado por Dipierri (2005). En Venezuela se obtuvo una lista con los 40 apellidos más comunes publicada en un estudio genético (Rodríguez-Larralde et al., 2000) y proveniente del registro de electores de 1991 del Consejo Supremo Electoral. Para Estados Unidos se utilizaron dos fuentes del US Census Bureau; una lista de apellidos frecuentes del Censo de 1990 (con 88.799 apellidos y su frecuencia a nivel nacional), y una lista con “apellidos hispanos” en el Censo de 1990 (con 25.276 apellidos, frecuencia, e indicador de “hispanidad”) cuya creación es discutida en Word & Perkins (1996). Ambas fuentes están disponibles en (US Census, 2006), y se combinaron de la siguiente manera para obtener un único listado de “apellidos hispanos” en Estados Unidos. Se unieron ambas listas a través del campo “apellido” eligiendo por tanto los que aparecen en ambas. Se seleccionaron solamente aquellos apellidos que Word & Perkins (1996) consideran como “Fuertemente Hispanos” (Heavily Hispanic -categorías ‘01xx’) y “Generalmente Hispanos” (Generally Hispanic -categorías ‘02xx’), ya que el resto corresponden a apellidos mucho más mezclados con la población general de Estados Unidos y presentan extremas alteraciones de transcripción o corresponden con apellidos de origen Portugueses, y en todo caso con un bajo porcentaje de respuestas a la pregunta de “origen hispano” en el Censo de 1990. Como resultado se obtuvieron 4.745 apellidos, para los que se calcularon las frecuencias relativas dividiendo el número de “hogares” con cada apellido (única información de frecuencia disponible) entre el número total de hogares hispanos (141.115 hogares “fuertemente y generalmente” hispanos). De esta manera se pudo elaborar una lista de apellidos hispanos con frecuencias relativas para la sub-población hispana en Estados Unidos, de manera que pudiera ser comparada en términos relativos con la población del resto de los países.

3. Metodología y Análisis

En primer lugar se procesaron las distintas fuentes de información de nombres y apellidos de los 6 países generando tablas de frecuencias de apellidos a nivel nacional. Con dichas tablas construyó una base de datos de apellidos que se procesó con el paquete estadístico SPSS y en un Sistema de Información Geográfica (ArcGIS). En la Figura 1 se representan las curvas de frecuencia acumulada para los 5 países, distinguiendo entre la población general de Estados Unidos y la población Hispana (resultando en 6 variables). Se calcularon correlaciones entre la distribución de cada apellido en cada país, para aquellos apellidos en la base de datos con presencia en más de un país. Los resultados se representaron mediante gráficos de dispersión que se ofrecen en las Figuras 2 y 3. Asimismo, se realizó un análisis de clustering con el método *K-means* para clasificar los apellidos en un total de 5 clusters en función de sus frecuencias en cada país.

4. Resultados

La Figura 1 muestra el gráfico de frecuencia acumulada de los 100 apellidos más frecuentes en los 5 países más “Estados Unidos – Población Hispana” como sexta variable, ordenados según la posición de cada apellido en el ranking del país. Se puede constatar la alta asimetría de la distribución de apellidos en todos los países, ya que solamente los 10 primeros apellidos de cada país, que suponen alrededor de 0.005% del universo nacional de apellidos, representan entre el 5% y el 25% de la

población de cada país. Este es un hecho bien conocido en estudios cuantitativos de apellidos y se ha demostrado inherente al sistema de apellidos y en gran medida es independiente del país (Tucker, 2002). No obstante, las diferencias entre países son claras, y denotan una significativa mayor mezcla de tipos de apellidos en Argentina y Estados Unidos, que en el resto de países. La cercanía entre las curvas de México y de Estados Unidos- Población Hispana refleja el hecho de que esta última está principalmente compuesta por emigrantes del país vecino. En el gráfico superior derecho de la Figura 2, pueden verse los apellidos en que las frecuencias son más altas en Estados Unidos que en México, por lo que han de provenir de otros países de habla hispana con presencia en Estados Unidos, y donde son más bajas pueden representar a zonas de México con una emigración relativa menor.

El análisis de clusters permitió clasificar un total de 3000 apellidos con frecuencia en dos o más países en 5 grupos o clusters según las diferencias de sus frecuencias en cada país. Los resultados (que no se ofrecen en el texto por limitaciones de espacio) indican que al menos 3 de estos clusters si son significativamente distintos entre sí; a) apellidos con alta frecuencia en España y no en ninguno de los demás países, b) apellidos Latinoamericanos más frecuentes en varios países que en España, y c) apellidos Argentinos sólo presentes en la población general de Estados Unidos (No hispanos). Estos tres grupos representan a comunidades originarias de población suficientemente distinguibles; la primera aquellos colectivos o regiones de España que no emigraron a América, la segunda los que sí lo hicieron (o sus apellidos) y con una expansión mucho mayor que en España, y la tercera emigrantes europeos a Argentina y Estados Unidos distintos de españoles.

5. Conclusión

El análisis aquí presentado demuestra la utilidad de una técnica de clasificación apellidos en grupos de origen para América Latina, donde la gran mayoría de los nombres han sido heredados, junto con parte de sus poblaciones, de otras partes del mundo en los últimos siglos. Los procesos migratorios, tanto históricos como actuales, pueden ser parcialmente desentrañados mediante el análisis de la distribución geográfica actual de los nombres y apellidos de manera comparada entre regiones y países. Este estudio se ha limitado al estudio de frecuencias de apellidos a nivel nacional para los apellidos más comunes, analizando a grandes rasgos cuales son las diferencias y similitudes entre las características actuales de los apellidos en 5 países, y qué procesos migratorios, de poblamiento y de influencia cultural pueden estar detrás de ellas. Ulteriores análisis tomando todos los apellidos y a nivel regional dentro de cada país, arrojarán sin duda resultados más desagregados y precisos que permitirán desglosar los principales patrones aquí encontrados, en procesos concretos dentro de cada país, a una escala espacio-temporal mucho más precisa que la realizada aquí.

Entre las principales limitaciones de este método se encuentra la dificultad para obtener acceso a registros de población a nivel suficientemente desagregados para varios países. Asimismo, como es bien sabido los apellidos se heredan generalmente vía paterna y por tanto sólo representan a una parte de la historia de la sociedad. Además, los procesos originales de generación o asignación de apellidos en América difieren mucho entre países, regiones y momentos históricos, lo cual no permite una generalización para todos los portadores actuales de un mismo apellido. No obstante, las diferencias entre apellidos y regiones pueden también revelar también indicios esta rica y compleja faceta de la historia de la población.

6. Referencias bibliográficas

- DIPIERRI, J. E., ALFARO, E. L., SCAPOLI, C., MAMOLINI, E., RODRIGUEZ-LARRALDE, A. & BARRAI, I. (2005) Surnames in Argentina: A population study through isonymy. *American Journal of Physical Anthropology*, 128, 199-209.
- JOBLING, M. A. (2001) In the name of the father: surnames and genetics. *Trends in Genetics*, 17, 353-357.
- LAUDERDALE, D. & KESTENBAUM, B. (2000) Asian American ethnic identification by surname. *Population Research and Policy Review*, 19, 283-300.
- NANCHAHAL, K., MANGTANI, P., ALSTON, M. & DOS SANTOS SILVA, I. (2001) Development and validation of a computerized South Asian Names and Group Recognition Algorithm (SANGRA) for use in British Health-related studies. *Journal of Public Health Medicine*, 23, 278.
- RODRIGUEZ-LARRALDE, A., MORALES, J. & BARRAI, I. (2000) Surname Frequency and the Isonymy Structure of Venezuela. *Am. J. Hum. Biol.*, 12, 352-362.
- TUCKER, K. (2002) Distribution of Forenames, Surnames, and Forename-Surname Pairs in Canada. *Names*, 50, 105-132.
- US CENSUS (2006) US Census Bureau Genealogy Resources. <http://www.census.gov/genealogy/www/>, Accessed 12/05/2006
- WORD, D. L. & PERKINS, R. C. (1996) Building a Spanish surname list for the 1990s a new approach to an old problem. *Technical Working Paper 13*. Washington DC, US Census Bureau, Population Division, <http://www.census.gov/population/documentation/twpno13.pdf>, Accessed 29/05/2005

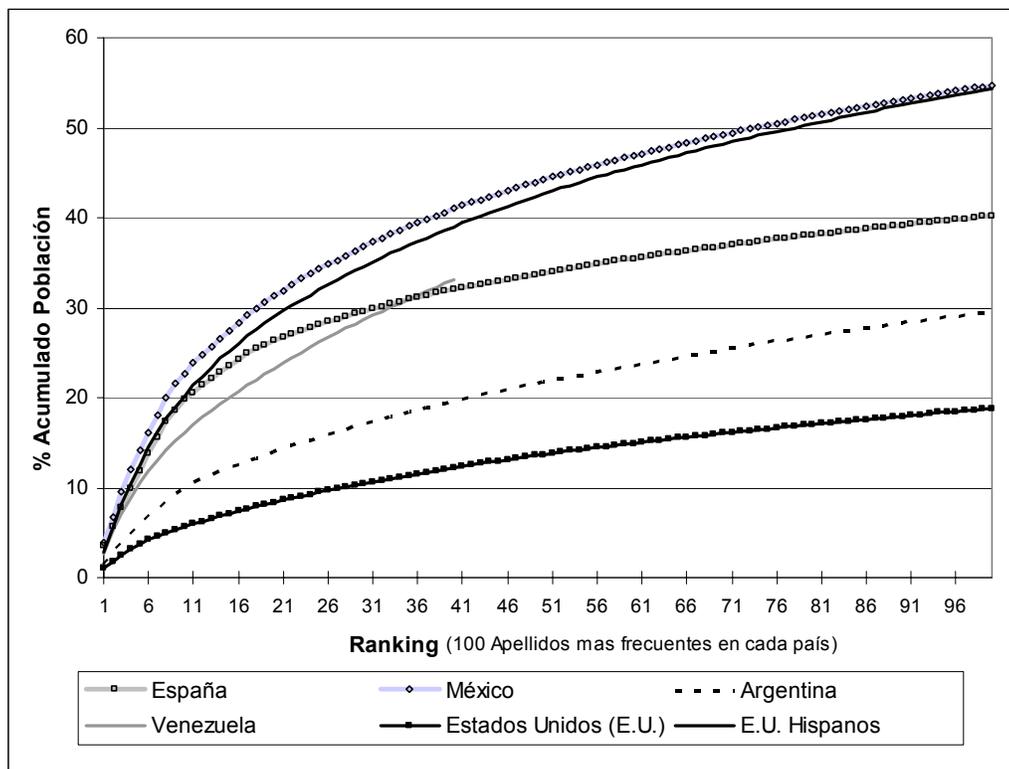


Figura 1: Frecuencia acumulada de los 100 apellidos más frecuentes en varios países de habla hispana. Para Venezuela solamente se incluyen los 40 apellidos más comunes.

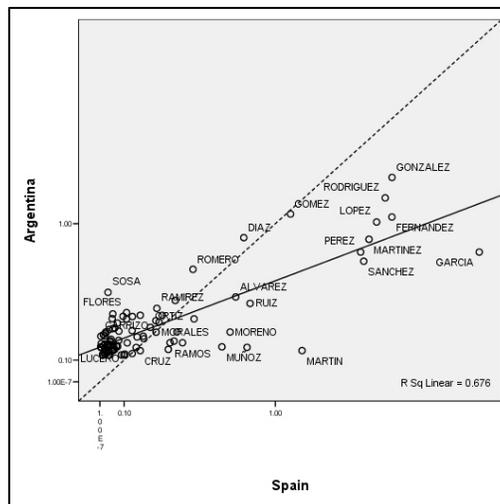
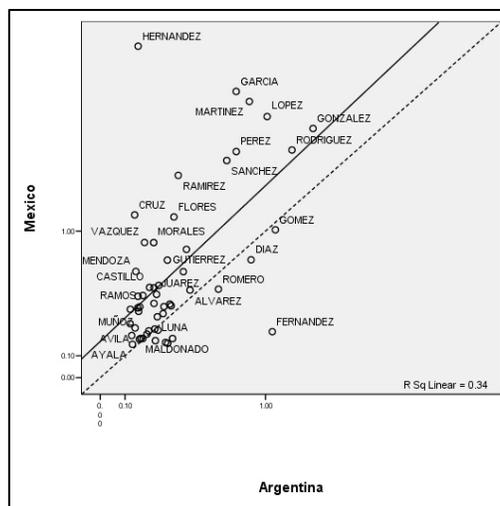
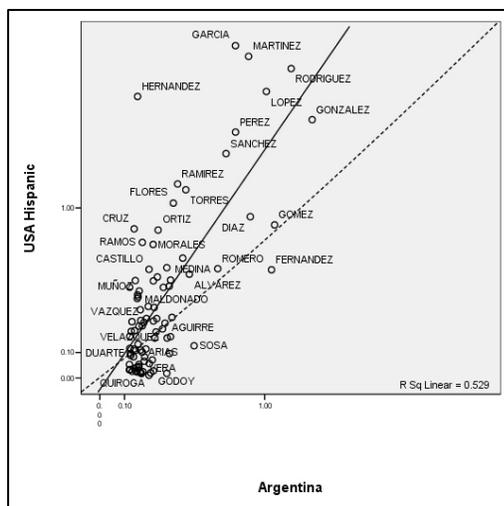
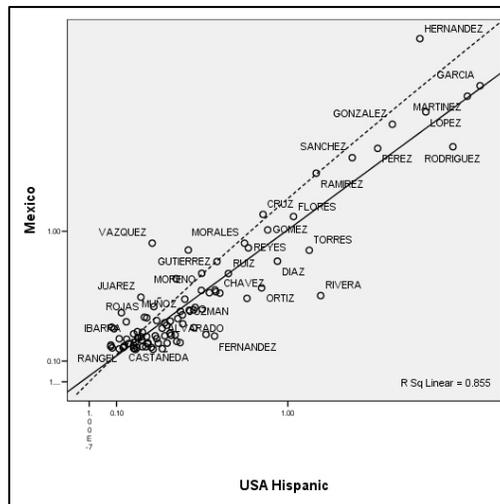
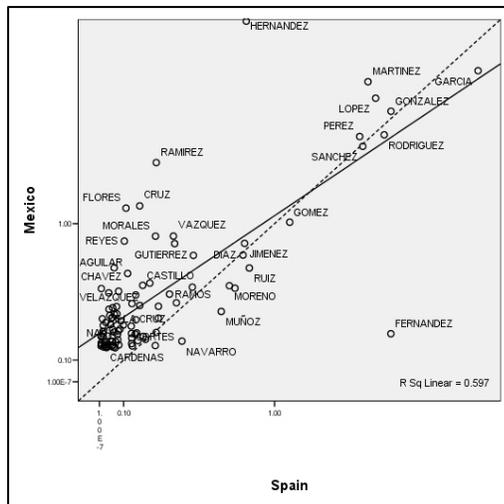


Figura 2: Gráficos de Dispersión entre frecuencias de apellidos por partes de países (I). Se utiliza una escala logarítmica (log-log). La línea punteada indica la diagonal de correlación donde todos los puntos tendrían la misma frecuencia en ambos países. La línea de color sólido es la recta de regresión lineal entre todos los puntos. Se ofrece el coeficiente R^2 de dicha regresión en la parte inferior derecha de cada gráfico.

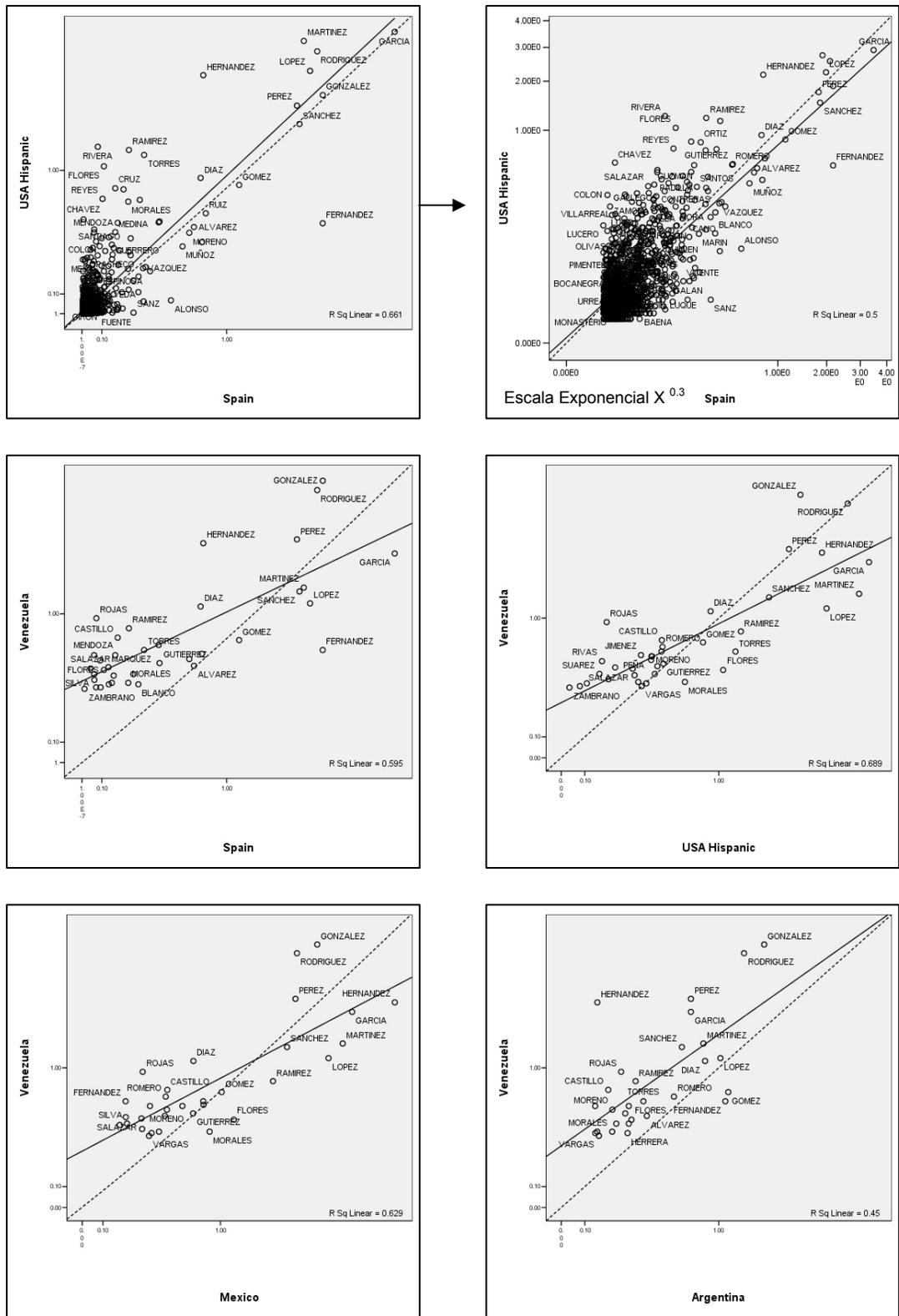


Figura 3: Gráficos de Dispersión entre frecuencias de apellidos por partes de países (II).

Se utiliza una escala logarítmica (log-log). La línea punteada indica la diagonal de correlación donde todos los puntos tendrían la misma frecuencia en ambos países. La línea de color sólido es la recta de regresión lineal entre todos los puntos. Se ofrece el coeficiente R^2 de dicha regresión en la parte inferior derecha de cada gráfico.