

DE GENTES ANONIMAS Y ARCHIVOS NOMINALES; UNA EXPERIENCIA DE INTEGRACION ENTRE ANALISIS ESTADISTICO Y PROSOPOGRAFICO

*Mario Samper K.**

I. Contexto historiográfico

En la Costa Rica de los años setenta, como en Francia y otras partes del mundo académico durante el medio siglo anterior, la investigación histórica comenzó a alejarse de la mera reconstrucción narrativa de acontecimientos únicos e irrepetibles y de una visión que privilegiaba las acciones de los personajes destacados. Buscamos nuestra "cientificidad" en el estudio de procesos estructurales, tendencias y fluctuaciones, o grandes conglomerados humanos. Esto, a su vez, requirió de procedimientos descriptivos y demostrativos apegados -con mayor o menor rigor e inventiva- a la ló-

* Doctor en Historia, profesor e investigador en la UCR y en la UNA.

gica estadística, y condujo a la búsqueda, recopilación y utilización de datos seriables o al menos de información agregada comparable a lo largo del tiempo. Para ello, se recurrió primordialmente a métodos cuantitativos para el tratamiento de la información, a la definición precisa de variables y su medición, a la construcción de series temporales y cortes sincrónicos, a la búsqueda de regularidades y rupturas, y a la demostración fundamentada en cifras. Con poca frecuencia, pero con una trascendencia mayor, se introdujeron a la pesquisa histórica modelos de comportamiento social más o menos formalizados, se avanzó hacia el análisis comparado sistemático con parámetros objetivos y se fundamentó la explicación del cambio social en el estudio de interrelaciones entre múltiples variables cuantificadas.

La renovación de nuestros modos de historiar conllevó, pues, la adopción de un determinado paradigma científico, un conjunto de innovaciones teórico-metodológicas fundamentales y una redefinición del quehacer profesional de historiadores e historiadoras. Aunque no siempre fue aprovechado todo el potencial de este tipo de investigación, pues a menudo nos limitamos a la descripción estadística más elemental, la cuantificación histórica comenzó a demostrar su utilidad para introducir una mayor rigurosidad en el planteamiento de los problemas, en el diseño de estrategias metodológicas y en el modo de responder a las preguntas planteadas. El recurso de la cuantificación, con sus respectivos supuestos epistemológicos explícitos o implícitos, fue incorporado a la historia, al igual que lo había sido anteriormente a diversas ciencias naturales y algunas de las sociales, como herramienta metodológica fundamental para la prueba de hipótesis mediante la recopilación, sistematización y procesamiento de datos medibles y verificables. Con ello habría de superarse, en el estudio de la historia, tanto el narrativismo meramente descriptivo como el subjetivismo interpretativo. Al mismo tiempo, se procuraba establecer plenamente la intersubjetividad del conocimiento histórico, vale decir, la posibilidad de que otros investigadores pudiesen corroborar no solamente la precisión de fechas, lugares y personajes, sino también los procedimientos estadísticos y analíticos empleados, así como la validez de las explicaciones propuestas.

La historia serial y, de modo más general, las aplicaciones de la estadística a la historia, tendieron a difundirse como modalidades de procesamiento de información sobre el pasado social. Si bien su utilización fue inicialmente ilustrativa y luego descriptiva, poco a poco se avanzó hacia la demostración cuantitativa como elemento central de estrategias de investigación construidas sobre el modelo hipotético-deductivo. El énfasis en el dato numérico se tradujo, asimismo, en formas de presentación de los resultados de la pesquisa, típicamente mediante la combinación de texto analítico-descriptivo, gráficos y cuadros, u ocasionalmente recurriendo a coeficientes diversos y -en forma excepcional- a modelos matemáticos formalizados.

Con el advenimiento de la cuantificación histórica en nuestro país se logró, ciertamente, relegar a un segundo plano (pero no desplazar del todo, por su función sociopolítica e ideológica medular y por la perdurable resistencia al cambio dentro del gremio) el relato de acontecimientos al estilo denominado -equivocamente- "positivista". Las biografías de próceres y otras personalidades eminentes -casi siempre masculinas, blancas y pertenecientes a la élite valle-centralina- cedieron su lugar preeminente (pero no su mercado) a los estudios sobre procesos económicos y sobre grandes conglomerados sociales. Se comprendieron mejor las estructuras productivas y la composición de la población, las tendencias y las fluctuaciones macrosociales, aunque hubo pocos progresos en el esclarecimiento del devenir microsocial (por ejemplo, en el estudio histórico de procesos decisivos al interior de las unidades productivas, de las familias, de las comunidades o del sistema político).

Mediante la cuantificación sistemática, la historia -pensada como saber en proceso de cientifización- ganó en precisión y en rigor demostrativo. Se incorporaron áreas temáticas fundamentales que habían sido descuidadas hasta entonces, aunque también se descuidaron algunas otras, o se dejaron en manos de investigadores metodológicamente tradicionales, algunos de ellos engañosamente jóvenes. La historia económica y social, sobre todo, adquirió, por medio del análisis cuantificado una mayor profundidad interrogativa y capacidad de explicación, superiores sin duda alguna a

la historia de viejo cuño, centrada en acontecimientos y figuras.

No cabe hacer aquí una referencia detallada a los numerosos estudios que desde los años setenta aplicaron, de alguna manera y con mayor o menor suceso, procedimientos estadísticos al estudio de la historia económica y social. Cualquier lista tendría que incluir una serie de tesis de licenciatura y otras investigaciones centradas en sectores productivos específicos, como el tabaco o el cacao, la producción bananera, cañera o cafetalera, la ganadería, la agroindustria y la minería. Abarcaría, asimismo, los estudios agrarios regionales, los recuentos demográficos locales y sobre todo los esfuerzos por interrelacionar variables socioeconómicas y poblacionales, las historias de empresas de diversa índole y los análisis históricos de relaciones mercantil-crediticias. Contemplaría, sin duda, las reconstrucciones históricas del comercio exterior y en general del comportamiento macroeconómico, como también los de tenencia de la tierra, transacciones inmobiliarias y concentración del capital, al igual que las caracterizaciones más generales de la estructura social urbana y rural. Debe anotarse, sin embargo, que la mera presentación de cuadros estadísticos no supone la adopción del paradigma científico asociado a la cuantificación, pues a menudo se trata más bien de ilustraciones numéricas de un discurso fundamentalmente narrativo, cuya finalidad es la descripción monográfica de una secuencia de acontecimientos.

En contraste con la profusión de monografías locales, regionales o nacionales que recurren en forma importante a la estadística, han sido pocos los estudios comparados ya de regiones costarricenses o respecto de otros casos pertinentes. Menos, aun, son los que incorporan en forma significativa la cuantificación para precisar los términos de la comparación o para explicar convergencias y divergencias.²

Al empleo de métodos cuantitativos en la investigación histórica contribuyó, en forma importante, el diálogo interdisciplinario que condujo a la incorporación de modelos teórico-metodológicos e instrumentales técnicos de otras ciencias sociales, como la economía y la demografía. En Costa Rica, algunos economistas y profesionales en otras

disciplinas, además de hacer contribuciones relevantes al conocimiento de nuestra historia, aportaron saludables cuestionamientos teórico-metodológicos a los procedimientos estadísticos, analíticos y demostrativos empleados en algunas investigaciones sobre historia económica y social.³

Cabe advertir, sin embargo, que los avances de la cuantificación histórica en nuestro medio han sido desiguales y limitados en sus alcances. Sus logros principales han sido en historia de la población (estudios parroquiales agregados, modelos de comportamiento demográfico, migraciones),⁴ en historia económica (coyunturas y tendencias, comercio exterior, sectores productivos)⁵ y en historia social (estructuras y procesos de diferenciación, tenencia de la tierra, grupos ocupacionales).⁶ Aun en esos tres campos -con algunas excepciones notables- su papel ha sido con frecuencia descriptivo y estadísticamente elemental, e incluso en ocasiones poco riguroso en los procedimientos y en el análisis de los datos resultantes.

En historia política o en la de mentalidades y vida cotidiana los progresos de la cuantificación sistemática en Costa Rica son todavía más limitados, circunscribiéndose a usos rudimentarios como el análisis comparado de datos electorales y socioeconómicos por cantones, o al análisis de frecuencias absolutas y relativas de comportamientos delictivos, bienes culturales u otros.⁷

En síntesis, la cuantificación histórica se ha limitado las más de las veces al empleo descriptivo e ilustrativo de la estadística. Más allá de este uso bastante generalizado, sobre todo en la historia económica y social, son pocos los ejemplos de prueba cuantitativa de hipótesis, y aun menos los de procesamiento multivariado.⁸

Algunas de las limitaciones de los estudios históricos costarricenses basados sustancialmente en información estadística se derivan de la falta de claridad sobre el propósito y papel de la cuantificación, como también de la vaguedad de las hipótesis de trabajo, de la falta de correspondencia entre objetivos y estrategias metodológicas, y -todavía- del escaso dominio de las técnicas respectivas. Con frecuencia, aun en el empleo de estadísticas elementales, se cometen serios errores básicos por la aplicación de procedimientos inapro-

piados (v.g. los falsos promedios o las estadísticas derivadas de muy pocos casos) o por fallas en la interpretación (confusión entre asociación y causalidad, atribución de significado a diferencias no significativas, etc.).

Los primeros pasos de la cuantificación histórica se dieron mediante engorrosos procedimientos apoyados en rudimentarias máquinas calculadoras. Pronto se contó -para proyectos de cierta envergadura y con suficiente respaldo financiero- con las grandes computadoras tipo "main-frame", inicialmente de tarjeta perforada, y cuyo uso requería de entrenamiento especializado. Las aplicaciones estadísticas desarrolladas para dichos equipos, y posteriormente adaptadas a las nuevas microcomputadoras, estaban fuertemente sesgadas hacia el manejo de grandes volúmenes de datos cuantitativos o, al menos, codificados numéricamente.

El abaratamiento y difusión de los ordenadores personales, aunados a la relativa facilidad de uso de los "paquetes" estadísticos (tipo SPSS), como también de versátiles programas graficadores, permitieron el acceso de un creciente número de investigadores a estas tecnologías y procedimientos computadorizados. A la vez, crearon el riesgo muy real de su utilización sin pleno conocimiento de la lógica estadística implícita en ellos, y sin una clara comprensión de los alcances y limitaciones de los resultados numéricos o gráficos así obtenidos. Con demasiada frecuencia, pese a la aparente precisión numérica de los cuadros y a la excelente presentación de los gráficos, los procedimientos demostrativos son imprecisos, las asociaciones entre variables no se exploran adecuadamente, y las conclusiones no se desprenden de la información presentada.

Aunque la cuantificación es indudablemente un poderoso instrumento demostrativo para ciertos fines de investigación histórica, la fascinación acrítica por las cifras ha producido engañosas imágenes de exactitud numérica artificial y abundantes falacias argumentales. Por otra parte, su utilización juiciosa, bien informada y acompañada tanto de sentido común como de claridad analítica, ha abierto la posibilidad de saltos cualitativos en el estudio histórico de la población y de la producción, de las es-

estructuras y fuerzas sociales, e incluso -en menor grado hasta ahora- del comportamiento político colectivo y de ciertos planos socioculturales.

La búsqueda de cientificidad en los estudios históricos, con todo lo que ha representado efectiva o potencialmente en rigor y sistematización, no fue gratuita: su precio, en el corto plazo, fue marcado por la “despersonalización” de la historia. Esta devino el producto de fuerzas anónimas que actuaban -sabiéndose tales o sin saberlo- como conglomerados: “poblaciones”, “clases sociales” o “categorías socioprofesionales”, “movimientos reivindicativos” o “electorados”. Las masas traspasaron el umbral de la historia escrita, pero sus nombres y apellidos -sus dimensiones incuantificables, sus vivencias individuales y sus relaciones interpersonales, sus motivaciones subjetivas, sueños y frustraciones- no se consignaron al entrar a la casa de Clío. Lo que más nos ha interesado de ellas es precisamente su comportamiento colectivo, estadísticamente analizable; su acción productiva o reproductiva agregada; sus conductas medibles y susceptibles de tratamiento cuantitativo.

La cuantificación ha sido, es y seguirá siendo importante para el estudio de nuestra historia. Lejos de abandonarla porque algunas veces se ha abusado de ella, ya por inexperiencia en su manejo o por ligereza interpretativa, debemos -todos nosotros- aprender a utilizarla mejor para abordar con seriedad problemas de investigación que difícilmente pueden resolverse sólo a escala microsocia o mediante procedimientos cualitativos. Pero los métodos cuantitativos, por sí solos, son insuficientes para responder a las preguntas medulares e, incluso, para plantearlas. Siempre hay, o debiera haber, un contrapunteo entre lo cuantificado y lo cualitativo en el proceso de investigación, aunque algunas veces prive lo uno o lo otro según la naturaleza del problema, las fuentes disponibles, el enfoque metodológico o las preferencias y habilidades de quien investiga.

En el decenio de los ochentas, a la vez que aplicamos con mayor o menor destreza determinados procedimientos estadísticos y recurrimos de modo cada vez más generalizado a la microcomputación, se produjo una reacción en sentido inverso. Por inclinación personal o por convicción, por

recelo de las matemáticas o por desconfianza de resultados obtenidos mediante la reducción de fenómenos sociales complejos a conjuntos de variables, por rechazo explícito al cuantitativismo o por el interés de rescatar testimonios populares, comenzó a desarrollarse una línea de investigación basada primordialmente en métodos cualitativos.

La irrupción de la investigación cualitativa en el ámbito de la historia costarricense conllevó, de hecho, un rechazo -generalmente implícito- de la cuantificación como herramienta para la verificación de hipótesis y como técnica para el procesamiento de la información. También -aunque no siempre de modo expreso ni consciente- significó una ruptura con los supuestos paradigmáticos acerca de la "cientificidad" de la historia, y condujo a la búsqueda de maneras alternativas de investigar el pasado-presente de nuestra sociedad. La historia oral, y especialmente la técnica de las historias de vida, fue el basamento de una manera de indagar sobre lo pretérito que cuestionó, más en la práctica investigativa que en la reflexión epistemológica, al paradigma cuantitativista.⁹

Paralelamente, se desarrollaron proyectos con entrevistas semi-estructuradas, ya como componente principal o como pieza importante en el engranaje de la investigación histórico-social. En tales casos, el testimonio y el tratamiento cualitativo de la información fueron, las más de las veces, técnicas aplicadas para un fin específico de acuerdo con los objetivos de la pesquisa, sin que hubiese un rompimiento radical ni con el paradigma hipotético-deductivo ni con el recurso a la estadística. De hecho, sin que se explicitasen sus bases epistemológicas, tendieron a combinarse técnicas cualitativas y cuantitativas en forma un tanto ecléctica, aunque ciertamente eficaz.¹⁰

Como en el caso de la cuantificación, se trataba de la adopción demorada pero localmente innovadora de técnicas y enfoques metodológicos desarrollados en la primera mitad del siglo por sociólogos, etnólogos y otros investigadores sociales norteamericanos y europeos, y enriquecidos luego por contribuciones específicamente latinoamericanas en campos como la investigación participativa y el rescate de memoria popular.

Sin que se explicitaran claramente las diferencias conceptuales y epistemológicas, la historia oral costarricense -como la de otros lares- tendió a separarse de la que procuraba su cientificidad por vía hipotético-deductiva y mediante la cuantificación. La diferenciación fue patente en lo concerniente a proyectos individuales y colectivos, docencia metodológica, publicaciones y otras esferas de la vida académico-profesional. Con ello se abrieron nuevos derroteros temáticos y metodológicos para la investigación, y se planteó la posibilidad -hasta ahora poco aprovechada- de confrontar perspectivas epistemológicas y enfoques investigativos. Al mismo tiempo, el sesgo implícitamente anti-estadístico impidió a quienes optaron por la historia oral o por la investigación cualitativa con documentos testimoniales, en forma excluyente, aprovechar desde un principio las oportunidades de retroalimentación entre ambos enfoques, y limitó en consecuencia la complementariedad creativa -según los requerimientos de cada estudio- entre los modos cuantitativos y cualitativos de investigar.

En algunas otras disciplinas como la investigación educativa o quizás en la sociología, hubo también en nuestro medio una ruptura o cuando menos una divergencia entre enfoques que privilegiaban el testimonio o el estudio de casos a profundidad, y otros que priorizaban la cifra o el tratamiento cuantitativo de grandes volúmenes de información. Ciertas disciplinas como la antropología y en particular la etnografía continuaron siendo básicamente cualitativas, mientras que algunas otras -v.g. la demografía y la psicología experimental- tuvieron sesgos fuertemente cuantitativos en su enfoque metodológico.

La reivindicación del testimonio y de lo cualitativo en la investigación sobre el pasado más o menos reciente -dos o tres generaciones, salvo en casos donde hubiese una fuerte tradición oral comunitaria- se expresó principalmente en una serie de experiencias con métodos y técnicas como la entrevista a profundidad, la semi-estructurada y en general la historia oral. Bajo esta nueva modalidad, se recopiló por diversos medios y para distintos fines una serie de historias de vida y testimonios varios de artesanos, agricultores y otros grupos ocupacionales, de etnias autóctonas e inmi-

grantes, o más recientemente de los ex-combatientes de la guerra civil. La autobiografía adquirió, asimismo, renovada prominencia como documento oral o escrito para la historia política y social.¹¹

En ciertos proyectos pioneros como los concursos de autobiografías campesinas y obrero-artesanales, que emulaban las experiencias polacas de varias décadas atrás, se trataba fundamentalmente de reunir narraciones de la propia vida, tal como quisieron contarla quienes decidieron escribirlas o -en algunos casos- relatarlas oralmente.

Algunas veces, el recurso a informantes se constituye en el modo principal de recabar información y explorar las vivencias de quienes presenciaron o participaron en coyunturas decisivas o situaciones y cambios que interesa analizar. Otras veces, los testimonios orales son una fuente importante pero complementada por otras, sobre todo de índole documental. El primer tipo de estudios recurre usualmente a las historias de vida, profundiza los recuerdos y valoraciones de uno o varios informantes destacados, y evidencia la riqueza de este tipo de entrevista a profundidad para ahondar en experiencias personales, percepciones, motivaciones y evaluaciones retrospectivas de participantes o testigos de procesos históricos relevantes. El segundo tipo de estrategia tiende a relacionar más directa y estrechamente a las fuentes orales y documentales, recurre a un mayor número de informantes seleccionados a menudo con criterios tipológicos, y suele emplear entrevistas semi-estructuradas u otras técnicas para recabar un conjunto de testimonios sobre aspectos específicos atinentes a la problemática estudiada. Profundiza menos en los casos individuales, a fin de lograr una mayor cobertura social; se apoya en el análisis comparado y procura integrar las perspectivas micro- y macro-analíticas.

El estudio histórico por vía cualitativa también se trajo -más tardía y limitadamente- en el uso sistemático de documentos testimoniales de períodos anteriores, así como en el estudio de casos a profundidad y la prosopografía.¹² Para ello se recurrió al seguimiento de individuos y familias en múltiples fuentes, como también al desarrollo de tipologías y a la sustitución del concepto de representatividad es-

tadística por el de representatividad cualitativa. Casi subrepticamente, sin una discusión teórico-metodológica abierta, comenzaron a replantearse las bases epistemológicas del estudio de la historia. Algunos supuestos del modelo hipotético-deductivo fueron abandonados, entre ellos el "principio fundador" de la prueba cuantificada de hipótesis, y el paradigma cuantitativo se vio sometido de hecho a un cuestionamiento saludable aunque no siempre bien entendido por los practicantes de una u otra manera de investigar. Se esbozaron ciertas propuestas para el desarrollo de la historia oral e, implícitamente, de formas cualitativas de investigación sobre el pretérito social. Se habló incluso de investigación-acción y de recuperación participativa de la memoria popular, aunque apenas se han gestado unas pocas iniciativas de este tipo. No ha resultado fácil que proyectos de investigación concebidos desde la academia y generados por una entidad de educación superior devengan realmente participativos, pese a la voluntad expresa de avanzar en esa dirección. Lo que sí se ha incrementado es la preocupación por devolver conocimiento a sectores específicos de la sociedad mediante esfuerzos divulgativos y -en algunos casos- actividades en las cuales es factible un mayor grado de participación.

Si bien en Costa Rica no se explicitó claramente el trasfondo epistemológico y social de la cuestión, en otras latitudes venían planteándose como opciones excluyentes el recurso sistemático a la cuantificación y al dato anónimo, de un lado, o a la fuente oral y el documento testimonial por otro. En tal perspectiva, no se trataba simplemente de una escogencia de fuentes o técnicas en función de cierto problema de análisis, sino de formas contrapuestas de investigación social. De un lado, el modelo hipotético-deductivo de las ciencias, y un rol protagónico del investigador; del otro, la investigación cualitativa, más exploratoria e interactiva, que procuraría redefinir la relación entre sujeto y objeto de estudio, hasta desarrollar formas de investigación participativa. La dicotomía, llevada al extremo, hizo difícil relacionar clara y sistemáticamente los planos micro- y macro-sociales, y evidenció la necesidad de modelos teórico-metodológicos, estrategias investigativas y estudios concretos que permitiesen esclarecer dicha interrelación.

Claro está que la escisión metodológica a que hemos hecho referencia no fue completa y, de hecho, ha continuado la exploración de formas no dicotómicas de investigación histórico-social. En Costa Rica hubo una serie de proyectos específicos en los cuales se procuró combinar de modos más o menos creativos, aunque ciertamente eclécticos, el recurso a la cuantificación y la indagación de índole cualitativa. No siempre se logró su plena imbricación en estrategias que permitiesen una eficaz retroalimentación entre ambas modalidades, más yuxtapuestas que entrelazadas. En historia, no se trataba en modo alguno de una “triangulación” sistemática (a veces llamada “metodología convergente”), sino más bien de la pervivencia de una fuerte tradición metodológica que valoraba el recurso a fuentes históricas diversas y su tratamiento en formas complementarias, sin adscribirse necesariamente a un modelo determinado sobre lo que debía ser la investigación. A fin de cuentas, los historiadores nos consideramos artesanos, cada quien tiene “su método”, mezcla los ingredientes a su manera para lograr determinada “sazón”. Además, nos nutrimos del buen o mal ejemplo de nuestros mentores en última instancia, autores de las grandes obras de la historiografía universal, muchos de los y las cuales han recurrido frecuentemente al más franco y fructífero eclecticismo en el terreno de las estrategias investigativas. El análisis cuantitativo y cualitativo, en particular, han sido integrados de modos diversos y creativos, aun por quienes se inclinan en una u otra dirección metodológica por sus preferencias personales, objetivos específicos o fuentes disponibles.

En Costa Rica, sin alcanzar obviamente tal brillantez, la combinación más o menos empírica y limitada de datos cuantificados y testimonios orales o documentales ha sido frecuente en tesis de licenciatura o maestría en historia. Unas pocas de ellas, como también determinadas tesis doctorales realizadas por costarricenses o extranjeros y ciertas otras obras históricas -aunque rara vez explicitan su fundamentación epistemológica o detallan su estrategia investigativa- logran un mayor refinamiento y una integración más estrecha entre métodos cuantitativos y cualitativos. Las maneras de hacerlo han sido variadas e imaginativas, pero tam-

bién poco sistemáticas y difícilmente reproducibles. Fieles a nuestra tradición artesanal, cada historiadora o historiador construía a su manera individual bases de datos para uso exclusivo en su investigación, y desarrollaba una manera muy personal de integrar los resultados del procesamiento estadístico y de la exploración cualitativa. La creación de bases de datos personalizadas evidenció el potencial de la informática para agilizar la recopilación y procesamiento de los datos, y permitió efectuar análisis más complejos. Demostró, por otra parte, el riesgo de la dispersión de esfuerzos y la dificultad de "socializar" tales bases de datos, entre otras razones por la incompatibilidad entre los criterios o procedimientos aplicados.¹³

En tal contexto, a fines de los años ochenta se iniciaron algunos esfuerzos colectivos o coordinados por crear bases de datos históricas de índole numérico-nominal, con información proveniente de padrones censales u otras fuentes que reuniesen, de modo homogéneo, nombres y datos cuantitativos o cuantificables. Algunos de esos esfuerzos tuvieron su punto de partida en proyectos individuales que luego buscaron compatibilizar de alguna manera sus procedimientos de recopilación, codificación, digitado y almacenamiento de la información. El éxito en estos casos fue parcial, pues en la práctica resultó difícil o imposible conformar una base de datos unificada, aunque sí se logró cierta comparabilidad en el tipo de información y en los resultados del análisis.¹⁴ Otras iniciativas tuvieron o adquirieron un carácter más institucional, a partir de la temprana confluencia de intereses de investigación de individuos que coordinaron su labor en un equipo de trabajo formal o informal, y contaron con apoyo técnico y presupuestario para impulsar una propuesta de alcance supra-individual.¹⁵

El desarrollo informático, especialmente con microcomputadoras cada vez más potentes y programas versátiles -v.g. las bases de datos jerárquicas y relacionales- contribuyó a posibilitar el manejo más eficiente y flexible de archivos que reunían datos nominales, otras variables cualitativas y un conjunto de variables numéricas o codificadas. Concretamente, permitió al investigador moverse más fácil y ágilmente entre los casos individuales y los datos agregados.

Asimismo, simplificó desde el punto de vista técnico la creación de uno o más planos intermedios para el manejo de información tanto cualitativa como cuantitativa. Ello no sólo abre nuevas posibilidades de análisis prosopográfico a diversa escala, sino que invita a desagregar los grandes conglomerados humanos (v.g. la "clase") en las categorías y subcategorías que la conforman (v.g. los "grupos ocupacionales"); al mismo tiempo, facilita el proceso de generalización progresiva a partir de casos particulares comparados entre sí, contrastados con otros, reagrupados y confrontados nuevamente con grupos más amplios, con lo cual se evitan riesgosos y algunas veces arbitrarios "saltos" directos desde lo individual o micro-grupal hasta lo general y macro-social.

Las nuevas tecnologías informáticas no sólo posibilitaron, pues, el tratamiento rápido, eficiente y relativamente barato de grandes volúmenes de información cuantitativa, sino que permitieron explorar novedosas formas de integrar el procesamiento estadístico con los estudios de caso. Para ello habrían de ser fundamentales las bases de datos numérico-nominales, no como recurso único ni mucho menos como recetario metodológico, sino como infraestructura informática para múltiples y muy diversas estrategias de investigación histórico-social, acordes con los objetivos e intereses específicos de cada persona o equipo de trabajo.

Por otra parte, la localización de nuevas fuentes con información cuantificada y nominal, concretamente en la forma de padrones o boletas censales, invitaba a un proceso cooperativo para convertirlas en bases de información computadorizada, recopilada y digitada de modo uniforme aunque para fines diversos. Una vez creadas, se pondrían inicialmente a disposición del colectivo de investigadores que hubiesen contribuido al digitado (mediante un mismo programa de captura) y a la "limpieza" de los datos alfabéticos y cifrados. Dichas bases numérico-nominales se utilizarían tanto para el procesamiento estadístico como para el seguimiento prosopográfico, y eventualmente para el pareo o enlace de archivos. Tal enlace permitirá no sólo trascender las limitaciones de una fuente única o aislada, sino potenciar cada fuente informatizada cuyo valor para la investi-

gación se incrementará considerablemente al formar parte de un todo muy superior a sus partes.

Las bases de datos creadas del modo descrito podrían más adelante duplicarse fácilmente y -previo acuerdo interinstitucional o entre investigadores- intercambiarse para multiplicar a mediano plazo las posibilidades de entrelazamiento. A más largo plazo, tal proceso permitiría el establecimiento de un sistema de información numérico-nominal de dominio público, accesible mediante consulta directa, distribución de copias autorizadas, o incluso -quizás- por vía telemática, para su utilización remota por parte de múltiples personas o entidades.

Uno de estos esfuerzos de creación de tales bases de datos, importante por su magnitud e influencia metodológica, resultó de la confluencia de intereses de varios investigadores y estudiantes de posgrado en historia de la Universidad de Costa Rica. Tuvo raíces en otros proyectos individuales previos, pero se articuló finalmente en el proyecto "Transformación del censo de población de 1927 en base de datos nominales", que tuvo como principal apoyo institucional el Centro de Investigaciones Históricas. Posteriormente, entró a formar parte de un proceso de coordinación y colaboración entre dicho centro y la Escuela de Historia de la Universidad Nacional, con miras al intercambio y entrelazamiento de bases de datos numérico-nominales.

El presente ensayo se propone efectuar, sucintamente, un balance de la experiencia del proyecto aludido, con sus aciertos pero también con sus errores, tal como fue evaluada por quienes participamos en ella. Abarca el período desde 1988 hasta fines de 1993, y procura reflejar de alguna manera el aprendizaje de métodos y procesos por parte de un grupo cada vez más amplio de investigadores, coinvestigadores, asistentes y tesarios, vale decir, el colectivo de los autodenominados investigadores "censuales" del '27.¹⁶

II. Breve reseña del proyecto

En las páginas siguientes se explicará sucintamente la gestación del proyecto 1927, como lo designaremos en ade-

lante, así como los principales criterios y decisiones en lo concerniente a selección de zonas y muestreo, digitado de los datos y depuración de archivos.

a) Orígenes

Tal como se explicó más detalladamente en una publicación anterior,¹⁷ en octubre de 1988 se inició oficialmente, bajo la coordinación del autor, el proyecto institucional del Centro de Investigaciones Históricas (C.I.H.) para crear una base de datos numérico-nominales a partir de los padrones censales de 1927, que poco tiempo antes habían sido puestos a disposición de cualquier persona interesada en el Archivo Nacional. Dicho proyecto fue, en realidad, una reformulación de otro anterior, de índole más individual, y fruto de un proceso previo en el cual habían participado otros investigadores. Conviene reseñar este tránsito desde intereses de investigación individuales hacia un esfuerzo concertado, por su pertinencia para la creación de otras bases de datos semejantes.

Los antecedentes más directos del proyecto 1927 fueron otros dos, también inscritos en el C.I.H.: el proyecto "Estructura socio-ocupacional costarricense en 1927", iniciado en enero de 1988 y posteriormente reformulado por el autor, y el proyecto "Migración y desarrollo regional, 1883-1963", a cargo de Arodys Robles, que concluyó en setiembre de ese mismo año cuando su investigador principal salió del país a continuar estudios de posgrado. Los lineamientos generales del proyecto reformulado, junto con una serie de pasos metodológicos previos, fueron discutidos ampliamente por esos dos investigadores.

En la fase precedente de elaboración de un formulario para la recopilación de datos, el diseño del formato fue responsabilidad principal de Arodys Robles, en consulta con quien escribe. En las pruebas de versiones preliminares del mismo participaron algunos estudiantes de la Escuela de Historia de la Universidad Nacional, en el marco de una práctica metodológica dirigida por el autor. Posteriormente se incorporaron varios tesarios y tesarias de la Maestría

Centroamericana en Historia, de la Universidad de Costa Rica, cuyos aportes se especificarán más adelante.

Al mismo tiempo que se avanzaba gradualmente desde proyectos individuales a un proceso coordinado entre investigadores, con apoyo institucional, se logró aunar esfuerzos con otro proyecto metodológicamente afín, la actividad "Creación de Bases de Datos para la Historia Agraria Centroamericana", de la Escuela de Historia de la Universidad Nacional (EHUNA). La afinidad metodológica provenía de que en dicho proyecto se informatizaría el contenido numérico y nominal de los padrones microfilmados del censo agropecuario de 1955, por lo cual debían resolverse problemas de concepción y de procedimiento bastante similares. Ello planteó la posibilidad inmediata de compartir experiencias y obtener asesoría conjunta, como también la oportunidad futura de entrelazar dichas bases de datos.

Una vez que se dieron los primeros pasos en ambos esfuerzos de creación de bases de datos, se obtuvo gracias a una gestión de Héctor Pérez Brignoli la valiosa asesoría de Robert McCaa, experimentado investigador estadounidense y responsable de un proyecto semejante con los padrones de Parral, en México. El profesor McCaa ofreció sugerencias muy atinadas sobre diversos aspectos, incluyendo los relativos al empleo de microcomputadoras portátiles para digitar directamente; la importancia de conservar lo más fielmente que fuese posible el formato de la fuente original, reduciendo al mínimo la codificación al momento de digitar; la necesidad de evaluar realistamente la posibilidad de incluir en un plazo razonable toda la información, y la conveniencia de recurrir al muestreo.

Los primeros esfuerzos colaborativos y de gestión de recursos externos contaron con el oportuno respaldo de Lowell Gudmundson, y se tradujeron en un financiamiento complementario de la Universidad de Pittsburg. El mismo sirvió para respaldar la adquisición de microcomputadoras y ciertos gastos de operación, en apoyo tanto al proyecto 1927 como al de 1955.

Otro paso fundamental a este respecto fue la negociación y firma de una carta de entendimiento entre la Universidad de Costa Rica y la Universidad Nacional, para la crea-

ción e intercambio de bases de datos para la investigación histórica. Ello no sólo evidenció un encomiable espíritu de colaboración por encima de cualquier recelo institucional, sino que aseguró a futuro la posibilidad de compartir y entrelazar éstas y otras bases de datos, en lugar de duplicar absurdamente los costos de su creación. Significó, asimismo, una muestra importante del tipo de logros que es posible alcanzar cuando una visión académica compartida antepone el progreso del conocimiento y la racionalización de esfuerzos al aislacionismo contraproducente y al recelo infundado.

b) Selección de zonas

En la fase inicial del proyecto 1927, fueron seleccionadas varias zonas urbanas y rurales para recopilar en ellas la información de todas las personas reportadas en los padrones localizados en el Archivo Nacional. El levantamiento de la información completa en zonas específicas permitiría evaluar la cobertura de los padrones y determinar con mayor certeza la duración y el costo de la recopilación y digitado de los datos, además de facilitar la realización de estudios específicos tanto locales o regionales como comparados. También brindaría la oportunidad de corregir oportunamente errores en dicha fase, antes de emprender un esfuerzo a escala nacional.

Los criterios para la selección de zonas para el digitado completo fueron de naturaleza cualitativa. Respondían a los problemas de investigación planteados en el proyecto original y a su confluencia con los intereses de otros investigadores e investigadoras que habían manifestado su intención de participar. Guardaban relación, primordialmente, con la estructura socioproductiva y los procesos de colonización agrícola. Se tomó como unidades base las de la división territorial administrativa de 1928, a nivel de cantón y distrito.

En un principio, tales zonas incluirían la ciudad de San José y eventualmente otras áreas urbanas, de las cuales hasta la fecha se han digitado los distritos centrales de Ala-

juela y Liberia, y se está invitando a personas interesadas a estudiar los cascos urbanos de Heredia y Cartago, así como los puertos de Limón y Puntarenas. En el ámbito rural del Valle Central, se escogió los principales cantones cafetalero-cañeros del oriente y occidente del mismo, y en la periferia del país las principales regiones de inmigración desde el interior. Allí se concentraría inicialmente el trabajo del equipo de investigación vinculado laboralmente al proyecto.

Al mismo tiempo, se dejó abierta la posibilidad de que se incorporasen nuevas zonas para el digitado del 100% de la población empadronada, en función de las áreas que fuesen pertinentes para otras personas que expresaran su intención de incorporarse a este esfuerzo colectivo. En tales casos, personal del proyecto suministraría el programa de captura de datos, brindaría información y asesoría, y participaría en la limpieza de los datos, sin menoscabo de otras formas de colaboración que podrían convenirse de mutuo acuerdo.

Como se desprende de lo ya explicado, se invitó a otros investigadores profesionales y a tesarios o tesarias de Licenciatura y Maestría en Historia a participar en el proceso de creación de la base de datos, de acuerdo con sus propios intereses de investigación. Ello significaba que cada cual definiera en forma independiente no sólo el problema y la estrategia de investigación, sino también el uso que daría a la información de 1927 y el ámbito geográfico en el cual recopilaría datos de los padrones censales. Algunas personas se incorporaron bajo distintas modalidades de colaboración y asesoría a un trabajo compartido en zonas de interés común, mientras que otras personas escogieron áreas distintas de las priorizadas originalmente.

En el cuadro número 1 se detallan los cantones y distritos en que se digitó el 100% de los individuos registrados en los padrones censales que se conservan en el Archivo Nacional. Ello incluye las zonas seleccionadas inicialmente y áreas correspondientes a otros proyectos o tesis de grado y posgrado. En algunos casos, la población en dichos padrones no corresponde al total de la población censal de 1927, porque ciertos padrones aún no han sido localizados.

Cuadro 1
Censo de Población de 1927
Zonas digitado 100%

Provincia	Cantón	Distrito	Digitado	Limpieza	Procesamiento
San José	San José	Hospital	Sí	Sí	Sí
San José	San José	Carmen	Sí	Sí	Sí
San José	San José	Merced	Sí	Sí	Sí
San José	San José	Catedral	Sí	Sí	Sí
San José	Desamparados	Completo	Sí	Sí	
San José	Montes de Oca	San Pedro	Sí	Sí	
San José	Puriscal	Santiago	Sí	Sí	
Alajuela	Alajuela	Alajuela	Sí	En proceso	
Alajuela	Palmares	Completo	Sí	Sí	
Alajuela	Grecia	Completo	Sí	Sí	
Alajuela	San Carlos	Completo	Sí	En proceso	
Cartago	Turrialba	Completo	En proceso		
Guanacaste	Nicoya	Nicoya	Sí	Sí	
Guanacaste	Santa Cruz	Santa Cruz	Sí	Sí	
Guanacaste	Santa Cruz	Bolsón	Sí	Sí	

c) Muestreo

Aunque se espera incrementar gradualmente las áreas en que la base de datos incluye al 100% de los empadronados, el costo y duración de tal empresa hizo evidente que no sería factible completarla en un plazo razonable, además de que la inexistencia de algunos padrones impediría contar -de todas formas- con la totalidad de la población nacional. Tomando en cuenta los objetivos del proyecto y la viabilidad de las metas propuestas, así como otras experiencias afines en el exterior, se tomó la decisión de digitar una muestra de 10% para el resto del país. También se definió que el método de muestreo debería permitir aumentar fácilmente el tamaño de la muestra, v.g. a 20%, para análisis

más detallados, o completar el total de los empadronados para estudios locales.

Con la asesoría del estadístico Miguel Gómez, se diseñó un sistema de muestreo por familias y por folios, el cual ya fue explicado en un producto metodológico anterior citado previamente. En forma sucinta, el procedimiento consiste en digitar los hogares completos de la cara A de cada quinto folio, comenzando por el primer jefe de familia de dicha cara y completando la última familia, de ser necesario, en la cara B.

El propósito de este sistema fue facilitar la labor de futuros investigadores o proyectos que deseen incrementar la muestra o digitar el 90% restante de la población en determinadas zonas, para sus propios fines. En el cuadro núm. 2 se incluye la información pertinente sobre los folios contenidos en la muestra, y en el Centro de Investigaciones Históricas está a disposición de los interesados su numeración detallada para el propósito antedicho.

Cuadro 2

Censo de Población de 1927
Zonas Digitado 10%

Provincia	Cantón	Distrito	Digitado	Limpieza	Procesamiento
San José	San José	Zapote	Sí	Sí	
San José	San José	San Fco. 2 Ríos	Sí	Sí	
San José	San José	Curridabat	Sí	Sí	
San José	San José	La Uruca	Sí	Sí	
San José	San José	Mata Redonda	Sí	Sí	
San José	San José	Las Pavas	Sí	Sí	
San José	San José	Hatillo	Sí	Sí	
San José	San José	San Sebastián	Sí	Sí	
San José	Escazú	Completo	Sí	Sí	
San José	Puriscal	Mercedes Norte	Sí	Sí	
San José	Puriscal	Barbacoas	Sí	Sí	
San José	Puriscal	Grifo Alto	Sí	Sí	
San José	Puriscal	San Rafael	Sí	Sí	
San José	Puriscal	Candelarita	Sí	Sí	
San José	Puriscal	Desamparaditos	Sí	Sí	

San José	Tarrazú	Completo	Sí	Sí
San José	Aserrí	Completo	Sí	Sí
San José	Mora	Completo	Sí	Sí
San José	Goicoechea	Completo	Sí	Sí
San José	Santa Ana	Completo	Sí	Sí
San José	Alajuelita	Completo	Sí	Sí
San José	Coronado	Completo	Sí	Sí
San José	Acosta	Completo	Sí	Sí
San José	Tibás	Completo	Sí	Sí
San José	Moravia	Completo	Sí	Sí
San José	Montes de Oca	Sabanilla	Sí	Sí
San José	Turrubares	Completo	Sí	Sí
San José	Dota	Completo	Sí	Sí
Alajuela	Alajuela	Resto Cantón	Sí	En proceso
Alajuela	San Ramón	Completo	Sí	Sí
Alajuela	San Mateo	Completo	Sí	Sí
Alajuela	Atenas	Completo	Sí	Sí
Alajuela	Naranjo	Completo	Sí	Sí
Alajuela	Poás	Completo	Sí	Sí
Alajuela	Orotina	Completo	Sí	Sí
Alajuela	Alfaro Ruiz	Completo	Sí	Sí
Cartago	Cartago	Completo	Sí	Sí
Cartago	Paraíso	Completo	Sí	Sí
Cartago	La Unión	Completo	Sí	Sí
Cartago	Jiménez	Completo	Sí	Sí
Cartago	Alvarado	Completo	Sí	Sí
Cartago	Oreamuno	Completo	Sí	Sí
Heredia	Heredia	Completo	Sí	Sí
Heredia	Barva	Completo	Sí	Sí
Heredia	Sto. Domingo	Completo	Sí	Sí
Heredia	Sta. Bárbara	Completo	Sí	Sí
Heredia	San Rafael	Completo	Sí	Sí
Heredia	San Isidro	Completo	Sí	Sí
Heredia	San Belén	Completo	Sí	Sí
Heredia	San Flores	Completo	Sí	Sí
Guanacaste	Liberia	Completo	Sí	Sí
Guanacaste	Nicoya	Completo	Sí	Sí
Guanacaste	Sta. Cruz	27 de abril	Sí	Sí
Guanacaste	Sta. Cruz	Tempate	Sí	Sí
Guanacaste	Bagaces	Completo	Sí	Sí
Guanacaste	Carrillo	Completo	Sí	Sí
Guanacaste	Cañas	Completo	Sí	Sí
Guanacaste	Abangares	Completo	Sí	Sí
Guanacaste	Tilarán	Completo	Sí	Sí

Puntarenas	Puntarenas	Completo	Sí	Sí
Puntarenas	Esparta	Completo	Sí	Sí
Puntarenas	De Osa	Completo	Sí	Sí
Puntarenas	Montes de Oro	Completo	Sí	Sí
Limón	Limón	Completo	Sí	Sí
Limón	Pococí	Completo	Sí	Sí
Limón	Siquirres	Completo	Sí	Sí

Por el tamaño absoluto de las poblaciones y el porcentaje muestral, como también por la inexistencia de algunos folios y por el mismo sistema de muestreo, los resultados obtenidos por esta vía se presumen válidos (esto es, con un margen de error aceptable pero que deberá verificarse) para la muestra nacional y para grandes regiones multicantonales. Su validez está siendo evaluada actualmente, para precisar el grado de confiabilidad para distintos niveles de desagregación en ciertas zonas en que se digitó el 100% de la población, como también por otros procedimientos estadísticos. Asimismo, se especificarán las precauciones indispensables en el procesamiento e interpretación de los resultados obtenidos a partir del muestreo.

En principio, salvo ciertos cantones y unos pocos distritos muy populosos y homogéneos, se ha considerado que la utilidad de la muestra a escala local será sumamente limitada. En cambio, será de gran utilidad para el país en su conjunto y -de acuerdo con la regionalización que se está elaborando- para zonas multicantonales con una serie de características en común. Quizás podrán hacerse algunas comparaciones y ciertos análisis específicos para subregiones en que la muestra refleja bien los datos desagregados disponibles en el censo publicado. Este proyecto no generará resultados estadísticos basados en la muestra para localidades con escasa población o donde ésta es muy heterogénea, por cuanto los datos podrían ofrecer una imagen bastante inexacta y no podrían extrapolarse a la respectiva población distrital o cantonal. En cambio, sí podrán ser útiles, como información nominal, para ciertos estudios cualitativos, seguimientos prosopográficos y pareos con otras fuentes o bases de datos nominales.

Con el mismo procedimiento ya explicado, se han creado muestras para aquellas zonas en que se digitó desde un principio la totalidad de la población empadronada. Ello permitirá no solamente completar la muestra nacional de 10% sino también comparar los datos resultantes del procesamiento de la muestra con los correspondientes a toda la población en padrones, y de ambas respecto del censo publicado.

Los resultados obtenidos hasta la fecha con el procesamiento estadístico regional de la muestra de un décimo de los empadronados se han confrontado con ciertos datos agregados del censo nacional que se publicó en su oportunidad.¹⁸ Otro tanto se hará con la muestra nacional recientemente completada. También se hará, para zonas en que se digitó el 100%, un análisis comparado de los datos resultantes de procesar muestras de diverso tamaño y el total. A ello se sumará un conjunto de procedimientos demográficos para evaluar la muestra nacional, validar las submuestras regionales y precisar límites de confiabilidad para desagregaciones menores. Varios de ellos formarán parte de la tesis de maestría en estadística de Saray Castro, computóloga del proyecto, y otros serán aplicados por Arodys Robles y Mario Samper.

ch) Recolección manual y digitado directo de la información

Con la intención de simplificar el procesamiento y garantizar la comparabilidad de resultados, la recopilación de datos numérico-nominales del censo de 1927 por parte de todas las personas participantes en este esfuerzo coordinado se efectuó de manera uniforme, con un formato previamente establecido. Como es lógico, otros investigadores utilizaron los padrones originales de ese censo en forma independiente y diversa, recopilando únicamente aquellos datos que requerían para sus propios fines.¹⁹ Ambos usos son, por supuesto, perfectamente legítimos, y cada quien toma en consideración los objetivos de su estudio, el tiempo necesario para recopilar de una u otra manera la información requerida, y los beneficios a obtener. Es justo decir que en al-

gunos casos, las demoras provenientes de ajustarse a los requerimientos del proceso de creación de una base de datos han sido sustanciales, aunque sin duda se ha podido aprovechar en mayor grado el potencial de esta fuente, analizando tanto a la población total como subconjuntos de ella y procesando estadísticamente volúmenes de información apreciables. Al mismo tiempo, se ha conservado el detalle correspondiente a cada caso individual en la base numérico-nominal, tanto para seguimientos prosopográficos como para el estudio de casos familiares.

En un principio, la información se recopiló en formularios parcialmente precodificados, con espacios para anotar información nominal sobre la persona censada, como también para el lugar de trabajo y las categorías socio-ocupacionales. La labor de codificación posterior de tales categorías resultó sumamente compleja, ardua y lenta, con el agravante de que pese a los esfuerzos constantes por unificar criterios hubo algunas divergencias que fue necesario corregir mediante una minuciosa revisión. La información codificada y revisada fue introducida al sistema de cómputo central por digitadores profesionales del Centro de Informática. Este paso introdujo nuevos y abundantes errores que también debieron enmendarse mediante una segunda revisión manual y redigitado. Todo ello consumió un tiempo excesivo, e hizo evidente la urgencia de adoptar otro procedimiento.

A raíz de la experiencia antedicha, y con la asesoría del Dr. McCaa, se optó por crear un programa de captura de datos para digitar directamente en el Archivo Nacional (o con fotocopias y microfilmes de los padrones) la información respectiva. El programa "Cen27" fue diseñado y mejorado por la computóloga Saray Castro, quien efectuó consultas con los investigadores del proyecto así como otros usuarios.

Las características principales de la versión final de dicho programa son la fidelidad a las variables alfanuméricas y numéricas del padrón original, la facilidad de uso y los controles de error. Previo acuerdo sobre el uso posterior de los archivos creados, es posible obtener en el C.I.H. copia gratuita de dicho programa y asesoría para su empleo. El

C.I.H. cuenta también con procedimientos informáticos para la revisión y limpieza posterior de los archivos creados con este programa de captura.

El uso del programa Cen27 por parte del equipo del proyecto y otras personas agilizó notablemente la recopilación de información y redujo significativamente los errores. En la fase de limpieza es necesario retornar ocasionalmente al Archivo Nacional, aunque ello puede obviarse con fotocopias de los padrones.

III. Utilización de la base de datos de 1927

a) El procesamiento estadístico general de la información contenida en la base de datos del '27, una vez hecha su revisión y limpieza, se ha efectuado en varias fases:

a.1 Inicialmente, los esfuerzos se concentraron en la generación de estadística descriptivas univariadas y bivariadas.

El investigador principal, en consulta con otros tesisarios e investigadores interesados, elaboró una guía detallada para el procesamiento estadístico en SPSS. La misma incorporó los principales requerimientos de los proyectos en curso, pero también incluyó otros cruces de variables en función de usos adicionales previsibles. El programa de procesamiento fue elaborado por Saray Castro, y actualmente puede ser ejecutado con alguna asesoría inicial por cualquier persona que domine los fundamentos del SPSS. Los cuadros resultantes de dicho procesamiento básico se incluyen en los volúmenes correspondientes a cada distrito y cantón en las zonas de digitado completo, como también para el país y para las principales regiones en base a la muestra.²⁰

La autoría institucional de los volúmenes de cuadros resultantes responde tanto a la naturaleza del proyecto como al deseo de reconocer el valioso y constante respaldo del Centro de Investigaciones Históricas, a la vez que procura enfatizar el carácter colectivo de una empresa que su-

peraba ampliamente las posibilidades individuales. En cada volumen se reconoce explícitamente todos los aportes de quienes participaron directamente en la recopilación, limpieza y procesamiento de los datos.

Se encuentra en proceso de elaboración otro conjunto de cuadros resultantes del procesamiento estadístico, relativos a subconjuntos de la población y atinentes a objetivos de investigación más específicos de personas ligadas directa o indirectamente al proyecto. Entre ellos se cuentan los concernientes a las mujeres en general y a las trabajadoras en particular, como también a los principales grupos ocupacionales. Se contempla asimismo una caracterización estadística más pormenorizada de los hogares, al igual que su diferenciación por variables socioeconómicas y geográficas.

a.2 Procesamiento de la muestra

Como se indicó anteriormente, fue creada también una muestra nacional que incluye aproximadamente un 10% de la población en los padrones de 1927. Dicha muestra combina archivos creados por digitado selectivo de un décimo de las familias en la mayor parte del país, y otros archivos extraídos de la base de datos censal -siguiendo el mismo criterio muestral- en zonas donde se digitó la totalidad de la población.

El proceso de validación de la muestra nacional está siendo efectuado coordinadamente por Saray Castro, Arodys Robles y Mario Samper, quienes publicarán resultados evaluativos.

Una vez completada y sometida a crítica y prueba la regionalización, se producirán las submuestras regionales definitivas.

Tanto la muestra nacional como las submuestras por regiones se procesarán estadísticamente de modo análogo a lo hecho con distritos y cantones en que se digitó la totalidad de la población.

Más adelante, se espera incorporar la información de la base de datos muestral a un sistema de mapeo computarizado.

a.3 Los resultados del procesamiento estadístico básico, tanto en zonas de digitado completo como de muestra, se han incorporado ya -de modos creativos y diversos- a varias tesis de grado y de posgrado presentadas, como también a algunas otras que se encuentran en etapa final de preparación. Entre las primeras cabe mencionar, además de la ya citada del M.Sc. Ronny Viales sobre la desestructuración del enclave bananero en la provincia de Limón, la del M.Sc. José Manuel Cerdas acerca de los trabajadores urbanos.²¹ De las varias tesis en curso, con una amplia gama temática y estrategias metodológicas muy variadas que integran los datos de ésta y otras fuentes, podemos señalar aquí:

- la tesis de maestría en estadística, con énfasis en población, en etapa final de preparación por Saray Castro, la cual se basa fundamentalmente en la muestra nacional de 1927;

- la tesis de maestría en historia, próxima a concluirse, de Wilder Sequeira, sobre las unidades productivas campesinas en la península de Nicoya;

- una tesis de posgrado de Nancy Worsfold, en torno a la temática de mujer y familia en el San José de principios de siglo;

- otra de Virginia Mora, acerca de las mujeres trabajadoras urbanas entre fines del siglo XIX y primeras décadas del XX;

- un interesante estudio ya muy avanzado de José William Solano, también como tesis de maestría, relativo a las relaciones entre empresas agroindustriales y sus "clientes" en Turrialba;

- la tesis de posgrado de Rafael Cordero, sobre la producción cañera en el cantón de Grecia a partir de 1914.

La información de la base de datos construida cooperativamente también ha sido o está siendo utilizada por otras personas tesieras de grado y posgrado que participaron en el proceso, en forma individualizada según los requerimientos de cada proyecto. Así, por ejemplo, se incorpora en algunas tesis sobre categorías ocupacionales específicas, como la de Franciso J. Rojas sobre carpinteros y ebanistas, y está siendo utilizada por Doryam Carvajal y Gina

Rivera en una tesis histórico-demográfica de licenciatura, para la ciudad de Liberia. El programa de captura de datos de 1927 también ha sido utilizado por grupos de estudiantes del Trabajo Comunal Universitario, bajo la dirección de Francisco Enríquez, para distritos aledaños a la ciudad de San José.

Otros investigadores vinculados de alguna forma a este esfuerzo conjunto han utilizado o se proponen emplear datos resultantes del procesamiento, como también la base misma, en sus respectivos proyectos. La información relativa a la ciudad de Alajuela, por ejemplo, fue digitada por asistentes del Dr. Víctor Hugo Acuña y una vez completada su revisión y limpieza será utilizada en un estudio de los grupos artesanales en ella. También se ha facilitado información limitada sobre grupos muy particulares o para fines claramente demarcados a otras personas, v.g. al investigador norteamericano Stephen Palmer para su estudio sobre el consumo de drogas entre los artesanos de principios de siglo; o a Lowell Gudmundson para un estudio comparado de Desamparados y Santo Domingo, aun inconcluso.

Los resultados del procesamiento estadístico de la muestra nacional y de las submuestras regionales, como también de zonas específicas, se incorporarán asimismo a estudios comparados y análisis de índole más general.

Se espera iniciar próximamente un análisis interrelacionado de variables múltiples, mediante la producción de tabulados que combinen diversas variables para subgrupos poblacionales. También se ensayarán -en función de problemas de investigación concretos- análisis multivariados con el fin de esclarecer y precisar, para el conjunto de la población muestral, las relaciones menos aparentes pero potencialmente significativas entre variables demográficas y socioeconómicas.

b) Estudio sobre sectores ocupacionales y redes de parentesco

Durante varios años, el proceso de creación de la base de datos de 1927, con el consiguiente aprendizaje metodológico y múltiples labores asociadas tanto a él, así como la

decisión de dar prioridad a tesaros y tesaras participantes en el uso de la información recopilada conjuntamente, llevó a postergar la producción de resultados analíticos propios del proyecto. Cumplida en lo fundamental dicha etapa y producida ya la muestra nacional, se retornó a los propósitos iniciales de integrar análisis estadístico y seguimiento de casos para un estudio sobre la estructura socio-ocupacional en 1927, el cual a su vez se insertaría en un estudio diacrónico al respecto. Lógicamente, se replantearon tanto el enfoque general y los contenidos específicos como la estrategia metodológica, en base a la experiencia adquirida.

b.1 Análisis cualitativo/cuantitativo de sectores ocupacionales

En función de los objetivos originales del proyecto, así como sus reformulaciones posteriores, se han creado -a partir de la base de datos de 1927- archivos diferenciados por sector ocupacional. Estos incluyen a todos los miembros de aquellas familias en que algún miembro ejercía cualquiera de los oficios que conforman el respectivo grupo de ocupaciones afines.²²

Mediante una combinación de procedimientos exploratorios, mediciones estadísticas y seguimiento de casos, se espera relacionar un conjunto de variables ligadas al trabajo con aquellas vinculadas al parentesco. Algunos de los lineamientos preliminares para ello fueron expuestos por Acuña y Samper en su ponencia al Taller de Intercambio de Experiencias Metodológicas "Fuentes numérico-nominales e investigación histórica".²³

El propósito específico es efectuar un análisis cualitativo y cuantitativo de redes familiares en las cuales existe una tradición ocupacional. La revisión exploratoria de los listados impresos ha permitido formular nuevas hipótesis sobre cada sector ocupacional, para su verificación posterior. También ha sido la base para el seguimiento en otras fuentes y para la identificación de informantes potenciales. El procesamiento estadístico de los subconjuntos poblacionales permite efectuar una caracterización pormenorizada de

cada uno y su análisis comparado, así como relacionarlos con la población mayor.

b.2 Prosopografía y seguimiento de casos con apoyo en la base de datos nominal

Una vez generados los listados por sector ocupacional a que se hizo referencia en el numeral anterior, se efectúa una selección de potenciales casos familiares en los cuales varios miembros ejercieron un mismo oficio u otro afín. Lejos de ser excepcional, ello era frecuente, lo que ha hecho posible contar con un adecuado número de hogares con tales características.

Seguidamente se prioriza a los grupos familiares por su idoneidad para los propósitos del seguimiento prosopográfico. Los criterios principales para ello son el número de personas que ejercían dichos oficios y su edad en 1927, a fin de establecer si hay posibilidad de localizar informantes.

El paso siguiente es la revisión sucesiva de tales casos en el Registro Civil, para determinar si las personas han fallecido o no. En caso de estar vivas, se procura obtener información sobre su lugar de residencia, y se procede a localizarlas. Previa autorización de su parte y valoración de su potencial como informantes, se hace un seguimiento genealógico de la familia respectiva en el Registro Civil, el cual se complementa con una búsqueda en el archivo de la Curia Metropolitana, en el Archivo Nacional y en el Registro de la Propiedad. Con ello y la información nominal del censo de 1927, se elabora un genograma (vale decir, un diagrama genealógico con datos adicionales pertinentes al estudio). En el mismo se destaca la información sobre tradiciones ocupacionales dentro de la respectiva red de parentesco.

Hasta el momento, estos seguimientos se han hecho "a pie" (en forma no automatizada), aunque con apoyo en sistemas computadorizados de búsqueda tanto en nuestra base de datos como en la del Registro Civil. Actualmente se están dando los primeros pasos para automatizar y por ende simplificar considerablemente estos procesos, lo cual permitiría reconstrucciones genealógicas más exhaustivas.

La información obtenida mediante el seguimiento prosopográfico se profundiza a través de una entrevista a profundidad, bajo el formato de historia de vida temática. Esta es videograbada y posteriormente transcrita y analizada.

Los resultados de la prosopografía familiar/ocupacional se integran luego con los del análisis estadístico y con la recopilación de información complementaria en fuentes secundarias como también en fuentes primarias impresas, principalmente periodísticas y publicaciones oficiales.

Ya se efectuó una primera serie de entrevistas, con una guía flexible derivada de los objetivos de investigación y del estudio bibliográfico y documental sobre un sector ocupacional específico: el de los trabajadores de tipografía e imprenta. Actualmente está completándose el análisis estadístico y cualitativo, cuyos resultados preliminares se presentarán en una sesión de discusión metodológica con investigadores interesados, y en una entrevista colectiva con miembros de la respectiva red familiar/ocupacional, así como con viejos y jóvenes tipógrafos. Paralelamente, se ha iniciado el estudio de un segundo sector ocupacional, el de trabajadoras y trabajadores del ramo textil y de la vestimenta. También se está evaluando la experiencia metodológica inicial, basada en la estrecha integración del análisis estadístico y prosopográfico, y en un modelo de búsqueda focalizada de información, sin pretensiones de exhaustividad, por cuanto se espera abarcar en el lapso de dos años los principales sectores ocupacionales tanto urbanos como rurales, y existe cierta premura por la avanzada edad de los y las potenciales informantes.

En el caso específico de la investigación en curso sobre redes de parentesco con una o más tradiciones ocupacionales, el seguimiento de casos mediante reconstrucción genealógica, análisis documental e historias de vida temáticas permite, por ejemplo, profundizar en la trama de relaciones interpersonales y de trabajo al interior del grupo familiar amplio, estudiar los mecanismos de transmisión de la tradición laboral compartida de generación en generación, explorar las vivencias subjetivas y las valoraciones retrospectivas de quienes formaron parte de esas redes, y conocer la vida cotidiana tanto familiar como laboral. Al mismo

tiempo, se responden otras interrogantes e hipótesis derivadas del análisis cuantitativo y cualitativo del sector ocupacional y de las redes de parentesco, y se formulan nuevas preguntas y respuestas tentativas a algunas de ellas, a partir del caso particular y de su comparación con otros.

En un sentido más general, la prosopografía y el seguimiento longitudinal de casos en fuentes múltiples, referidas a su vez a una tipología u otra categorización social y a datos agregados, facilitan la “triangulación” metodológica y el análisis comparado, pero sobre todo la reintegración de perspectivas micro y macrosociales, de lo que sucede al interior de un grupo específico y de la sociedad en su conjunto.

c) Consulta nominal y pareo

El aprendizaje metodológico a lo largo del proceso de creación de la base de datos de 1927, como también en la fase inicial de la investigación específica a la cual se hizo referencia en el numeral anterior, se ha confrontado con otras experiencias de investigación con fuentes similares dentro y fuera del país. Se han generado diálogos fructíferos entre personas interesadas en el uso de este tipo de bases de datos, así como relaciones de colaboración entre instituciones que cuentan con ellas o las están produciendo actualmente.

Una de las conclusiones a que hemos llegado es que se hace necesario diseñar sistemas de consulta ágiles y “amigables” que puedan ser utilizados por múltiples usuarios locales, ya sea con la base de datos original o con eventuales copias de uso público que puedan ponerse a disposición de investigadores en el futuro.

Se encuentra en su fase inicial el desarrollo de un sistema que facilitará la búsqueda y localización de personas (por nombre y apellidos) en la base de datos censal de 1927. Ello permitirá efectuar en ella, sin necesidad de apoyo por parte de un programador, el seguimiento de casos nominales provenientes de otras bases de datos o de muy diversas fuentes documentales e incluso orales. Con ello po-

drá invertirse la dirección de las búsquedas y seguimientos efectuados hasta la fecha, que han partido de listados provenientes de esta base de datos. Tal acceso simplificará considerablemente la incorporación de la base de datos censal de 1927, como fuente complementaria, en estrategias investigativas prosopográficas que tienen su punto de origen en cualquier otra fuente nominal.

La segunda conclusión a que se ha llegado es que el momento actual es propicio para emprender el enlace semi-automatizado de archivos entre ésta y otras bases de datos. Para ello se inscribió en la Vicerrectoría de Investigación de la Universidad de Costa Rica, y ya dio inicio, un nuevo proyecto denominado precisamente "Enlace de archivos nominales para la investigación histórico-social". También se han estrechado lazos de cooperación con otras instituciones que cuentan con bases de datos entrelazables con la del '27, en particular la Escuela de Historia de la Universidad Nacional y el Registro Civil. Paralelamente, se han efectuado talleres de discusión metodológica sobre el pareo computadorizado, y fue formulado un plan de trabajo coordinado conducente al enlazamiento de varias bases de datos nominales a mediano plazo.²⁴

En consecuencia, se están efectuando las labores previas que permitirán establecer un sistema de enlace de archivos para el pareo ("record linkage") parcialmente automatizado de bases numérico-nominales y otras afines. En primera instancia este enlace o pareo se efectuará con la base de datos del Registro Civil, en lo concerniente a nacimientos, matrimonios y defunciones, que pese a ser incompleta para el período 1888-1927, ofrece la posibilidad de reconstruir redes familiares amplias, efectuar análisis histórico-demográficos que antes no eran factibles, agilizar considerablemente el seguimiento de casos y, sobre todo, entrelazar dos importantes bases de datos poblacionales. Pese a que la cobertura de esa base de datos registral decrece cuantitativa y cualitativamente al retroceder en el tiempo, al reestructurar su sistema de consulta y cruzar la información nominal con la de 1927 se podrán efectuar análisis demográficos, genealógicos, prosopográficos y de subconjuntos de la población que hasta el momento no eran factibles. Además, cabe

la posibilidad de completar gradualmente la información contenida en dicha base de datos, a fin de mejorar su cobertura y calidad.

Posteriormente, en coordinación con otros proyectos de investigación, se desarrollarán experiencias de enlace de archivos entre la población urbana de San José en 1927 y 1904, así como para zonas rurales en 1927 y 1955. El objetivo metodológico de más largo plazo es la creación de una red entrelazada de bases de datos nominales, que podrá servir como apoyo a múltiples y muy variados tipos de pesquisa histórico-social.

IV. Productos

El proyecto "Transformación del censo de población de 1927 en base de datos nominales" produjo, entre otros resultados concretos, los siguientes:

- Una base de datos numérico-nominales con la información relativa a un número aproximado de veinte mil hogares costarricenses, o cerca de cien mil individuos. De ellos, poco más de la mitad corresponde a zonas en que se digitaron los datos de toda la población, y el resto abarca las demás áreas del país, en que se digitó una muestra de la forma ya explicada.

- Un conjunto de volúmenes con los cuadros que contienen la información estadística básica (univariada y bivariada) para los distritos y cantones en que se digitó el 100% de la población.

- La muestra nacional con alrededor de 10% de la población costarricense, debidamente revisada y en proceso de evaluación final. A muy corto plazo estarán disponibles también las submuestras regionales, y se hará el procesamiento respectivo a escala nacional y regional.

- El aporte metodológico de la experiencia específica del proyecto como también de las discusiones formales e informales entre investigadores e investigadoras sobre estrategias de investigación que incorporan fuentes numérico-nominales. Las publicaciones y actividades ya fueron mencionadas anteriormente.

- Una serie de productos analíticos concluidos o en proceso de elaboración, en los cuales se recurre a información derivada de la base de datos, del procesamiento estadístico o del seguimiento de casos nominales. Esto incluye varias tesis de grado y posgrado ya presentadas, y otras en etapa de elaboración final. En lo concerniente a este proyecto, los productos analíticos en proceso o previstos incluyen una caracterización descriptiva de la población costarricense en 1927, con diferenciaciones regionales y por sectores económicos; un estudio conjunto de Arodys Robles y Mario Samper sobre hogar y familia en 1927; el estudio monográfico y comparado de sectores ocupacionales y redes de parentesco con tradición ocupacional; y la comparación diacrónica para zonas urbanas y rurales, a partir de pareos entre la base de datos censal de 1927 y otras bases de datos existentes.

- Acuerdos para la creación y entrelazamiento de bases de datos nominales con otras instituciones de educación superior y entidades estatales, a fin de racionalizar esfuerzos y garantizar el acceso recíproco de investigadores a dichas bases de información.

- Un nuevo proyecto institucional al cual ya se hizo referencia, denominado "Enlace de archivos nominales para la investigación histórico-social", vigente en el Centro de Investigaciones Históricas, así como un proceso de concertación y retroalimentación entre personas que se interesan por aplicaciones metodológicas afines.

V. Balance general de la experiencia metodológica del proyecto de 1927

Entre los aspectos más positivos del proceso reseñado, cabe mencionar el desarrollo de nuevas estrategias de investigación a partir de una fuente numérico-nominal, enlazada con otras manualmente (o con apoyo en consultas alfabéticas computadorizadas). Ello ha permitido enriquecer el análisis sincrónico y diacrónico, reuniendo y procesando información complementaria sobre individuos, grupos familiares u ocupacionales, y grandes conglomerados humanos.

El dominio, adaptación y diseño de herramientas computacionales flexibles para el manejo de este tipo de fuentes sólo fue posible gracias al apoyo de especialistas en informática, pero los programas y procedimientos se han estandarizado considerablemente, de modo que puedan ser aplicados por personas con poco entrenamiento en ese campo. La producción de sistemas de consulta y pareo “amigables” al usuario debe seguir siendo una prioridad para facilitar el acceso a este tipo de tecnologías y a las bases de datos mismas.

Por otra parte, la combinación de información nominal y otros datos cualitativos sobre cada individuo con datos numéricos o cuantificados fue una invitación clara a combinar el análisis estadístico y el seguimiento de casos individuales o grupales. Gracias a la facilidad de agregar o desagregar la información, hemos podido movernos con agilidad creciente entre varios planos: el de los grandes grupos humanos, el de subsectores específicos, el de la biografía colectiva y el de los estudios de caso a profundidad mediante reconstrucciones genealógicas e historias de vida, complementadas por el estudio documental, estadístico y bibliográfico.

Lejos de renegar de un saludable “eclecticismo creativo”, la experiencia concreta de este proyecto permite valorar su pertinencia y utilidad como una de las vías para avanzar hacia la reintegración de los planos micro- y macro-analíticos, del estudio de caso y de la sociedad en su conjunto. Inserto en una muy necesaria reflexión epistemológica, ello ofrece nuevas oportunidades para trascender el divorcio entre “cuantitativistas” y “cualitativistas”, sin desconocer la validez de sus divergencias y contraposiciones en cuanto a epistemología, método, técnica y relación sujeto/objeto de estudio.

Otro logro medular fue la coordinación de esfuerzos entre un equipo de investigación con respaldo institucional y un conjunto de estudiantes de grado y posgrado, cada cual con su propia agenda de investigación pero con necesidades convergentes. El salto hacia nuevos modos de investigación requirió y requerirá de una acción colaborativa que trascienda el marco estrecho de cada proyecto individual. Esto fue posible, en parte, gracias al sostenido respal-

do institucional, pese a las dificultades presupuestarias, la inestabilidad en algunos de los nombramientos y otras limitaciones.

Entre las lecciones negativas, igualmente importantes, extraemos las siguientes:

Las estimaciones de costo en recursos humanos y materiales, pero sobre todo de la probable duración del proceso de creación de la base de datos censal de 1927 fueron excesivamente optimistas. Fue necesario hacer ajustes sobre la marcha para redimensionar el proyecto, v.g. mediante el recurso al muestreo. Aun así, el proyecto duró más de lo previsto, y todavía quedan ciertas tareas inconclusas que deberá retomar el proyecto sucesor. Las demoras afectaron también a tesarios o tesarias y a otros investigadores con quienes se coordinaron esfuerzos.

En el caso de este proyecto, la recopilación en cuestionarios, la codificación de todas las variables y el digitado posterior (inicialmente en el Centro de Informática) resultaron ser formas ineficientes y contraproducentes de introducir la información de los padrones a una base de datos. Ello causó pérdida de tiempo, triplicación de esfuerzos y numerosos errores e incongruencias, que sólo fue posible subsanar mediante un arduo trabajo posterior. La situación únicamente mejoró cuando se tomó la difícil decisión de echar marcha atrás y adoptar otro enfoque, mediante la captura directa de los datos en su formato original, con microcomputadoras portátiles. Sin embargo, debe reconocerse que en las condiciones iniciales del proyecto y del propio Centro de Investigaciones Históricas, la recopilación en planillas y posterior digitado en el Centro de Informática parecían constituir la única opción viable en ese momento.

La revisión, limpieza y procesamiento de la información digitada también fueron procesos mucho más dispendiosos de lo previsto. No sólo se subestimó el tiempo requerido, sino que los recursos humanos e informáticos para tales labores eran limitados, tornándolas excesivamente lentas.

A su vez, la prolongación de la fase de limpieza y del procesamiento mismo, aunado al deseo de respetar la "primicia" de los tesarios y tesarias participantes, demoró considerablemente la generación de productos analíticos por

parte del proyecto mismo. Afortunadamente se contó con el respaldo decidido y sostenido de la dirección y Consejo Científico del Centro de Investigaciones Históricas y de la propia Vicerrectoría de Investigación, gracias a la evidente importancia metodológica e historiográfica del dilatado proceso de creación de esta base de datos para su posterior utilización colectiva.

VI. Construcción cooperativa de una red entrelazada de bases de datos nominales

Al concluir oficialmente el proyecto para la transformación del censo de 1927 en base de datos, a fines de 1993, e iniciarse el de enlace de archivos, terminó una primera etapa, cuya duración y costo en recursos humanos y materiales superó nuestras previsiones originales. Durante ella se creó, mediante un esfuerzo mancomunado, una gran base de datos nominal para múltiples y polifacéticas investigaciones históricas sobre la población costarricense. También se gestó una experiencia metodológica invaluable, se procesó estadísticamente la información fundamental, y se hicieron estudios sectoriales y prosopográficos como también seguimientos de casos individuales en otras fuentes.

La tarea institucional que ha de emprenderse ahora no consiste en producir nuevas bases de datos nominales mediante el digitado de grandes volúmenes de información a un costo considerable y en el transcurso de varios años. Ya han sido creadas, en otros proyectos, algunas bases de datos que tienen, asimismo, carácter nominal. Es el caso de la base de datos agropecuaria de 1955, la de población josefina en 1904, o la construida desde los padrones de 1844. También han sido creadas, para fines específicos, otras bases de información nominal computadorizada en diversos proyectos individuales, especialmente para tesis de grado y posgrado. Hay, asimismo, valiosas fuentes nominales informatizadas que han sido creadas por instituciones que por una u otra razón efectúan labores registrales a escala masiva.

En principio, sería posible acceder a varias bases de datos nominales, según una estrategia que permita priorizarlas y llegar a acuerdos mutuamente beneficiosos que multiplicarán el potencial de tales bases de datos al entrelazarlas. Conviene aprovechar colectivamente las ricas posibilidades que ofrece el gradual ensanchamiento de un sistema de archivos nominales enlazados entre sí, para muy variadas opciones de pesquisa histórica, demográfica y social.

Claro está que la mayoría de las bases de datos existentes fueron creadas para fines distintos a los que tendría el propuesto enlace de archivos. Su estructura misma puede no ser idónea para el tipo de pareo que se efectuaría, o pueden resultar inadecuados los actuales sistemas de consulta. Pero si logramos aunar objetivos claros y compartidos, una estrategia metodológica apropiada y la indispensable capacidad técnica, más cierta dosis de realismo y sentido común, debería ser factible adaptar algunas de esas valiosísimas bases numérico-nominales para entrelazarlas y explotar sus potencialidades.

Cada base de datos adicional que se integre planificadamente al sistema sugerido abriría nuevas vías para abordar cuestiones que interesan a un amplio conjunto pluridisciplinario de investigadores. A modo de ejemplo, la interconexión entre la base censal de 1927 y la del Registro Civil, pese a que ninguna de ellas abarca la totalidad de la población, ensanchará muy considerablemente el ámbito de los análisis demográficos, genealógicos e histórico-sociales que es posible hacer actualmente, y lo hará a un costo relativamente bajo a la vez que maximiza el uso de las tecnologías informáticas actualmente disponibles. Lo que es igualmente importante, desde el punto de vista metodológico, es que ampliará y extenderá la trocha por la cual podrán transitar, si lo desean, quienes se interesen actualmente o en el futuro por desarrollar estrategias de investigación en las cuales haya una provechosa retroalimentación entre modos cuantitativos y cualitativos de tratar la información, pero sobre todo de explorar formas de reintegrar los planos analíticos micro- y macro-sociales.

El proyecto "Enlace de archivos nominales para la investigación histórico-social" se inició, a principios de 1994,

principalmente con el fin de dar apoyo al tipo de esfuerzo conjunto y coordinado que se desprende de las sugerencias anteriores. Completará algunas tareas pendientes del proyecto precedente y se cimentará en la base de datos censal de 1927 para entrelazarla con otras ya existentes, sin perjuicio de que más adelante se evalúe la conveniencia de crear complementariamente una o más bases adicionales, si hubiese interés y disponibilidad de recursos. Pero durante algún tiempo, parece más racional y conveniente, en términos de costo y eficacia, centrar nuestra atención y dedicar nuestros limitados recursos a desarrollar un sistema de pareo informático entre bases nominales ya existentes.

Las dificultades iniciales del enlace de archivos en fuentes múltiples no son despreciables, y será necesario tanto un nuevo aprendizaje colectivo de las metodologías y técnicas disponibles, como el diseño conjunto de estrategias y la invención de nuevos modos de investigar de acuerdo con nuestra problemática y posibilidades. Por más que nos asesoremos, nadie podrá decirnos cómo hacerlo ni desaparecerán los dilemas epistemológicos, historiográficos y sociales subyacentes en diversas estrategias de investigación. No se trata solamente de adquirir determinadas herramientas, programas, equipos o destrezas informáticas, sino también de reflexionar dialogadamente sobre el tipo de investigación que hemos hecho y aspiramos a hacer, sobre lo que indagamos, cómo formulamos y respondemos nuestras preguntas, por qué, para qué fines y para quiénes investigamos. Más allá de la mera acumulación y procesamiento de datos, la exploración propuesta nos interroga acerca de nuestros modos de conocer la historia, y nos invita a dar un salto metodológico con implicaciones que trascienden la simple -aunque no siempre fácil- aplicación de nuevas tecnologías.

Es probable, como ocurrió en la etapa anterior, que la inversión de tiempo y recursos genere pocos dividendos tangibles inmediatos, pero a mediano plazo habremos multiplicado nuestra capacidad de abordar interrogantes que quizá hubiésemos rehuido anteriormente. Estaremos menos limitados que ahora en cuanto a las preguntas que podemos plantearnos y abordar con alguna posibilidad de responder-

las satisfactoriamente. Pero nada sustituirá la imaginación del historiador o la historiadora, como ninguna base de datos relegará a segundo plano la pertinencia conceptual e historiográfica de las cuestiones planteadas, y ningún sistema computarizado obviará la necesidad de combinar la mayor rigurosidad analítica con la máxima creatividad intelectual.

Finalmente, esperamos que esta relación del proceso de creación de la base de datos de 1927, su evaluación autocrítica y la contextualización historiográfica retrospectiva y prospectiva, sirvan de invitación a otras personas y equipos de investigación a compartir y debatir con la mayor franqueza sus propias experiencias de investigación. Dada la enorme distancia existente entre los manuales y la práctica de la pesquisa histórico-social, tal confrontación de perspectivas y enfoques a partir de proyectos concretos será de mucha utilidad para quienes ejercemos ya o se preparan para ejercer este curioso oficio de “andar averiguando vidas ajenas” para comprender nuestro pasado y presente sociales.

Notas

1. Agradezco los atinados comentarios y sugerencias de Arodys Robles y José Antonio Fernández al borrador de este texto. Soy, por supuesto, único responsable de cualquier desatino, y muy consciente de mis propias limitaciones. Espero, no obstante, que la contextualización historiográfica y la cándida exposición de una experiencia metodológica particular generen algún debate e inviten a otros y otras a presentar su propia visión y experiencia.
2. Uno de los primeros estudios cuantitativos comparados del caso costarricense respecto de otro latinoamericano, fue el de Héctor Pérez, “El ciclo en las economías agrícolas de exportación de América Latina (1880-1930): hipótesis para un estudio”, en *Revista de Historia*, núm. 5, 1977, pp. 9-46. Casi simultáneamente vio la luz el análisis comparado sistemático de las economías centroamericanas, publicado por Ciro Cardoso y Héctor Pérez, *Centroamérica y la economía occidental* (San José, Editorial Universidad de Costa Rica, 1977), que incluyó y amplió uno anterior de Cardoso sobre la historia económica del café en el istmo. Poco después apareció su segunda obra conjunta *Historia económica de América Latina* (Barcelona, Editorial Crítica, 1979). De alcance más modesto son los estudios comparados de este autor que se apoyan en la confrontación de estadísticas cafetaleras de Costa Rica y otros países; ver M. Samper, “Caficultura, producción familiar y haciendas, 1920-1936: Análi-

sis comparado a partir del caso costarricense y colombiano", en *Estudios rurales latinoamericanos*, vol. 12, núm. 13, 1989, y "El significado social de la caficultura costarricense y salvadoreña: análisis comparado", en Héctor Pérez y Mario Samper (compiladores), *Tierra, café y sociedad* (San José, FLACSO, 1994).

3. Quizás el ejemplo más prominente sea el del economista Roger Churnside, quien aportó a la discusión sobre concentración de la tierra una mayor precisión y rigurosidad en la definición de términos y en el manejo de la información cuantitativa. Cf., entre otros trabajos suyos, "Concentración de la tierra en Costa Rica en 1935 y 1800-1850: algunas consideraciones de tipo metodológico", en *Documentos* (Instituto de Investigaciones en Ciencias Económicas, Universidad de Costa Rica), setiembre de 1978.
4. Aquí habría que considerar dieciséis tesis de licenciatura presentadas entre 1976 y 1986 en la Universidad de Costa Rica en historia demográfica, basadas en registros parroquiales. Todas éstas siguen un mismo esquema básico para la recopilación, procesamiento y análisis de la información, e incluso para la presentación de resultados. A ellas habría que sumar varias tesis de licenciatura en historia presentadas en la Universidad Nacional y algunos otros trabajos histórico-demográficos sobre coyunturas específicas, que si bien reúnen y procesan información similar sobre nacimientos, matrimonios y defunciones, proveniente de archivos parroquiales, la integran a estudios más integrales de la historia local o nacional. Por otra parte, hay algunos análisis cuantitativos más generales efectuados mediante la comparación de las series locales, notablemente los de Héctor Pérez B, v.g. "Las variables demográficas en una economía de exportación: el ejemplo del Valle Central de Costa Rica, 1800-1950", *Avances de Investigación* (Escuela de Historia y Geografía, Universidad de Costa Rica), núm. 7, 1978. Las migraciones también han sido estudiadas por varios autores, más recientemente y con mayor refinamiento metodológico por Arodys Robles, "Patrones de población en Costa Rica, 1860-1930", en *Avances de Investigación* (Centro de Investigaciones Históricas de la Universidad de Costa Rica), núm. 14, 1986.
5. Hay varias tesis de licenciatura en historia sobre comercio exterior, presentadas en la Universidad de Costa Rica y la Universidad Nacional, así como una valiosa obra en prensa del economista Jorge León sobre el transporte marítimo y el comercio exterior costarricense, que verá la luz próximamente como tomo de la Historia General de Costa Rica. Habría que añadir una serie de análisis macroeconómicos, algunos hechos por historiadores y otros por economistas. En los primeros, la estadística suele ser descriptiva e ilustrativa, como en la obra de Carlos Araya P., *Historia económica de Costa Rica, 1821-1971* (San José, Editorial Fernández-Arce, 1982). Los estudios retrospectivos hechos por economistas suelen limitarse al período posterior a 1950, pero algunos de ellos incorporan un análisis teórico y empírico más afinado, v.g. Reinaldo Carcanholo, *Desarrollo del capitalismo en Costa Rica* (San José, EDUCA, 1981). En historia agraria la producción ha sido abundante, y en términos generales

incorpora datos estadísticos descriptivos o, en ocasiones, información cuantitativa para la demostración de ciertas hipótesis; las tendencias generales y contribuciones más relevantes fueron evaluadas por José Antonio Salas, Mario Samper y Jorge Mora en sus contribuciones a "El agro costarricense en perspectiva histórica: un balance", en *Revista de Historia* núm. 19, enero-junio 1989.

6. El uso principal de la cuantificación en historia social ha sido para el análisis estadístico de la estructura social, generalmente a partir de censos publicados o -con menor frecuencia- de padrones censales. Cabe mencionar, en primer lugar, la obra de Lowell Gudmundson, principalmente *Costa Rica antes del café: sociedad y economía en vísperas del boom exportador* (San José, Editorial Costa Rica, 1990); la de Roger Churnside, *Formación de la fuerza laboral costarricense* (San José, Editorial Costa Rica, 1985), y -a un nivel más bien rudimentario- la tesis de licenciatura de este autor. Hay asimismo una serie de trabajos sobre tenencia de la tierra que recurren de una u otra manera a la cuantificación para fundamentar sus conclusiones en pro o en contra de la concentración de la propiedad fundiaria en distintos períodos; cf. la reseña crítica de Mario Ramírez, "La polémica de la concentración de la tierra en Costa Rica: mitos e ideologías sobre el desarrollo capitalista (1850-1930)", en *Materiales sobre estructura socio-económica de Costa Rica* (Departamento de Sociología, Universidad de Costa Rica), 1978. Marc Edelman y Mitchell Seligson efectuaron recientemente un estudio sobre el uso de datos censales y del Registro de la Propiedad para estudiar la desigualdad fundiaria, cuya versión definitiva será publicada en el futuro cercano.
7. Cf. Orlando Salazar, *El apogeo de la República Liberal en Costa Rica, 1870-1914* (San José, Editorial Costa Rica, 1990), y Mario Samper, "Fuerzas sociopolíticas y procesos electorales en Costa Rica, 1921-1936", en *Revista de Historia*, número especial 1988. En cuanto a mentalidades colectivas y vida cotidiana, hay varios trabajos sobre delictividad dirigidos o efectuados por José Gil Z. en la Universidad Nacional, en especial su tesis doctoral *Homicidio, asociación y conflicto en la provincia de Heredia (1895-1915)* (tesis de doctorado en Historia, Universidad Autónoma de Barcelona, Bellaterra, 1994). Cabe mencionar también los de Iván Molina y algunos estudiantes de maestría en historia de la Universidad de Costa Rica, sobre diversos aspectos de la vida cultural costarricense; cf. Iván Molina y Steven Palmer (compiladores), *Héroes al gusto y libros de moda. Sociedad y cambio cultural en Costa Rica (1750/1900)* (San José, Editorial Porvenir y Plumsock Mesoamerican Studies, 1992). En ambos casos, se trata de estudios muy sugerentes con una riqueza de información cualitativa y un tratamiento estadístico elemental de información se-riable, algunas veces para un reducido número de casos.
8. La situación que describimos para mediados de la década del noventa no es muy disímil de la expuesta por Héctor Pérez en una entrevista sobre el tema de la historia cuantitativa, en *Revista de Historia*, núm. 15, enero-junio de 1987, pp. 13-14. A la fecha, los estudios cuantitativos más rigurosos han sido los ya citados de Pérez Brigno-

li, en historia económica y demográfica, y los de Arodys Robles sobre migración. El psicólogo e historiador Alfonso González efectuó recientemente uno de los pocos análisis históricos factoriales, con datos de las dispensas matrimoniales, en *Mujer y familia en la vida cotidiana de la segunda mitad del siglo XIX (una aproximación desde la psicohistoria)* (tesis de maestría en Historia, Universidad de Costa Rica, 1993). En el proyecto Historia Social de la Tecnología Cafetalera se han efectuado asimismo análisis factoriales y se han aplicado otros procedimientos multivariados con datos del censo agropecuario de 1955, con participación de Héctor Pérez, Margarita Rojas y Mario Samper. Otro tanto se hará próximamente con datos de 1927 sobre estructura familiar, en el proyecto Enlace de archivos nominales para la investigación histórico-social, por parte de Arodys Robles y quien escribe.

9. Los estudios etnográficos fueron, desde un principio, de índole cualitativa, y permitieron una exploración y descripción detallada de la vida cotidiana y las costumbres. Algunos estudios antropológicos recurrieron a la tradición oral de pueblos indígenas y afrocaribeños. Entre las primeras investigaciones en aplicar la técnica de las historias de vida a la indagación sobre nuestra historia social, destacan las de Víctor Hugo Acuña, "La ideología de los pequeños y medianos productores cafetaleros costarricenses (1900-1961)," en *Revista de Historia*, núm. 16, julio-diciembre de 1987, y "Vida cotidiana, condiciones de trabajo y organización sindical: el caso de los zapateros en Costa Rica (1934-1955)", en *Revista de Historia*, número especial, 1988. Diversas perspectivas disciplinarias se vieron reflejadas en el *Primer seminario de tradición e historia oral* (San José, Universidad de Costa Rica, 1988), que incluye contribuciones históricas de Patricia Badilla, Yamileth González y Mario Samper. Para un conjunto de aportes y reflexiones en torno a la relación entre historia y recuperación de la memoria colectiva a través de modalidades de investigación más participativas, ver *Memoria y cultura popular costarricense* (San José, Centro Nacional de Acción Pastoral, 1986).
10. La entrevista semi-estructurada a profundidad fue aplicada a sesenta informantes en la investigación doctoral del autor presentada en 1987; cf. *Generations of settlers. Rural households and markets on the Costa Rican frontier, 1850-1935* (Boulder, Westview Press, 1990). Dos proyectos en curso que recurren de modo importante a variantes de la entrevista a profundidad y la integran con fuentes documentales son el de Patricia Badilla sobre "La participación de pequeños productores y jornaleros en la guerra civil de 1948 en la provincia de Cartago y la Zona de Los Santos" (Escuela de Historia y Geografía de la UCR) y el proyecto colectivo "Historia Social de la Tecnología Cafetalera" (Escuela de Historia de la UNA). Sobre la metodología de uno y otro proyecto, cf. Patricia Badilla, "Una estrategia de acercamiento al estudio de la cultura política en sectores campesinos costarricenses", y el texto de Brunilda Hilje, Carlos Naranjo y Mario Samper, "Cambio tecnológico y cambio social en zonas cafetaleras: estudio cuantitativo-cualitativo", ambos en la memoria del taller *Fuentes numérico-nominales e investigación histórica* (San José, CIH-UCR, en prensa, 1994).

11. Aquí no podemos dejar de mencionar los cinco volúmenes de *Autobiografías campesinas* (Heredia, EUNA, 1979 a 1981), como tampoco la colección más completa de los textos originales del concurso respectivo denominado CONAUCA, que se conservan en la biblioteca de la Universidad Nacional. Los contenidos fueron indizados con un sistema de codificación que permite localizar un conjunto de temáticas en los respectivos textos, aunque éstos no fueron digitados. También hubo un concurso de autobiografías obreras y artesanales, el CONOA, que si bien dejó un conjunto menor de textos, tiene indudable valor testimonial. Una versión reelaborada de uno de ellos se publicó recientemente en la *Revista de Historia* núm. 27, correspondiente a enero-junio de 1993, bajo el título "Juan Rafael Morales Alfaro. Autobiografía". Ambos tipos de testimonios autobiográficos han sido utilizados en algunos estudios académicos, pero su potencial no ha sido plenamente aprovechado. Una de las dificultades proviene de la heterogeneidad temática de tales relatos autobiográficos; no obstante, encierran una gran riqueza testimonial. Otro tipo de autobiografía, las memorias de dirigentes políticos, tiene una utilidad histórica restringida por su obvia intencionalidad justificatoria; sin embargo, algunas de ellas constituyen pese a la inevitable reevaluación retrospectiva, testimonios valiosos, v.g. la de Jaime Cerdas Mora, *La otra vanguardia* (San José, EUNED, 1993).
12. La prosopografía política más conocida y exhaustiva es, por supuesto, la obra sucesivamente refinada y expandida de Samuel Stone: *Los cafetaleros: Une étude de la classe des grands planteurs de café au Costa Rica* (tesis doctoral de tercer ciclo, Universidad de París, la Sorbona, 1968); punto de partida para *La dinastía de los conquistadores* (San José, EDUCA, 1975), y su estudio comparado respecto de otras élites centroamericanas en *The heritage of the Conquistadors* (Lincoln, E.E.U.U., University of Nebraska Press, 1990). La prosopografía en general se ha nutrido también de la genealogía, campo en el cual ha habido notables contribuciones desde Monseñor Sanabria hasta Rafael Obregón y Eduardo Fournier, entre otros. Las biografías de individuos prominentes, muchas de ellas publicadas por el Ministerio de Cultura, destacan la vida y obra del individuo, pero hasta el momento se han utilizado poco para reinterpretar la historia de la sociedad costarricense.
13. En algunos casos, fue necesario digitar nuevamente la información en un formato uniforme y de mayor utilidad colectiva, v.g. con padrones de 1844, 1927 y 1955, casos en los cuales pese a que el investigador estuviese dispuesto a compartir bases de datos que había creado y utilizado para sus propios fines, la información se había recopilado y organizado en formas que hacían muy difícil su utilización para otros fines. En ciertas oportunidades, los intentos por unificar formatos de entrada de datos fueron infructuosos, o sólo parcialmente exitosos, y en ocasiones el costo de recopilar y digitar nuevamente la información era prohibitivo.
14. Quizás el ejemplo más claro de ello sean las mortuales, fuente de la cual se extrajo información con un formato base que fue sufriendo modificaciones según los intereses de cada investigador o investiga-

dora. Las bases de datos respectivas no han podido unirse, pero se recogió cierta información básica comparable.

15. Los dos casos más prominentes, en estos últimos años, han sido la producción de una base de datos con la información por fincas del censo agropecuario de 1955, en la Escuela de Historia de la UNA, y el proceso de construcción conjunta de la base censal de 1927, en el Centro de Investigaciones Históricas de la UCR, que se reseña en la siguiente sección.
16. La experiencia metodológica que aquí se resumirá fue colectiva en más de un sentido. En primer lugar, el equipo de investigadores e investigadoras incluyó a varias personas: inicialmente -aunque su jornada correspondía a otro proyecto con objetivos un tanto distintos- participó muy activamente Arodys Robles, sobre todo en la fase de elaboración y prueba del formulario de recopilación manual de datos para su posterior digitado. Durante el lapso de casi un año en que el autor estuvo fuera del país, hacia 1991, la coordinación estuvo a cargo de Héctor Pérez B., quien también ha brindado otras formas de apoyo y asesoría. A todo lo largo del proyecto original y reformulado, gran parte del peso del digitado y otra serie de responsabilidades cada vez más complejas recayó sobre Marielos Acuña, quien comenzó como asistente y terminó -tras graduarse de Licenciada- como co-investigadora. Durante varios años, hasta el presente, el desarrollo informático fue responsabilidad de Saray Castro, computólogo y ahora también especialista en estadística. Varios asistentes de investigación no sólo prestaron invaluable y mal remunerados servicios, sino que participaron en mayor o menor grado de las discusiones metodológicas. Quiero reconocer muy especialmente el trabajo siempre responsable y excelente de nuestro "orillero", José William Solano, hoy profesor en la sede de la UCR en Turrialba, y más recientemente de los estudiantes de maestría en Historia Rafael Cordero y Kevin Ávalos. Participaron también como asistentes, en la fase inicial, Alicia Osorno y Juan Carlos González, como en otro momento Jenny Peña y Mauricio Guillén. De diversas maneras y previo acuerdo explícito al respecto, participaron asimismo algunos estudiantes de cursos metodológicos en historia, tanto de la Universidad de Costa Rica como de la Universidad Nacional, v.g. en la fase de prueba de formularios, en el digitado para determinada localidad o en el seguimiento de casos. El aporte muy significativo de las personas tesoreras se explica en el texto. Cabe mencionar, adicionalmente, el fructífero intercambio de perspectivas metodológicas con otros investigadores, y en particular con quienes participaron en el taller "Fuentes numérico-nominales e investigación histórica", en 1992.
17. Mario Samper K. (compilador), *El censo de población de 1927: Creación de una base nominal computadorizada* (Serie Trabajos de Metodología Núm. 2, Centro de Investigaciones Históricas, Universidad de Costa Rica, 1990).
18. El primer procesamiento regional con datos de la muestra se hizo, en el Centro de Investigaciones Históricas, para la provincia de Li-

món. Buena parte de la información resultante fue incorporada a la tesis de maestría en Historia presentada por Ronny Viales, quien participó activamente en el digitado de la información respectiva y en otras actividades relacionadas con el proyecto; cf. *La región atlántica costarricense y el enclave bananero: del esplendor a la crisis (1927-1950)* (tesis de Maestría en Historia, Universidad de Costa Rica, 1993).

19. El primer producto publicado que integra de modo muy interesante datos tomados del padrón de 1927 con otras fuentes fue el artículo de L. Gudmundson, "Campesino, granjero, proletario: formación de clase en una economía cafetalera de pequeños propietarios, 1850-1950", en *Revista de Historia*, Núm. 21-22, enero-diciembre de 1990 (hay versión inglesa anterior en *Hispanic American Historical Review*). Entre los tesarios que han utilizado en forma independiente dichos padrones, cabe mencionar el interesante trabajo aún inconcluso de Guillermo Rosabal sobre los trabajadores de panadería, y el cruce manual entre un listado de votantes del Partido Reformista y el padrón josefino de 1927, sugerencia bien acogida por Victoria Ramírez en su tesis de posgrado *El partido Reformista: alternativa política de las clases trabajadoras costarricenses en la década de 1920* (tesis de Maestría en Historia, UCR, 1987) (hay edición posterior en forma de libro, bajo el título *Jorge Volio y la Revolución Viviente*).
20. La referencia genérica para dichos volúmenes, con especificación de provincia, cantón y distrito, como también del volumen respectivo y del lugar y año de edición, es: Centro de Investigaciones Históricas, Universidad de Costa Rica, *Procesamiento de datos censales de 1927: cruces de variables*. Hasta la fecha, se han producido los volúmenes correspondientes a los cuatro distritos centrales del cantón central de la provincia de San José (versión preliminar), y se ha efectuado o está próximo a concluir el procesamiento para los cantones seleccionados para digitar el 100%; los volúmenes respectivos se producirán -una vez revisados los cuadros preliminares- en el transcurso del segundo semestre de 1994 e inicios de 1995. Actualmente, se está completando la limpieza y evaluación de la muestra nacional, que se procesará seguidamente del mismo modo. Otro tanto se hará próximamente con las submuestras regionales.
21. José Manuel Cerdas, *Condiciones de vida de los trabajadores manufactureros de San José 1930-1960* (tesis de Maestría en Historia, Universidad de Costa Rica, 1994).
22. Cabe advertir que tal procedimiento clasificatorio conllevaría la duplicación de numerosos casos familiares en distintos sectores ocupacionales, dado el ejercicio de más de un oficio por sus miembros adultos. Como segundo criterio de clasificación se consideró el oficio ejercido por mayor número de miembros del hogar, a fin de discriminar grupos familiares ubicados en dos sectores.
23. Marielos Acuña y Mario Samper, "Trabajo y familia: un intento de historia social a partir del censo de 1927", en Mario Samper (compilador), *Fuentes numérico-nominales e investigación histórica* (San

José, Centro de Investigaciones Históricas, UCR, Serie Trabajos de Metodología Núm. 4, en prensa, 1994).

24. Concretamente, en 1992 y 1993 se reunieron con este fin investigadores del Centro de Investigaciones Históricas de la UCR y la Escuela de Historia de la UNA, en el marco del "Acuerdo para la creación e intercambio de bases de datos". Pueden consultarse las minutas de las reuniones respectivas, disponibles en los archivos de ambas instituciones o mediante solicitud al autor, donde se detallan las opciones de enlace allí discutidas, se programan las etapas preliminares y se evalúan logros y limitaciones.