

INDICE

	<u>Páginas</u>
Lista de tablas	7
Lista de figuras	8
PRÓLOGO, por José M. ^a Labeaga Azcona	9
AGRADECIMIENTOS	13
INTRODUCCIÓN	15
1. DISEÑO Y MUESTREO	21
1.1. Bases de microdatos: estructuras y tipos de paneles	21
1.1.1. Bases de datos de sección cruzada	22
1.1.2. Bases de datos de serie temporal	22
1.1.3. <i>Pool</i> de datos (<i>Pooled Data</i>)	23
1.1.4. Paneles de Datos	25
1.2. Diseño del Panel de Declarantes de IRPF 1999-2007	31
1.2.1. Antecedentes	31
1.2.2. Principales características	34
1.2.3. Proceso de creación	38
2. CARACTERÍSTICAS BÁSICAS	49
2.1. Fichero de «principales»	49
2.1.1. Variables no monetarias	50
2.1.2. Variables monetarias	56
2.2. Fichero de «cónyuges»	57
2.3. Relación entre el fichero de «principales» y el de «cónyuges»: el concepto de hogar fiscal	57
3. CREACIÓN DE VARIABLES ADICIONALES Y SEGUIMIENTO DE HOGARES Y DECLARANTES	65
3.1. Creación de variables adicionales	65
3.1.1. Variables de hogar	65
3.1.2. Variables individuales	68
3.2. Proceso de seguimiento de declarantes y hogares	71
3.2.1. Proceso de seguimiento de declarantes	72
3.2.2. Proceso de seguimiento de hogares	72
4. ANÁLISIS DE CALIDAD Y REPRESENTATIVIDAD	77

4.1.	Estimación poblacional de las principales variables económicas del IRPF.....	77
4.2.	Desgaste del Panel (<i>attrition</i>)	89
4.3.	Anomalías detectadas en las variables del Panel	90
4.3.1.	Fichero de «principales»	91
4.3.2.	Fichero de «cónyuges»	92
4.4.	Anomalías detectadas en las estructuras relacionales del Panel	93
5.	ESTRUCTURA, CONTENIDOS Y UTILIZACIÓN.....	103
5.1.	Estructura de los ficheros y variables incluidas	103
5.2.	Utilización del Panel	108
5.2.1.	Cálculos de población	108
5.2.2.	Cálculo de agregados	110
5.2.3.	Cálculo de medias	112
5.2.4.	Cálculo de medidas de distribución	113
5.2.5.	Trabajo con hogares	114
5.2.6.	Seguimiento de observaciones	117
5.2.7.	Panel puro	118
	REFERENCIAS BIBLIOGRAFICAS	119
	ANEXOS:	
1.	Diseños de registro	123
2.	Tamaño muestral de los estratos	249
3.	Frecuencias anuales: variables originales, fichero de principales .	261
4.	Frecuencias anuales: variables originales, fichero de cónyuges	289
5.	Frecuencias anuales: variables creadas, fichero de principales	303
6.	Frecuencias anuales: variables creadas, fichero de cónyuges	349
7.	Frecuencias anuales: variables creadas, fichero de hogares	363
8.	Frecuencias anuales: variables creadas, fichero de declarantes	379
9.	Estadística descriptiva: variables originales, fichero de principales .	391
10.	Estadística descriptiva: variables originales, fichero de cónyuges .	405
11.	Estadística descriptiva: variables creadas, fichero de principales .	415
12.	Estadística descriptiva: variables creadas, fichero de cónyuges	419
13.	Resumen de la normativa del IRPF español 1999-2007	423

LISTA DE TABLAS

	<i>Páginas</i>
Tabla 1.1. Tramos de ingresos brutos	40
Tabla 2.1. Variables no monetarias incluidas en el Panel	50
Tabla 2.2. Delegaciones de Hacienda	52
Tabla 2.3. Comunidades Autónomas	53
Tabla 2.4. Definición de la renta de muestreo para cada año	54
Tabla 2.5. Tramos de renta de muestreo	54
Tabla 2.6. Definición de los ingresos del trabajo para cada año	55
Tabla 2.7. Equivalencias entre las principales variables monetarias	58
Tabla 2.8. Estructuras de las declaraciones según tipo de hogar fiscal	63
Tabla 3.1. Valores de la variable <i>estructurahogar</i>	67
Tabla 3.2. Valores de la variable <i>tipohogar</i>	67
Tabla 3.3. Valores de la variable <i>declaporhogar</i>	68
Tabla 3.4. Definición de la variable <i>categoriadecla</i>	69
Tabla 3.5. Definición de la variable <i>minimosdecla</i>	70
Tabla 3.6. Definición de la variable <i>rentamuestreo</i>	70
Tabla 3.7. Definición de la variable <i>ingresotrabajo</i>	71
Tabla 3.8. Cambios de tipo de hogar y seguimiento entre dos años	74
Tabla 4.1. Principales magnitudes del IRPF 1999. Comparación Panel-Badespe.....	78
Tabla 4.2. Principales magnitudes del IRPF 2000. Comparación Panel-Badespe.....	79
Tabla 4.3. Principales magnitudes del IRPF 2001. Comparación Panel-Badespe.....	80
Tabla 4.4. Principales magnitudes del IRPF 2002. Comparación Panel-Badespe.....	81
Tabla 4.5. Principales magnitudes del IRPF 2003. Comparación Panel-Badespe.....	82
Tabla 4.6. Principales magnitudes del IRPF 2004. Comparación Panel-Badespe.....	83
Tabla 4.7. Principales magnitudes del IRPF 2005. Comparación Panel-Badespe.....	84
Tabla 4.8. Principales magnitudes del IRPF 2006. Comparación Panel-Badespe.....	85
Tabla 4.9. Principales magnitudes del IRPF 2007. Comparación Panel-Badespe.....	80
Tabla 4.10. Desgaste del Panel (permanencia de declarantes)	89
Tabla 4.11. Desgaste del Panel (permanencia de hogares)	90
Tabla 4.12. Tipología de estructuras de hogar (variable <i>estructurahogar</i>)	95
Tabla 4.13. Número de casos para cada tipo de <i>estructurahogar</i> (fichero de «principales»)	97
Tabla 4.14. Número de casos para cada tipo de <i>estructurahogar</i> (fichero de «cónyuges»)	98
Tabla 4.15. Número de casos con <i>incluirdecla</i> = 1 para cada tipo de <i>estructurahogar</i> (fichero de «principales»)	99
Tabla 4.16. Número de casos con <i>incluirdecla</i> = 1 para cada tipo de <i>estructurahogar</i> (fichero de «cónyuges»)	100
Tabla 4.17. Número de casos para cada tipo de <i>estructurahogar</i> (fichero de «hogares») .	101
Tabla 5.1. Estructura de de ficheros de Panel, unidades de observación y variables en enlace	105
Tabla 5.2. Variables incluidas en los ficheros de «Principales» y de «Cónyuges» ..	106
Tabla 5.3. Variables incluidas en los ficheros de «Hogares»	107
Tabla 5.4. Variables incluidas en los ficheros de «Declarantes»	108

LISTA DE FIGURAS

	<i>Páginas</i>
Figura 1.1. <i>Attrition</i> y panel puro	27
Figura 1.2. Expansión y panel puro	30
Figura 1.3. Tamaño muestral por CC.AA. y tramo de renta, para contribuyentes con rentas de trabajo menor del 50% del total de renta	43
Figura 1.4. Tamaño muestral por CC.AA. y tramo de renta, para contribuyentes con rentas de trabajo mayor del 50% del total de renta	44
Figura 4.1. Comparación Panel –BADESPE (rendimientos y ganancias y pérdidas patrimoniales) Diferencias en %	87
Figura 4.2. Comparación Panel –BADESPE (mínimos y bases impositivas) Diferencias en %	88
Figura 4.3. Comparación Panel –BADESPE (cuotas impositivas) Diferencias en % ..	88
Figura 5.1. Estructura de de ficheros de Panel, unidades observacionales y variables en enlace	104

PRÓLOGO

La aprobación en 2005 del «Acuerdo de la Comisión Delegada del Gobierno para Asuntos Económicos sobre mejoras en la transparencia en el ámbito de la información económica y estadística proporcionada por el Gobierno» supuso un espaldarazo a la difusión de la información estadística. La exposición de motivos de este Acuerdo contempla que la disponibilidad de la información recibida de los poderes públicos es un elemento esencial para hacer posible el derecho de los ciudadanos a participar en los asuntos públicos (artículo 23.1 de la Constitución) y para que la Administración sirva con objetividad los intereses generales y actúe de acuerdo con los principios de eficacia, jerarquía, descentralización, desconcentración y coordinación, con sometimiento pleno a la Ley y al Derecho (artículo 103 de la Constitución).

También contempla el citado Acuerdo de transparencia que, dentro de esta información, la información estadística es esencial para conocer la realidad económica y social del país, por lo que el Estado, en colaboración con Comunidades Autónomas y Entes Locales, es quien debe velar porque esta información estadística se suministre a los agentes en tiempo y forma adecuada, sin ninguna restricción, salvo aquella que marque el secreto estadístico y protección de datos, con la máxima calidad posible, asegurando que su acceso sea universal y evitando cualquier tipo de discriminación. Asimismo, considerándose la relevancia de la información económica, el Acuerdo trata de vincular a todos los centros que la ofrecen con los criterios de máxima difusión y plena accesibilidad, adquiriendo la exigencia de transparencia de los poderes públicos la categoría de principio que debe guiar la actuación de tales poderes.

Este principio de transparencia afecta no sólo a la cantidad de información económica y estadística suministrada, sino sobre todo, a la calidad de la misma. Dicho conjunto de información debe ser publicado de manera periódica y previsible, cumpliéndose unos calendarios que los ciudadanos deben conocer. Además, la información debe ser publicada de forma que resulte útil al ciudadano, es decir, con la suficiente desagregación como para dar la idea más exacta del funcionamiento de los poderes públicos, y, también, facilitarse en soportes accesibles y adecuados a las nuevas tecnologías. Por otro lado, la información económica y estadística debe estar disponible con la mayor brevedad posible. En este sentido, se deben reducir lo máximo posible dos plazos: el de desfase y el de publicación. Para reducir al máximo dichos plazos es necesario fijar calendarios concretos y estrictos de publicación de los distintos datos económicos y estadísticos. Asimismo, la información se debe presentar de manera que sea lo más útil posible, facilitando su análisis y su estudio, lo que debe incluir una explicación sobre fuentes utilizadas, notas metodológicas y modelos empleados.

El propio Acuerdo de transparencia contempla los compromisos asumidos por el Ministerio de Economía y Hacienda, sus Organismos adscritos y otras entidades, en cuanto a la difusión de la información. En el Foro Mundial de estadística, conocimiento y política celebrado en Estambul en 2007, la Comisión Europea, la Organización para la Cooperación y el Desarrollo Económico, la Organización de la Conferencia Islámica, el Programa de Naciones Unidas para el Desarrollo y el Banco Mundial adoptaron la denominada Declaración de Estambul en la que en relación con las estadísticas se señala, entre otras cosas: «Las estadísticas oficiales son un bien público que fomenta el progreso de las sociedades». También se afirma: «Para promover el bienestar de las sociedades, se debe promover, en todos los niveles, una cultura que base sus decisiones en la evidencia». También se asevera: «La disponibilidad de indicadores sobre los resultados económicos, sociales, del medio ambiente y su difusión al público puede contribuir a promover un buen gobierno y el mejoramiento del proceso democrático».

En este marco de transparencia se estableció en el año 2006 un convenio para el suministro de información con fines estadísticos entre el Instituto de Estudios Fiscales (IEF) y la Agencia Estatal de Administración Tributaria (AEAT). Mediante este convenio, la AEAT facilita periódicamente información básica en microdatos al IEF para su análisis y difusión normalizada. Dentro de esta información se encuentra la relativa al Panel de Renta, registros administrativos que contienen información del Impuesto sobre las Personas Físicas (IRPF), cuya metodología y análisis se presenta en este libro. Cumpliendo con los objetivos del Acuerdo de transparencia, la información del Panel de Renta se difundirá anualmente en formato electrónico adecuado, lo que la hace accesible de forma sencilla. La petición de esta información se puede realizar de manera abierta desde la página Web del IEF, adecuándonos, asimismo, al cumplimiento del objetivo del Gobierno de España de implantar la Administración electrónica. Los datos se difunden con la correspondiente metainformación de la que forma parte esta publicación.

Por otro lado, el Panel de Renta del IEF constituye actualmente una operación estadística del Plan Estadístico Nacional, lo que supone dotarlo de una continuidad en el tiempo, además de la obligación de realizar la difusión de acuerdo a los compromisos adquiridos en el Proyecto Técnico defendido ante el Consejo Superior de Estadística y dictaminado favorablemente por este organismo, después de ser informado por el Instituto Nacional de Estadística y la Confederación Española de Organizaciones Empresariales. La Comisión del Consejo Superior de Estadística, en base a los informes presentados, valoró favorablemente la elaboración del proyecto del Panel de Renta, cuyo objetivo es facilitar información para realizar estudios de carácter longitudinal sobre la renta de las personas físicas y su distribución y concentración a partir del seguimiento de individuos a lo largo del tiempo. Con este Panel se dispondrá de información de rentas fiscales de personas y hogares de una población representativa de sujetos obligados a tributar en el Impuesto sobre la Renta de las Personas Físicas (IRPF), en el denominado Territorio Fiscal Común. Además, los resultados de este Panel vienen a satisfacer una parte de la demanda de información

fiscal requerida no sólo por las Administraciones Públicas sino también por los investigadores y los agentes sociales. Desde el punto de vista técnico, se considera muy positiva la utilización del panel expandido que permite su comparabilidad temporal manteniendo la representatividad de la muestra a lo largo del tiempo y, al mismo tiempo, incorporando los cambios derivados de las peculiaridades específicas de los nuevos contribuyentes.

La citada Comisión resaltó el interés y la utilidad del Panel de Renta en el diseño de política fiscal, porque ofrece una información completa y fiable sobre el IRPF que permite evaluar no sólo las políticas actuales sino también realizar simulaciones de reformas y estudiar ver sus efectos en el tiempo. Es decir, el formato de panel va a permitir evaluar los efectos de las políticas públicas en relación con el IRPF, así como los derivados de las posibles reformas normativas del impuesto, dado que será posible analizar el impacto sobre la población afectada por los cambios, tanto sincrónica como diacrónicamente. Destacó también la Comisión que el proyecto del Panel de Renta es un claro ejemplo del potencial informativo de los registros administrativos y su utilidad para la elaboración de estadísticas. Con la única utilización de registros administrativos y, por lo tanto, sin aumentar la carga estadística a los hogares, se proporciona, anualmente, información de la distribución y concentración de la renta de las personas físicas, a lo largo del tiempo. Además, en el caso concreto de la información fiscal, se evitan en gran medida algunos problemas de las encuestas, como falta de respuesta o de ocultación de información.

Otra utilidad que creemos que se debe resaltar es que el Instituto de Estudios Fiscales, una vez que dispone de los datos y realiza los necesarios procesos de depuración y preparación, los pone periódicamente a disposición de la comunidad académica dedicada a la investigación o a analistas institucionales entre los que, obviamente, están los pertenecientes al Ministerio de Economía y Hacienda. En este último grupo se encuentran los investigadores que realizan su labor en el propio Instituto de Estudios Fiscales, utilizando los datos tanto para realizar trabajos académicos como para alimentar algunos instrumentos de evaluación de políticas, de los que el Instituto es un actor relevante desde hace mucho tiempo. Es obvio que de esta forma estamos también cumpliendo el espíritu de la Declaración de Estambul puesto que los análisis que se realizan (o realizamos) utilizando estadísticas (datos administrativos) deben permitir tomar decisiones sobre la base de la evidencia en el intento de generar buenos gobiernos y mejoras en los procesos democráticos.

Este libro que ahora editamos recoge los resultados del laborioso proceso de elaboración del nuevo Panel de IRPF, desarrollado en el seno del Instituto de Estudios Fiscales por un equipo de trabajo integrado por César Pérez López, Vocal Asesor de Estadística del Instituto de Estudios Fiscales y Profesor de Estadística de la Universidad Complutense de Madrid; Jorge Onrubia Fernández, Profesor de Economía Aplicada de la Universidad Complutense de Madrid, y Fidel Picos Sánchez, Profesor de Economía Aplicada de la Universidad de Vigo.

Una vez realizada por la AEAT la selección de las muestras anuales de microdatos a partir del diseño estadístico realizado, la puesta a disposición de los usuarios del Panel suponía llevar a cabo un proceso complejo previo, dirigido a facilitar su aplicabilidad bajo condiciones de sencillez y eficiencia en el uso de la información contenida en la base de datos. Entre otras tareas, en este proceso resultaba fundamental definir una estructura relacional entre los distintos ficheros de datos lo más operativa posible, así como establecer las necesarias identificaciones de los declarantes y de los hogares fiscales que constituyen las observaciones del Panel, creando aquellas variables necesarias para permitir los seguimientos transversales y longitudinales de las mismas. Asimismo, también resultaba indispensable ofrecer a los analistas una presentación detallada de los aspectos técnicos del diseño estadístico del Panel, necesarios para conocer convenientemente las potencialidades y limitaciones del mismo en cuanto a su representatividad.

En definitiva, la obra constituye, en mi opinión, una guía detallada y rigurosa para la adecuada utilización del Panel de Renta, imprescindible dada la complejidad que entraña el diseño y construcción de una base de microdatos como ésta. Además, estoy seguro que este libro contribuirá a una mayor divulgación de la propia base de datos, en la medida que el trabajo incluye un primer análisis estadístico de los datos contenidos en el Panel, así como un ejercicio de validación de los mismos.

Por último, es de destacar el trabajo conjunto entre el personal del Instituto de Estudios Fiscales y de sus colaboradores y el personal del Departamento de Informática Tributaria de la Agencia Estatal de Administración Tributaria, así como las facilidades que los máximos responsables de la AEAT han mostrado siempre para facilitar los datos con la mayor rapidez y eficiencia. Sin este entendimiento entre las instituciones y la eficacia de esta colaboración conjunta hubiera sido imposible la elaboración de una herramienta de estudio y análisis tan potente como es el Panel de Renta.

JOSÉ MARÍA LABEAGA AZCONA

Director General del Instituto de Estudios Fiscales.

INTRODUCCIÓN

Las bases de microdatos procedentes de registros administrativos desempeñan un papel cada vez más relevante en el análisis económico aplicado y especialmente en el ámbito de la economía pública, al permitir el acceso a fuentes de información hasta hace relativamente poco tiempo no utilizables y cuyo potencial de explotación es enorme. La captación sistemática de la información a nivel individual mediante procedimientos administrativos implantados en la gestión de los programas de ingresos y gastos de las administraciones pública permite disponer de los datos con regularidad y precisión, y sin tener que incurrir en los importantes costes que acarrea el inevitable trabajo de campo necesario para construir cualquier base de microdatos basada en encuestas. Asimismo, la posibilidad de contar con los valores poblacionales de referencia para las variables incluidas en las bases de microdatos permite realizar diseños muestrales donde los márgenes de error estadístico al inferir resultados poblacionales son muy reducidos. Otra de las grandes ventajas que aporta la utilización de estas fuentes de información es la posibilidad de contar con bases de datos de un tamaño muestral muy grande a un coste de producción estadística bastante reducido. Por último, el desarrollo y tratamiento informático de estos registros administrativos permite incluso la conexión entre distintas bases de datos a partir de la identificación en las mismas de los individuos u hogares a los que se refiere la información registrada.

En el ámbito de los ingresos públicos, en los países anglosajones y nórdicos existe una importante tradición en la utilización de información procedente de fuentes tributarias para la construcción de bases de microdatos. Un antecedente de obligada referencia es el caso de Estados Unidos, donde su agencia federal de administración tributaria, el *Internal Revenue Service (IRS)*, elabora muestras anuales con información individualizada de las declaraciones del impuesto sobre la renta personal desde el año 1916¹. Inicialmente, la construcción de estas bases de datos estaba motivada por la necesidad de desarrollar análisis de la política tributaria federal desde la propia administración pública, si bien sus propiedades estadísticas las convirtieron más tarde en una fuente de información fundamental en el campo de la investigación socioeconómica².

Esta preocupación por disponer de bases de datos con información individualizada procedente de los registros tributarios se vio reforzada por la gran revolución

¹ Esta base de datos pionera fue el *Individual Complete Report File*.

² Véase WEBER y BRYANT (2005).

de la informática en la década de los años setenta del pasado siglo. La posibilidad de procesamiento computacional de cantidades enormes de observaciones en tiempos muy reducidos, que brindaba la tecnología de los ordenadores personales y sus programas de análisis estadístico, impulso decididamente el diseño de este tipo de bases de microdatos. Así, a finales de los setenta, la *Statistics of Income Division (SOI)* del IRS elaboró los primeros paneles de microdatos, aunque limitados en cuanto a las variables incluidas³. El primer panel de microdatos completo procedente del registro de las declaraciones del impuesto sobre la renta personal (*Federal Individual Income Tax*) fue elaborado por la SOI para el ejercicio fiscal de 1987 y su horizonte temporal se extendió hasta el ejercicio de 1996. Se trataba de una base de microdatos con estructura de panel expandido, en la que los distintos cortes transversales anuales, extraídos mediante muestreo estratificado, eran representativos de la población declarante del impuesto sobre la renta personal federal norteamericano. Posteriormente, en 2001 vio la luz un nuevo panel que incorpora los ejercicios impositivos desde 1999 hasta 2007, el último disponible.

En Europa, los antecedentes de la creación de bases de microdatos procedentes de registros administrativos tributarios los encontramos en países como Alemania y Suecia. En Alemania, fue una Ley Federal de Estadística de 1981 la que reconoció la necesidad de construir este tipo de bases estadísticas, cuyo ejemplo más representativo es el «Income tax scientific-use-file 1995» que elabora el *Research Data Center (RDC)* de la *Federal Statistical Office (Destatis)* con la colaboración de las oficinas estadísticas de los Laenders desde 1995 con una periodicidad trienal⁴. Las muestras incluyen un 10% de las declaraciones del impuesto federal sobre la renta personal alemán (unos 3 millones de observaciones muestrales)⁵. Los países nórdicos también cuentan con una amplia tradición en la construcción de bases de microdatos con información de origen tributario, destacando los casos de Suecia y Dinamarca, donde en algunos casos estas muestras son enlazadas con otras bases de microdatos con información sobre prestaciones sociales o contratos laborales.

En el campo de la construcción de bases de microdatos de origen tributario puede decirse que España no ha permanecido ajena a esta corriente liderada por Estados Unidos y los países del norte y centro de Europa. En el ámbito de la imposición sobre la renta de las personas físicas, el Instituto de Estudios Fiscales, organismo dependiente del Ministerio de Economía y Hacienda abordó a finales de la década de los años ochenta el primer Panel de Declarantes de IRPF, que inicialmente abarcó los ejercicios impositivos comprendidos entre 1982 y 1994, siendo posteriormente prolongado, en 2006, hasta el ejercicio de 1998⁶. Se trataba de un panel expandido,

³ Inicialmente estas bases de microdatos incluían información enlazada de la *Continuous Work History Sample (CWHS)* y de las *Individual Income Tax Return Cross Sectional Sample* anuales. Ver WEBER (2004).

⁴ Véase IZA (2009).

⁵ Véase MERZ *et al.* (2005).

⁶ Véase AYALA *et al.* (2004) y ONRUBIA *et al.* (2007).

representativo de la población declarante por el IRPF en cada año, y construido mediante extracción aleatoria proporcional (2% de las declaraciones presentadas por los contribuyentes de este impuesto ante la administración tributaria). En la primera mitad de los años noventa, fruto de la colaboración entre el Instituto de Estudios Fiscales y la Agencia Estatal de Administración Tributaria también se construyó un panel con microdatos correspondientes a las declaraciones presentadas por el Impuesto de Sociedades, que comprendía los ejercicios de 1991 a 1994⁷.

La reconocida utilidad de este tipo de bases de datos para el análisis económico aplicado, bien en su explotación directa, bien a través de su fusión estadística con otras bases de microdatos procedentes de encuestas o de otros registros administrativos, impulsó el interés del Ministerio de Economía y Hacienda por continuar, incluso ampliar esta oferta⁸. Así, en la segunda mitad de esta década y nuevamente fruto de la colaboración entre el Instituto de Estudios Fiscales y la Agencia Estatal de Administración Tributaria (AEAT), ha visto la luz una nueva serie de muestras anuales de microdatos de IRPF, iniciada con el ejercicio de 2002 y que hasta el momento incluye los ejercicios comprendidos hasta 2006⁹.

La experiencia obtenida por investigadores y analistas en la utilización del Panel de Declarantes de IRPF 1982-1998, junto con el auge creciente de los trabajos internacionales que utilizan bases de microdatos fiscales con estructura de panel, ponían de manifiesto la necesidad de dar continuidad al proyecto iniciado en el IEF hace veinte años. La potencialidad de las muestras anuales de declarantes del IRPF es muy grande, pero su carácter de muestras transversales (*cross-section samples*) limitan la realización de determinados análisis longitudinales que requieren la disposición de muestras de panel puro. Esta realidad ha llevado al Instituto de Estudios Fiscales, con la inestimable ayuda del Servicio de Estudios Tributarios y Estadísticas de la AEAT, a diseñar y construir un nuevo Panel de Declarantes de IRPF que cubra el periodo que se inicia en 1999 —año de entrada en vigor de la Ley 40/1998, que reformó con profundidad el IRPF español— y que abarque hasta el último ejercicio para el que se dispone de microdatos, 2007.

Los casi veinte años transcurridos desde el diseño del anterior Panel de IRPF y el momento en que se comenzó el proceso de construcción del nuevo Panel han permitido incorporar mejoras en el nuevo diseño estadístico. Asimismo, la generalización de la confección automática de las declaraciones de IRPF a partir de la implantación del programa de ayuda PADRE, junto con los avances informáticos en el registro de las declaraciones por parte de la AEAT, han permitido también contar con una fuente de información poblacional mucho más fiable y consistente en cuanto a la existencia de errores de cumplimentación que la utilizada en los ejercicios del an-

⁷ Véase ROMERO (2001).

⁸ Sobre la fusión de muestras de microdatos tributarios de IRPF con bases de microdatos procedentes de encuestas puede verse PICOS (2006).

⁹ Véase PEREZ y DIAZ DE SARRALDE (2009).

tigo Panel, en cuyos primeros años, la informatización de la administración tributaria estaba en su fase inicial. La utilidad de contar con una base de datos como el nuevo Panel de IRPF ha sido reconocida por el Consejo Superior de Estadística, lo que ha supuesto la inclusión del mismo como una Operación Estadística del Plan Estadístico Nacional. Este hecho supone dotar al proyecto de una continuidad en el tiempo, además de la obligación de realizar la difusión de acuerdo a los compromisos adquiridos.

El interés del Instituto de Estudios Fiscales por divulgar con la máxima amplitud el nuevo Panel de IRPF, especialmente entre los analistas socioeconómicos y los investigadores, está en el origen de la elaboración de este libro. Los principios de transparencia y generalización en el acceso a la información estadística han presidido desde los momentos iniciales la realización del proyecto. Por ello, se ha sido consciente de la necesidad de contar con una publicación que incorporase de forma detallada las especificaciones técnicas del diseño estadístico realizado, así como una descripción lo más completa posible de su estructura y contenido informativo. También es indudable que la amplia utilización de esta base de microdatos, menos conocida que las procedentes de las encuestas principales elaboradas por el Instituto Nacional de Estadística (INE), depende en gran medida de la difusión de sus potencialidades y de la clarificación de cómo llevar a cabo su explotación, al menos, en un nivel básico del análisis. Por ello, este libro incluye también una guía de utilización del Panel que cubre distintos aspectos, como la estimación poblacional de la población de referencia, el cálculo de magnitudes agregadas a partir de los microdatos, los cálculos poblacionales de las medidas estadísticas fundamentales y el cálculo de medidas de distribución habituales en la utilización de este tipo de bases de microdatos. La posibilidad de vincular las declaraciones de IRPF de los cónyuges que, de acuerdo con la legislación del IRPF español, optan por tributar separadamente dentro del matrimonio, abre una interesante aproximación de estudio, cambiando la unidad observacional de análisis del declarante al hogar fiscal. Esta guía de utilización aborda esta cuestión, lo mismo que la forma de trabajar longitudinalmente con las muestras, empleando la posibilidad de seguir ejercicio a ejercicio tanto a los declarantes del impuesto como a los hogares fiscales en los que éstos se encuadran.

En cuanto a su contenido, el libro se organiza de la siguiente manera. Tras esta introducción, el primer capítulo está dedicado a presentar el diseño estadístico del Panel de IRPF, justificando la elección realizada en cuanto al tipo de base de datos y la estructura de sus ficheros. En este capítulo también se describe el proceso de extracción de las muestras tanto en el año base, el ejercicio 2003, como en el resto de años incluidos, tanto retrospectivamente (de 1999 a 2002) como prospectivamente (de 2004 hasta 2007). El segundo capítulo de la obra recoge las características básicas del Panel, describiendo el contenido de los ficheros y sus variables, tanto monetarias como no monetarias. También se incluye en el mismo la definición del hogar fiscal como unidad observacional alternativa a la declaración, que fue la unidad de muestreo utilizada en la construcción del Panel. El interés por construir una base

de microdatos lo más rica posible en cuanto a su potencial de explotación ha llevado a los autores del proyecto a incorporar una serie de variables adicionales construidas a partir de la información original aportada por la AEAT en sus declaraciones. El tercer capítulo describe la creación de estas variables, así como el procedimiento diseñado para facilitar el seguimiento longitudinal, con «entradas» y «salidas» tanto de declarantes como de hogares fiscales. El capítulo cuarto del libro tiene por objeto ofrecer un análisis de la representatividad de las muestras del Panel. Para ello se analiza su capacidad de inferencia poblacional para las principales variables económicas del IRPF. Este capítulo incorpora además un estudio sobre el «desgaste» (*attrition*) del Panel, así como un análisis detallado de las anomalías detectadas en la identificación de algunas observaciones originales, especialmente referidas a las relaciones entre declarantes y en la conformación de los hogares. Junto con la guía de utilización del Panel antes reseñada, el capítulo quinto incluye también la presentación de la estructura final de los ficheros y de las variables incluidas en los mismos, con los valores que pueden tomar éstas.