

Capítulo 5

Duración del periodo de transición hacia el primer empleo significativo

“El tiempo lineal es aquel que coincide con la narración personal de cada uno, en el que vamos escribiendo nuestras ilusiones y decepciones. A veces ese tiempo se acelera y otras, vencidos por la tristeza, se detiene y parece que nunca más va a arrancar”

Autor desconocido

5.1 Consideraciones generales

En el presente capítulo realizamos una aproximación al proceso de transición hacia un empleo significativo de los jóvenes que salen del sistema educativo durante la década de los noventa a través de un *análisis de duración* de este periodo transitorio, con la finalidad de detectar qué factores influyen en la longitud de este periodo, clave en la trayectoria laboral de este colectivo. Es decir, buscamos determinar cuáles son las causas que provocan que, una vez que el joven ha decidido integrarse en el mundo laboral dando por finalizada su formación, su proceso se prolongue durante más o menos tiempo.

Con anterioridad a la propuesta y estimación de un modelo que mejor explique el comportamiento de nuestros datos de duración, hemos realizado un repaso a algunas

de las cuestiones de importancia en el *análisis de duración*, expuestas en el capítulo metodológico, y que conviene retomar para hacer una valoración de cuáles son las características técnicas de los datos procedentes del módulo. Estas cuestiones, como se señaló, condicionan el *análisis de duración* tanto en su planteamiento inicial como en su desarrollo posterior. En este sentido, y a diferencia de otros análisis estadísticos, la fuente de datos, y en concreto, el diseño de la misma, determinan en gran medida esta etapa estadística. En particular, se repasan aspectos tales como la censura, el origen de los datos, el diseño del experimento, etc.

Censura

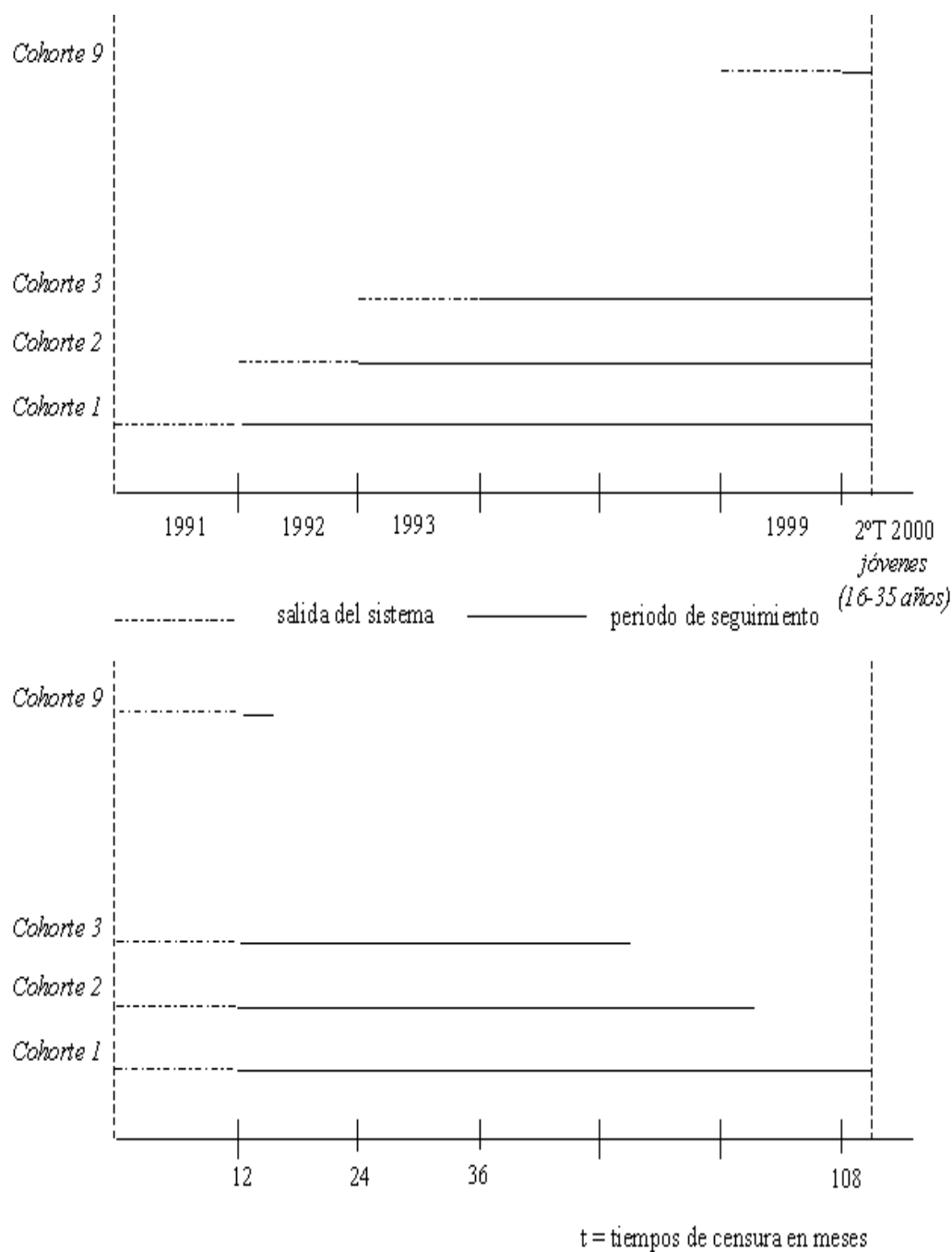
En primer lugar, analizamos el tipo de censura que encontramos en los datos. La censura, que caracteriza a nuestros datos, está provocada porque algunos de los jóvenes todavía no habían encontrado su primer empleo significativo en el momento en el que son encuestados (2º trimestre del 2000). Para estos individuos, por lo tanto, nos es desconocida la duración exacta de su proceso de transición aunque sí que sabemos que ésta sobrepasa un determinado valor. Este tipo de censura es la que hemos denominado *censura por la derecha*. Siendo aún más precisos, podemos indicar que se trata de censura por la derecha *tipo I* ya que se observan diferentes valores para los tiempos de censura (*multiply censored*), tal y como se aprecia en la figura 5.1.

Por otro lado, y como ya se aventuró, la censura es no informativa dado que viene impuesta por el diseño del experimento. En nuestro caso, la causa que ha inducido la censura no está directamente relacionada con el proceso de transición en sí, sino con el procedimiento utilizado para seleccionar la muestra y extraer la información sobre la duración. Por lo tanto, para la estimación de los modelos de duración no es necesario conocer la distribución de los tiempos de censura (Neumann, 1997).

Selección de la muestra

En cuanto al esquema de selección, hemos de señalar que la estructura del cuestionario del módulo permite hacer un estudio retrospectivo de los jóvenes que entran en el mercado de trabajo tras su salida inicial del sistema educativo. En el análisis de este proceso de transición, por lo tanto, estamos teniendo en cuenta a aquellos individuos que *entran* en el estado cuya duración se quiere estudiar (estado transitorio entre el sistema educativo y el empleo estable) y a los que se les sigue durante un periodo concreto (hasta

Figura 5.1: Estructura de la censura y periodo de seguimiento de cada cohorte



el 2º trimestre del 2000). En consecuencia, y atendiendo a la clasificación de Lancaster (1990), se trata de una selección flujo de la muestra. Este esquema de selección simplifica en gran medida la aplicación de los procedimientos estadísticos que no requieren de correcciones que habrían de utilizarse ante otros esquemas de selección alternativos.

De esta descripción del procedimiento de selección, y del posterior seguimiento de los individuos, se deduce que el tiempo de observación no es el mismo para todos. En concreto, aquellos individuos que pertenecen a las primeras cohortes que salen del sistema educativo son observados durante un periodo de tiempo más largo. La razón está en que el módulo incluye varias cohortes de individuos, en particular, las formadas por jóvenes que salen del sistema educativo entre los años 1991 y 2000, a las que, independientemente de cuando se inició su seguimiento, se les continúa observando hasta el momento de la encuesta (2º trimestre del 2000). En principio, cabría esperar que la proporción de los que finalmente transitan hacia un empleo significativo sea mayor para aquellos individuos que permanecen más tiempo en el estudio. Desde un punto de vista descriptivo, la diferencia en el éxito en el proceso de transición se muestra en la tabla 5.1 donde se recoge el porcentaje de jóvenes de cada cohorte que consigue situarse en el mercado de trabajo, cifra que varía entre el 46,5% y el 65,3% correspondientes a las cohortes de 1999 y 1991 respectivamente¹.

Tabla 5.1: Distribución del éxito en la consecución de un empleo significativo según la cohorte de salida del sistema educativo

| | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | Total |
|----------|---------------|----------------|----------------|----------------|----------------|----------------|---------------|---------------|---------------|----------------|
| no éxito | 503 (34,7) | 658 (38,9) | 639 (37,4) | 696 (38,6) | 708 (38,9) | 686 (40,3) | 685 (41,5) | 678 (46,3) | 742 (53,5) | 5995 (40,9) |
| éxito | 948 (65,3) | 1034 (61,1) | 1067 (62,5) | 1107 (61,4) | 1111 (61,1) | 1016 (59,7) | 967 (58,5) | 785 (53,7) | 646 (46,5) | 8681 (59,1) |
| Total | 1451 | 1692 | 1706 | 1803 | 1819 | 1702 | 1652 | 1463 | 1388 | 14676 |

Nota: Dentro de la categoría “éxito” se han incluido los individuos que transitan directamente sin la existencia de un periodo intermedio entre los dos estados.

Fuente: Elaboración propia a partir del módulo de la EPA 2º trimestre del 2000

La conclusión que se extrae de este resultado es que encontramos un mayor porcentaje

¹Recordamos que los individuos que finalizaron, abandonaron o interrumpieron sus estudios iniciales en el mismo año de la encuesta han sido eliminados del análisis pues no han tenido tiempo material suficiente para iniciar su proceso de transición.

de censura en las cohortes más jóvenes. Pero, los procedimientos estadísticos utilizados manejan de forma correcta este tipo de censura por lo que en principio no debería resultar un problema la existencia de una mayor censura. Por otro lado, las frecuencias recogidas en la tabla 5.1 se refieren al conjunto del periodo, y las probabilidades que vamos a analizar concretan la ocurrencia del suceso en un momento dado, realizando un análisis condicional del proceso de transición. Esto es, no interesa tanto la probabilidad de que se produzca el suceso a lo largo de todo el periodo de observación sino cuál es la probabilidad de que ocurra una vez que el individuo ha permanecido un tiempo en esa situación.

Tipología del modelo de duración

Siendo el suceso que se está estudiando tan específico, hemos reducido de antemano las posibilidades del análisis estadístico. En realidad, y dada la imposibilidad de que se repita el suceso que se analiza², estamos ante los llamados modelos de *vida-muerte* en los que únicamente se observa una duración del suceso considerado por individuo.

Sin embargo, cabe la posibilidad de que la salida del periodo transitorio, aun siendo a un primer empleo significativo, sea diferente si se tiene en cuenta la calidad del mismo en términos de su adecuación al nivel educativo del joven. Esto es, otra posible fuente de diferenciación en los tiempos de espera hasta la consecución de ese primer empleo puede ser la calidad del empleo aceptado. Algunos jóvenes, por motivos diversos, tienden a no demorar su entrada a la vida activa esperando un empleo acorde con su formación, y presentan procesos de transición más cortos. Por ello, aunque en este trabajo no se aborda esta cuestión, y nos ceñiremos al primer empleo significativo sin hacer ninguna distinción, queda abierta una futura vía de investigación en la que diferenciaremos entre empleos significativos en los que existe afinidad entre la ocupación y la formación recibida y aquéllos en los que se detecta un desajuste entre los estudios poseídos y los requeridos por el empleo. Por lo tanto, este análisis en el que se observan diferentes estados de salida se encuadraría dentro de la metodología de los modelos de riesgos en competencia.

Una vez establecidas las características más relevantes de los datos de duración que vamos a analizar, pasamos sin más preámbulos al estudio de la duración del proceso de transición del sistema educativo a un empleo significativo. En nuestro caso, el análisis de duración de este periodo de transición lo vamos a realizar en dos etapas.

²Los futuros empleos de estos jóvenes nunca volverán a ser un *primer* empleo significativo.

En primer lugar, nos hemos planteado una estimación no paramétrica de las funciones de supervivencia para el conjunto de la población y para los subgrupos que se obtienen al considerar diferentes variables. Este análisis nos va a permitir obtener estimaciones tanto de la función de supervivencia como de la función de riesgo, a partir de las cuales identificaremos algunos de los factores que afectan a la duración del proceso de transición del sistema educativo a un empleo significativo.

En una segunda etapa, y avanzando en la complejidad de los modelos, estimaremos modelos semiparamétricos y paramétricos en los que se incluyen como variables explicativas aquellos factores disponibles y con posible relevancia en la duración de este proceso. Por lo tanto, en esta etapa analizamos, en mayor detalle, cuáles son los factores más significativos que rigen el periodo de transición entre la finalización de los estudios y el primer empleo significativo y cuantificamos los efectos de dichos factores. En particular, tratamos de comprobar qué características individuales, así como qué factores del entorno económico inciden sobre la probabilidad de encontrar un primer empleo significativo tras la salida del sistema educativo. Dentro de esta etapa, dedicamos un apartado específico al estudio de la dependencia de la duración, cuestión de gran relevancia por sus importantes implicaciones desde un punto de vista de las medidas de política económica encaminadas a facilitar el acceso al mercado de trabajo de los jóvenes.

La forma de actuar, como se deduce de la lectura del capítulo, está supeditada a los objetivos que en cada caso se pretenden.

5.2 Análisis no paramétrico de la duración del periodo transitorio entre la salida del sistema educativo y el primer empleo significativo

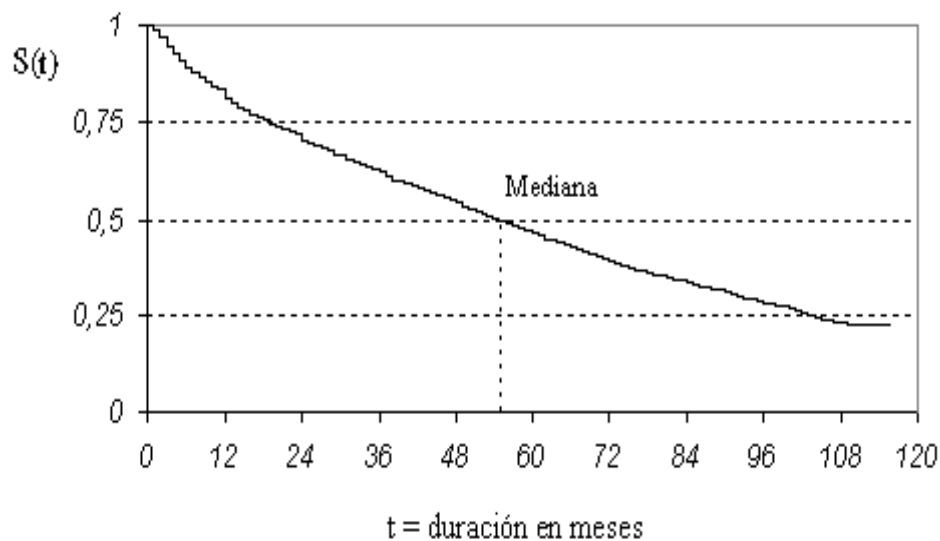
5.2.1 El estimador de Kaplan-Meier

Una primera aproximación al estudio de la duración de este periodo de tránsito hacia un primer empleo significativo consiste en estimar la función de supervivencia mediante el estimador no paramétrico de Kaplan-Meier, también llamado estimador del producto-límite (Kaplan y Meier, 1958). Esta función de supervivencia, como ya se ha señalado, recoge la probabilidad de que la duración del proceso de transición sea mayor

a un número determinado de meses³.

La estimación de la función de supervivencia, para algunos valores de t , aparece en la tabla 5.2 y una representación gráfica de la misma en la figura 5.2⁴. En la tabla, junto con la función de supervivencia que se encuentra en la columna 5, se incluyen otras columnas que recogen para cada mes: el número de *supervivientes* al inicio (*Total*), el número de jóvenes que consiguen un empleo significativo (*Encuentran*) y el de aquellos que no lo han conseguido durante ese tiempo y dejan de ser observados (*Censurados*). Finalmente, las dos últimas columnas muestran una estimación de la función de riesgo acumulada Nelson-Aalen (Nelson, 1972 y Aalen, 1978) y sus correspondientes errores estándar.

Figura 5.2: Estimación no paramétrica de la función de supervivencia (método Kaplan-Meier)



Una lectura detenida del comportamiento de la curva representada en la figura 5.2 pone de manifiesto que la función de supervivencia se caracteriza por descender muy

³En nuestro caso la función de supervivencia está referida al mes porque ha sido la unidad de medida del tiempo elegida en este trabajo (véase a este respecto el apartado 3.1.1 del capítulo 3).

⁴El cálculo de la función de supervivencia se realiza con las observaciones que no presentan duraciones nulas, esto es, la muestra se reduce a 14287 individuos tras eliminar las 389 observaciones cuya duración es igual a cero.

Tabla 5.2: Función de supervivencia de Kaplan-Meier y función de riesgo acumulada Nelson-Aalen

| Duración (meses) | Total (n_j) | Encuentran (d_j) | Censurados (m_j) | Función de supervivencia | Std. error | Función de riesgo acumulada | Std. error |
|---------------------|--------------------|-------------------------|-------------------------|-----------------------------|---------------|--------------------------------|---------------|
| 1 | 14287 | 199 | 0 | 0,986 | 0,001 | 0,014 | 0,001 |
| 2 | 14088 | 270 | 0 | 0,967 | 0,002 | 0,033 | 0,002 |
| 3 | 13818 | 345 | 0 | 0,943 | 0,002 | 0,058 | 0,002 |
| 4 | 13473 | 267 | 0 | 0,924 | 0,002 | 0,078 | 0,002 |
| 5 | 13206 | 237 | 0 | 0,908 | 0,002 | 0,096 | 0,003 |
| 6 | 12969 | 219 | 0 | 0,892 | 0,003 | 0,113 | 0,003 |
| 7 | 12750 | 214 | 36 | 0,877 | 0,003 | 0,130 | 0,003 |
| 8 | 12500 | 155 | 13 | 0,867 | 0,003 | 0,142 | 0,003 |
| 9 | 12332 | 163 | 23 | 0,855 | 0,003 | 0,155 | 0,003 |
| 10 | 12146 | 162 | 140 | 0,844 | 0,003 | 0,169 | 0,004 |
| 11 | 11844 | 142 | 3 | 0,834 | 0,003 | 0,180 | 0,004 |
| 12 | 11699 | 297 | 43 | 0,812 | 0,003 | 0,206 | 0,004 |
| 18 | 10233 | 99 | 8 | 0,757 | 0,004 | 0,276 | 0,005 |
| 24 | 9459 | 179 | 55 | 0,705 | 0,004 | 0,347 | 0,006 |
| 30 | 8321 | 90 | 7 | 0,664 | 0,004 | 0,407 | 0,006 |
| 36 | 7686 | 78 | 1 | 0,621 | 0,004 | 0,472 | 0,007 |
| 42 | 6617 | 59 | 7 | 0,580 | 0,004 | 0,540 | 0,007 |
| 48 | 6104 | 58 | 3 | 0,544 | 0,004 | 0,605 | 0,008 |
| 54 | 5093 | 64 | 2 | 0,503 | 0,005 | 0,683 | 0,009 |
| 60 | 4626 | 59 | 2 | 0,465 | 0,005 | 0,761 | 0,010 |
| 72 | 3298 | 50 | 0 | 0,393 | 0,005 | 0,928 | 0,012 |
| 84 | 2148 | 31 | 0 | 0,332 | 0,005 | 1,095 | 0,014 |
| 96 | 1255 | 22 | 2 | 0,280 | 0,005 | 1,265 | 0,018 |
| 108 | 524 | 6 | 53 | 0,230 | 0,006 | 1,459 | 0,025 |
| 111 ^(*) | 428 | 4 | 416 | 0,222 | 0,006 | 1,494 | 0,027 |
| 116 | 1 | 0 | 1 | 0,222 | 0,006 | 1,494 | 0,027 |
| | | 8292 | 5995 | | | | |

Nota: Las estimaciones de la función de supervivencia ($\hat{S}(t) = \prod_{t_{(j)} \leq t} \frac{n_j - d_j}{n_j}$) y de la función de riesgo acumulada ($\hat{H}(t) = \sum_{t_{(j)} \leq t} \frac{d_{(j)}}{n_{(j)}}$) han sido calculadas para el conjunto de valores de duración pero sólo se presentan algunos de ellos en esta tabla. Asimismo el total de las columnas *Encuentran* y *Censurados* es el que se obtiene incluyendo todas las duraciones.

(*) Último valor para el que se tienen duraciones no censuradas.

lentamente y por no alcanzar, como era de esperar debido a la presencia de censura, el valor mínimo teórico de esta función⁵. Únicamente el intervalo correspondiente a los dos primeros años muestra una pendiente más pronunciada, en el que la función cae algo más rápidamente. Este descenso lento nos está indicando que los jóvenes encuentran importantes dificultades antes de encontrar ese primer empleo significativo. De hecho, la probabilidad de conseguirlo antes del primer año es tan sólo de un 19%, un 50% tarda al menos cuatro años y medio⁶ y la probabilidad de que el proceso se prolongue más allá de los 96 meses (ocho años) es de un 28%.

5.2.2 La probabilidad instantánea de transitar o función de riesgo

En relación a la función de riesgo no paramétrica que recoge, desde un punto de vista dinámico, la probabilidad de que la transición ocurra en un momento dado, o lo que es lo mismo, la tasa de riesgo de que se produzca a lo largo del tiempo, sus valores no exceden del 2% (véase figura 5.3), alcanzándose esta cifra una vez han transcurrido dos meses desde la salida del sistema educativo⁷. Los valores son ligeramente inferiores a estimaciones de esta función en otros países⁸. La función de riesgo no presenta, sin

⁵Téngase en cuenta que estamos suponiendo que la transición al empleo (significativo) tendrá lugar en algún momento del tiempo. Esta suposición es ciertamente válida en el caso de los jóvenes varones y no resulta tan desacertada entre las mujeres si se acepta, que en la actualidad, la incorporación de la mujer al mundo laboral a esas edades es prácticamente total.

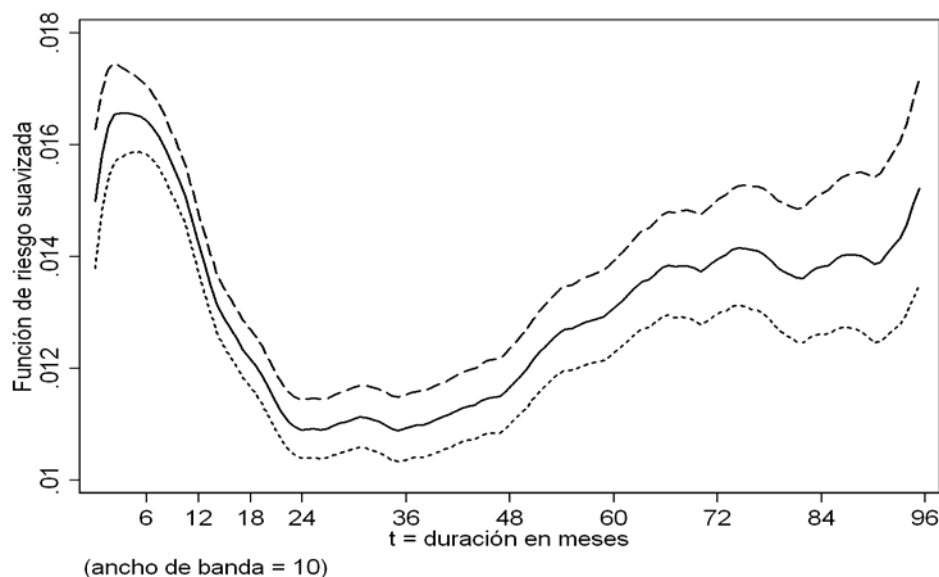
⁶El cálculo de la media de los tiempos de supervivencia en el análisis de duración no tiene en cuenta que algunos de los valores de duración están incompletos y, por lo tanto, infraestima la duración media. Es por ello que la mediana, calculada a partir de la función de supervivencia Kaplan-Meier, se ha convertido en la medida de tendencia central más utilizada, tomándose su valor como representante de toda la población.

⁷Para estimar esta función se ha calculado primeramente la función de riesgo acumulada -utilizando el estimador de Nelson-Aalen que aparece en la tabla 5.2- y a continuación se ha aplicado una técnica de suavizado que permite obtener una aproximación de esta función. El procedimiento de suavizado ha consistido en la aplicación de una función Kernel (la de Epanechnikov) tomando como amplitud de banda el valor 10, una vez se ensayaron varios anchos de banda. Véase Klein y Moeschberger (1997) para una descripción de la técnica usada.

⁸Serneels (2002a, pág. 20) puntualiza que la función de riesgo no llega a superar el 10% en los países de la OCDE. No obstante, ha de tenerse en cuenta que la mayoría de los estudios se centran en el análisis de la salida a un empleo tras la finalización de los estudios y no específicamente en la salida a un empleo significativo. Esta distinción es importante pues, como parece obvio, la salida a cualquier empleo en un momento dado ha de tener una probabilidad superior que la salida a un empleo al que se le imponen unas condiciones más restrictivas. Por otra parte, algunos estudios revelan importantes diferencias respecto a otros países europeos en cuanto al tiempo que tardan los jóvenes en integrarse

embargo, picos muy pronunciados en determinados meses ya que el procedimiento de suavizado los ha eliminado⁹.

Figura 5.3: **Estimación no paramétrica de la función de riesgo (método de suavizado Kernel)**



El estudio del comportamiento en el tiempo de esta función de riesgo estimada nos permite hacer una primera valoración de la *dependencia de la duración*, o sea, determinar si la probabilidad condicionada de que el periodo transitorio finalice resulta menor o mayor a medida que la longitud de dicho periodo es mayor. En este sentido, hemos de decir que la función muestra poca variabilidad aunque su comportamiento no es monótono sino que más bien tiene forma de U¹⁰.

En concreto, un análisis pormenorizado nos lleva a detectar en los inicios del proceso laboralmente, lo que en términos de la función de riesgo se traduciría en unas tasas de salida más bajas.

⁹Ya se advirtió una tendencia más o menos clara al redondeo en la duración del proceso de transición lo que asociamos al procedimiento de recogida de la información. Esta situación es bastante habitual cuando se apela a la memoria del entrevistado (Ahn y Ugidos, 1995b; Congregado y García Pérez, 2002; González Betancor *et al.*, 2004 ; Bover y Gómez, 2004). Sin embargo, en ocasiones, los picos son consecuencia de la verdadera distribución del suceso aunque no siempre es fácil distinguirlo.

¹⁰A excepción del primer tramo creciente que podría venir explicado por la existencia de un periodo inicial, relativamente corto, que incluye por un lado, un tiempo en el que el joven se da un respiro antes de iniciar su búsqueda de un empleo -sobre todo si el final de los estudios coincide con una etapa vacacional- y, por otro, un periodo de aprendizaje durante el cual el joven se ejercita en las técnicas

de transición una ligera pendiente decreciente, que reduce el valor de la función de riesgo casi a la mitad, y que sugiere una pérdida en las posibilidades de emplearse a medida que transcurre el tiempo sin que se haya conseguido el empleo significativo, esto es, se detectaría la presencia de una débil dependencia de la duración negativa. Tras ese tramo inicial, que abarcaría aproximadamente los dos primeros años posteriores a la salida del sistema educativo, se observa un periodo en el que la función se mantiene más o menos constante, por lo que en este tramo no existe dependencia de la duración. Finalmente, en la última etapa de este proceso, la función de riesgo muestra un repunte, sin alcanzar en cualquier caso sus valores iniciales, lo que indicaría que para aquellos individuos que aún no han transitado después de tres años, la probabilidad instantánea de que realicen el tránsito va creciendo paulatinamente (dependencia de la duración positiva).

En resumen, y a excepción de los jóvenes que inician su vida laboral con un empleo significativo de forma casi inmediata tras la salida del sistema educativo, parece que es necesario experimentar, en los inicios de la actividad laboral, un periodo caracterizado por un alto grado de inseguridad e inestabilidad, en el que las combinaciones de desempleo y de empleos temporales muy cortos son muy diversas, antes de alcanzar cierta estabilidad laboral. La duración mínima de este periodo estaría entre los dos y tres años.

Ante este resultado nos hemos planteado realizar un breve repaso de la literatura teórica y empírica sobre la dependencia de la duración que nos ayude a determinar cuáles son los motivos que pueden justificar este comportamiento.

El término *dependencia de la duración* se acuña para recoger un hecho que, en principio, puede resultar paradójico: la probabilidad de que la búsqueda de un empleo sea exitosa puede verse afectada por el tiempo de búsqueda, reduciendo o aumentando esta probabilidad^{11,12}. A este respecto, el modelo básico de búsqueda de empleo

de búsqueda para mejorar sus cualidades como candidato, lo que va a redundar en un incremento del número de ofertas que recibe tras su primer contacto con el mundo laboral. Algunos estudios detectan un comportamiento similar en la función de riesgo que viene a significar que la probabilidad instantánea de salida de la situación transitoria es muy alta en los primeros momentos de salida del sistema educativo (McVicar y Podivinsky, 2001, Nguyen y Taylor, 2003 o Carroll, 2004).

¹¹Ryan (2001) considera que el tipo de dependencia de la duración no puede ser una característica nacional sino que sea cual sea su comportamiento (positivo, negativo o nulo), éste debe ser el mismo para todos los países. Por su parte, Serneels (2002b) es de la opinión contraria al abogar por diferentes explicaciones, entre ellas algunas institucionales, que pueden determinar la dependencia de la duración.

¹²Cuando el efecto de la duración se traslada a periodos posteriores, es decir, incrementa (o disminuye) la probabilidad de experimentar en un futuro más periodos de desempleo, se habla de dependencia

infiere que la duración del periodo de búsqueda es independiente del tiempo que se está buscando empleo. Las posteriores adaptaciones de este modelo para recoger unos supuestos más realistas parecen plantear la existencia de dependencia al tener en cuenta factores variables en el tiempo pero no muestran resultados concluyentes.

La existencia de esta dependencia tiene relevancia en la implementación de las políticas laborales dirigidas a fomentar la participación laboral y a mejorar las expectativas de los desempleados, principalmente, de cara a decidir cuál es el momento óptimo de aplicación de las mismas para que su efectividad sea máxima. Bien, una vez que el individuo ha sido clasificado como parado de larga duración (o como inactivo), como es el caso de aquellas políticas orientadas a reducir el desempleo de larga duración¹³, o bien desde el inicio del periodo de búsqueda de empleo mediante políticas que fomenten la formación o el reciclaje para evitar que el capital humano adquirido se deteriore o que el desánimo aparezca.

En general, en las primeras investigaciones que analizan con cierto detalle esta cuestión se detecta dependencia de la duración *negativa* -aunque ésta no es muy acusada- que implica que cuánto más tiempo se está buscando, menores son las probabilidades de encontrarlo (Franz *et al.* (1997) y Chuang (1999))¹⁴.

En el caso de la búsqueda desde el desempleo¹⁵, esta dependencia suele explicarse bien por la pérdida de capital humano que se produce al estar desempleado o por la percepción por parte del empresario de que el desempleo es un indicador negativo de las cualidades laborales del trabajador¹⁶. Esta justificación tiene su validez siempre y cuando el individuo no aproveche la ocasión para reciclar sus conocimientos en el caso

de la duración *retardada* (*lagged dependence duration*) o dependencia de la ocurrencia (*occurrence dependence*).

¹³Obsérvese que estas políticas se fundamentan en la creencia de que existe dependencia de la duración negativa y que, por tanto, es conveniente el uso de incentivos en el caso de personas que permanecen largo tiempo desempleadas para que salgan de esa situación, de la que difícilmente saldrían sin ese apoyo institucional.

¹⁴Machin y Manning (1999, pág. 17) realizan un estudio de la dependencia de la duración dentro de la zona europea, concluyendo que no parece existir una fuerte dependencia negativa una vez que se han controlado algunas características y que, posiblemente, un mayor control de las características explicativas puede conducir a dependencia de la duración positiva.

¹⁵En este estudio a los periodos de inactividad se les da el mismo tratamiento que a los periodos de desempleo por la dificultad de separar para este colectivo las dos situaciones.

¹⁶También se señala la pérdida de confianza del propio trabajador sobre sus propias habilidades como un factor que produce dependencia negativa.

de desempleados adultos o para completar su formación en el caso de los jóvenes. Por otro lado, y desde la perspectiva del empleo¹⁷, las investigaciones empíricas corroboran que la probabilidad de abandonarlo es menor a medida que la duración del empleo es mayor, produciéndose a su vez una mayor estabilidad en el empleo. Este resultado requiere de ciertas matizaciones, sobre todo, tras la masiva utilización de la contratación temporal con especial incidencia entre los jóvenes que ha reducido la duración del empleo (García-Serrano y Jimeno, 1999).

Ahora bien, como se muestra en la tabla 5.3 en la que se recogen algunos de los trabajos en los que se estudia la dependencia de la duración del desempleo (en particular, el primer desempleo juvenil)¹⁸, no siempre los estudios empíricos han ratificado la existencia de dependencia negativa¹⁹. En algunos casos, la explicación de este reverso en los resultados está en la dificultad de separar la dependencia de la duración cuando hay un efecto de selección o existen factores que no se han podido tener en cuenta (sesgo de selección o heterogeneidad no observada). En este sentido, tal como vimos en el capítulo 3, la dependencia de la duración negativa puede estar motivada por la existencia de heterogeneidad no observada que no ha sido controlada adecuadamente en el estudio, y que cuando se controla reduce o elimina el efecto negativo del propio devenir del tiempo sobre la duración, como así ocurre en los trabajos de Andrews *et al.* (2002) y Nguyen y Taylor (2003).

No obstante, el cambio en la tendencia de los resultados en las investigaciones empíricas ha obligado a buscar argumentos en contra de la dependencia de la duración negativa o al menos a favor de que ésta no tenga un comportamiento monótono durante toda la duración²⁰. Las argumentaciones han sido varias. Entre ellas, la existencia de cobertura por desempleo limitada en el tiempo, la efectividad de las políticas activas de empleo dirigidas a los desempleados de larga duración (van den Berg y van Ours, 1994),

¹⁷Esta distinción parece tener cierto sentido ya que no es lo mismo estar buscando un empleo mejor (desde una posición de empleado) que estar buscando un primer empleo significativo.

¹⁸El hecho de si la dependencia es más fuerte entre los jóvenes que entre los adultos es un tema aún poco explorado. Aunque los jóvenes parecen a primera vista más resistentes, también son el colectivo más inmaduro y vulnerable. Algunas evidencias al respecto las podemos encontrar en Fougère y Kamionka (1992), en Allaire, Cahuzar y Tahar (1995) y Cockx y Dejemeppe (2002) no habiendo encontrado resultados concluyentes en este sentido.

¹⁹Serneels (2002b) realiza una revisión más exhaustiva de la dependencia de la duración sin circunscribirse al colectivo de jóvenes que inician su actividad laboral.

²⁰La dependencia positiva, según la teoría de la búsqueda de empleo, suele estar asociada a una caída del salario de reserva.

la distorsión que produce el ciclo económico (Arulampalam y Stewart, 1995 y van den Berg y van der Klaauw, 2000) o la segmentación del mercado de trabajo (García Blanco, 1995).

Tabla 5.3: Estudio de la dependencia de la duración en varios países

| Autores | País | Modelo sin factor de heterogeneidad | Modelo con factor de heterogeneidad (<i>distribución</i>) |
|--|------------|-------------------------------------|---|
| Wolpin (1987) | USA | negativa | |
| Franz <i>et al.</i> (1997) | Alemania | negativa | - |
| D'Addio (1999) | Francia | negativa | ligeramente negativa (<i>gamma</i>) |
| Chuang (1999) | Taiwan | negativa | negativa (<i>gamma</i>) |
| Andrews <i>et al.</i> (2002) | Inglaterra | negativa | no dependencia (<i>varias</i>) |
| McVicar y Podivinsky (2001) | Irlanda | negativa | * |
| Nguyen y Taylor (2003) | USA | negativa | positiva (<i>varias</i>) |
| Cockx y Dejemeppe (2002) | Bélgica | | no dependencia (<i>gamma y no paramétrica</i>) |
| Cañada <i>et al.</i> (1998) | España | negativa | |
| Alba (1998b) | España | negativa | |
| Aguilar (2002) | España | positiva/negativa | positiva (<i>gamma</i>) |
| Basulto <i>et al.</i> (2002) | España | negativa | |
| Congregado y García Pérez (2002) | España | negativa | * (<i>no paramétrica</i>) |
| González Betancor <i>et al.</i> (2004) | España | - | positiva (<i>gamma y gaussiana</i>) |

Nota: * significa que la heterogeneidad no observada resultó no significativa.

Fuente: Elaboración propia

Así, por ejemplo, la existencia de cobertura por desempleo implica que a medida que se va acercando la fecha de finalización de dicha cobertura, el desempleado intensifique su búsqueda de un empleo o reduzca su salario de reserva para incrementar las posibilidades de encontrarlo.

Por lo que respecta a la teoría de los mercados de trabajo segmentados, que establece

la existencia de dos tipos de trabajos (buenos y malos), la probabilidad de conseguir un trabajo “bueno” decrece con el tiempo porque las habilidades para realizar este tipo de trabajo se van perdiendo en la medida que no se utilizan. En este sentido, en el inicio de la búsqueda el trabajador irá rechazando aquellos empleos que no son buenos -en términos de salarios- hasta un determinado momento a partir del cual aceptará cualquier empleo (bueno o malo)²¹.

Esta podría ser una explicación plausible a los resultados obtenidos en este análisis no paramétrico pues como ya se expuso existen evidencias contrastadas de que en España se ha producido una importante segmentación del mercado de trabajo que ha afectado al empleo de los jóvenes (Bentolila y Dolado, 1994), aunque para una confirmación de la misma se requiere un estudio más exhaustivo de los datos que se llevará a cabo en los siguientes apartados²².

5.2.3 Comparativa de curvas de supervivencia

Dado que la duración del proceso de transición puede venir determinada por las características familiares y personales del individuo así como por factores ambientales e institucionales, el siguiente paso ha sido indagar sobre cuál es la influencia de algunos de estos factores. El estimador de Kaplan-Meier es un estimador no paramétrico de ahí que, para realizar dicha tarea, haya sido necesario obtener las funciones de supervivencia para los diferentes subgrupos definidos a partir de las categorías de las variables explicativas. El inconveniente de este método es que no permite controlar la existencia de variables correlacionadas, sin embargo, puede ser útil para identificar cuáles son los principales determinantes.

²¹García Blanco (1995) emplea un argumento similar para explicar la dependencia de la duración del desempleo retardada, señalando que las características no observadas de los trabajadores son suplidas por los empleadores por el número y la duración de los episodios de desempleo experimentados previamente por los candidatos. De tal forma que los individuos con peores registros tienen también menores oportunidades de emplearse.

²²Las otras teorías apuntadas no parecen adecuadas al suceso que aquí se analiza, la transición de los jóvenes al mercado laboral. En primer lugar, y con respecto a la cobertura por desempleo, se trata de un colectivo que no participa del sistema de prestaciones por desempleo ya que, en general no reúne los requisitos mínimos de cotización, al menos según la legislación sobre prestación por desempleo en España. Por lo que se refiere a las políticas activas dirigidas a los jóvenes, cuyo estudio histórico se realizó en el capítulo 1, éstas estaban dirigidas a combatir el desempleo entre los jóvenes y no exclusivamente a favorecer el empleo estable, por lo que más que ayudar a que el proceso de transición se acortara, han podido tener el efecto contrario.

En la figura 5.4 se representan las funciones de supervivencia teniendo en cuenta el sexo, la edad de salida del sistema educativo, el nivel de estudios, el sector de estudios, el máximo del nivel de estudios de los padres y el tiempo máximo de búsqueda de un empleo. Para complementar este análisis gráfico también se han calculado las duraciones medianas del proceso de transición para los colectivos resultantes de seleccionar las diferentes categorías de cada variable considerada, que aparecen recogidas en la tabla 2 del anexo 3.

La mera observación de la posición de las curvas evidencia que las funciones de supervivencia de algunas de las categorías de cada variable muestran diferencias significativas frente al resto. A continuación vamos, entonces, a realizar un estudio individualizado de estos gráficos. El análisis al diferenciar por género nos revela que la probabilidad de supervivencia es siempre mayor en el caso de las mujeres para cualquier valor de duración, al encontrarse su curva en todo momento por encima de la de los hombres. Esto significa que la probabilidad de mantenerse en esa situación transitoria hasta que se encuentra el primer empleo significativo es superior para el colectivo de mujeres durante todo el proceso de transición²³.

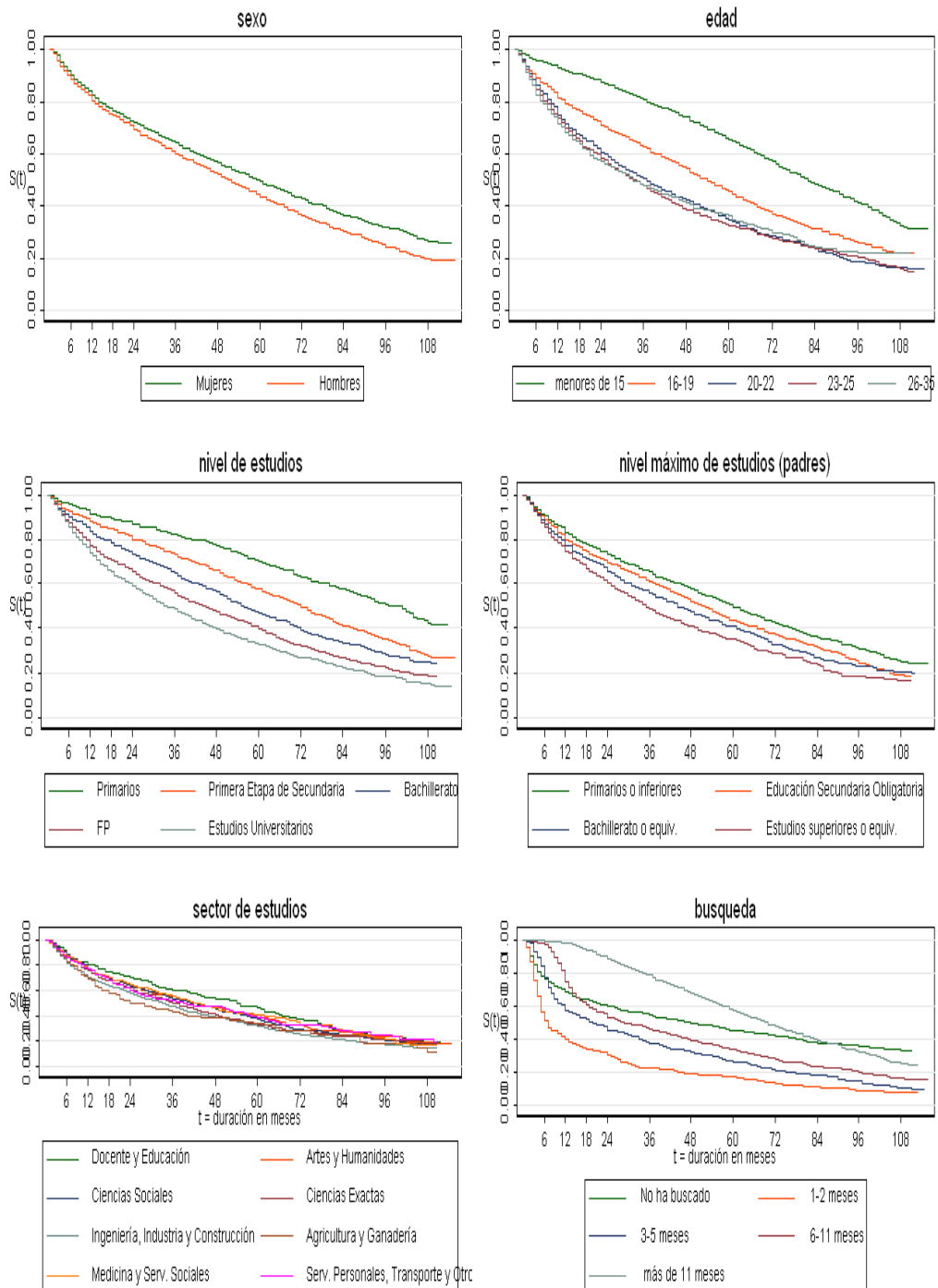
Las diferencias entre las curvas de supervivencia resultan aún más evidentes cuando se considera la edad a la que se sale del sistema educativo. Los jóvenes más adultos presentan unas probabilidades de supervivencia menores que los más jóvenes (menores de 15 años y 16-19 años), indicándonos que los jóvenes que inician antes su andadura en el mercado laboral tienen mayores dificultades para encontrar su primer empleo significativo.

Por lo que se refiere al nivel de estudios con el que se accede al mercado laboral, la posición de las curvas de supervivencia, que en ningún momento se entrecruzan y que están situadas de abajo a arriba según el orden inverso del nivel de estudios, sugiere una correspondencia negativa entre la probabilidad de supervivencia y el nivel de estudios, de tal forma que cuanto mayor es el nivel de estudios menores son esas probabilidades.

La formación de los padres y, en concreto, el máximo del nivel de estudios de ambos, parece tener influencia en las probabilidades de supervivencia recayendo sobre los jóvenes que tienen padres con únicamente estudios primarios las probabilidades más altas mientras que las tasas menores están asociadas a jóvenes cuyos progenitores cuentan con

²³Las diferencias encontradas, en principio, cabe esperar que sean mayores ya que no se ha tenido en cuenta que los jóvenes varones, al menos una parte de ellos, cumplen el servicio militar (o la prestación social sustitutoria) tras la finalización de sus estudios.

Figura 5.4: Estimación no paramétrica de las funciones de supervivencia según diferentes variables (método de Kaplan-Meier)



estudios superiores.

Asimismo, observamos que para el sector en el que se inscriben los estudios, aunque algunas categorías no presentan diferencias significativas entre sí, si que existen sectores de estudios que llevan asociados unas tasas de supervivencia más elevadas que otros, como es el caso de los estudios encuadrados dentro del sector de la Docencia y la Educación frente a las especialidades de Agricultura y Veterinaria o Ingeniería, Industria y Construcción.

La posición de las curvas de supervivencia según la duración de la búsqueda de empleo muestra un resultado ciertamente demostrativo de las consecuencias directas de la actividad desarrollada por el individuo para conseguir un empleo. Si bien, los que no han buscado nunca presentan inicialmente una probabilidad menor de sobrevivir, que estaría asociada al hecho de que algunos jóvenes no tienen necesidad de realizar ningún tipo de búsqueda para encontrar empleo, este resultado se invierte cuando el individuo, tras esa etapa inicial, sigue sin intervenir de una forma activa en el proceso de búsqueda de empleo²⁴.

Por último, si atendemos al año de salida del sistema educativo (figura 5.5), también encontramos diferencias en las probabilidades de supervivencia. En este caso, se han dibujado las curvas de supervivencia una vez que se han censurado artificialmente los datos de forma que todas las observaciones con una duración superior a los 60 meses han sido tomadas como censuradas (siguiendo a Congregado y García-Pérez, 2002). El objetivo de tal actuación es conseguir una visualización más clara de la separación existente entre las curvas de supervivencia resultantes²⁵. La figura muestra que las mayores dificultades para acceder a un empleo estable las han encontrado los jóvenes pertenecientes a las primeras cohortes, verificándose también que cuanto mayor es la duración, la distancia entre las curvas se va agrandando. Esto significa que las diferencias encontradas no se producen en los inicios de la salida del sistema educativo sino con posterioridad, una vez que los jóvenes empiezan a ser tachados como jóvenes con dificultades (parados o inactivos de larga duración) por el mercado.

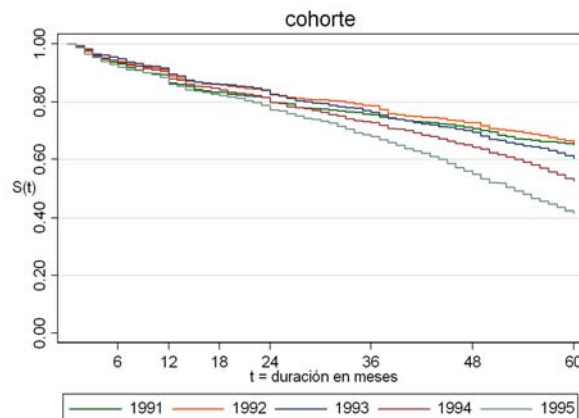
En resumen, los factores explicativos analizados parecen conducir a diferencias

²⁴La intersección de las curvas de supervivencia es un síntoma evidente del incumplimiento de la hipótesis de proporcionalidad, como veremos más adelante.

²⁵En el gráfico no se han incluido todas las cohortes pues las últimas cohortes presentan duraciones más cortas debido a que su seguimiento es menor lo que de nuevo produciría distorsiones en la lectura del gráfico. Sin embargo, el resultado es extrapolable.

significativas en las probabilidades de supervivencia de los jóvenes siendo las diferencias observadas mayores cuando se tiene en cuenta el nivel de estudios, la edad de salida del sistema educativo y la duración de la búsqueda. Como es costumbre en estos casos, la confirmación de estas diferencias se ha llevado a cabo mediante la realización de un contraste estadístico que permita determinar si las curvas de supervivencia son iguales, estadísticamente hablando. Entre las múltiples opciones, que se comentaron en el capítulo 3, se han considerado inicialmente el contraste del logaritmo del rango (Mantel y Haenszel, 1959) y el contraste de Wilcoxon (Gehan, 1965 y Breslow, 1970). Este último, al ponderar por el número de individuos sujetos a riesgo, da un mayor peso a las diferencias localizadas en la parte inicial de las curvas. En cualquier caso, los dos tests no miden las diferencias para un valor de t concreto sino que son tests globales.

Figura 5.5: **Estimación no paramétrica de las funciones de supervivencia para cada cohorte (método de Kaplan-Meier)**



Además, dado que ambos tests nos han llevado a las mismas conclusiones, nos centraremos en los resultados del test del logaritmo del rango²⁶. Un ejemplo de como funciona este test, en el caso en el que se comparan las curvas de supervivencia de hombres y mujeres, aparece en la tabla 5.4. En esta tabla se recoge el valor del estadístico -distribuido como una χ^2 con $k - 1$ grados de libertad (donde k el número de categorías de la variable en cuestión)- que se obtiene al comparar el número esperado y el número

²⁶El test del logaritmo del rango se utiliza en aquellos casos en los que las curvas muestran proporcionalidad mientras que cuando existen sospechas de que las curvas no son proporcionales el test de Wilcoxon es el más indicado.

observado de “fallos”, esto es, al comparar el número de individuos de cada sexo que han logrado un empleo significativo con el que se espera bajo la hipótesis de que las curvas son iguales. Obsérvese que, en este caso, mientras para los jóvenes varones se produce un número de “fallos” mayor del esperado, en las mujeres se detecta la situación contraria, *sobreviviendo* más que los del sexo opuesto.

Tabla 5.4: **Contraste no paramétrico de igualdad de curvas de supervivencia (test del logaritmo del rango)**

| sexo | Nº de sucesos | Nº de sucesos |
|--------|---------------|---------------|
| | observados | esperados |
| mujer | 3639 | 3970,72 |
| hombre | 4653 | 4321,28 |
| Total | 8292 | 8292,00 |

$$\chi^2(1) = 53,96$$

$$\text{Pr} > \chi^2 = 0,0000$$

En la tabla 5.5 mostramos, de forma resumida, los resultados obtenidos al aplicar este test (y el de Wilcoxon) a cada una de las variables explicativas consideradas. A tenor del p -valor asociado, podemos concluir que se rechaza en todos los casos la hipótesis nula de que la curvas sean estadísticamente equivalentes.

Tabla 5.5: **Resumen de los contrastes no paramétricos de igualdad de curvas de supervivencia**

| Variable | Test del logaritmo del rango | | Test de Wilcoxon | |
|--|------------------------------|----------------------|------------------|----------------------|
| | χ^2 | $\text{Pr} > \chi^2$ | χ^2 | $\text{Pr} > \chi^2$ |
| Sexo | 53,96 | 0,00 | 34,94 | 0,00 |
| Edad de salida del sistema educativo | 879,25 | 0,00 | 1044,69 | 0,00 |
| Nivel de estudios | 799,34 | 0,00 | 819,34 | 0,00 |
| Sector de estudios | 706,18 | 0,00 | 765,86 | 0,00 |
| Máximo del nivel de estudios de los padres | 173,19 | 0,00 | 174,68 | 0,00 |
| Comunidad Autónoma ^(a) | 765,18 | 0,00 | 644,14 | 0,00 |
| Año de salida del sistema educativo ^(b) | 254,94 | 0,00 | 202,45 | 0,00 |

Nota: ^(a) El gráfico de las funciones de supervivencia para la variable “comunidad autónoma” no se ha presentado en el análisis debido a que el elevado número de categorías no permite apreciar claramente las diferencias encontradas; ^(b) El contraste en el caso de la variable “año de salida” se ha realizado con una muestra censurada artificialmente para el periodo 1991-1995.

5.3 Análisis de los factores determinantes de la duración del periodo de transición entre el sistema educativo y el primer empleo significativo: estimación de un modelo semiparamétrico

5.3.1 Modelo de duración de Cox

Como acabamos de ver, el análisis no paramétrico procura estimaciones útiles tanto de la función de supervivencia como de la función de riesgo, y nos ha permitido identificar algunos de los factores que afectan a la duración del proceso de transición del sistema educativo a un empleo significativo. Sin embargo, no proporciona una valoración cuantitativa del efecto de cada uno de esos factores sobre la probabilidad de encontrar un primer empleo significativo. Este objetivo se logra cuando utilizamos métodos paramétricos o semiparamétricos para estimar la función de supervivencia o la función de riesgo²⁷. Por ello, una vez finalizado el análisis no paramétrico, pasamos a estudiar cuáles son los factores más importantes que rigen el periodo de transición entre la finalización de los estudios y el primer empleo significativo, y en qué medida afectan a este proceso.

Como se ha señalado anteriormente, el modelo más utilizado en las aplicaciones prácticas ha sido el modelo de riesgos proporcionales de Cox (Cox, 1972), en el que la función de riesgo viene especificada

$$h(t; x) = h_0(t)\varphi(x)$$

por lo que el modelo asume que el efecto de las variables, x , sobre la función de riesgo base ($h_0(t)$) es multiplicativo, teniendo por lo tanto como consecuencia un desplazamiento hacia arriba o hacia abajo de la misma -dependiendo del signo de la función $\varphi(x)$ - siempre en la misma proporción (hipótesis de proporcionalidad).

La ventaja de este modelo es que no requiere especificar previamente la forma de la función de riesgo (base), de ahí que se le catalogue como método semiparamétrico. Esto

²⁷Recuérdese la discusión presentada en el capítulo metodológico de modelos de duración sobre la elección entre modelos en tiempo discreto o continuo. En este apartado, se ha optado por un modelo en tiempo continuo, a pesar de que la duración del periodo de transición ha sido medida en meses. Más adelante, y con la finalidad de concretar el comportamiento de la función de riesgo, se utilizarán modelos discretos.

significa que no se impone ninguna restricción en la estimación salvo la de que sea cual sea el comportamiento de la función de riesgo, siempre es el mismo para cualquier tipo de individuo²⁸. Esta cualidad está íntimamente relacionada con el método de estimación que utiliza, el método de máxima verosimilitud parcial, que no presta atención a los valores de la variable duración sino al orden en que aparecen. Sin embargo, esta forma de proceder puede añadir ciertas complicaciones al proceso de estimación cuando se da la circunstancia de que existen muchas duraciones idénticas (empates) para las que se desconoce el orden exacto en el que se han producido.

Teniendo presente la posibilidad de que éste fuera un problema real en nuestros datos, hemos hecho un análisis previo sobre cuál es el tamaño del mismo. En particular, lo que interesa conocer es el número de empates en relación al número de individuos sujetos a riesgo para cada duración. De los 14287 individuos que componen la muestra, 8292 han conseguido encontrar un primer empleo significativo. Estos individuos se reparten entre un total de 116 valores de duración, lo que nos lleva a una media de 71,5 “fallos” por tiempo de “fallo”, no superando el porcentaje de empates en relación al conjunto de individuos sujetos a riesgo el valor de 2,7% (véase tabla 5 en anexo 3). Esto nos induce a sospechar que los dos procedimientos aproximativos para el tratamiento de los empates, el método de Breslow (Breslow, 1974) y el de Efron (Efron, 1977), no han de mostrar grandes divergencias por lo que la decisión sobre cuál de ellos utilizar no es una cuestión relevante en nuestro caso²⁹. Además, en esta ocasión ha sido posible estimar el modelo de Cox recurriendo a un método exacto, el método *marginal* a pesar del tamaño del fichero, por lo que serán estas últimas estimaciones las que utilizaremos de cara a comentar los resultados obtenidos

Como va siendo habitual en los trabajos empíricos, las estimaciones se presentan en la forma de “*hazard ratios*” ($HR = e^{\hat{\beta}}$) dado que su interpretación es más sencilla al indicar el efecto proporcional de la variable sobre la función de riesgo. En particular, un valor igual a 0,5 ($\hat{\beta} < 0$), en el caso de una variable ficticia, supone que la función de riesgo (o tasa de salida al empleo significativo) evaluada en cada valor de t se vea reducida a la mitad para la categoría señalada frente a la de referencia, lo que conlleva en consecuencia un proceso de duración más largo. En el caso de una variable continua, ese mismo valor de HR se traduce en un cambio en esa cuantía en la función de riesgo

²⁸Tal ventaja lleva aparejada una cierta pérdida de eficiencia (Singer y Willett, 2003 pág. 521).

²⁹En la tabla 6 del anexo 3 se muestran los resultados de la estimación del modelo de Cox bajo estas dos alternativas.

ante un cambio unitario de la variable en cuestión. Por el contrario, un valor del *hazard ratio* por encima de uno ($\hat{\beta} > 0$), por ejemplo de 1,5, supone un desplazamiento hacia arriba de la función de riesgo, obteniéndose para cada valor de t un valor de esta función un 50% superior al valor de la función correspondiente a la categoría de referencia. En términos de duración, supone que los individuos con esa característica experimenten procesos de transición más cortos.

En cuanto a las variables discretas no dicotómicas, éstas se han convertido en múltiples variables dicotómicas, eliminándose en cada caso aquella que reflejaba una mayor disparidad en el estudio no paramétrico. De esta forma, el individuo de referencia viene caracterizado por aquellas categorías eliminadas para cada variable discreta. En relación a las variables continuas, una cuestión que no va a tener ninguna importancia de cara a la estimación pero que nos va a ayudar a mejorar la interpretación de los resultados es el origen que se utiliza para medir estas variables. Eligiendo correctamente este origen, estaremos construyendo un individuo de referencia con unas características con verdadero significado para la población que se está analizando. Por esta razón, algunas de las variables han sido transformadas realizando un cambio de origen de tal manera que al dibujar la función de riesgo (o de supervivencia) base ésta tenga un sentido claro en relación a la población objeto de estudio³⁰. En definitiva, el individuo de referencia presenta las siguientes características: es mujer, de 16 años, con estudios primarios y sin ningún tipo de especialización, la formación de los padres es la correspondiente a estudios primarios y la región de procedencia es Andalucía, Ceuta o Melilla.

Por último, en la selección de las variables que finalmente han pasado a formar parte del modelo que comentaremos a continuación, se han tenido en cuenta varios aspectos. Por una parte, se incorporan aquellas variables que la teoría económica y las aportaciones empíricas hasta la fecha han determinado como relevantes, teniendo en cuenta, en cualquier caso, la restricción de la disposición de información sobre las mismas. Por otro lado, el análisis llevado a cabo en el apartado no paramétrico, nos ha servido para afianzar nuestras expectativas iniciales sobre las variables allí consideradas, incluyéndolas directamente en el modelo de partida que ha sido sometido a un conjunto de técnicas estadísticas de valoración.

En este sentido, también se ha tenido en cuenta la existencia de factores explicativos de la duración no observados sobre todo teniendo presente la revisión de condicionantes

³⁰Por ejemplo, la transformación realizada en la variable *edad de salida del sistema educativo* conlleva que el individuo de referencia sea un joven que finalizó su formación cuando tenía 16 años.

que se realizó en el capítulo 4. A pesar de que algunos autores son de la opinión de que la gran flexibilidad en la especificación de la función de riesgo base que caracteriza al modelo de Cox mitiga en gran medida los efectos de la existencia de heterogeneidad no observada, hemos considerado oportuno estimar un modelo incluyendo también un factor de heterogeneidad con distribución gamma, del que se deduce que ésta no es significativa³¹. Más adelante, en la presentación de los modelos paramétricos, en los que también se ha incluido un factor de heterogeneidad no observada, se trata esta cuestión de forma más profunda.

5.3.2 Análisis e interpretación de los resultados del modelo de Cox

Con anterioridad a la estimación del modelo de Cox, se ha efectuado un análisis descriptivo de la muestra atendiendo a la cohorte a la que pertenecen para comprobar si existen importantes diferencias demográficas entre las distintas cohortes de salida del sistema educativo (véase tabla 4 en el anexo 3). Como resultado de este análisis, se han detectado algunas de las tendencias que han caracterizado la evolución del sistema educativo y del mercado de trabajo en la década de los noventa y que ya se expusieron en el capítulo 1. En particular, se observa una variación en la proporción de mujeres que salen del sistema educativo que ha hecho que incluso se rompa la tendencia de un leve predominio del sexo masculino frente al femenino. Asimismo, y dada la expansión educativa que ha identificado esta época, también advertimos un mayor nivel de estudios en las cohortes más cercanas a la fecha de la entrevista, lo que lleva parejo que también se aprecie un ligero crecimiento en la edad a la que se sale del sistema educativo.

El objetivo de este análisis previo era determinar la viabilidad de un análisis conjunto de la duración del periodo de transición sin particularizar para cada cohorte, lo que significaría que la probabilidad de que se inicie el proceso de transición es independiente del tiempo (calendario) y que su longitud presenta la misma distribución independientemente de cuando se inició³². Aunque, como ya hemos indicado, existen algunas diferencias en las características generales de estas cohortes, finalmente no se ha

³¹En concreto, se ha estimado un modelo en el que la heterogeneidad se supone que varía dentro de cada uno de los grupos que se derivan de considerar tanto la edad como la comunidad autónoma (véase tabla 8 en el anexo 3).

³²En definitiva, estamos haciendo referencia a lo que Heckman y Singer (1984b) definieron como condiciones iniciales que permite, en el caso de que se verifique, que las duraciones puedan ser comparadas.

considerado la realización de un análisis por separado ya que las diferencias apreciadas están básicamente relacionadas con variables que han sido controladas en la estimación³³.

Una vez cerradas estas cuestiones preliminares, hemos pasado a la estimación del modelo de Cox. A continuación mostramos los resultados obtenidos así como algunas consideraciones técnicas que han ido surgiendo conforme se avanzaba en la especificación definitiva del modelo. Asimismo, a lo largo de la exposición se presenta un estudio comparativo con los resultados de otros estudios estrechamente relacionados con el tema y se intenta igualmente proporcionar explicaciones teóricas a los resultados encontrados.

En primer lugar se ha estimado el *modelo 1* (véase tabla 5.7) en el que se han incluido todas las variables³⁴. Partiendo de este modelo hemos evaluado, tanto de forma individual como conjunta, el cumplimiento de la hipótesis de proporcionalidad³⁵. Asumir que las funciones de riesgo de dos individuos cualesquiera son proporcionales entre sí para cualquier valor de t significa, por ejemplo, en el caso de la variable “nivel de estudios”, que el efecto sobre la función de riesgo de acceder al mercado de trabajo con estudios superiores frente a no tenerlos es el mismo cuando ha transcurrido un año desde la incorporación que cuatro años después de la misma, es decir, el efecto se mantiene constante en el tiempo³⁶.

³³En este sentido, algunas investigaciones preliminares sí que tienen en cuenta el hecho de que la muestra está formada por varias cohortes estudiando específicamente el comportamiento diferente de la función de riesgo según la cohorte (Betts, Ferral y Finnie, 2000). Los estudios nacionales, sin embargo, no analizan esta casuística (Congregado y García Pérez, 2002 y Albert *et al.*, 2003b) ni siquiera incluyendo un factor explicativo que recogiera el año en el que el joven se incorpora al mercado laboral.

³⁴Se ha detectado una correlación entre el nivel y el sector de estudios que nos ha obligado a trabajar con una variable que recoge, de forma conjunta, ambas características en el caso de aquellos estudios para los que es posible conocer la especialización (Formación Profesional y Estudios Universitarios). En la tabla 7 del anexo 3 se incorpora la estimación de los modelos de Cox en los que se han incluido estas variables por separado y conjuntamente. De la comparación de estos modelos, observamos como los estudios dejan de ser significativos cuando se incorpora el sector de especialización.

³⁵La gran popularidad que ha adquirido el modelo de Cox dentro de las ciencias sociales ha hecho que en muchas aplicaciones se recurra a su estimación para analizar la duración de un suceso sin pararse a comprobar si se verifica la hipótesis de proporcionalidad que le caracteriza. Algunos autores como Singer y Willett (1993, pág. 186) señalan que esta hipótesis es más una excepción que una regla por lo que su violación es muy frecuente. Desde un punto de vista de la teoría económica, sólo en el caso de que la estrategia de búsqueda del individuo sea “miope”, la hipótesis de proporcionalidad tendría pleno sentido (van den Berg, 2000, pág. 29). Esta actitud suele plantearse cuando la búsqueda es repetida, en el sentido de que el individuo tiene una segunda oportunidad (Seernels, 2002a, pág. 26).

³⁶El resultado contrario ocurriría si la pérdida de capital humano que se produce durante el periodo en el que el individuo no participa activamente en el mercado de trabajo dependiera del nivel de estudios

Uno de los procedimientos habitualmente utilizado para verificar el cumplimiento de esta hipótesis consiste en analizar los residuos *Schoenfeld* dado que éstos, bajo el supuesto de proporcionalidad, son independientes del tiempo. En consecuencia, una representación de estos residuos frente al tiempo no debería mostrar ninguna tendencia sino que habrían de estar dispuestos aleatoriamente en torno al cero³⁷. En base a esta consideración, también es posible valorar la hipótesis de proporcionalidad incluyendo en el modelo interacciones de cada variable con una función del tiempo y contrastando si éstas son significativas. En la tabla 5.6 se han incorporado los resultados obtenidos cuando se utiliza este último procedimiento para evaluar la hipótesis de proporcionalidad tanto de forma individual, para cada una de las variables, como de forma conjunta³⁸. En ella se recoge, junto al valor del estadístico de contraste para cada variable, el *p*-valor asociado que nos indica el nivel de significación con el que rechazaríamos la hipótesis nula de proporcionalidad. A la vista de esta tabla, la aceptación de esta hipótesis parece tener sentido para la mayoría de las variables, a excepción de la variables sexo, búsqueda de empleo, edad y la variación anual en la tasa de paro³⁹. Este resultado del rechazo de la hipótesis de proporcionalidad para algunas de las variables del modelo conduce al resultado global de este contraste, que también rechaza la proporcionalidad cuando se evalúa de forma conjunta.

La violación de esta hipótesis no supone, a priori, ningún problema grave siempre que se asuma que el efecto de esas variables no es constante en el tiempo. Para ello, la metodología de los modelos de duración aporta varias propuestas. De entre ellas hemos optado por, por un lado, estratificar el modelo atendiendo a la variable búsqueda de empleo y, por otro lado, estimar este modelo estratificado por separado para cada sexo.

alcanzado. Así, por ejemplo, el efecto de los estudios en el tiempo no sería el mismo si los jóvenes que adquirieron unas habilidades específicas al completar unos estudios con una mayor especialización, las pierden más rápidamente que los individuos con estudios más generales al no ejercitarlas de forma inmediata (o al menos, así es percibido por los empresarios).

³⁷En realidad, para facilitar la interpretación de estos gráficos se representan los residuos *Schoenfeld* corregidos o escalados, que son una modificación de los anteriores con el fin de que no valgan cero para las observaciones censuradas.

³⁸El uso de los gráficos de los residuos *Schoenfeld* no ha sido posible por las dificultades encontradas a la hora de interpretarlos debido al tamaño muestral con el que trabajamos.

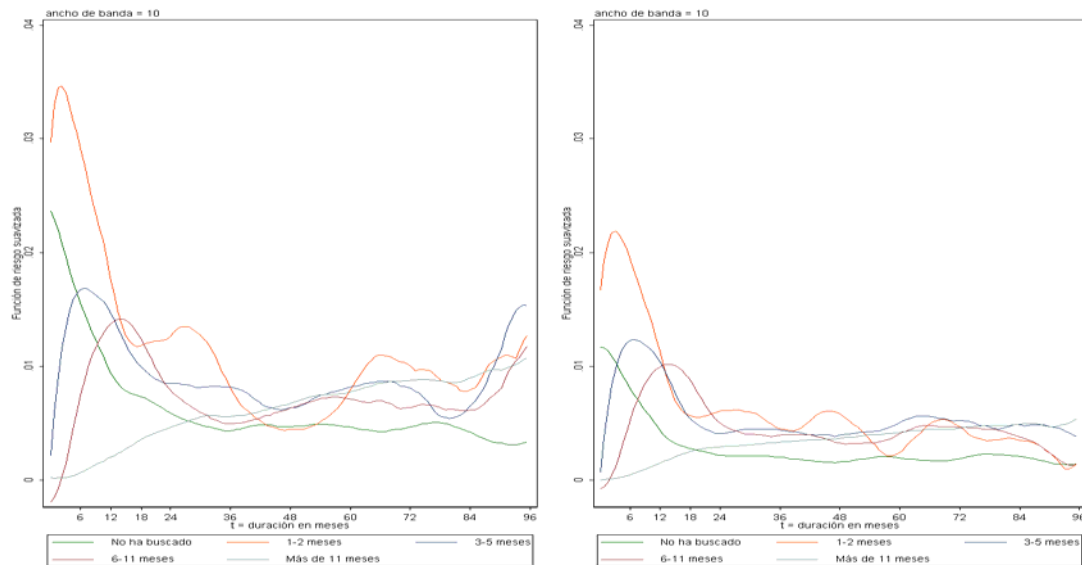
³⁹En relación a las variables que recogen el entorno económico, recordamos que la inclusión de variables dependientes del tiempo, como las que estamos considerando, rompe con la hipótesis de proporcionalidad. Ya que al suponer que la variable cambia con el tiempo -no así su efecto-, la función de riesgo también se ve modificada como se puede apreciar en la siguiente expresión: $\frac{h(t; x_1(t))}{h(t; x_2(t))} = e^{(x_1(t) - x_2(t))\beta} \neq cte$.

Tabla 5.6: Contraste de la hipótesis de proporcionalidad en el modelo de Cox

| | χ^2 | Prob> χ^2 | | χ^2 | Prob> χ^2 |
|---|----------|----------------|--|----------|----------------|
| <i>Sexo</i> | | | <i>Nivel y sector de estudios</i> | | |
| Hombres | 9,56 | 0,002 | Primera Etapa de Secundaria | 0,00 | 0,985 |
| <i>Búsqueda de empleo</i> | | | Bachillerato | 5,49 | 0,019 |
| 1-2 meses | 5,86 | 0,016 | Formación Profesional | | |
| 3-5 meses | 48,37 | 0,000 | Docencia y Artes y Humanidades | 3,34 | 0,068 |
| 6-11 meses | 141,18 | 0,000 | Ciencias sociales | 2,71 | 0,100 |
| más de 11 meses | 1168,82 | 0,000 | Ciencias exactas | 1,46 | 0,227 |
| <i>Edad de salida del sistema educativo</i> | | | Ingeniería | 0,80 | 0,372 |
| edad | 130,64 | 0,000 | Industria y producción | 0,00 | 0,960 |
| edad ² | 59,10 | 0,000 | Arquitectura y construcción | 2,22 | 0,136 |
| <i>Entorno económico</i> | | | Agricultura y Veterinaria | 0,09 | 0,760 |
| vartparo | 16,78 | 0,000 | Medicina y Servicios Sociales | 5,69 | 0,017 |
| varocupa | 0,05 | 0,823 | Otros servicios | 1,12 | 0,290 |
| <i>Región de procedencia</i> | | | Estudios Universitarios | | |
| Aragón | 7,72 | 0,006 | Docencia | 8,88 | 0,003 |
| Asturias | 2,55 | 0,111 | Artes | 0,23 | 0,632 |
| Baleares | 17,05 | 0,000 | Humanidades | 1,69 | 0,194 |
| Canarias | 18,71 | 0,000 | Ciencias sociales y Periodismo | 10,88 | 0,001 |
| Cantabria | 0,08 | 0,780 | Economía | 9,53 | 0,002 |
| Castilla y León | 1,21 | 0,271 | Derecho | 11,37 | 0,001 |
| Castilla La Mancha | 9,90 | 0,002 | Ciencias exactas | 1,73 | 0,189 |
| Cataluña | 6,21 | 0,013 | Informática | 3,38 | 0,066 |
| Com. Valenciana | 2,45 | 0,118 | Ingeniería | 3,56 | 0,059 |
| Extremadura | 1,82 | 0,177 | Arquitectura y construcción | 3,29 | 0,070 |
| Galicia | 0,41 | 0,523 | Agricultura y Veterinaria | 0,00 | 0,959 |
| Madrid | 0,36 | 0,547 | Medicina y Servicios Sociales | 4,21 | 0,040 |
| Murcia | 1,79 | 0,181 | Otros servicios | 1,11 | 0,291 |
| Navarra | 2,19 | 0,139 | <i>Nivel de estudios de los padres</i> | | |
| País Vasco | 2,08 | 0,149 | Secundaria Obligatoria o equiv. | 2,09 | 0,148 |
| La Rioja | 22,50 | 0,000 | Bachillerato o equiv. | 0,04 | 0,848 |
| | | | Superiores o equiv. | 0,02 | 0,876 |
| Test global | 1788,02 | 0,000 | | | |
| (grados de libertad) | (46) | | | | |

La diferencia entre estas dos soluciones radica en que en el primer caso, la estratificación, se presupone que el efecto de la variable “búsqueda” es el mismo para todos los individuos aunque la forma de la función de riesgo base es diferente, mientras que al estimar por separado, se está considerando que tanto la función de riesgo base como los efectos de las variables explicativas pueden ser distintos. La toma de decisión está basada en la relevancia que asignamos a la variable “sexo” frente a la variable “búsqueda” que consideramos más una variable de control. Por esta razón, asumimos el coste que tiene la utilización de los modelos estratificados al no cuantificar directamente el efecto de la variable por la cual se estratifica, en nuestro caso, la búsqueda de empleo. Las estimaciones correspondientes a este modelo (*modelo 2*) se presentan también en la tabla 5.7.

Figura 5.6: Comportamiento de la función de riesgo según el tiempo de búsqueda (modelo de Cox estratificado)



No obstante, se ha hecho un estudio, meramente gráfico, del comportamiento de la función de riesgo que se obtiene al estratificar según esta variable y que se presenta en la figura 5.6. En ella se advierten dos hechos resaltables. En primer lugar, la diferente forma funcional de esta función para aquellos que no han buscado empleo durante todo el periodo frente al resto de categorías, ya que su función decae constantemente

(dependencia estrictamente negativa)⁴⁰. En concreto, partiendo de una situación ligeramente más favorable que la mayor parte de los jóvenes, vemos como su posición va empeorando al alejarse su función de riesgo de las correspondientes a los individuos que señalan haber dedicado algún tiempo a la búsqueda activa de empleo. En segundo lugar, los individuos que más buscan parecen obtener su recompensa al final del periodo, colocándose su función de riesgo a la misma altura que las de aquellos que han dedicado menos tiempo a la búsqueda. Las diferencias apuntadas son más acusadas en el caso de los hombres. En particular, las curvas correspondientes al colectivo de mujeres no muestran el repunte final detectado en el caso de los hombres, considerando por tanto que su dificultad para encontrar un empleo significativo se mantiene prácticamente constante a lo largo del tiempo.

Una última consideración antes de comentar los resultados obtenidos, proviene del hecho de que se ha observado también que la presencia de la edad en el modelo rebaja de forma importante la significación y el efecto de las variables nivel y sector de estudios. Por esta razón, para la evaluación del efecto de estas variables se ha reestimado el modelo eliminando la edad (*modelo 3*).

A continuación resumimos los principales resultados que se obtienen al evaluar los modelos estimados (*modelo 1*, *modelo 2* y *modelo 3*).

En relación al sexo, deteniéndonos en la estimación del *modelo 1*, modelo conjunto para toda la muestra, observamos que el valor del *hazard ratio* está por encima de la unidad ($HR = 1,23$), lo que significa que los hombres tienen una probabilidad instantánea mayor que las mujeres de lograr un primer empleo significativo en cualquier momento del periodo. Sin embargo, el sexo del individuo tiene implicaciones mayores que las anteriormente expuestas pues como veremos al analizar los otros dos modelos (*modelo 2* y *modelo 3*), el resto de variables tienen un efecto distinto, no sólo en magnitud sino también incluso en dirección, sobre la duración del periodo de transición⁴¹.

Con respecto a la edad, en general, no está claro cuál es el efecto global de esta variable existiendo en la literatura empírica resultados contradictorios, e incluso estudios en los que no se detecta efecto alguno de esta variable sobre la transición, como es el caso del estudio de Albert *et al.* (1998, pág. 21) en el que se concluye que el proceso de inserción sigue una lógica biográfica independiente de la edad, de tal forma que tarda lo

⁴⁰Este resultado es el que está produciendo que se incumpla la hipótesis de proporcionalidad.

⁴¹En investigaciones a nivel internacional el sexo no resulta ser una variable que marque diferencias importantes en la transición al empleo (Russell y O'Connell, 2001).

mismo en integrarse un joven de 16 años que uno con 30 años.

Tabla 5.7: Estimación del modelo de riesgos proporcionales de Cox (*hazard ratios*)

| | Modelo (1) | | Modelo estratificado (2) | | | | Modelo estratificado (3) | | | | | |
|--|------------|---|--------------------------|---------|---------|---------|--------------------------|---|-------|---|-------|---|
| | Total | | Total | Mujeres | Hombres | Mujeres | Hombres | | | | | |
| Sexo (ref: mujeres) | | | | | | | | | | | | |
| Hombres | 1,234 | * | 1,208 | * | | | | | | | | |
| Búsqueda de empleo (ref: no ha buscado) | | | | | | | | | | | | |
| 1-2 meses | 2,547 | * | | | | | | | | | | |
| 3-5 meses | 1,589 | * | | | | | | | | | | |
| 6-11 meses | 1,179 | * | | | | | | | | | | |
| más de 11 meses | 0,669 | * | | | | | | | | | | |
| Edad de salida del sistema educativo centrada (Edad-16 años) | | | | | | | | | | | | |
| edad | 1,131 | * | 1,121 | * | 1,104 | * | 1,134 | * | | | | |
| edad ² | 0,992 | * | 0,993 | * | 0,994 | * | 0,992 | * | | | | |
| Nivel y sector de estudios (ref: estudios primarios y sin especialización) | | | | | | | | | | | | |
| Primera Etapa de Secundaria | 1,321 | * | 1,290 | * | 1,516 | * | 1,200 | * | 1,641 | * | 1,284 | * |
| Bachillerato | 1,089 | | 1,095 | | 1,388 | * | 0,970 | | 2,035 | * | 1,521 | * |
| Formación Profesional | | | | | | | | | | | | |
| Docencia y Humanidades | 1,314 | * | 1,324 | * | 1,600 | * | 1,197 | | 2,503 | * | 2,091 | * |
| Ciencias sociales | 1,309 | * | 1,298 | * | 1,685 | * | 1,122 | | 2,565 | * | 1,832 | * |
| Ciencias exactas | 1,317 | * | 1,326 | * | 1,869 | * | 1,120 | | 2,923 | * | 1,902 | * |
| Ingeniería | 1,352 | * | 1,356 | * | 2,094 | * | 1,180 | * | 3,234 | * | 1,870 | * |
| Industria y producción | 1,406 | * | 1,409 | * | 1,885 | * | 1,210 | | 2,782 | * | 1,893 | * |
| Arquitectura y construcción | 1,449 | * | 1,351 | | 1,959 | * | 1,128 | | 3,104 | * | 1,827 | * |
| Agricultura y Veterinaria | 1,489 | * | 1,448 | * | 1,536 | | 1,297 | | 2,764 | * | 2,093 | * |
| Medicina y Serv. sociales | 1,297 | * | 1,256 | * | 1,680 | * | 0,829 | | 2,605 | * | 1,395 | |
| Otros servicios | 1,487 | * | 1,429 | * | 1,953 | * | 1,084 | | 2,852 | * | 1,778 | * |
| Estudios Universitarios | | | | | | | | | | | | |
| Docencia | 1,163 | | 1,156 | | 1,605 | * | 0,718 | | 2,751 | * | 1,363 | |
| Artes | 1,027 | | 1,068 | | 1,243 | | 1,084 | | 2,151 | * | 2,039 | * |
| Humanidades | 1,103 | | 1,092 | | 1,385 | * | 0,954 | | 2,416 | * | 1,850 | * |
| Ciencias soc. y Periodismo | 1,188 | | 1,200 | | 1,506 | * | 1,074 | | 2,629 | * | 2,041 | * |
| Economía | 1,446 | * | 1,448 | * | 1,986 | * | 1,159 | | 3,438 | * | 2,213 | * |
| Derecho | 1,342 | * | 1,369 | * | 1,895 | * | 1,024 | | 3,326 | * | 1,963 | * |
| Ciencias exactas | 1,239 | | 1,229 | | 1,730 | * | 0,977 | | 3,001 | * | 1,849 | * |
| Informática | 1,904 | * | 1,739 | * | 2,545 | * | 1,492 | * | 4,432 | * | 2,84 | * |
| Ingeniería | 1,453 | * | 1,377 | * | 1,755 | * | 1,206 | | 3,048 | * | 2,277 | * |
| Arquitectura y construcción | 2,475 | * | 2,168 | * | 2,590 | * | 1,937 | * | 4,530 | * | 3,595 | * |
| Agricultura y Veterinaria | 1,487 | * | 1,423 | * | 1,759 | * | 1,247 | | 3,123 | * | 2,405 | * |
| Medicina y Serv. sociales | 1,325 | * | 1,320 | * | 1,731 | * | 1,041 | | 2,976 | * | 1,972 | * |
| Otros servicios | 1,308 | * | 1,359 | * | 1,848 | * | 1,034 | | 3,173 | * | 1,923 | * |

Tabla 5.7 (Continuación)

| | Modelo (1) | Modelo estratificado (2) | | | Modelo estratificado (3) | |
|---|------------|--------------------------|---------|---------|--------------------------|---------|
| | Total | Total | Mujeres | Hombres | Mujeres | Hombres |
| <i>Nivel máximo de estudios de los padres (ref: estudios primarios)</i> | | | | | | |
| Secundaria Obligatoria o equiv. | 1,077 * | 1,071 * | 1,000 | 1,128 * | 0,998 | 1,120 * |
| Bachillerato o equiv. | 1,061 | 1,038 | 1,006 | 1,055 | 1,015 | 1,085 |
| Superiores o equiv. | 1,128 * | 1,115 * | 1,165 * | 1,094 | 1,185 * | 1,128 * |
| <i>Entorno económico</i> | | | | | | |
| vartparo | 0,656 * | 0,694 * | 0,582 * | 0,822 * | 0,545 * | 0,779 * |
| varocupa | 1,244 * | 1,229 * | 1,099 | 1,749 * | 1,088 | 1,763 * |
| <i>Región de procedencia (ref: Andalucía, Ceuta y Melilla)</i> | | | | | | |
| Aragón | 1,932 * | 1,898 * | 2,003 * | 1,853 * | 2,010 * | 1,839 * |
| Asturias | 0,843 | 0,817 * | 0,829 | 0,858 | 0,855 | 0,921 |
| Baleares | 1,547 * | 1,601 * | 1,809 * | 1,423 * | 1,796 * | 1,423 * |
| Canarias | 1,386 * | 1,435 * | 1,539 * | 1,336 * | 1,562 * | 1,330 * |
| Cantabria | 0,995 * | 0,968 | 1,233 | 0,765 | 1,251 | 0,789 * |
| Castilla y León | 1,447 * | 1,469 * | 1,364 * | 1,548 * | 1,392 * | 1,539 * |
| Castilla La Mancha | 1,693 * | 1,695 * | 1,859 * | 1,602 * | 1,813 * | 1,558 * |
| Cataluña | 1,515 * | 1,555 * | 1,720 * | 1,438 * | 1,680 * | 1,399 * |
| Com. Valenciana | 1,654 * | 1,686 * | 1,807 * | 1,603 * | 1,768 * | 1,546 * |
| Extremadura | 0,925 | 0,972 | 0,999 | 0,970 | 1,002 | 0,923 |
| Galicia | 1,226 * | 1,299 * | 1,289 * | 1,316 * | 1,304 * | 1,312 * |
| Madrid | 1,306 * | 1,320 * | 1,494 * | 1,160 * | 1,486 * | 1,125 |
| Murcia | 1,237 * | 1,270 * | 1,323 * | 1,247 * | 1,297 * | 1,197 * |
| Navarra | 1,633 * | 1,611 * | 1,575 * | 1,670 * | 1,581 * | 1,604 * |
| País Vasco | 1,709 * | 1,699 * | 1,534 * | 1,839 * | 1,563 * | 1,874 * |
| La Rioja | 1,708 * | 1,673 * | 1,465 * | 1,853 * | 1,428 * | 1,738 * |
| Nº Individuos | 14287 | 14287 | 6797 | 7490 | 6797 | 7490 |
| Nº Observaciones | 64642 | 64642 | 30890 | 33752 | 30890 | 33752 |
| Log verosimilitud | -72367 | -61290 | -24225 | -31340 | -24253 | -31412 |
| AIC | 144840 | 122675 | 48545 | 62775 | 48596 | 62916 |

Nota: La estimación se ha llevado a cabo aplicando el método Breslow para el tratamiento de los empates y utilizando errores estándar robustos debido a la existencia de múltiples registros por individuo.

* significativo al 5%

En la especificación final hemos introducido la edad y la edad al cuadrado, siendo ambas altamente significativas tanto para los hombres como para las mujeres⁴². En nuestro caso, el *hazard ratio* asociado a la edad, superior a la unidad, nos está indicando que la probabilidad de lograr un empleo significativo aumenta con la edad. Ahora bien, el valor del *hazard ratio* que acompaña a la edad al cuadrado por debajo de uno nos permite detectar la existencia de un punto de inflexión en el efecto total de la edad sobre la probabilidad de salida al empleo significativo, que aproximadamente estaría situado en torno a los 24 años (véase gráfico 5.7). Por otro lado, también observamos que el efecto de la edad es algo mayor para los hombres que para las mujeres ($HR_H = 1,13$ vs. $(HR_M = 1,10)$).

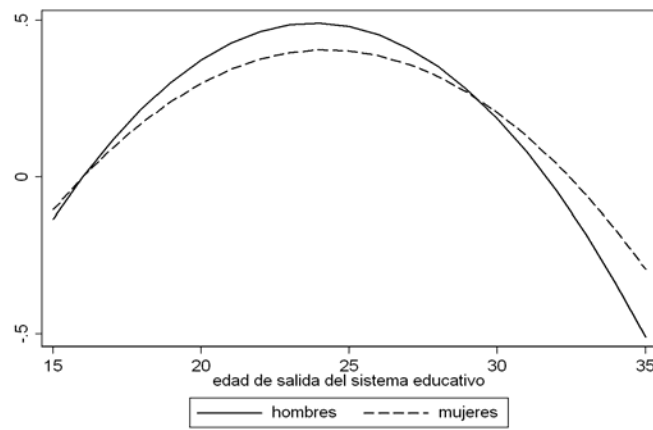
Otros investigadores llegan a resultados parecidos, por ejemplo, Albert *et al.* (2003b) introducen la edad agrupada en intervalos, y aunque comparado con el grupo de referencia (menores de 16 años) el resto de jóvenes tienen una probabilidad mayor de encontrar un empleo significativo, el coeficiente muestra una ligera caída para el último grupo de edad considerado (el de mayores de 26 años)⁴³. Asimismo, en este trabajo también se detecta un efecto diferente según el sexo, siendo mayor en el caso de los hombres (al menos para los jóvenes más adultos).

Una explicación a estos resultados es que la edad ejerce una mayor presión en la búsqueda de empleo, por lo que el individuo empieza a buscar de forma más intensa. Además, esta presión es mayor en el caso de los hombres que, por tradición, suelen asumir un papel más activo laboralmente, lo que hace que la necesidad de un empleo se vuelva aún más acuciante. Por otro lado, la ruptura de esta relación directa entre la edad y la probabilidad de conseguir un empleo significativo se puede deber a que la edad de finalización de los estudios también funciona como un indicador de las cualidades futuras del trabajador y aquellos jóvenes que permanecen más tiempo del esperado dentro del sistema educativo por dificultades en la superación de los distintos niveles, se convierten

⁴²En especificaciones alternativas hemos utilizado variables indicadoras para aquellos grupos de edad que se consideraron en la parte descriptiva de este estudio o en el análisis no paramétrico, siendo los resultados muy similares a los aquí presentados pero menos significativos.

⁴³Congregado y García Pérez (2002) obtienen que los universitarios experimentan tasas de salida menores a medida que su edad aumenta al incluir este nivel de estudios interactuado con la edad. Por otro lado, este resultado es también corroborado por aquellos estudios que se limitan a analizar el proceso de inserción de los jóvenes con titulación superior (Chuang, 1999 y Betts, Ferrall y Finnie, 2000) en los que se detecta un efecto negativo de la edad sobre la duración como era de esperar al centrarse en aquellas edades en las que se produce ese revés en el signo de esta relación.

Figura 5.7: Efecto de la edad según sexo en el modelo de Cox estratificado



en demandantes de empleo con perspectivas menos halagüeñas que sus compañeros igualmente formados pero en un menor tiempo⁴⁴.

Por lo que se refiere al nivel de estudios, según diversas teorías económicas, se trata de un factor determinante de la duración del proceso de transición al empleo, aunque aún queda por resolver cuál es el efecto último de este factor. Por un lado, la teoría del capital humano establece que la educación incrementa el capital humano y, por lo tanto, aumenta las posibilidades de acceder a un empleo, lo que nos llevaría a la existencia de una relación positiva entre el nivel de estudios y la probabilidad de conseguir un empleo (cualquier tipo de empleo, significativo o no). Sin embargo, la teoría de la búsqueda de empleo señala que un mayor nivel de estudios puede elevar el salario de reserva, lo que conduce directamente a un rechazo de todas aquellas ofertas que no superan ese salario de reserva y, en consecuencia, a un alargamiento de la búsqueda del empleo⁴⁵.

De nuevo, los trabajos empíricos ofrecen resultados no siempre en la misma línea.

⁴⁴Russell y O'Connell (2001) proponen como posible justificación a la relación entre la edad y la tasa de salida del desempleo el hecho de que la edad esté funcionando como una *proxy* de otros factores no disponibles como la experiencia laboral previa. Sin embargo, en nuestro caso, esta explicación no es válida pues la edad ha sido tratada como una variable fija recogida en el momento en el que se sale por primera vez del sistema educativo lo que significa que en ese momento la experiencia laboral de los jóvenes es prácticamente nula.

⁴⁵Los jóvenes que acceden al mercado de trabajo con niveles de estudios superiores son más exigentes a la hora de aceptar una oferta laboral y están dispuestos a esperar pacientemente a que aparezca la oferta de empleo que mejor se adapte a sus expectativas.

Para algunos, el proceso de inserción laboral no parece depender del nivel de estudios alcanzado, es decir, tarda lo mismo en integrarse en el mercado de trabajo un joven con escasa formación que uno que haya adquirido un mayor nivel de estudios (Albert *et al.*, 1998, pág. 24)⁴⁶, mientras que para otros, la variable estudios es especialmente discriminatoria en el proceso de integración en el mercado de trabajo.

Los resultados obtenidos, globalmente, parecen confirmar que la teoría del capital humano prevalece, verificándose que un mayor nivel de estudios tiende a reducir la duración del proceso de transición. Dado que el modelo finalmente estimado contiene interacciones entre el nivel de estudios y el sector de estudios, los comentarios que presentamos a continuación irán encaminados a distinguir, en la medida de lo posible, los efectos de ambas variables. En primer lugar todos los niveles educativos muestran un efecto positivo significativo frente a los estudios primarios, siendo en todos los casos este efecto superior en las mujeres que en los hombres. Así, un joven que haya superado la primera etapa de enseñanza secundaria tiene una probabilidad de alcanzar una situación estable en el mercado de trabajo un 26% superior a la de un joven con tan sólo estudios primarios. El nivel de estudios que proporciona estabilidad laboral más rápidamente es el correspondiente a los estudios universitarios con especialización en Arquitectura ($HR_M = 4,53$ y $HR_H = 3,6$) o Informática ($HR_M = 4,43$ y $HR_H = 2,84$) seguido de algunas especializaciones en Formación Profesional (Ingeniería, Industria y Otros Servicios)⁴⁷.

No obstante, el desglose planteado para los niveles de estudios de Formación Profesional y Estudios Universitarios según el sector de estudios, nos ha permitido observar que no siempre un mayor nivel de estudios proporciona unas mejores expectativas laborales, y que hay algunos sectores de estudios que no mejoran las posibilidades de acceso a un empleo estable como veremos más adelante⁴⁸.

⁴⁶Ahora bien, aunque la duración pueda resultar similar, esto no quiere decir que la “calidad” de la integración sea igual tanto en términos de permanencia como de estabilidad.

⁴⁷En nuestro caso no hemos distinguido, dentro de los estudios superiores, entre carreras de ciclo corto y largo. Sin embargo, algunas investigaciones ofrecen evidencias de que existen diferencias en las duraciones del proceso de transición dentro de este nivel de estudios, no apreciándose un mejor situación en los jóvenes que optan por completar unos estudios más largos (Lassibille *et al.*, 2001 y Albert *et al.*, 2003b).

⁴⁸La formación universitaria condiciona la empleabilidad de dos formas muy distintas. En función de la propia empleabilidad de la titulación, teniendo ésta un impacto significativo en la capacidad de inserción laboral de un titulado. Sin embargo, a medida que avanza la carrera profesional de un individuo, la importancia de los estudios como condicionante de la empleabilidad es cada vez menor y cobran mayor

El hecho de que los menos formados se enfrenten a duraciones más largas podría deberse a que tienen menos oportunidades de empleo como consecuencia del cambio tecnológico que se ha producido en las economías y que ha dado lugar a una redistribución de las ofertas de empleo, aumentando el número de empleos para los que se requiere cierta cualificación. Asimismo, si apelamos a la teoría expuesta por Dolado *et al.* (2000) que admite la existencia de un efecto de desplazamiento originado por el exceso de oferta de trabajadores con estudios superiores⁴⁹, -que ha echado por tierra la creencia de que los empresarios son poco partidarios de contratar a trabajadores sobrecualificados- además de esa menor oferta de puestos de trabajo con escasa cualificación también se enfrentan a una competencia mayor por parte de un colectivo con una mejor preparación.

Como se ha señalado, el nivel de estudios es un factor más determinante para explicar la duración del proceso de transición en el caso de las mujeres que en el de los hombres, resultado que ya ha sido constatado en investigaciones anteriores tanto en el caso del primer desempleo como en la transición al empleo (Ahn y Ugidos, 1995b; Alba, 1998b y Albert *et al.*, 2003b). En esencia, la educación funciona como un indicador de las habilidades y capacidades del individuo de forma más acusada para las mujeres, lo que les exige alcanzar cotas más altas de formación para demostrar verdaderamente su valía. Por esta razón, el efecto del nivel educativo se manifiesta más claramente.

La comparación de la importancia del nivel educativo sobre la probabilidad de transitar con otros países es difícil pues como ya se advirtió muchos de los trabajos que estudian este proceso de transición restringen la muestra a un colectivo con un determinado nivel de estudios. Aún así, y en términos generales, se puede decir que los estudios también juegan un papel importante en la duración del proceso de transición aunque su efecto es menor y no hay diferencias tan significativas entre los dos sexos (Franz *et al.*, 1997 y Russell y O'Connell, 2001)

Retomando el sector de estudios, de la tabla 5.7 se desprende que las diferencias

importancia factores como la experiencia o la coherencia en la carrera profesional.

⁴⁹Estos autores confirman la existencia de un efecto de desplazamiento o *crowding-out* que ha provocado que los jóvenes más preparados decidan, ante un mercado laboral precario, aceptar ofertas laborales por debajo de su formación, al menos temporalmente, desplazando a los jóvenes sin cualificación o con escasa cualificación. En este sentido, Delgado y Díaz Rodríguez (2002) señalan que, en la actualidad, los titulados superiores tienden a no demorar su entrada a la vida activa esperando un empleo acorde a la formación recibida. De producirse tal situación, los jóvenes titulados que se inician en la búsqueda de un empleo establecerían un salario de reserva por debajo de sus expectativas, lo que les lleva a aceptar salarios inferiores a los que se corresponderían con su formación.

son más acusadas dentro de los Estudios Universitarios donde la decisión de elegir una u otra carrera parece tener importantes implicaciones de cara a conseguir un empleo significativo. En efecto, la variabilidad en el efecto de las distintas titulaciones es mayor que en los estudios de Formación Profesional. Así, los individuos con estudios de Arquitectura e Informática tienen mayores posibilidades de conseguir un empleo significativo más rápidamente que los que cuentan con una formación en Artes y Humanidades.

Por lo que respecta a los estudios de Formación Profesional, las especialidades que encabezan la lista de las mejor situadas en términos de un proceso de transición más corto son las inscritas en el ámbito de la Ingeniería y la Arquitectura seguidas de cerca por los estudios relacionados con la Docencia, Artes y Humanidades y Agricultura y Veterinaria.

Las diferencias encontradas en cuanto al sector de estudios no se manifiestan por igual cuando tenemos en cuenta el género. Mientras que para los hombres, la especialización no parece tener tanta importancia, para las mujeres el sector de estudios adquiere un valor determinante. Casi la totalidad de los sectores considerados tienen un efecto significativo sobre la probabilidad de conseguir un empleo estable, tanto en el caso de los estudios de Formación Profesional como en el de los Estudios Universitarios. Las jóvenes que se titulan en Arquitectura o Ingeniería (dentro de la Formación Profesional) o en Arquitectura, Informática o Economía (dentro de los Estudios Universitarios) son las que consiguen alcanzar antes un empleo significativo.

Esta descripción de los sectores con más posibilidades está reflejando, de alguna forma, la distribución de la demanda laboral existente en el mercado entre las distintas titulaciones. El grupo de titulaciones con una mayor demanda es el de las titulaciones técnicas, seguido de las titulaciones del área jurídico-social y, con volúmenes mucho menores, las titulaciones científico-sanitarias y las de humanidades⁵⁰. Por otro lado, las titulaciones jurídico-sociales se caracterizan por un elevado número de estudiantes más que por una baja demanda de estos titulados por parte de las empresas, lo que posiblemente está produciendo una saturación del mercado. Esta puede ser la razón de que las perspectivas de encontrar un empleo significativo a corto plazo de los jóvenes con

⁵⁰La demanda de trabajadores poco cualificados y de trabajadores de la industria manufacturera se ha visto reducida en contraposición al incremento del empleo en la industria de los servicios y en ocupaciones profesionales y técnicas.

formación específica en el grupo denominado *Ciencias Sociales* sean menos favorables⁵¹.

Figura 5.8: Efecto del sector de estudios en el modelo de Cox estratificado

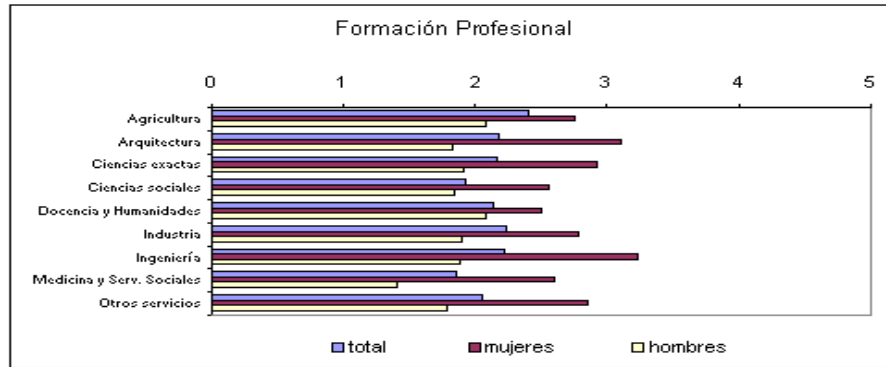
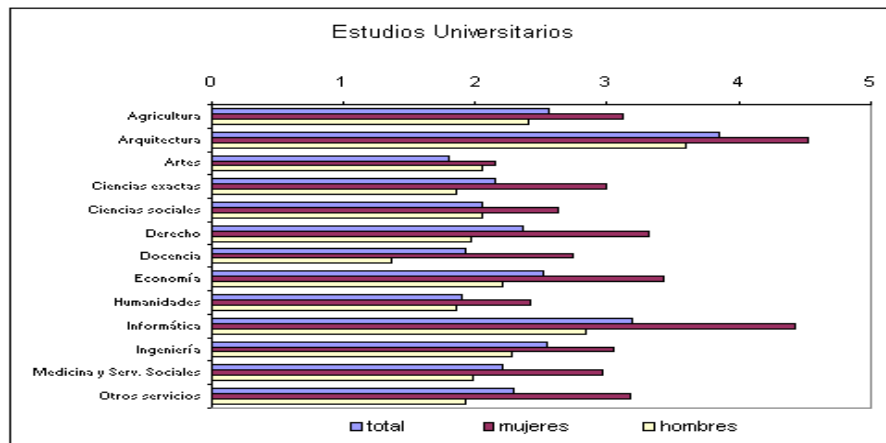


Figura 5.9: Efecto del sector de estudios en el modelo de Cox estratificado



Asimismo, las características de los mercados laborales propios de cada especialización pueden ayudarnos a explicar también estas diferencias. En este sentido, algunas de las titulaciones que aparecen en los primeros puestos del ranking en cuanto a la integración laboral, rápida y de forma estable, de los jóvenes que las eligen, se caracterizan porque son especialidades en las que una de las salidas laborales más habitual es el autoempleo, bien mediante la puesta en marcha de un negocio propio o bien trabajando

⁵¹Estas titulaciones, por otra parte, son más propias del colectivo femenino lo que en parte puede explicar sus mayores problemas para encontrar un empleo estable y también el porqué de las diferencias entre éstas jóvenes y las que eligen especialidades técnicas.

como autónomo para otra empresa (este es el caso, por ejemplo, de los titulados en Arquitectura). Por otro lado, aquellas titulaciones que ofrecen pocas posibilidades laborales en el sector privado muestran periodos de transición más largos debido a que la oferta del sector público, además de requerir tiempo para la preparación de la oposición, está supeditada al cumplimiento de unos determinados plazos administrativos.

Siguiendo con las características microeconómicas, en un modelo de oferta de trabajo familiar en el que las decisiones se toman conjuntamente, algunos miembros de la familia tienen una probabilidad menor de trabajar (y mayor de no participar) según la situación del resto de integrantes del hogar. Esta circunstancia se hace más palpable en el caso de los jóvenes españoles para los que la presencia y soporte constante de la familia alivia las dificultades que tienen para encontrar empleo y el alto grado de inestabilidad al que se ven sometidos en sus primeras incursiones laborales. Por otro lado, la facilidad para acceder al mercado de trabajo de un joven también puede verse mejorada por la situación laboral del resto de miembros de la familia, ya que las oportunidades de empleo, en una gran mayoría, provienen del entorno familiar. Detrás de esta afirmación subyace la evidencia de que el conocimiento del mercado de trabajo es imperfecto de tal forma que aspectos como el nivel de formación o la situación laboral de los padres pueden incrementar el conocimiento del mismo y, en consecuencia, influir en la probabilidad de encontrar un empleo.

Por estas razones, en el estudio de la transición de los jóvenes al mercado laboral hemos tenido en cuenta el papel que juegan las familias. Como se ha señalado, el nivel de formación y/o la situación profesional de los padres suelen tomarse como variables para detectar la importancia que dentro de las características propias del individuo tiene el número de contactos que le pueden ofrecer sus progenitores.

En este caso, hemos tratado de captar los aspectos anteriormente descritos a través del nivel de estudios de los padres, definiendo, en particular, una variable que recoge el nivel máximo de los estudios de ambos progenitores. Al igual que en otros estudios la variable ha resultado ser significativa. Más concretamente, el valor significativo del *hazard ratio* superior a la unidad asociado a aquellos jóvenes cuyos padres cuentan con estudios universitarios pone de relieve que éstos tienen una probabilidad de acceder a un empleo significativo un 12% superior a la que corresponde a jóvenes con padres que no han superado el nivel primario en sus estudios. El efecto del entorno familiar es algo más alto para las mujeres, resultado que concuerda con el de González Betancor *et al.* (2004) en su análisis sobre la duración del primer desempleo de jóvenes con formación

profesional, no así con otros estudios, específicos de la duración del proceso de transición a un empleo estable como el del Albert *et al.* (2003b), para los que el papel de la familia no es relevante en el caso de los hombres.

A nivel macroeconómico, es posible detallar un número concreto de condicionantes que pueden ayudar a que la transición al empleo significativo sea más o menos rápida. Entre ellos, señalamos la demandada agregada que vendrá determinada por aspectos tales como las políticas laborales encaminadas a mejorar la empleabilidad e incrementar la demanda laboral efectiva o el nivel de liberalización de los mercados laborales facilitando un mayor o menor grado de movilidad laboral.

En relación a esta cuestión, está claro que las ofertas de empleo que recibe un individuo no sólo dependen de sus características individuales y familiares sino que también el componente de demanda actúa como elemento diferenciador de la probabilidad de conseguir un empleo. Este componente puede ser captado a través de diferentes variables agregadas como la tasa de crecimiento del PIB, la tasa de desempleo de la zona de residencia del individuo, etc. En nuestro caso, tras la estimación de varios modelos con distintas alternativas, las variables que mejor han funcionado han sido la variación de la tasa de paro y la tasa de variación del número de ocupados⁵² evaluadas para cada sexo, grupo de edad y comunidad autónoma.

Por otra parte, dado que estas variables tienen una evolución en el tiempo que, por supuesto, afecta a la facilidad para conseguir un empleo significativo no sólo en el momento en que se inicia el proceso de incorporación al mercado de trabajo sino durante todo el proceso de transición, no se han tomado sus valores en la fecha de salida del sistema educativo sino que están referenciados a cada momento del tiempo⁵³. Se trata, pues, de variables dependientes del tiempo (calendario).

Como era de esperar, el efecto de la variación de la tasa de paro sobre la tasa de salida al empleo significativo es negativo mientras que el de la variación del número de ocupados es positivo. En el primer caso, la consecución de un empleo estable se alcanza con mayor rapidez en los periodos en los que la economía se encuentra en expansión y, en el segundo caso, cuanto mayor es la demanda, las posibilidades de situarse de forma rápida en el mercado laboral crecen. Destacar tan sólo que esta última no parece ser relevante cuando nos centramos en el grupo de las mujeres.

⁵²En el capítulo 1 realizamos un análisis básico del comportamiento de estas dos variables para el periodo de estudio.

⁵³Véase capítulo 4 para una explicación más completa de cómo se han construido.

Por último, al prestar atención a la variable “región de procedencia” se llega a la conclusión de que, en la actualidad, algunos factores como la educación tienen un desigual valor en las distintas regiones españolas a la hora de transitar a la vida activa, pues las oportunidades de empleo parecen depender más de las características productivas de sus respectivas comunidades de residencia que del grado de formación alcanzado. Así es, estudios sobre la especialización de cada región según el número de sectores muestran que la distribución regional de la oferta de empleo, sobre todo cualificado, se caracteriza por el papel relevante que juegan algunas comunidades y, dentro de éstas, por la concentración en las capitales de provincia.

Es más, teniendo en cuenta que las diferencias en la demanda ya han sido medidas parcialmente a través de las variables macroeconómicas, se deduce que éstas no son suficiente para explicar las divergencias existentes entre las regiones en cuanto a las posibilidades de acceder a un empleo estable, pues los coeficientes siguen siendo significativos (Russell y O’Connell, 2001).

En particular, las posibilidades de acceder a un empleo estable más rápidamente las encontramos en las comunidades de Aragón, Baleares, Comunidad Valenciana, Navarra y La Rioja, mientras que el proceso de transición es más largo en comunidades como Andalucía (incluyendo Ceuta y Melilla), Asturias, Cantabria y Extremadura, lo que viene a corroborar que las comunidades que tradicionalmente se enfrentan a mayores problemas de paro también ofrecen menores oportunidades para estabilizar el empleo de los jóvenes. El resto de comunidades se encontrarían en una situación intermedia. Las diferencias no son tan importantes -en relación a las encontradas para la variable estudios- pues el *hazard ratio* varía entre 1,89 para la comunidad con mejores perspectivas para los jóvenes (Aragón) frente a un 0,82 de Asturias. En cuanto a la distinción por sexo, no se advierte una pauta de comportamiento clara. En general, los jóvenes varones tienen un proceso de transición más corto en aquellas comunidades que favorecen en mayor medida la inserción estable que sus homónimas.

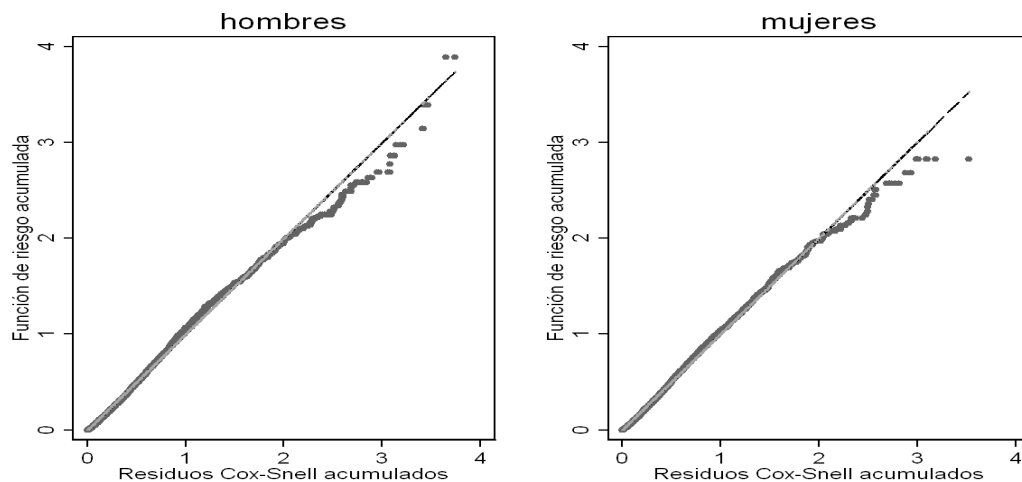
En general, los resultados obtenidos resultan bastante consistentes con las predicciones del modelo de búsqueda de empleo y con la mayoría de los trabajos empíricos sobre la transición al primer empleo (significativo). Es más, hemos encontrado varios elementos comunes entre la duración del desempleo juvenil, la duración del primer desempleo y la duración del periodo transitorio entre la salida del sistema educativo y el primer empleo significativo, resultado que era presumible teniendo en cuenta que para algunos jóvenes se trata del mismo suceso y para otros el periodo de desempleo es un

subconjunto del periodo de transición

5.3.3 Evaluación del modelo de Cox

Las anteriores conclusiones ofrecen una inestimable herramienta para introducir modificaciones en el diseño de los programas de estabilidad laboral que los poderes públicos dirigen hacia el colectivo de los jóvenes, ya que señalan cuáles son las características que les hace más vulnerables a la exclusión laboral. No obstante, la extrapolación de estos resultados a la aplicación de políticas públicas requiere que los modelos estimados sean robustos. En este apartado se busca, por tanto, determinar si el modelo presentado posee buenas propiedades.

Figura 5.10: **Evaluación del modelo de Cox: Función de riesgo acumulada de los residuos Cox-Snell**



Comenzaremos estudiando la bondad del ajuste mediante un análisis de los residuos *Cox-Snell* (Cox y Snell, 1968). Este procedimiento permite valorar de forma gráfica si el modelo presenta un buen ajuste. Partiendo de los residuos *Cox-Snell* ($e = \hat{H}_0(t)e^{x\hat{\beta}}$), se obtiene una estimación empírica de la función de riesgo acumulada suponiendo que estos residuos desempeñan el papel de la variable duración, ($\hat{H}(e)$), pero manteniendo la misma variable de censura. La representación gráfica de esta función de riesgo

acumulada frente a los residuos se aproxima, cuando el modelo presenta un buen ajuste, a la recta de pendiente igual a uno. En la figura 5.10 se puede ver cuál es el comportamiento de estos residuos obtenidos tras la estimación del modelo de Cox y, en concreto, se observa que no se alejan mucho de esa recta al menos en el caso de los residuos de tamaño más pequeño⁵⁴. Ciertas desviaciones son normales especialmente para los residuos más altos debido a que la estimación de esta función en este tramo está sometida a una mayor variabilidad como consecuencia de la pérdida de unidades en la muestra por los “fallos” que se van produciendo y, secundariamente, por la censura⁵⁵.

Tabla 5.8: **Contraste de error de especificación en el modelo de Cox estratificado**

| t | Coef. | Std. Error | z | $P > z $ |
|--------------------------------|--------|------------|-------|--------------|
| Modelo estratificado (Mujeres) | | | | |
| $x\hat{\beta}$ | 1,187 | 0,127 | 9,33 | 0,000 |
| $(x\hat{\beta})^2$ | -0,036 | 0,063 | -0,57 | 0,569 |
| Modelo estratificado (Hombres) | | | | |
| $x\hat{\beta}$ | 1,090 | 0,076 | 14,27 | 0,000 |
| $(x\hat{\beta})^2$ | 0,051 | 0,057 | 0,90 | 0,369 |

Para completar este análisis, realizamos a continuación un contraste global sobre la especificación definitiva del modelo, tal como recomiendan Cleves, Gould y Gutiérrez (2002), que consiste básicamente en la incorporación de nuevas variables para posteriormente verificar su significatividad. En particular, el contraste estima un modelo auxiliar en el que se incluyen como variables $x\hat{\beta}$ y su cuadrado⁵⁶. Bajo la hipótesis nula de que el modelo está correctamente especificado, $(x\hat{\beta})^2$ no tiene poder explicativo por lo que un contraste de significación individual de esta variable debería llevarnos a no rechazar la hipótesis nula. Este test es de gran utilidad para, al menos, cerrar la especificación del modelo en el sentido de que las variables que contiene son las que tienen

⁵⁴En realidad, se han utilizado los residuos *Cox-Snell* acumulados ya que la expansión de los datos, con el fin de incorporar las variables dependientes, hace que no tengamos una única observación por individuo.

⁵⁵Dado que los residuos se construyen a partir de estimaciones de β y $H_0(t)$, las desviaciones que encontramos en estos gráficos pueden parcialmente deberse a la incertidumbre asociada a las estimaciones.

⁵⁶Como se puede apreciar, este contraste no es más que una versión adaptada del test Reset a los modelos de duración.

que estar. Como se observa en la tabla 5.8, la realización de este contraste nos lleva a rechazar un posible error de especificación por omisión tanto en el modelo estratificado para el grupo de las mujeres como en el de los hombres (modelo estratificado (2) de la tabla 5.7).

5.4 Análisis de la dependencia de la duración del periodo de transición entre la salida del sistema educativo y la consecución de un empleo significativo

5.4.1 Dependencia de la duración mediante modelos paramétricos

Tras la estimación del modelo de Cox, caracterizado porque no se impone ningún tipo de restricción en la forma de la función de riesgo, se ha planteado la realización de una serie de estimaciones con diversas opciones sobre el comportamiento de dicha función a través de la utilización de modelos paramétricos⁵⁷. Por lo tanto, en este apartado estamos más interesados en la dependencia de la duración que en la cuantificación de los efectos de las variables explicativas y, en concreto, estamos buscando aquella distribución que recoja mejor el comportamiento en el tiempo de la función de riesgo.

La elección de la distribución más adecuada es una decisión, en la mayoría de las ocasiones, complicada en cuanto que la teoría económica no establece cuál es el comportamiento esperado de la forma de la función de riesgo⁵⁸. Además, en nuestro caso la literatura empírica sobre el tiempo que transcurre entre la salida del sistema educativo y el primer empleo significativo, al estar en sus albores, no aporta aún resultados concluyentes en un único sentido sobre la dependencia de la duración. Frecuentemente se ha recurrido para el análisis de la dependencia de la duración del primer desempleo a la especificación de funciones tipo Weibull (Chuang, 1999 o Cañada *et al.*, 1998), que asumen tasas de riesgo monótonas crecientes o decrecientes en el tiempo. Esta función aunque presenta la ventaja de su facilidad matemática no tiene porque ser necesariamente la que mejor se acomode a la distribución de los tiempos de espera hasta

⁵⁷Como señalábamos en el capítulo metodológico, si la especificación elegida es la correcta, entonces las estimaciones obtenidas bajo este tipo de modelos tienen mejores propiedades que las del modelo de Cox ya que son más eficientes.

⁵⁸Salvo en el caso del modelo básico de búsqueda de empleo del que se deriva la inexistencia de dependencia de la duración y, por lo tanto, presupone que la función de riesgo sigue una distribución exponencial.

la consecución de un primer empleo significativo.

Por esta razón, en este apartado no se ha desechado ninguna de las distribuciones que proporciona la metodología de los modelos de duración para recoger el comportamiento de variables no negativas⁵⁹. En concreto, estas distribuciones son: exponencial, Weibull, Gompertz, gamma generalizada, log-normal y log-logística. El número y la variedad de distribuciones consideradas es suficientemente amplio aunque, como veremos, no siempre es posible captar comportamientos muy específicos de la función de riesgo.

En esta ocasión, hemos partido de una especificación más general que la adoptada finalmente en la modelización semiparamétrica ya que en el modelo de Cox fue necesario estimar un modelo por separado para cada sexo y estratificar por la variable búsqueda, tras detectar que estas dos variables incumplían claramente la hipótesis de proporcionalidad. Por el contrario, la mayor parte de los modelos paramétricos aquí planteados no parten de la presunción de proporcionalidad.

La estimación de estos modelos se presenta en la tabla 5.9, que muestra los valores de los coeficientes estimados tanto en el caso de los modelos de riesgos acelerados (AFT) como en el de riesgos proporcionales (PH), para facilitar su comparación. Por otra parte, dado que los modelos exponencial y Weibull admiten las dos especificaciones, se ha optado por presentar los coeficientes correspondientes a la versión que supone riesgos acelerados. De este modo, prácticamente todos los modelos son comparables directamente a excepción del modelo Gompertz que es un modelo de riesgos proporcionales exclusivamente.

El coeficiente de un modelo de riesgos acelerados mide de forma directa la relación existente entre la duración y la variable, como se advierte en la expresión general de estos modelos ($\ln t = x\beta + \epsilon$)⁶⁰. Por lo tanto, un $\beta < 0$ ($\beta > 0$) tendrá el efecto de acelerar (retrasar) el tiempo de tal forma que el suceso a examen, en nuestro caso la transición al empleo significativo, ocurrirá antes (más tarde) de lo previsto para los individuos con esa característica⁶¹.

⁵⁹Es verdad que la función de riesgo obtenida en la estimación no paramétrica puede proporcionar algunas pistas sobre la forma funcional correcta (Kiefer, 1988) y que nos permitiría ser más selectivos en las distribuciones. Sin embargo, la función de riesgo estimada en el apartado 5.2.2, al no tener en cuenta los factores explicativos, recoge de forma bruta la dependencia de la duración.

⁶⁰El tipo de distribución del término de error, ϵ , determina el modelo paramétrico de duración (véase tabla 3.3).

⁶¹Téngase en cuenta que el efecto acelerador al que aluden estos modelos no es proporcional a lo largo del tiempo sino que el efecto marginal se amplifica cuanto mayor es el valor de t . Por ejemplo, el

A primera vista, y tras un examen de los coeficientes se desprende que los diferentes modelos estimados no muestran mucha disparidad, existiendo consistencia en los signos de las estimaciones de todos los modelos al menos para aquellas variables que resultan ser significativas, con la excepción de la variable búsqueda de empleo⁶². Así, las diferencias son casi despreciables en aquellos modelos que verifican la hipótesis de proporcionalidad (modelos exponencial, Weibull y Gompertz), siendo estas divergencias más importantes en el caso de los modelos de riesgos acelerados. Por lo tanto, como primera conclusión cabría decir que nuestras estimaciones son bastante robustas no apreciándose una especial sensibilidad a la forma funcional de la función de riesgo base elegida.

Nuestra preocupación, en este apartado, por verificar la posible generación de situaciones de dependencia nos lleva a analizar con detalle la relación entre la prolongación del proceso de transición y los cambios en la función de riesgo (la dependencia de la duración). Esta dependencia, en los modelos paramétricos, viene determinada por el valor del parámetro de *forma* de la distribución. A continuación vamos a analizar más detenidamente la forma de la función de riesgo bajo cada uno de los modelos atendiendo al valor obtenido de este parámetro. Los resultados se han incorporado al final de la tabla 5.9.

En primer lugar, y como ya se explicó, la distribución exponencial se caracteriza por su falta de memoria lo que se traduce en que la función de riesgo tenga un comportamiento constante en el tiempo. En consecuencia, exponenciando el término constante de este modelo se obtiene el valor de la función de riesgo de esta distribución para el individuo de referencia ($x = 0$), siendo este valor, en este caso, $e^{-5,012} = 0,0067$, constante para todo t ⁶³.

Dentro de la lista de distribuciones, dos de ellas se caracterizan porque su función de riesgo tiene un comportamiento monótono en el tiempo, la distribución Weibull y

coeficiente negativo de la variable sexo indica que para los jóvenes varones el proceso de transición es más corto, pero además que la diferencia entre las duraciones de hombres y mujeres será cada vez mayor a medida que va pasando el tiempo.

⁶²Las dudas que suscita esta variable en cuanto a su impacto en la duración del periodo de transición, ya detectadas en el modelo de Cox, provienen posiblemente de su definición en el módulo ya que se ha planteado como una variable no directamente asociada al empleo significativo que se analiza sino a cualquier periodo dentro del proceso de transición. En este sentido, se debería incluir en el modelo como una variable dependiente del tiempo indicando en cada momento del proceso cuál es el tiempo acumulado de búsqueda y no el máximo total observado para todo el periodo.

⁶³Se trata de la función de riesgo base obtenida cuando todas las variables valen cero o la que correspondería al individuo de referencia.

Tabla 5.9: Estimación de diferentes modelos de duración paramétricos (sin heterogeneidad no observada)

| | Exp. | Weibull | Gompertz | Gamma | LogN | LogL |
|---|----------|----------|----------|----------|----------|----------|
| | (aft) | (aft) | (ph) | (aft) | (aft) | (aft) |
| <i>Sexo (ref: mujeres)</i> | | | | | | |
| Hombres | -0,204 * | -0,198 * | 0,208 * | -0,194 * | -0,201 * | -0,187 * |
| <i>Búsqueda de empleo (ref: no ha buscado)</i> | | | | | | |
| 1-2 meses | -0,949 * | -0,930 * | 0,970 * | -0,820 * | -0,887 * | -1,057 * |
| 3-5 meses | -0,465 * | -0,455 * | 0,477 * | -0,167 * | -0,251 * | -0,385 * |
| 6-11 meses | -0,165 * | -0,169 * | 0,181 * | 0,345 * | 0,225 * | 0,065 * |
| más de 11 meses | 0,420 * | 0,408 * | -0,422 * | 1,105 * | 0,959 * | 0,788 * |
| <i>Edad de salida del sistema educativo centrada (Edad-16 años)</i> | | | | | | |
| edad | -0,120 * | -0,119 * | 0,126 * | -0,146 * | -0,144 * | -0,141 * |
| edad ² | 0,008 * | 0,008 * | -0,008 * | 0,009 * | 0,009 * | 0,009 * |
| <i>Nivel y sector de estudios (ref: estudios primarios y sin especialización)</i> | | | | | | |
| Primera Etapa de Secundaria | -0,275 * | -0,265 * | 0,278 * | -0,235 * | -0,245 * | -0,243 * |
| Bachillerato | -0,088 | -0,077 | 0,078 | -0,029 | -0,037 | -0,038 |
| Formación Profesional | | | | | | |
| Docencia y Humanidades | -0,282 * | -0,265 * | 0,273 * | -0,119 | -0,159 | -0,191 |
| Ciencias sociales | -0,270 * | -0,255 * | 0,265 * | -0,181 * | -0,203 * | -0,214 * |
| Ciencias exactas | -0,274 * | -0,262 * | 0,273 * | -0,190 | -0,204 | -0,203 |
| Ingeniería | -0,305 * | -0,289 * | 0,300 * | -0,226 * | -0,248 * | -0,263 * |
| Industria | -0,344 * | -0,328 * | 0,342 * | -0,285 * | -0,302 * | -0,309 * |
| Arquitectura | -0,364 * | -0,358 * | 0,377 * | -0,169 | -0,205 | -0,182 |
| Agricultura | -0,404 * | -0,386 * | 0,401 * | -0,366 * | -0,396 * | -0,418 * |
| Medicina y Serv. sociales | -0,259 * | -0,249 * | 0,261 * | -0,129 | -0,151 | -0,136 |
| Otros servicios | -0,400 * | -0,385 * | 0,402 * | -0,288 * | -0,314 * | -0,319 * |
| Estudios Universitarios | | | | | | |
| Docencia | -0,158 | -0,143 | 0,145 | 0,008 | -0,025 | -0,032 |
| Artes | -0,029 | -0,021 | 0,020 | 0,134 | 0,095 | 0,054 |
| Humanidades | -0,112 | -0,095 | 0,092 | -0,008 | -0,029 | -0,021 |
| Ciencias soc. y Periodismo | -0,176 | -0,161 | 0,166 | 0,015 | -0,023 | -0,050 |
| Economía | -0,370 * | -0,350 * | 0,364 * | -0,198 * | -0,237 * | -0,263 * |
| Derecho | -0,296 * | -0,274 * | 0,284 * | -0,102 | -0,145 | -0,159 |
| Ciencias exactas | -0,222 | -0,206 | 0,212 | -0,027 | -0,068 | -0,094 |
| Informática | -0,661 * | -0,636 * | 0,660 * | -0,415 * | -0,474 * | -0,471 * |
| Ingeniería | -0,383 * | -0,365 * | 0,379 * | -0,215 * | -0,253 * | -0,269 * |
| Arquitectura y construc. | -0,932 * | -0,905 * | 0,940 * | -0,586 * | -0,674 * | -0,702 * |
| Agricultura y Veterinaria | -0,426 * | -0,401 * | 0,410 * | -0,196 | -0,253 | -0,294 * |
| Medicina y Serv. sociales | -0,284 * | -0,266 * | 0,275 * | -0,173 | -0,206 * | -0,206 * |
| Otros servicios | -0,280 | -0,264 | 0,272 | -0,162 | -0,191 | -0,245 |

Tabla 5.9 (Continuación)

| | Exp. | Weibull | Gompertz | Gamma | LogN | LogL |
|---|----------|-------------|------------------|-------------------|----------------|----------------|
| | (aft) | (aft) | (ph) | (aft) | (aft) | (aft) |
| <i>Nivel máximo de estudios de los padres (ref: estudios primarios)</i> | | | | | | |
| Secundaria Obligatoria o equiv. | -0,069 * | -0,070 * | 0,075 * | -0,047 | -0,052 | -0,058 |
| Bachillerato o equiv. | -0,061 | -0,062 | 0,065 | -0,036 | -0,045 | -0,056 |
| Superiores o equiv. | -0,118 * | -0,117 * | 0,124 * | -0,100 * | -0,107 * | -0,104 * |
| <i>Entorno económico</i> | | | | | | |
| vartparo | 0,456 * | 0,403 * | -0,405 * | 0,250 * | 0,279 * | 0,314 * |
| varocupa | -0,233 * | -0,222 * | 0,228 * | -0,386 * | -0,369 * | -0,365 * |
| <i>Región de procedencia (ref: Andalucía, Ceuta y Melilla)</i> | | | | | | |
| Aragón | -0,653 * | -0,634 * | 0,665 * | -0,621 * | -0,639 * | -0,648 * |
| Asturias | 0,164 | 0,160 | -0,170 | 0,109 | 0,120 | 0,136 |
| Baleares | -0,422 * | -0,405 * | 0,427 * | -0,365 * | -0,382 * | -0,378 * |
| Canarias | -0,314 * | -0,304 * | 0,321 * | -0,228 * | -0,239 * | -0,243 * |
| Cantabria | 0,006 | 0,006 | -0,006 | 0,129 | 0,104 | 0,082 |
| Castilla y León | -0,365 * | -0,351 * | 0,368 * | -0,330 * | -0,341 * | -0,347 * |
| Castilla La Mancha | -0,514 * | -0,497 * | 0,523 * | -0,440 * | -0,457 * | -0,456 * |
| Cataluña | -0,411 * | -0,396 * | 0,415 * | -0,432 * | -0,43 * | -0,432 * |
| Com. Valenciana | -0,497 * | -0,478 * | 0,501 * | -0,479 * | -0,488 * | -0,486 * |
| Extremadura | 0,078 | 0,076 | -0,079 | 0,106 | 0,109 | 0,087 |
| Galicia | -0,208 * | -0,197 * | 0,204 * | -0,230 * | -0,223 * | -0,235 * |
| Madrid | -0,266 * | -0,254 * | 0,265 * | -0,266 * | -0,27 * | -0,283 * |
| Murcia | -0,205 * | -0,196 * | 0,206 * | -0,208 * | -0,204 * | -0,176 * |
| Navarra | -0,478 * | -0,463 * | 0,487 * | -0,488 * | -0,49 * | -0,461 * |
| País Vasco | -0,534 * | -0,517 * | 0,542 * | -0,472 * | -0,492 * | -0,472 * |
| La Rioja | -0,509 * | -0,492 * | 0,521 * | -0,384 * | -0,421 * | -0,376 * |
| constante | 5,012 * | 4,976 * | -5,108 * | 3,926 * | 4,166 * | 4,273 * |
| parámetro de forma | $p=1$ | $p=1,043 *$ | $\gamma=0,003 *$ | $\kappa=-0,231 *$ | $\sigma=1,237$ | $\gamma=0,704$ |
| Nº de individuos | 14287 | 14287 | 14287 | 14287 | 14287 | 14287 |
| Nº de observaciones | 64642 | 64642 | 64642 | 64642 | 64642 | 64642 |
| Log L | -17639 | -17628 | -17624 | -17243 | -17255 | -17265 |
| AIC | 35386 | 35366 | 35358 | 34598 | 34620 | 34640 |

Nota: La estimación se ha llevado a cabo utilizando errores estándar robustos debido a la existencia de múltiples registros por individuo.

* significativo al 5%

la Gompertz (aparte de la distribución exponencial). Esto significa, que un ajuste a cualquiera de estas dos distribuciones produce una función de riesgo que muestra bien dependencia de la duración positiva o bien negativa, pero nunca se altera la dirección de esta dependencia. En el caso de la distribución Weibull, hemos obtenido un parámetro de forma (p) igual a 1,043 lo que supone una ligera dependencia de la duración positiva al ser este valor superior a la unidad ($p > 1$). Es más, teniendo en cuenta que el parámetro no sobrepasa el valor de 2, estamos ante una función creciente que crece a un ritmo decreciente. Por su parte, el valor obtenido del parámetro auxiliar de la distribución Gompertz es de 0,003 que también implica una función creciente en el tiempo ($e^{0,003} > 1$). En ambos casos, el parámetro es significativamente distinto de cero por lo que la suposición de una función de riesgo constante queda descartada atendiendo a estos contrastes. Esto es, se acepta una cierta dependencia de la duración positiva.

Ahora bien, teniendo en cuenta los resultados del estudio no paramétrico, parece que no podemos restringir este análisis a distribuciones con un comportamiento monótono. Dentro de las distribuciones no monótonas, la de mayor interés es la gamma por su gran versatilidad englobando algunas de las anteriores distribuciones cuando los parámetros que la definen toman determinados valores. En concreto, los modelos exponencial, Weibull y lognormal son modelos anidados de la gamma. Esta característica nos permite seleccionar entre estos modelos anidados mediante la realización de los contrastes oportunos sobre el valor de los parámetros. Así, la significatividad del parámetro κ nos lleva a preferir el modelo gamma frente al modelo lognormal y, por otro lado, al rechazar la hipótesis de $\kappa = 1$, tampoco consideramos apropiada la distribución Weibull (véase tabla 5.10).

Tabla 5.10: Selección entre modelos paramétricos anidados

| Modelo 1 <i>vs.</i> modelo 2 | Hipótesis nula | Decisión |
|---|-----------------------------|--------------------|
| Weibull <i>vs.</i> exponencial | $\ln p = 0$ | weibull |
| Gamma generalizada <i>vs.</i> lognormal | $\kappa = 0$ | gamma generalizada |
| Gamma generalizada <i>vs.</i> Weibull | $\kappa = 1$ | gamma generalizada |
| Gamma generalizada <i>vs.</i> exponencial | $\kappa = 1$ y $\sigma = 1$ | gamma generalizada |

Cuando los modelos que constituyen el grupo de selección no están anidados, el criterio de información de Akaike (AIC) es un instrumento de selección más adecuado⁶⁴.

⁶⁴El estadístico de AIC viene definido por $AIC = -2 \log L + 2(c + p + 1)$ donde c es el número de

La elección, siguiendo este criterio, es de nuevo el modelo gamma generalizado que presenta un menor valor de la función de verosimilitud y también el valor más pequeño cuando se penaliza esta función por el número de parámetros auxiliares de cada modelo, como establece el criterio de AIC (véase tabla 5.9).

También se pueden utilizar los residuos de los modelos estimados para comprobar la bondad del ajuste de las estimaciones. Para ello, tras la estimación se calculan los residuos *Cox-Snell*, tal como vimos en la evaluación del modelo de Cox, de tal forma que si el modelo estimado se ajusta a los datos, estos residuos deberían tener una distribución exponencial censurada estándar con una tasa de riesgo igual a uno. Los residuos cuya función de riesgo acumulada se asemeja más a una distribución exponencial unitaria son los generados a través del modelo gamma. Por lo que, en general, cualquiera de los procedimientos utilizados parece conducir a un mismo modelo paramétrico, el modelo con distribución gamma generalizada, cuya función de riesgo aparece recogida en la figura 5.12⁶⁵. El comportamiento de esta función muestra un rápido crecimiento en los primeros meses, alcanzando su máximo en torno a los 15 meses, para posteriormente comenzar una descenso continuado.

La comparación de esta función de riesgo con la que se obtuvo en el análisis no paramétrico, nos lleva a reafirmarnos en nuestra sospecha de que el comportamiento de la función de riesgo de este proceso de transición no se adapta a ninguna de las distribuciones paramétricas al uso. De forma que, la imposición de una forma determinada puede conducir a resultados bastante contradictorios, como es el caso de las distribuciones Weibull y Gompertz en las que la exigencia de que la función tenga un comportamiento monótono tiene como resultado la presencia de dependencia de la duración positiva, consecuencia de la pretensión de recoger de forma conjunta un crecimiento y un decrecimiento de la función que da lugar a un parámetro estimado que promedia las dos situaciones.

Para finalizar este análisis paramétrico se ha contemplado la posibilidad de que la población siga presentando un grado de heterogeneidad de dimensión considerable aún después de clasificarla según el listado de variables explicativas contenidas en el modelo⁶⁶.

variables y p es el número de parámetros auxiliares del modelo.

⁶⁵Se ha procedido a estimar igualmente estos modelos paramétricos para cada género, siendo el modelo gamma el elegido en ambos casos.

⁶⁶Algunos autores suavizan las consecuencias de la heterogeneidad no observada sobre la dependencia de la duración cuando el modelo sin heterogeneidad presenta dependencia de la duración positiva, aunque no es éste el caso (Singer y Willett, 2003).

Figura 5.11: Evaluación de los modelos paramétricos: Función de riesgo acumulada de los residuos Cox-Snell

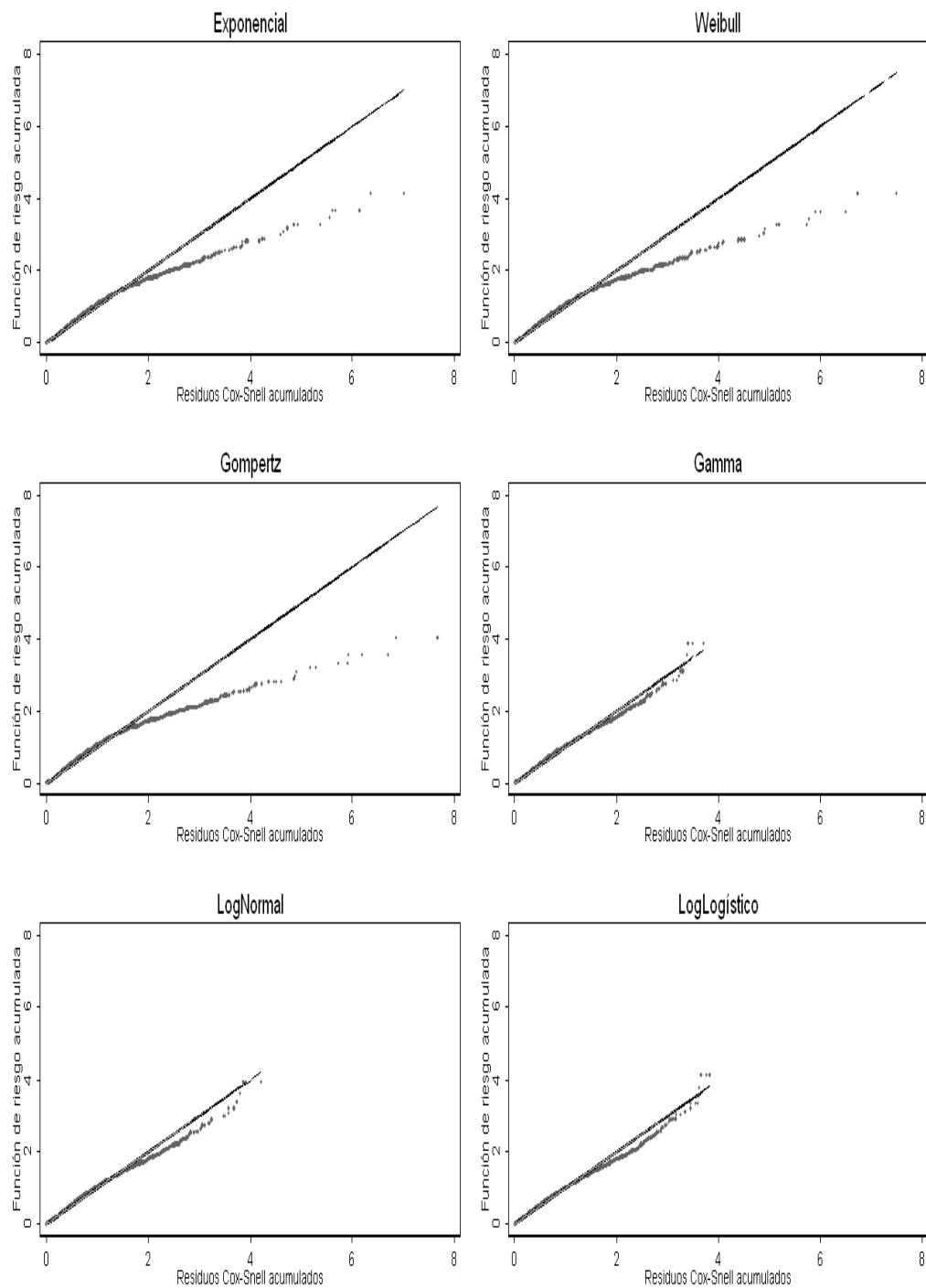
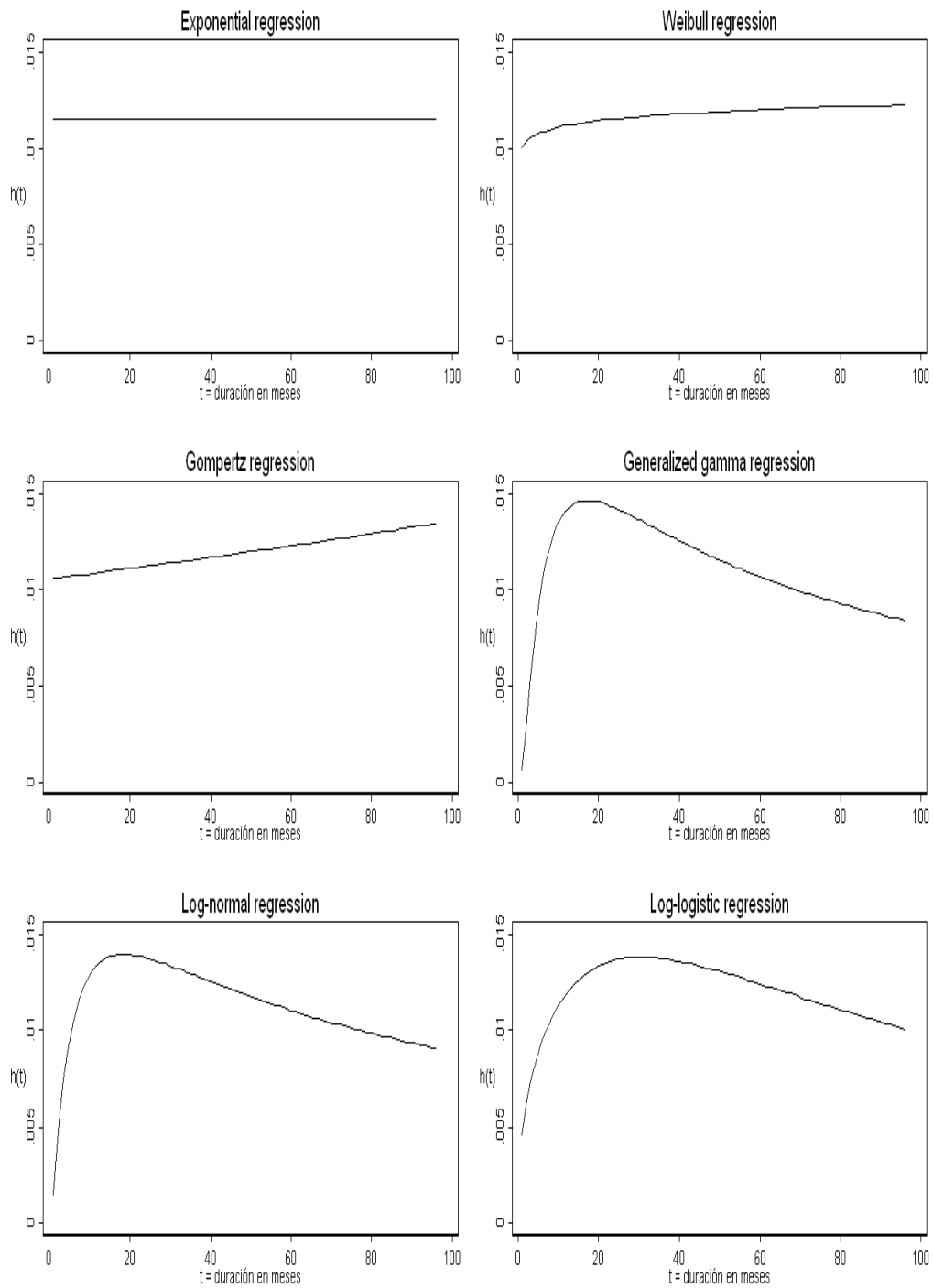


Figura 5.12: Funciones de riesgo de los modelos paramétricos estimados



La existencia de esta heterogeneidad no observada significa que existe información, de la que no disponemos, que puede estar afectando de forma diferente a los individuos⁶⁷. En este sentido, algunos de los jóvenes del estudio tal vez permanecen más tiempo en esa situación intermedia antesala del primer empleo significativo, por razones que están más relacionadas con características que no han sido observadas y no realmente porque la propia duración provoque una dependencia negativa. En principio, el objeto de introducir variables explicativas en un modelo es reducir, en parte, la heterogeneidad del colectivo estudiado. No obstante, es difícil hacer una completa especificación de las fuentes de heterogeneidad por diversas razones y esta falta de control de la posible heterogeneidad produce, en estos modelos, sesgos en los parámetros estimados, tanto en aquellos que recogen la dependencia de la duración como en los que evalúan los efectos de las variables explicativas⁶⁸.

Con respecto a la dependencia de la duración, está demostrado que produce dependencia negativa espúrea que se deriva del hecho de que si los individuos con las características no observadas tienden en general a presentar duraciones más largas, el conjunto de *individuos sujetos a riesgo* con el paso del tiempo estará formado por un mayor número de individuos con dificultades para transitar, lo que hace que la probabilidad instantánea de que se produzca el tránsito sea cada vez menor, esto es, se observe una dependencia de la duración negativa⁶⁹. La consecuencia inmediata de lo expuesto es que se complica la interpretación de tasas de salida hacia el empleo decrecientes, ya que la caída podría atribuirse (al menos en parte) al cambio en la composición del grupo a medida que avanza la duración.

Como señalábamos en la parte metodológica, de entre las distintas alternativas que la literatura econométrica ha desarrollado para controlar la heterogeneidad no observada, la que más se ha popularizado ha sido la consistente en una especificación paramétrica

⁶⁷Además de la existencia de factores no observados, Lancaster (1990) incluye como posible fuente generadora de heterogeneidad los errores de medición en las variables. En este sentido, el hecho de que la duración se construya apelando a la memoria del entrevistado puede haber producido ciertas incorrecciones en su cálculo.

⁶⁸Dolton y van der Klaauw (1995) afirman que la significatividad de la heterogeneidad no observada es más fiable en modelos que tienen una especificación más flexible de la función de riesgo base como es el caso del modelo de Cox. No obstante, en este caso el efecto de la mala especificación tiene como resultado infraestimar los parámetros, lo que en principio no invalida los contrastes de significación de las variables.

⁶⁹Varios de los trabajos empíricos que estudian la duración del primer desempleo confirman este hecho, como ya se puso de relieve en la tabla 5.3.

de la misma mediante la introducción de un factor en el modelo, incorrelacionado con el resto de variables explicativas, cuya distribución tiene media igual a uno y varianza desconocida, θ^{70} . La falta de información sobre la caracterización apropiada de esta distribución obliga a menudo a realizar distintos supuestos, siendo las distribuciones gamma y la gaussiana inversa las que en la práctica se han utilizado más comúnmente⁷¹.

Bajo esta consideración, se han reestimado los modelos presentados en la tabla 5.9 pero, en esta ocasión, incluyendo un término de heterogeneidad no observada con distribución gamma⁷², que ha requerido de la estimación de un parámetro más, la varianza del factor de heterogeneidad⁷³.

Los resultados aparecen en la tabla 5.11, incorporando en la misma una medida del grado de importancia de esa heterogeneidad no observada, capturada a través de la varianza del factor. Dado que el parámetro θ es estadísticamente distinto de cero para cualquiera de las distribuciones consideradas, la heterogeneidad debe ser tomada en cuenta en el análisis, por lo que el análisis de la dependencia de la duración realizado previamente ha de ser mirado con cierta cautela. En particular, los efectos de esta especificación incorrecta del modelo -debido a la no inclusión de un factor de heterogeneidad- son los esperados. Por un lado, se sobreestima (infraestima) la dependencia de la duración negativa (positiva) y, por otro lado, los efectos de las variables explicativas son minorados por la heterogeneidad no observada.

En concreto, la heterogeneidad parece más importante en los modelos con riesgos proporcionales, lo que nos ha inducido a pensar que ésta está más bien relacionada con una incorrecta especificación del modelo en cuanto a su forma funcional, que con un error por omisión de variables relevantes.

Repasemos de forma sucinta qué implicaciones tiene el tomar la distribución gamma

⁷⁰La otra alternativa (véase Heckman y Singer, 1984) a pesar de resultar más atractiva al no imponer restricciones en el comportamiento de la heterogeneidad no observada, presenta algunos inconvenientes. Entre ellos, graves problemas de cálculo que hacen difícil su puesta en práctica en numerosas ocasiones.

⁷¹La utilización de la distribución gaussiana inversa viene respaldada por la consideración de que el término de heterogeneidad recoge un gran número de características no observables. Por su parte, la primera justificación del uso de la distribución gamma para recoger la heterogeneidad no observada fue propuesta por Abbring y van den Berg (2001).

⁷²Resultados muy similares se han obtenido cuando se ha considerado la distribución gaussiana inversa.

⁷³El uso de esta técnica no está exenta de críticas provenientes de la posibilidad de que la permanencia en el estado transitorio provoque cambios en las variables no observadas, como podría ser el caso de la motivación individual, lo que produciría que al tratar de corregir la heterogeneidad no sólo no se eliminara el sesgo en la estimación de la dependencia sino que se introdujeran otros.

Tabla 5.11: Estimación de diferentes modelos de duración paramétricos (con heterogeneidad no observada gamma)

| | Exp. | | Weibull | | Gompertz | | Gamma | | LogN | | LogL | |
|---|--------|---|---------|---|----------|---|--------|---|--------|---|--------|---|
| | (aft) | * | (aft) | * | (ph) | * | (aft) | * | (aft) | * | (aft) | |
| <i>Sexo (ref: mujeres)</i> | | | | | | | | | | | | |
| Hombres | -0,219 | * | -0,146 | * | 0,295 | * | -0,143 | * | -0,164 | * | -0,142 | * |
| <i>Búsqueda de empleo (ref: no ha buscado)</i> | | | | | | | | | | | | |
| 1-2 meses | -1,061 | * | -0,804 | * | 1,138 | * | -0,810 | * | -0,751 | * | -0,891 | * |
| 3-5 meses | -0,479 | * | -0,107 | * | 0,284 | * | -0,104 | * | -0,065 | * | -0,179 | * |
| 6-11 meses | -0,103 | * | 0,449 | * | -0,472 | * | 0,445 | * | 0,485 | * | 0,342 | * |
| más de 11 meses | 0,541 | * | 1,267 | * | -1,958 | * | 1,255 | * | 1,281 | * | 1,139 | * |
| <i>Edad de salida del sistema educativo centrada (Edad-16 años)</i> | | | | | | | | | | | | |
| edad | -0,130 | * | -0,146 | * | 0,298 | * | -0,146 | * | -0,147 | * | -0,148 | * |
| edad ² | 0,009 | * | 0,009 | * | -0,018 | * | 0,009 | * | 0,009 | * | 0,009 | * |
| <i>Nivel y sector de estudios (ref: estudios primarios y sin especialización)</i> | | | | | | | | | | | | |
| Primera Etapa de Secundaria | -0,290 | * | -0,206 | * | 0,447 | * | -0,207 | * | -0,215 | * | -0,216 | * |
| Bachillerato | -0,097 | | -0,013 | | 0,067 | | -0,016 | | -0,015 | | -0,017 | |
| Formación Profesional | | | | | | | | | | | | |
| Docencia y Humanidades | -0,313 | * | -0,044 | * | 0,125 | * | -0,054 | | -0,061 | | -0,071 | |
| Ciencias sociales | -0,292 | * | -0,147 | * | 0,307 | * | -0,145 | * | -0,147 | * | -0,159 | * |
| Ciencias exactas | -0,282 | * | -0,168 | * | 0,328 | * | -0,168 | | -0,168 | | -0,167 | |
| Ingeniería | -0,337 | * | -0,182 | * | 0,348 | * | -0,189 | * | -0,189 | * | -0,202 | * |
| Industria | -0,375 | * | -0,275 | * | 0,496 | * | -0,282 | * | -0,261 | * | -0,281 | * |
| Arquitectura | -0,316 | * | -0,091 | * | 0,204 | * | -0,092 | | -0,103 | | -0,088 | |
| Agricultura | -0,451 | * | -0,349 | * | 0,612 | * | -0,344 | * | -0,323 | * | -0,349 | * |
| Medicina y Serv. Sociales | -0,244 | * | -0,065 | * | 0,187 | * | -0,062 | | -0,080 | | -0,053 | |
| Otros servicios | -0,409 | * | -0,251 | * | 0,491 | * | -0,246 | * | -0,247 | * | -0,246 | * |
| Estudios Universitarios | | | | | | | | | | | | |
| Docencia | -0,151 | | 0,059 | | -0,065 | | 0,063 | | 0,064 | | 0,055 | |
| Artes | -0,022 | | 0,127 | | -0,181 | | 0,131 | | 0,160 | | 0,104 | |
| Humanidades | -0,123 | | 0,086 | | -0,103 | | 0,081 | | 0,060 | | 0,078 | |
| Ciencias soc. y Periodismo | -0,171 | | 0,038 | | -0,009 | | 0,044 | | 0,059 | | 0,031 | |
| Economía | -0,389 | * | -0,143 | * | 0,294 | * | -0,144 | * | -0,144 | * | -0,162 | * |
| Derecho | -0,314 | * | -0,020 | * | 0,101 | * | -0,020 | | -0,030 | | -0,024 | |
| Ciencias exactas | -0,223 | | 0,024 | | 0,041 | | 0,026 | | 0,028 | | 0,012 | |
| Informática | -0,670 | * | -0,274 | * | 0,477 | * | -0,278 | * | -0,302 | * | -0,281 | * |
| Ingeniería | -0,394 | * | -0,136 | * | 0,231 | * | -0,141 | * | -0,149 | * | -0,153 | * |
| Arquitectura y construc. | -0,946 | * | -0,436 | * | 0,697 | * | -0,437 | * | -0,446 | * | -0,459 | * |
| Agricultura y Veterinaria | -0,461 | * | -0,110 | * | 0,247 | * | -0,110 | | -0,114 | | -0,132 | * |
| Medicina y Serv. Sociales | -0,310 | * | -0,124 | * | 0,232 | * | -0,094 | | -0,116 | * | -0,118 | * |
| Otros servicios | -0,315 | | -0,164 | | 0,327 | | -0,165 | | -0,143 | | -0,203 | |

Tabla 5.11 (Continuación)

| | Exp. | Weibull | Gompertz | Gamma | LogN | LogL |
|---|----------|-------------|------------------|------------------|----------------|------------------|
| | (aft) | (aft) | (ph) | (aft) | (aft) | (aft) |
| <i>Nivel máximo de estudios de los padres (ref: estudios primarios)</i> | | | | | | |
| Secundaria Obligatoria o equiv. | -0,061 * | -0,047 * | 0,096 * | -0,042 | -0,042 | -0,052 |
| Bachillerato o equiv. | -0,056 | -0,045 | 0,085 | -0,043 | -0,032 | -0,047 |
| Superiores o equiv. | -0,116 * | -0,077 * | 0,146 * | -0,079 * | -0,083 * | -0,079 * |
| <i>Entorno económico</i> | | | | | | |
| vartparo | 0,545 * | 0,189 * | -0,312 * | 0,197 * | 0,204 * | 0,205 * |
| varocupa | -0,311 * | -0,369 * | 0,632 * | -0,368 * | -0,389 * | -0,372 * |
| <i>Región de procedencia (ref: Andalucía, Ceuta y Melilla)</i> | | | | | | |
| Aragón | -0,712 * | -0,539 * | 1,042 * | -0,548 * | -0,571 * | -0,559 * |
| Asturias | 0,164 | 0,109 | -0,220 | 0,112 | 0,102 | 0,117 |
| Baleares | -0,462 * | -0,275 * | 0,623 * | -0,276 * | -0,308 * | -0,274 * |
| Canarias | -0,315 * | -0,183 * | 0,427 * | -0,189 * | -0,207 * | -0,196 * |
| Cantabria | 0,023 | 0,193 | -0,260 | 0,185 | 0,170 | 0,162 |
| Castilla y León | -0,395 * | -0,292 * | 0,577 * | -0,295 * | -0,307 * | -0,307 * |
| Castilla La Mancha | -0,539 * | -0,371 * | 0,768 * | -0,377 * | -0,397 * | -0,382 * |
| Cataluña | -0,457 * | -0,404 * | 0,764 * | -0,405 * | -0,426 * | -0,420 * |
| Com. Valenciana | -0,542 * | -0,434 * | 0,851 * | -0,433 * | -0,449 * | -0,433 * |
| Extremadura | 0,084 | 0,085 | -0,131 | 0,088 | 0,092 | 0,076 |
| Galicia | -0,239 * | -0,260 * | 0,503 * | -0,263 * | -0,255 * | -0,279 * |
| Madrid | -0,303 * | -0,246 * | 0,479 * | -0,254 * | -0,257 * | -0,272 * |
| Murcia | -0,212 * | -0,146 * | 0,370 * | -0,151 * | -0,180 * | -0,148 * |
| Navarra | -0,515 * | -0,376 * | 0,712 * | -0,387 * | -0,434 * | -0,376 * |
| País Vasco | -0,562 * | -0,372 * | 0,703 * | -0,379 * | -0,410 * | -0,380 * |
| La Rioja | -0,511 * | -0,236 * | 0,544 * | -0,240 * | -0,277 * | -0,231 * |
| constante | 4,897 * | 3,406 * | -5,024 * | 3,563 * | 3,639 * | 3,708 * |
| parámetro de forma | $p=1$ | $p=1,990 *$ | $\gamma=0,047 *$ | $\kappa=0,624 *$ | $\sigma=1,048$ | $\gamma=0,577 *$ |
| | | | | $\sigma=0,708$ | | |
| varianza (θ) | 0,239 * | 2,684 * | 2,630 * | 1,440 * | 0,484 * | 0,539 * |
| Nº de individuos | 14287 | 14287 | 14287 | 14287 | 14287 | 14287 |
| Nº de observaciones | 64642 | 64642 | 64642 | 64642 | 64642 | 64642 |
| Log L | -17584 | -17165 | -17065 | -17170 | -17203 | -17185 |
| AIC | 35277 | 34440 | 34240 | 34452 | 34515 | 34479 |

Nota: La estimación se ha llevado a cabo utilizando errores estándar robustos debido a la existencia de múltiples registros por individuo.

* significativo al 5%

como base para recoger el comportamiento de la función de riesgo. Dados los parámetros de esta distribución, el aspecto de la función es el que se recoge en la figura 5.14. La función tiene forma campaniforme no simétrica, con un tramo creciente en el que el crecimiento es muy rápido y a partir de ese máximo situado en los 20 meses, va decreciendo pero a un ritmo más lento que en la fase de crecimiento. En contraposición, la función de riesgo obtenida en la estimación paramétrica (sin imposición de forma funcional ninguna), presentaba forma de U.

De nuevo, volvemos a subrayar la dificultad de representar la estructura de la dependencia de la duración que presentan estos datos mediante la utilización de un modelo paramétrico. En realidad, lo más que se puede esperar es que recojan el comportamiento de al menos una parte de esta función. En este sentido, el modelo paramétrico seleccionado (el modelo gamma) lo que parece ajustar bien es la primera “joroba” de la función de riesgo.

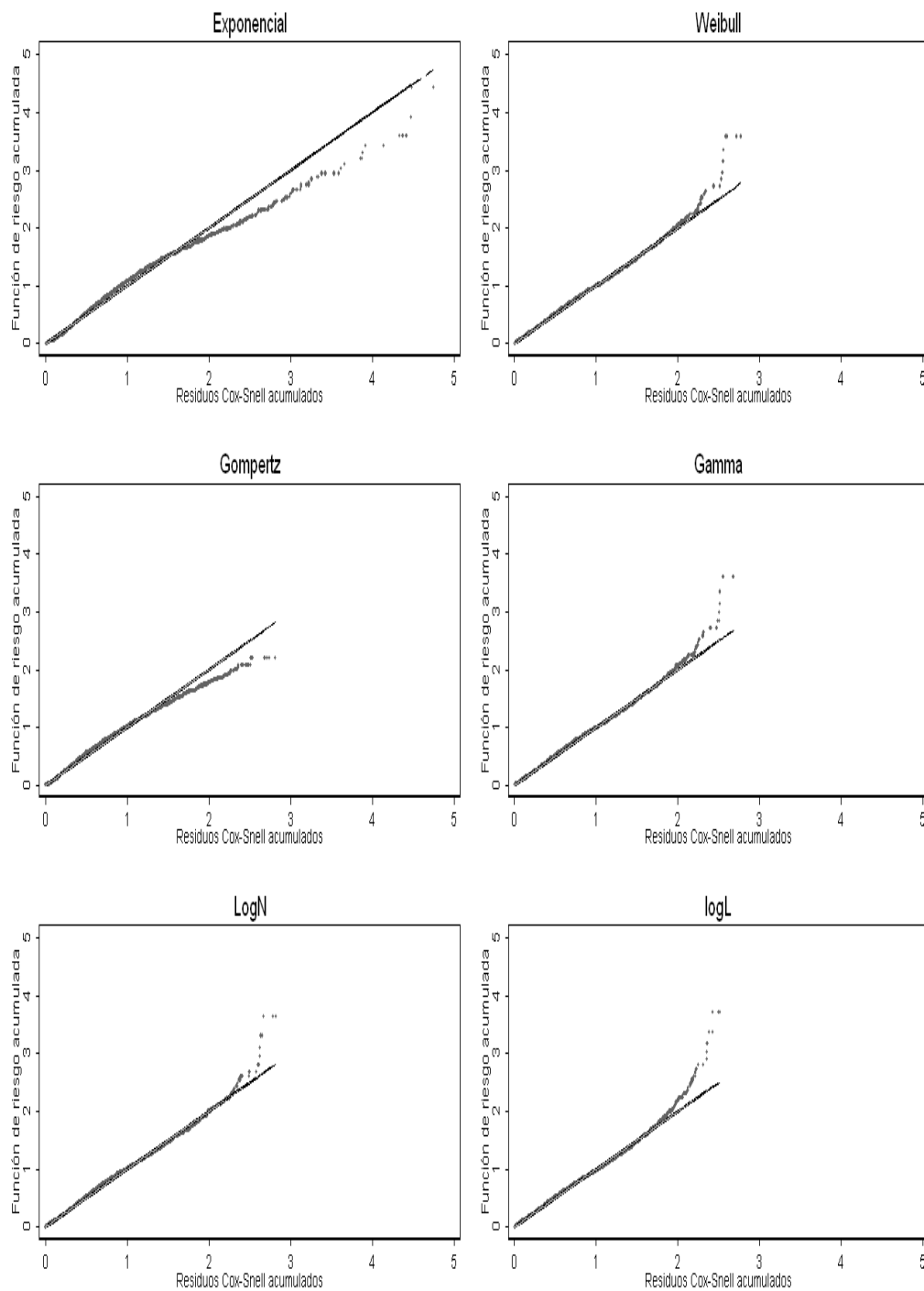
Tabla 5.12: **Efecto de la heterogeneidad no observada sobre la dependencia de la duración**

| | <i>Modelo sin heterogeneidad</i> | <i>Modelo con heterogeneidad gamma</i> |
|--------------------|--------------------------------------|--|
| Exponencial | <i>constante</i> = 5,012 | <i>constante</i> = 4,897 |
| Weibull | $p = 1,043$ | $p = 1,990$ |
| Gompertz | $\gamma = 0,003$ | $\gamma = 0,047$ |
| Gamma generalizada | $\kappa = -0,231$ | $\kappa = 0,624$ |
| | $\sigma = 1,274$ | $\sigma = 0,708$ |
| Log-normal | $\sigma = 1,237$ | $\sigma = 1,048$ |
| Log-logística | $\gamma = 0,704$ | $\gamma = 0,577$ |

5.4.2 Dependencia de la duración mediante la estimación de un modelo discreto

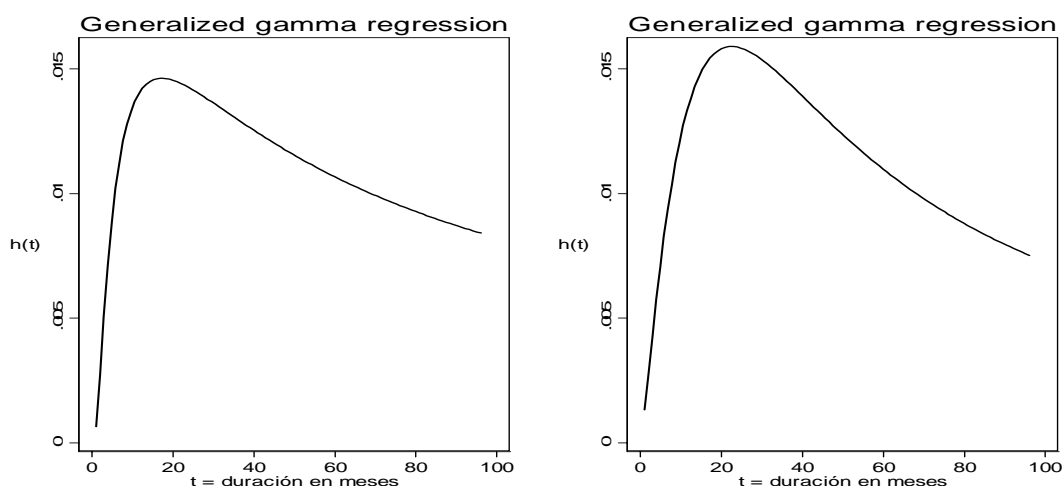
En algunos de los trabajos revisados en el capítulo 2, que han estudiado específicamente la duración de este periodo transitorio, se ha preferido la utilización de modelos de duración discretos (por ejemplo, Andrews *et al.*, 2002 dentro de los estudios internacionales y Congregado y García Pérez, 2002 en el caso de los trabajos nacionales). Varios han podido ser los motivos que han inducido a sus autores a tomar tal decisión:

Figura 5.13: Evaluación de los modelos paramétricos con heterogeneidad no observada gamma: Función de riesgo acumulada de los residuos Cox-Snell



la observación del proceso de forma discreta, la existencia de variables dependientes del tiempo, un elevado número de empates, el requerimiento de un estudio más profundo de la función de riesgo base, etc⁷⁴.

Figura 5.14: Función de riesgo del modelo Gamma con y sin heterogeneidad no observada



La primera de estas justificaciones podría ser de aplicación en nuestro caso, ya que la duración viene medida en meses como consecuencia del procedimiento de cálculo y, en última instancia, de la información que suministra el módulo, a pesar de que la transición al empleo significativo puede tener lugar cualquier día del mes, una vez el individuo sale del sistema educativo. En definitiva, aunque la naturaleza del proceso es realmente continua nosotros tan sólo lo observamos de forma discreta. Pese a esta cuestión básicamente técnica, el hecho de que la longitud del proceso de transición varíe de forma significativa de un país a otro, nos lleva a pensar que no siempre la unidad de medida óptima ha de ser la misma y que, en este caso, dependerá de si estamos analizando periodos de transición cortos o largos⁷⁵.

⁷⁴Allison (1995) propone una guía orientativa para tomar la decisión sobre cuál es la opción más conveniente según las características de los datos y los objetivos perseguidos.

⁷⁵Así, por ejemplo, la duración ha sido medida en semanas en el estudio de Andrews *et al.* (2002) sobre la transición al empleo en la zona de Lancashire en el Reino Unido. En este caso, con un periodo de observación más corto, el porcentaje de censura no ha excedido el valor de 20%, lo que pone de relieve que para esta población el proceso de transición es más corto que en el caso español.

En cuanto al segundo motivo, el número de variables dependientes del tiempo incluidas en la especificación ha sido tan sólo de dos aunque su incorporación, mediante un proceso de expansión de los datos, ha incrementado de forma significativa el tamaño final del fichero. Aún así, su manejo ha sido relativamente razonable. Por último, en relación al número de empates presentes en los datos tampoco ha constituido una dificultad insalvable, pues tras una larga espera, fue posible estimar el modelo de Cox utilizando un método exacto para el tratamiento de los empates.

Por lo tanto, al tener en cuenta cada uno de esos supuestos por separado, nada nos induce a sospechar que los resultados obtenidos en la modelización continua queden comprometidos en modo alguno. No obstante, la necesidad de profundizar en la estructura de la dependencia de la duración y la combinación en nuestros datos de un número considerable de empates y la existencia de variables dependientes del tiempo, nos ha impulsado a ensayar con varios modelos discretos. Es por ello que a continuación realizamos un *análisis de duración* discreto de tal forma que tengamos todos los puntos de vista posible⁷⁶.

En general, la estimación de un modelo de duración en tiempo discreto supone una estructura de los datos bastante diferente de la correspondiente a un modelo en tiempo continuo⁷⁷. En efecto, como se contemplaba en el capítulo metodológico, la función de verosimilitud se plantea como si de un modelo de elección discreta binario se tratara donde la variable dependiente -que requiere de una reorganización de los datos de tal manera que se pasa de tener una observación por individuo a tener tantas observaciones como meses ha estado sujeto a riesgo- toma el valor 1 si el individuo ha encontrado un empleo significativo en ese mes y cero en el resto de los casos. En este proceso de reorganización de los datos también se generan otras dos variables, una que identifica todas las observaciones correspondientes a un mismo individuo y otra que registra cada uno de los meses en que el individuo ha estado expuesto al riesgo de abandonar el estado transitorio previo al empleo significativo.

Además, la implementación del método discreto también necesita de una especi-

⁷⁶Otra alternativa, que mantendría el supuesto de proceso continuo en el tiempo pero que se ajustaría más a una valoración discreta de la dependencia de la duración, sería la estimación de un modelo *piece-wise* exponencial.

⁷⁷No ha sido exactamente así en nuestro caso al incorporar, en el modelo continuo, variables explicativas dependientes del tiempo aunque su variación no estaba asociada directamente al mes sino al año, por lo que el resultado de expandir la muestra no ha conducido al mismo número de observaciones que en el caso discreto.

ficación concreta de la función de riesgo, conociéndose dos posibilidades reales en la literatura: la función *log-log del complementario* y la función *logit*, cuyas expresiones son $h(t_j; x) = 1 - e^{-e^{\theta_j + x\beta}}$ y $h(t_j; x) = \frac{1}{1 + e^{-\theta_j - x\beta}}$, respectivamente. Mientras que la primera es la versión discreta del modelo de riesgos proporcionales de Cox, presuponiendo por tanto que se cumple la hipótesis de proporcionalidad, la segunda alternativa no considera esta hipótesis. En estas dos propuestas, el parámetro θ_j es el encargado de informar sobre la dependencia de la duración, pudiéndose especificar paramétrica (a través de una función de t) o no paramétricamente (mediante la creación de una variable ficticia por cada mes observado). Atendiendo a nuestra exigencia de buscar una forma funcional que se adapte lo mejor posible al comportamiento de la función de riesgo, hemos estimado varios modelos discretos con formas funcionales distintas para la función de riesgo base. En particular, se han ensayado diferentes polinomios en t y $\ln t$ con las dos especificaciones comentadas: *log-log del complementario* y *logit*.

Dado que recurriendo a criterios estrictamente estadísticos, como el criterio de información de AIC, ninguna de las dos especificaciones parece mostrar una mayor superioridad, nuestra decisión de utilizar finalmente la función *logit* ha estado basada exclusivamente en los resultados que se han obtenido en anteriores apartados, en particular, en la modelización continua⁷⁸.

En la tabla 5.13 se presentan algunos de los criterios de valoración utilizados para seleccionar entre las distintas formas funcionales. En particular, el modelo con un menor valor de AIC y, por tanto, que mejor se ajusta a los datos, se obtuvo considerando una función polinómica en $\ln t$ de orden 4, cuyos coeficientes recogemos en la tabla 5.14. Este polinomio replica casi perfectamente y de manera muy significativa el carácter decreciente y posteriormente creciente de la función de riesgo como se aprecia en la figura 5.15, donde también se ha incluido la función de riesgo que se obtuvo en la estimación no paramétrica y la función de riesgo que se extrae del ajuste a un modelo de Cox una vez se ha recogido el efecto de las variables explicativas.

Para finalizar, la forma de la función de riesgo estimada no entra en contradicción con otros resultados de estudios previos sobre el comportamiento de esta función. Por ejemplo, el trabajo de Congregado y García Pérez (2002) detectaba, tras truncar la duración en 36 meses, la existencia de dependencia de la duración negativa durante este

⁷⁸El resultado obtenido no es tan sorprendente si se tiene en cuenta que la representación de estas dos funciones produce unas curvas con un aspecto casi idéntico cuando la función de riesgo toma valores muy pequeños, inferiores a 0,2.

Tabla 5.13: Comparación de los modelos discretos según el criterio de Akaike (AIC): función de riesgo *logit*

| Especificación paramétrica de t | Nº parámetros | Log L | AIC |
|-----------------------------------|---------------|-----------|----------|
| en $\ln t$ | | | |
| función lineal | 1 | -42326,24 | 84760,48 |
| función cuadrática | 2 | -42312,15 | 84734,30 |
| función cúbica | 3 | -42246,64 | 84605,28 |
| polinomio de orden 4 | 4 | -42242,61 | 84599,22 |
| polinomio de orden 5 | 5 | -42242,02 | 84600,04 |
| en t | | | |
| función lineal | 1 | -42320,18 | 84748,36 |
| función cuadrática | 2 | -42285,38 | 84680,76 |
| función cúbica | 3 | -42264,16 | 84640,32 |
| polinomio de orden 4 | 4 | -42264,07 | 84642,14 |
| Nº de variables = 52 | | | |
| Nº de individuos = 14287 | | | |
| Nº de observaciones = 622048 | | | |

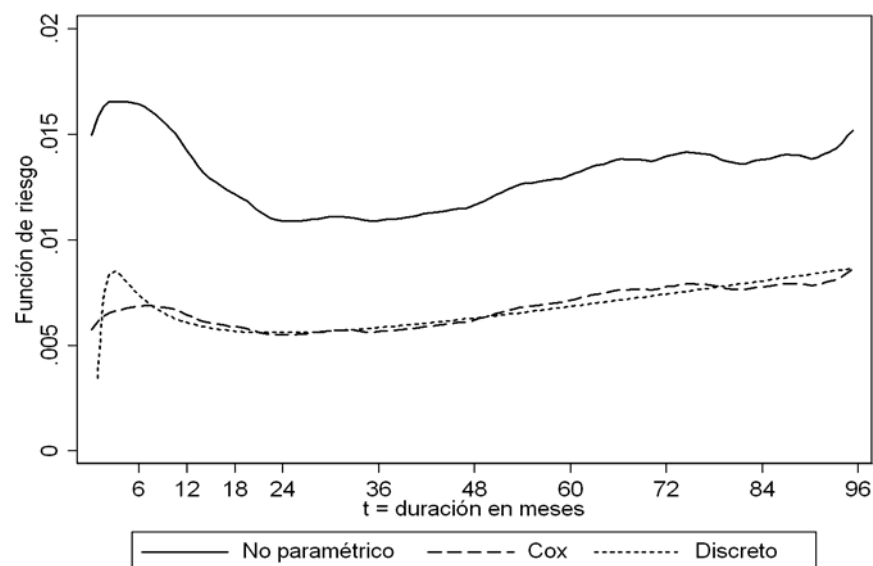
Tabla 5.14: Estimación paramétrica de la función de riesgo del modelo discreto *logit*

| Variable | Coefficiente | p-valor |
|-------------------|--------------|---------|
| $(\ln t)$ | 1,163 | 0,000 |
| $(\ln t)^2$ | -0,859 | 0,000 |
| $(\ln t)^3$ | 0,202 | 0,000 |
| $(\ln t)^4$ | -0,014 | 0,000 |
| Log L = -42242,61 | | |
| AIC = 84599,22 | | |

periodo, hecho que coincide con la caída inicial de la función aquí obtenida. Por otro lado, otros trabajos que restringen el comportamiento de la función de riesgo mediante la utilización de un modelo paramétrico, como el de Aguilar (2002) y el de González Betáncor *et al.* (2004), obtienen justamente el resultado contrario debido a que al ceñirse a una distribución monótona (distribución Weibull y Gompertz, respectivamente), es imposible captar los dos tramos de la función de riesgo, prevaleciendo en este caso el que se observa a partir de los 36 meses, cuando la dependencia es positiva.

En último lugar, la función de riesgo, como se indicó en el análisis no paramétrico, se caracteriza por presentar valores algo más reducidos que las estimaciones propias de esta función en otros países. En este sentido, la presunción de Russel y O'Connell (2001, pág. 16) de que la dependencia de la duración ha de ser menor en aquellos países en los que el mercado laboral juvenil se identifica por altas tasas de paro que obliga a los jóvenes a esperar pacientemente en la cola del paro durante un periodo de tiempo considerable antes de encontrar un empleo, parece confirmarse.

Figura 5.15: Comparativa de las funciones de riesgo



5.5 Conclusiones

En este capítulo se ha llevado a cabo un estudio empírico de la consecución de un primer empleo significativo durante la década de los noventa y, en particular, de los determinantes que pudieron afectar al tiempo que transcurrió desde que los jóvenes concluyeron su formación educativa hasta que alcanzaron el objetivo perseguido, disfrutar de una cierta estabilidad laboral. Este análisis nos ha permitido, por tanto, delimitar cuáles han sido los factores claves del éxito (o fracaso) en ese empeño por parte de los jóvenes de establecerse laboralmente. La atención, por tanto, se ha centrado en los factores que han incidido directamente en la duración de este periodo de transición, intentando cuantificar su efecto a través de diferentes propuestas contenidas en los modelos de duración, aprovechando en cada caso las ventajas específicas que ofrecen cada una de las técnicas. En particular, se han estimado tanto modelos semiparamétricos como paramétricos. Mientras que en la estimación paramétrica se requiere suponer una determinada distribución para recoger la forma de la función de riesgo, de tal manera que los parámetros que caracterizan a esa distribución son estimados a la par que los efectos de los factores explicativos, en la estimación semiparamétrica no se impone ningún tipo de restricción sobre el comportamiento de la función de riesgo.

Con anterioridad a esta etapa de modelización del proceso de transición, hemos realizado un análisis gráfico y numérico de las curvas de supervivencia -estimadas no paramétricamente mediante el estimador de Kaplan-Meier- de los distintos subgrupos de individuos que resultan de dividir a la muestra según las categorías de algunas de las variables explicativas. De este análisis hemos podido deducir que los factores explicativos considerados sí que conducen a diferencias significativas en las probabilidades de supervivencia de los jóvenes, siendo estas diferencias más relevantes cuando se tienen en cuenta factores como el nivel de estudios, la edad de salida del sistema educativo o la duración de la búsqueda.

Nuestra primera apuesta fue por la flexibilidad que caracteriza a la estimación semiparamétrica, planteando un modelo de riesgos proporcionales (o modelo de Cox). Dado que las propiedades de este modelo, su fácil interpretación e implementación, pueden verse trastocadas si la hipótesis en la que se sustenta, la hipótesis de proporcionalidad, no es apoyada por los datos, hemos realizado -como una etapa más en el proceso de especificación del modelo de Cox- una evaluación del grado de cumplimiento de la hipótesis, tanto para cada variable individualmente como para el modelo en su

conjunto. Tras comprobar que la sostenibilidad de dicha hipótesis quedaba en entredicho para algunas de las variables, se ha resuelto estimar el modelo por separado para cada sexo y estratificar según la variable búsqueda de empleo.

A continuación exponemos los resultados más significativos que se extraen de interpretar la especificación final del modelo de Cox. La primera diferencia detectada en relación al tiempo que se tarda en alcanzar una posición laboral estable viene establecida por el sexo del individuo: los jóvenes varones que se incorporaron al mercado de trabajo en la década de los noventa tenían una probabilidad mayor de obtener su primer empleo significativo más rápidamente que las jóvenes que tomaron esa misma decisión, independientemente de en que momento del proceso de transición se encontraran. Además, este factor ha condicionado el efecto de los otros factores analizados al considerarse más oportuno estimar el modelo de Cox por separado para cada sexo.

Con respecto a la variable edad, cuando se considera el conjunto de la muestra (tanto hombres como mujeres) el efecto de esta variable no es lineal, apareciendo un punto de inflexión alrededor de los 24 años, lo que significa que si bien la probabilidad de lograr un empleo significativo se incrementa con la edad, cumplidos unos años los individuos ven reducidas sus ventajas comparativas con respecto a los más jóvenes. Una explicación plausible a este hecho está fundamentada en la idea de que la edad puede estar también recogiendo otras cualidades de estos futuros trabajadores que los empleadores utilizan en el proceso de selección de personal, de tal manera que los jóvenes que dilatan su estancia en el sistema educativo son peor valorados y encuentran mayores dificultades para conseguir un empleo. Cuando se distingue por sexo, se observa que el efecto de la edad es algo mayor para los hombres que para las mujeres, posiblemente porque la edad ejerce una mayor presión dentro del colectivo de los varones que, por tradición, se les ha asignado un papel más activo en el mercado de trabajo.

El efecto del nivel educativo parece corroborar los postulados de la teoría del capital humano al verificarse que cuanto mayor es el nivel de estudios, menor es el tiempo que se tarda en realizar la transición. Por otro lado, este factor es más relevante dentro del colectivo de las mujeres para el que el nivel de estudios alcanzado condiciona en gran medida sus posibilidades de conseguir un empleo significativo. Es más, la importancia del bagaje educativo es doble pues no sólo tiene importancia el nivel alcanzado sino que también la especialidad determina significativamente la duración del proceso de transición. Resulta de interés también subrayar las diferencias encontradas en relación

a esta última variable cuando se consideran los Estudios Universitarios, de manera que la decisión personal sobre cuál carrera estudiar se ha convertido en un elemento de gran trascendencia para el futuro laboral de los jóvenes actuales. En general, la distribución de los especialidades que ofrecen mayores posibilidades de estabilidad laboral concuerda con la distribución sectorial de la demanda laboral.

Dentro de las variables familiares que pretenden recoger diversos aspectos como el grado de conocimiento y accesibilidad al mercado de trabajo, el nivel de estudios de los padres como una medida estimativa de las oportunidades laborales si se tiene en cuenta que los contactos familiares suelen ser una de las principales vías de búsqueda de empleo, explica también parte de las diferencias encontradas en los tiempos de espera hasta conseguir el empleo significativo. En particular, cuanto mayor es el nivel de estudios de los padres, mayor es el recorte en la duración del proceso de transición de sus progenitores.

Las condiciones económicas tienen el efecto esperado, la consecución de un empleo estable se alcanza con mayor rapidez en los periodos que la economía se encuentra en expansión y, por otra parte, cuanto mayor es la demanda, mayores son las posibilidades de afianzarse en el mercado laboral.

Por último, la localización geográfica de los individuos en la búsqueda de un empleo estable también muestra los resultados esperados. En particular aquellas comunidades más ricas presentan una estimación más favorable a la salida a un empleo significativo frente a las más pobres.

La segunda pretensión de este análisis de duración era la de calificar el tipo de dependencia de la duración que caracteriza a este proceso de transición. Para llevar a cabo este objetivo, se ha barajado la posibilidad de encontrar un modelo paramétrico que capte las especificidades observadas en la estimación no paramétrica de la función de riesgo. En concreto, la forma de esta función parece transmitir un mensaje claro y, en cierta manera razonable, en el sentido de que la trayectoria laboral de un joven, al igual que los comienzos en cualquier otra actividad, se aborda experimentando un periodo inicial de adaptación y aprendizaje, antes de conseguir asentarse. Esta situación acontece en la vida laboral de la mayor parte de los jóvenes, a excepción de aquellos afortunados que logran un primer empleo significativo en la primera oportunidad que les surge. Sin embargo, aunque el resultado no nos sorprende sí que resulta descorazonador que ese periodo de tiempo sea tan amplio, de al menos dos años.

La modelización paramétrica, haciendo uso de las distribuciones habituales en este

campo, no ha conseguido plasmar claramente los dos tiempos que se detectan en el comportamiento de la función de riesgo. Con la intención de avanzar en este campo sobre la dependencia de la duración y en aras de lograr una forma funcional que se adapte a la forma de esta función, se han implementado en la última parte de este capítulo modelos de duración discretos (en particular, un modelo logístico) en los que la función de riesgo se especifica mediante una función polinómica en $\ln t$. Después de estimar varios modelos incrementando de forma sucesiva el orden del polinomio, se logra obtener una función que replica casi perfectamente el carácter decreciente y posteriormente creciente de la función de riesgo.

A modo de resumen final de este capítulo, los resultados obtenidos son bastante coherentes con lo previsto y denotan que las diferencias en la duración del proceso de transición que existen tienen que ver principalmente con las características personales pero que también otros factores relacionados con el entorno familiar y el entorno económico conducen igualmente a tiempos de espera más o menos prolongados.