

UNIVERSIDAD DE COSTA RICA  
SISTEMA DE ESTUDIOS DE POSGRADO

Práctica Profesional I:

SELECCIÓN DE MODELOS POLITÓMICOS DE RESPUESTA AL ÍTEM  
PARA ANALIZAR EL CONOCIMIENTO SOBRE LOS MEDIOS DE  
TRANSMISIÓN DEL VIH

Práctica Profesional II:

EFFECTO DE LOS COMPONENTES EDAD, PERIODO Y COHORTE SOBRE  
EL NIVEL DE LA FECUNDIDAD EN COSTA RICA: UN ANÁLISIS DE LAS  
ENCUESTAS DE 1976 A 2015

Trabajo final de investigación aplicada sometida a la consideración de la Comisión  
del Programa de Estudios de Posgrado en Estadística para optar por el grado y  
título de Maestría Profesional en Estadística

ALEJANDRA GONZÁLEZ GÓMEZ

Ciudad Universitaria Rodrigo Facio, Costa Rica

2019

“Este trabajo final de investigación aplicada fue aceptado por la Comisión del Programa de Estudios de Posgrado en Estadística de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Profesional en Estadística”

- Comité evaluador de Práctica Profesional I -

---

Dr. Guaner Rojas Rojas

**Tutor**

---

Dr. Gilbert Brenes Camacho

**Lector**

---

Dra. Eiliana Montero Rojas

**Lectora**

---

Alejandra González Gómez

**Candidata**

- Comité evaluador de Práctica Profesional II -

---

Dr. Arodys Robles Soto

**Tutor**

---

Dr. Gilbert Brenes Camacho

**Lector**

---

Mag. María Isabel González Lutz

**Lectora**

---

Alejandra González Gómez

**Candidata**

# CONTENIDO

RESUMEN.....	vii
ÍNDICE DE CUADROS.....	viii
ÍNDICE DE GRÁFICOS.....	x
ÍNDICE DE FIGURAS.....	xi
SIGLAS Y ABREVIATURAS.....	xii
I PARTE.....	1
INTRODUCCIÓN.....	2
JUSTIFICACIÓN.....	4
MARCO DE REFERENCIA.....	12
1. Modelos politómicos de respuesta al ítem.....	16
1.1. Modelos politómicos de respuesta nominal (MPRN).....	17
1.2. Modelos politómicos de respuesta ordinal.....	19
1.2.1. Modelo de crédito parcial (MCP).....	20
1.2.2. Modelo de crédito parcial generalizado (MCPG).....	22
2. Métodos y criterios para la selección de modelos.....	22
2.1. Análisis factorial confirmatorio (AFC).....	23
2.2. Criterios para la selección de modelos: índices de ajuste.....	25
a. Test de razón de verosimilitud ( $G^2$ ).....	26
b. Criterio de Información de Akaike (AIC).....	27
c. Criterio Bayesiano de información (BIC).....	27
2.3. Funcionamiento diferencial de los ítems (DIF).....	28
METODOLOGÍA.....	35
1. Objetivos de estudio.....	35
1.1. Objetivo general.....	35
1.2. Objetivos específicos.....	35
2. Fuentes de información.....	35

3. Población de estudio.....	36
4. Variables de estudio.....	36
5. Análisis estadístico .....	39
RESULTADOS.....	41
1. Descripción de la población y variables de estudio.....	41
1.1. Características de la población de estudio .....	41
1.2. Distribución de respuestas y comportamiento de los ítems .....	42
2. Análisis factorial confirmatorio.....	46
3. Estimación de modelos .....	47
3.1. Parámetros de dificultad y discriminación de los modelos.....	48
3.2. Curvas características de categoría (CCC).....	51
3.3. Información de los ítems y del test.....	53
3.4. Estimación de las puntuaciones ( $\theta$ ).....	54
4. Selección de modelos y el principio de invarianza.....	55
4.1. Análisis de criterios de selección.....	55
4.2. Distribución de las puntuaciones.....	56
4.3. Funcionamiento diferencial de los ítems.....	57
CONCLUSIONES Y DISCUSIÓN.....	60
BIBLIOGRAFIA.....	66
II PARTE .....	73
INTRODUCCIÓN.....	74
JUSTIFICACIÓN .....	78
MARCO DE REFERENCIA.....	87
1. El descenso de la fecundidad en Costa Rica .....	87
2. Los estudios de Edad – Periodo – Cohorte.....	97
2.1. El problema de identificación .....	102
2.2. Métodos para el análisis de datos EPC .....	104
2.2.1. Modelo reducido de dos factores.....	104
2.2.2. Modelos lineales generalizados restringidos .....	106

2.2.3.	Función de penalización.....	107
2.2.4.	Transformaciones paramétricas no lineales.....	107
2.2.5.	Empleo de variables proxy.....	108
2.3.	Modelos lineales mixtos generalizados (MLMG).....	108
2.3.1.	Definición de cohorte y periodo como efectos fijos o de aleatorios.....	113
2.3.2.	Significancia de los componentes cohorte y periodo.....	116
METODOLOGÍA.....		122
1.	Objetivos de estudio.....	122
1.1.	Objetivo general.....	122
1.2.	Objetivos específicos.....	122
2.	Fuentes de información.....	122
3.	Población de estudio.....	123
4.	Variables de estudio.....	124
5.	Análisis de datos.....	127
RESULTADOS.....		132
1.	Características de la población y variables de estudio.....	132
1.1.	Descripción de las variables edad, periodo y cohorte.....	132
1.2.	Número de hijos tenidos.....	138
2.	MLMG para datos de fecundidad.....	144
CONCLUSIONES Y DISCUSIÓN.....		156
BIBLIOGRAFIA.....		165

## RESUMEN

Este documento presenta los informes generados en los Proyectos de Práctica Profesional I y Práctica Profesional II, ambos requisitos para optar al título de Maestría Profesional en Estadística.

El trabajo desarrollado como Práctica Profesional I titulado "*Selección de modelos politómicos de respuesta al ítem para analizar el conocimiento sobre los medios de transmisión del VIH*" toma como fuente de información la Encuesta Nacional de Salud Sexual y Reproductiva llevada a cabo entre la población costarricense en el año 2015. Partiendo de la premisa de que el conocimiento sobre los medios de transmisión del VIH permite que las personas puedan protegerse, este trabajo analiza la calidad psicométrica de una escala que pretende medir los conocimientos, actitudes y prácticas hacia el VIH, con el fin de identificar la idoneidad de los ítems de una escala cuya información se emplea para formular distintas estrategias de política pública. En este trabajo se estiman diferentes modelos politómicos de respuesta al ítem para determinar cuál se ajusta mejor a la distribución de los datos y se verifica el principio de invarianza de los parámetros de los ítems según sexo y edad, mediante el funcionamiento diferencial de ítems.

En el trabajo "*Efecto de los componentes Edad, Periodo y Cohorte sobre el nivel de la fecundidad en Costa Rica: un análisis de las encuestas de 1976 a 2015*", desarrollado en el ejercicio de la Práctica Profesional II, se ajusta un modelo lineal mixto generalizado para estimar el efecto neto de los componentes Edad, Periodo y Cohorte sobre el número promedio de hijos tenidos (como medida de fecundidad), por las mujeres entre 15 y 49 años de edad, controlando por el efecto de otras variables. Se toman como fuentes de información las encuestas de fecundidad y salud sexual realizadas en Costa Rica desde 1976 hasta 2015. La particularidad de este trabajo reside precisamente en la incorporación de esos tres componentes en un modelo de regresión, que logra lidiar con su relación linealmente dependiente y brindar información que permita describir los fenómenos involucrados en la compleja dinámica que determina la fecundidad.

## ÍNDICE DE CUADROS

Cuadro 1 Número y porcentaje de entrevistados, según características demográficas .....	42
Cuadro 2 Porcentaje de respuestas correctas, incorrectas y no sabe para cada ítem de conocimiento sobre las formas de transmisión del VIH, según sexo .....	44
Cuadro 3 Porcentaje de respuestas correctas, incorrectas y no sabe para cada ítem de conocimiento sobre las formas de transmisión del VIH, según edad .....	45
Cuadro 4 Matriz de correlaciones policóricas entre los ítems que componen la escala de conocimientos sobre el VIH .....	46
Cuadro 5 Parámetros de dificultad y discriminación de cada ítem, estimados para los modelos analizados.....	48
Cuadro 6 Prueba de razón de verosimilitudes entre los modelos anidados .....	56
Cuadro 7 Estadísticas de bondad de ajuste, AIC y BIC para los modelos estimados .....	56
Cuadro 8 Pruebas de razón de verosimilitudes y cambio en el coeficiente $\beta$ de la variable latente para analizar el DIF, según sexo .....	58
Cuadro 9 Pruebas de razón de verosimilitudes y cambio en el coeficiente $\beta$ de la variable latente para analizar el DIF, según grupos de edad.....	58
Cuadro 10 Tamaño de muestra y rango etario de las mujeres entrevistadas en las encuestas de 1976 al 2015.....	124
Cuadro 11 Número y porcentaje de mujeres que se analizan en las categorías de las variables edad, periodo y cohorte.....	133
Cuadro 12 Número y porcentaje de mujeres entrevistadas, según grupos de edad y año de encuesta. Costa Rica, 1976-2015.....	134
Cuadro 13 Número y porcentaje de mujeres entrevistadas, según cohorte de nacimiento y año de encuesta. Costa Rica, 1976-2015 .....	135
Cuadro 14 Características de las mujeres por año de encuesta y principales variables analizadas. Costa Rica, 1975-2015.....	136
Cuadro 15 Características de las mujeres por grupos de edad y principales variables analizadas. Costa Rica, 1975-2015 .....	138

Cuadro 16 Número promedio de hijos nacidos vivos que han tenido las mujeres según edad, periodo y cohorte. Costa Rica, 1976-2015 .....	140
Cuadro 17 Número promedio de hijos nacidos vivos por periodo, edad y cohorte, según variables de estudio. Costa Rica, 1976-2015 .....	143
Cuadro 18 Resultados de la estimación del modelo de efectos fijos, sin incluir las variables cohorte y periodo. Costa Rica, 1976-2015.....	144
Cuadro 19 Resultados de la estimación del MLMG. Costa Rica, 1976-2015 .....	146
Cuadro 20 Resultados de la prueba F para determinar la presencia de efectos aleatorios. Costa Rica, 1976-2015 .....	154

## ÍNDICE DE GRÁFICOS

Gráfico 1 Porcentaje de respuestas correctas, incorrectas y no sabe, para cada ítem de conocimiento sobre las formas de transmisión del VIH.....	43
Gráfico 2 Análisis factorial confirmatorio para analizar la unidimensional del conjunto de ítems que estiman el conocimiento sobre las formas de transmisión del VIH.....	47
Gráfico 3 Curvas características de las categorías para cada ítem, según método de análisis .....	49
Gráfico 4 Función de información del test para cada modelo estimado .....	53
Gráfico 5 Función de información de los ítems, para cada modelo estimado .....	54
Gráfico 6 Relación entre la estimación del conocimiento sobre los medios de transmisión del VIH ( $\theta$ ) mediante los tres métodos analizados.....	55
Gráfico 7 Distribución de las puntuaciones obtenidas en la escala sobre conocimiento en VIH .....	57
Gráfico 8 Tasa global de fecundidad, Costa Rica 1950-2017. Mujeres 15-49 años .....	87
Gráfico 9 Tasas de fecundidad específicas por edad, Costa Rica 1950-2017. Mujeres de 15 a 49 años .....	96
Gráfico 10 Distribución del número de hijos nacidos vivos para el total de la muestra y por periodo. Costa Rica, 1976-2015.....	139
Gráfico 11 Número promedio de hijos nacidos vivos según periodo y grupos de edad. Costa Rica, 1976 – 2015.....	142
Gráfico 12 Efecto de edad sobre la estimación del número de hijos nacidos vivos. Costa Rica, 1976-2015.....	148
Gráfico 13 Efecto de periodo sobre la estimación del número de hijos nacidos vivos. Costa Rica, 1976-2015 .....	149
Gráfico 14 Efecto de cohorte sobre la estimación del número de hijos nacidos vivos. Costa Rica, 1976-2015 .....	151

## ÍNDICE DE FIGURAS

Figura 1 Curvas características del ítem y de categorías cuando se presenta DIF uniforme y no uniforme .....	30
Figura 2 Afirmaciones planteadas en la ENSSR-15 relacionadas con el conocimiento de los medios de transmisión del VIH .....	37
Figura 3 Estructura de datos con los componentes edad – periodo - cohorte.....	99
Figura 4 Descripción y codificación de las variables analizadas.....	125

## **SIGLAS Y ABREVIATURAS**

AFC	Análisis factorial confirmatorio
AFE	Análisis factorial exploratorio
AIC	Criterio de información de Akaike
BIC	Criterio Bayesiano de Información
CCC	Curva característica de categorías
CCI	Curva característica del ítem
CFI	Índice comparativo de ajuste
DIF	Funcionamiento diferencial del ítem
EAP	Esperanza a posteriori
EPC	Edad-Período-Cohorte
INEC	Instituto Nacional de Estadística y Censos
MCP	Modelo de crédito parcial
MCPG	Modelo de crédito parcial generalizado
MH	Test de Mantel-Haenszel
MLMG	Modelo lineal mixto generalizado
MPRN	Modelo politómico de respuesta nominal
MV	Máxima verosimilitud
MVR	Máxima verosimilitud restringida
RMSEA	Raíz cuadrada del error medio de aproximación
sida	Síndrome de inmunodeficiencia adquirida
SRMR	Raíz cuadrada estandarizada residual
TCT	Teoría clásica de los test
TGF	Tasa global de fecundidad
TLI	Índice de Tucker-Lewis
TRI	Teoría de respuesta al ítem
VIH	Virus de inmunodeficiencia humana

## **I PARTE**

SELECCIÓN DE MODELOS POLITÓMICOS DE RESPUESTA AL ÍTEM  
PARA ANALIZAR EL CONOCIMIENTO SOBRE LOS MEDIOS DE  
TRANSMISIÓN DEL VIH

## INTRODUCCIÓN

En la medición de constructos comúnmente se emplean escalas para generar una medida empírica de esa variable no observable y de difícil definición (Bunge, 1973; Borsboom, Mellenbergh, y van Heerden, 2003 y Martínez, 2013). Se analizan elementos como habilidades (matemática, lógica), rasgos (extroversión, ansiedad) o características conductuales (tendencia a compras, agresividad, hiperactividad). La construcción y análisis de escalas requiere de marcos de referencia, entre las que destacan la Teoría Clásica de los Test (TCT) y la Teoría de Respuesta al Ítem (TRI), que permiten la asignación de puntuaciones a un conjunto de individuos.

Los modelos basados en la TRI se han vuelto populares en los últimos años (An y Yung, 2014; Kohli, Koran y Henn, 2015 y Martínez, Hernández y Hernández, 2006) y se han empleado para analizar variables de salud (Edelen y Reeve, 2007; Hays, Brown, J., Brown, L. Spritzer y Crall, 2006 y Hays, Morales y Reise, 2000), psicológicas (Fraleay, Waller y Brennan, 2000; González-Betanzos, Rivera-Heredia y Padrós-Blásquez, 2015; Gray-Little, Williams y Hancock, 1997 y Orlando, Sherbourne y Thissen, 2000), en la medición educativa (Debera y Nalbarte, 2006 y Jiménez y Montero, 2013), en investigación de mercados (De Jong, Steenkamp, Fox y Baumgartner, 2008) y en muchas otras áreas (Caviezel, Bertoli-Barsotti y Lozza, 2011 y Hagman et al., 2009).

El objetivo general de este trabajo es comparar modelos politómicos de respuesta al ítem empleando como ejemplo los datos de una escala sobre conocimiento de los medios de transmisión del VIH. Esta escala fue dirigida a la población costarricense entre los 15 y 64 años de edad y aplicada a una muestra representativa de personas, quienes respondieron a la Encuesta Nacional de Salud Sexual y Reproductiva del año 2015.

Se analizan y comparan modelos politómicos de respuesta nominal (MPRN), el modelo de respuesta graduada (MRG) y el modelo de crédito parcial (MPRG). Estos dos últimos corresponden a modelos de tipo ordinal cuya codificación estará dada por los resultados del primer modelo.

En cada caso se calcularán medidas de bondad de ajuste para determinar cuál modelo explica de forma más eficiente la escala utilizada. A continuación se compara la estimación del parámetro de habilidad ( $\theta$ ) y la función de información de los modelos.

Posteriormente se verifica el principio de invarianza de los parámetros de los ítems según sexo y edad. El análisis del Funcionamiento Diferencial de los ítems (DIF, por sus siglas en inglés) permite verificar la presencia de sesgos sistemáticos en la medición del conocimiento sobre VIH entre subpoblaciones. El patrón en la incidencia y prevalencia de la infección por VIH varía en función de algunas variables como el sexo y la edad (Carey, 2002 y Ministerio de Salud, 2016), las guías metodológicas sugieren analizar el conocimiento en materia de VIH en estos grupos poblacionales (Arce, 2014 y ONUSIDA, 2011; 2014).

Se espera que este trabajo brinde insumos para el análisis apropiado del conocimiento de los medios de transmisión del VIH en la población costarricense y en el desarrollo de políticas públicas basadas en información estadística de calidad, que resulta clave para definir estrategias de salud. El conocimiento y la prevención es vital para reducir los nuevos casos de personas infectadas por el virus. Si la medición de esta variable es deficiente, las estrategias de información y prevención pueden estar siendo mal canalizadas.

## JUSTIFICACIÓN

En la construcción y análisis de los ítems y las escalas se utilizan métodos y modelos que permitan generar evidencia de la validez y confiabilidad del instrumento. La teoría clásica de los test y la teoría de respuesta al ítem son dos de las técnicas que se utilizan en dicho proceso.

A pesar del uso generalizado y de la utilidad práctica de la TCT, esta teoría parte de supuestos generales débiles, de escasa plausibilidad real, que constituyen tanto su fuerza como su debilidad (Martínez, 1995). En la metodología de la TCT los parámetros de los ítems y de las personas son dependientes del test y la muestra seleccionada, respectivamente, razón por la cual la TRI podría visualizarse como una herramienta que posee mejores propiedades (An y Yung, 2014; Fan, 1998 y Kohli, Koran y Henn, 2015) para alcanzar el mismo fin.

La teoría de respuesta al ítem permite obtener mediciones en el constructo latente evaluado que no están en función del instrumento utilizado, donde las propiedades de los test y los ítems sean invariantes con respecto a las personas que son evaluadas (Muñiz, 1997). En ese último caso se obtiene un análisis detallado de la calidad técnica del test y del aporte de cada ítem, de tal manera que se pueda analizar la relación entre el nivel de habilidad de los sujetos y la dificultad o discriminación de cada ítem. En estos modelos existe una estimación del error de medición para cada valor de  $\theta$ , a diferencia de la TCT donde se asume un error constante.

Los modelos politómicos de respuesta al ítem, en comparación con los modelos dicotómicos, permiten obtener una mejor estimación del constructo analizado porque brindan mayor información a lo largo del intervalo de habilidad, al tiempo que es posible conocer cuál es la probabilidad de elegir cada una de las categorías de respuesta en función de la habilidad de los sujetos.

En este trabajo se toma una escala de conocimiento sobre los medios de transmisión del VIH como aplicación para comparar modelos politómicos de respuesta al ítem. El ejercicio pretende mostrar el desarrollo de distintos modelos para ítems politómicos, identificar el modelo que

brinda el mejor ajuste en escalas de este tipo y en convención con las características de los datos, además de generar evidencia empírica acerca de la validez interna de dicho conjunto de ítems.

Desde el inicio de la epidemia del VIH, entre mediados y finales de los años setenta (Avert, 2017), más de 76 millones de personas [65.2 – 88.0 millones] contrajeron la infección y alrededor de 35 millones de personas [28.9 – 41.5 millones] han muerto a causa de este virus (ONUSIDA; WHO, 2017). Se estima que cerca de 36.7 millones de personas vivían con VIH en todo el mundo para finales del año 2016. La carga de la epidemia varía considerablemente entre países y regiones. África Sub-Sahariana es la región más afectada, donde el 4.2% de la población vive con VIH y donde se concentran cerca de dos tercios de los casos cero-positivos registrados en el mundo (WHO, 2017).

En Costa Rica, la mayoría de las nuevas infecciones por el VIH se presentan a causa de la transmisión sexual (ONUSIDA, 2011). La transmisión vertical del virus, es decir, de madre a hijo, muestra una baja incidencia en el país a raíz del alto porcentaje de mujeres que se realizan la prueba del VIH. Según datos de ONUSIDA (2013), al menos el 80% de las embarazadas se realizan la prueba de VIH. Los protocolos nacionales de prevención de la transmisión materno-infantil establecen el tamizaje para todas las mujeres embarazadas (Ministerio de Salud, 2016).

Se ha registrado un aumento en la tasa de diagnósticos por VIH<sup>1</sup> a lo largo del último decenio (Ministerio de Salud, 2017), pasando de 12.3 personas infectadas por cada 100,000 habitantes, en el año 2005, a 20.0 en el año 2016. Cabe señalar, sin embargo, que el aumento de las tasas de diagnóstico no implica necesariamente el aumento de nuevos casos de VIH sino que simplemente podría referirse a un mayor registro de las personas infectadas. Al respecto, de acuerdo con datos de las últimas encuestas de Salud Sexual y Reproductiva realizadas en Costa Rica, el porcentaje de personas que alguna vez se había hecho la prueba de VIH pasó de 30.6% en el año 2010 a 37.2% en el año 2015 (valor-p < 0.01).

---

<sup>1</sup> Formalmente se refiere a incidencia de nuevos diagnósticos por VIH.

Los nuevos casos detectados se presentan básicamente entre los hombres. Para el periodo 2002-2016 se estimó que cerca del 80% de los nuevos diagnósticos correspondían a ese grupo poblacional y que la razón hombre/mujer era de 4:1 (Ministerio de Salud, 2016). En el año 2016, la tasa de diagnóstico entre la población masculina fue de 34.6, mientras que, entre las mujeres, la cifra estimada fue de 5.2, por cada 100,000 habitantes (Ministerio de Salud, 2017).

Existen también variaciones entre grupos etarios. En el periodo comprendido entre 2002 y 2016, las mayores tasas de diagnóstico se encuentran entre la población en edad reproductiva, mostrándose un incremento significativo entre la población de 15 a 19 años de edad (80.6), pero principalmente entre la población de 20 a 24 años (306.3) (Ministerio de Salud, 2017). Considerando los distintos grupos etarios, aún se presentan diferencias según sexo, siendo mayor la tasa de nuevos diagnósticos entre la población masculina.

De acuerdo con datos del Ministerio de Salud (2017), el sida<sup>2</sup> muestra una tendencia a la disminución y del mismo modo la mortalidad por sida, que presentó un descenso a partir del año 1998 a causa de la introducción de la terapia antirretroviral y al aumento gradual de su cobertura. A partir del año 2000 se presentó la estabilización de la mortalidad.

Las estrategias para mitigar la epidemia van desde el aumento de la cobertura y fortalecimiento de los programas enfocados a la administración y distribución de medicamentos antirretrovirales, la detección temprana del VIH, implementar y fortalecer estrategias de prevención entre las que destacan el uso de preservativos y la circuncisión médica masculina voluntaria, reducir el estigma, la discriminación y la violencia de género (cf. ONUSIDA, 2016).

La vulnerabilidad a la infección de VIH aumenta considerablemente a raíz de una serie de factores que reducen la capacidad de las personas para prevenir la infección del virus, entre las que figuran la falta de conocimientos y aptitudes necesarias para protegerse y proteger a otros,

---

<sup>2</sup> El término 'sida' se utiliza para referirse a los procesos clínicos que están vinculados con la inmunodepresión provocada por el VIH y que en algunos casos ponen en peligro la vida.

la calidad y cobertura de los servicios, entre otros factores de índole cultural (Ministerio de Salud, 2016 y ONUSIDA, 2008).

La prevención es un elemento esencial en la lucha contra el sida. Casi dos tercios de las personas que viven con VIH podrían transmitir el virus, a menos que se les proporcionen herramientas de prevención y se les capacite para utilizarlas (ONUSIDA, 2016). No obstante, aún se manifiestan problemas en la prevención del VIH en adultos. Se observan déficits en la percepción del riesgo personal, es decir, no se tiene la percepción de encontrarse en riesgo de contraer el VIH (ONUSIDA, 2016).

El conocimiento sobre los medios de transmisión se considera un determinante de las conductas de riesgo (Carey y Schroder, 2002). Es importante analizar el conocimiento sobre los medios de transmisión del VIH porque está ligado a la reducción de conductas de riesgo. Un sólido conocimiento sobre el VIH y el sida es un requisito esencial, si bien, en ocasiones, insuficiente, para adoptar comportamientos que reduzcan el riesgo de la transmisión del VIH (ONUSIDA, 2014).

A su vez, debe considerarse que el conocimiento apropiado sobre los medios de transmisión permite reducir el estigma y la discriminación hacia esta población. Desde los primeros años de la epidemia mundial de VIH esta enfermedad fue estigmatizada, pues los primeros casos se registraron en hombres que tienen sexo con hombres y entre la población que utiliza drogas intravenosas, aun cuando el VIH se propagó a todos los sectores de la población (Acosta, 2015). La problemática, desde sus inicios, se ha centrado en la falta de información sobre su forma de transmisión y sobre el virus en general, lo que generó miedo entre la población y rechazo hacia los infectados (Ministerio de Salud, 2017 y Moss y Bacchetti, 1989).

Los programas públicos diseñados para reducir la transmisión del virus y de las conductas de riesgo entre la población, usualmente incluyen la educación en materia de VIH como un componente de la intervención, de tal modo que la medición del conocimiento proporciona

retroalimentación sobre el impacto de las estrategias así como una guía del seguimiento de intervenciones que deben monitorearse (Carey y Schroder, 2002 y Kalichman, 1998).

De esa manera, será necesario contar con una medición confiable y válida para determinar el nivel de conocimiento sobre el VIH, la identificación de los medios de transmisión e ideas erróneas sobre la transmisión del virus (Carey y Schroder, 2002) y las variables asociadas con esos elementos.

Para la medición de constructos, como el conocimiento en VIH, es común el uso de escalas. Múltiples estudios han propuesto escalas o cuestionarios para la medición del conocimiento en VIH, cuya extensión es variable. Algunas propuestas van desde los 45 ítems (Carey, Morrison-Beedy y Johnson, 1996), 18 ítems (Carey y Schroder, 2002) o 5 ítems (ONUSIDA, 2008).

Los test que se valen de una larga serie de ítems para analizar el conocimiento en materia de VIH, aun cuando posean características psicométricas apropiadas, presentan como desventaja el aumento en la duración de su aplicación, particularmente cuando forman parte del cuestionario de una encuesta por muestreo, ya que la incorporación de un gran número de ítems aumenta el costo de la encuesta.

Usualmente este tipo de cuestionarios poseen ya de por sí una amplia serie de preguntas, no solamente relacionadas con el VIH y su transmisión, sino también sobre otras variables de salud. Tal es el caso de las encuestas en materia de salud sexual y reproductiva, en las que es común encontrar preguntas sobre VIH, como el conocimiento sobre los medios de transmisión, pero también sobre fecundidad, maternidad, acceso a servicios de salud, nupcialidad, entre muchos otros temas (cf. USAID, 2015).

Por otra parte, a medida que aumenta el número de ítems en una escala también aumenta el número de parámetros que deben estimarse, lo que implica que el tamaño de muestra deberá también aumentar para que el modelo converja y se puedan estimar estos parámetros con un menor error estándar (cf. De Ayala, 2006).

En principio, la meta es aplicar una escala compuesta por pocos ítems pero que aún brinde una medición precisa (Carey y Schroder, 2002). En aquellas escalas cortas, sin embargo, es necesaria la inclusión de ítems con características psicométricas apropiadas y analizar el principio de invarianza en la medición, puesto que la diferencia de un punto podría significar un falso positivo (González-Betanzos, Rivera-Heredia y Padrós-Blásquez, 2015).

Como parte de los indicadores que se presentan en los informes mundiales de avance en la lucha contra el sida<sup>3</sup> (ONUSIDA, 2011), así como para otros estudios en la materia (Arce, 2014; Espada et al, 2014; Acosta et al, 2015), se ha planteado un indicador que permita analizar la correcta identificación de las formas de prevenir la transmisión del virus y el rechazo de las principales ideas erróneas sobre su transmisión, generado precisamente a partir de una escala.

Para construir ese indicador, se plantea a los entrevistados un conjunto de afirmaciones que tratan sobre formas en las que se podría transmitir el VIH. Los entrevistados deben responder si cada afirmación es cierta o falsa, y eventualmente, pueden señalar no saber la respuesta. Cada una de estas opciones de respuesta se registra en una categoría distinta.

Algunas de esas ideas son efectivamente medios de transmisión del virus, mientras que otras son ideas incorrectas. Se determina que un entrevistado respondió correctamente a un ítem si la categoría elegida coincide con la afirmación planteada. Esta última etapa (de recodificación) se realiza durante el proceso de análisis de la información.

El objetivo último de este conjunto de afirmaciones es construir un indicador que mida el conocimiento sobre las formas de prevenir la transmisión del virus, a partir de la suma del número de afirmaciones que se respondieron correctamente.

---

<sup>3</sup> Se trata del indicador denominado "Porcentaje de jóvenes de 15 a 24 años de edad que identifica correctamente las formas de prevenir la transmisión sexual del VIH y rechaza las principales ideas erróneas sobre la transmisión del virus", presente en las guías metodológicas de ONUSIDA.

Comúnmente se ha definido que para el cálculo de este indicador todas las respuestas en las que el entrevistado indica que “no sabe” sean consideradas como respuestas incorrectas (ONUSIDA, 2011). La justificación al respecto es que el desconocimiento de los medios de transmisión del VIH pone en riesgo a los individuos de ser infectados por el virus, por lo que este riesgo subyacente se ve reflejado en las puntuaciones del indicador.

La asignación de dichas respuestas como incorrectas debería basarse en algún fundamento metodológico de carácter técnico, que determine la idoneidad de esta operacionalización. Sin embargo, en los distintos informes metodológicos publicados no se registran estos detalles técnicos (cf. ONUSIDA, 2011; 2014).

Dado un nivel de conocimiento sobre los medios de transmisión del VIH y suponiendo que las respuestas “no sabe” reflejan un conocimiento erróneo sobre este aspecto, las curvas de probabilidad que representan a las categorías “no sabe” e “incorrecto” deberían ser asintóticamente equivalentes. Es decir, los parámetros intercepto  $\gamma_{jk}$  y pendiente  $\alpha_{jk}$  estimados para estas categorías deberían ser iguales, por lo que podrían colapsarse en una sola categoría (Ayala, 2009a).

Con el fin de relacionar el conocimiento de los medios de transmisión del VIH con los patrones en las tasas de incidencia y prevalencia de la infección, la estimación del conocimiento suele ser desagregada por diferentes variables, entre las que destacan el sexo y la edad (ONUSIDA, 2011; 2014; Arce, 2014). Conocer esta información permite formular distintas estrategias de política pública, en miras a reducir el número de infecciones (ONUSIDA, 2014).

Uno de los principios de la TRI es la invarianza en la medida, que debe verificarse para garantizar inequidades en la estimación de la habilidad de los sujetos. Este principio se verifica mediante el funcionamiento Diferencial de los ítems (DIF). En el contexto de la medición del conocimiento en VIH, el análisis del DIF es importante porque las puntuaciones son desagregadas de acuerdo a las características de las personas: sexo, edad y zona de residencia, entre otras. En la práctica, las posibles divergencias entre grupos permiten definir programas

de política pública. Si las mediciones están sesgadas podrían tomarse decisiones que quizá no sean apropiadas.

Este trabajo se desarrolla con el fin de responder a dos grandes objetivos, en virtud de lo anteriormente expuesto. En primer lugar, se analiza la viabilidad de considerar las respuestas "no sabe" como respuestas incorrectas, para lo que se desarrollan diferentes modelos politómicos de respuesta al ítem que permitirían conocer la distribución de probabilidad de responder a cada categoría en función del parámetro de habilidad. Y en segundo lugar, se analiza el funcionamiento diferencial de estos ítems según sexo y edad, dos variables por las que frecuentemente se desagregan las puntuaciones de los sujetos.

## MARCO DE REFERENCIA

La teoría de respuesta al ítem (TRI) consiste en una serie de modelos matemáticos y estadísticos diseñados para describir la relación entre la habilidad de un sujeto en un determinado constructo y su desempeño en la serie de ítems que componen una prueba. Se trata de una herramienta que se emplea para (a) crear medidas que permitan definir ítems y escalas (b) analizar las propiedades de dichos ítems y escalas y (c) medir a sujetos en constructos o variables latentes. La estimación de puntuaciones en la TRI parte de que la probabilidad de responder correctamente a un ítem, o de dar la respuesta que implica presencia de la variable latente medida, es una función del atributo subyacente y que es representado por la letra griega *theta* ( $\theta$ ).

Los métodos de TRI muestran ventajas sobre la TCT. La TRI nace como un nuevo enfoque en la teoría de las pruebas, permitiendo superar algunas de las limitaciones de la TCT (Muñiz, 1997) ya que se centra más en las propiedades individuales de los ítems que en las propiedades globales del test (Jiménez y Montero, 2013). La TCT incorpora una serie de supuestos que no necesariamente pueden ponerse a prueba, mientras que la TRI ofrece la posibilidad de contar con modelos cuyos supuestos pueden ser juzgados empíricamente (Asún y Zúñiga, 2008).

En la TRI, la probabilidad que tiene un evaluado de contestar correctamente un ítem depende de su nivel de aptitud y de las características de los ítems. Los examinados ya no son calificados con base en el número de ítems que respondieron correctamente, como en la TCT, sino que se obtiene una estimación más precisa de su ubicación en el continuo de la habilidad, análoga a una estandarización de puntuaciones normales ( $z$ ) (Thomson, 2009).

La TRI provee un análisis más detallado del desempeño de un ítem a lo largo de la variable latente que se pretende medir (Hangman et al., 2009). Una de las limitaciones de la TCT es que las características de los sujetos y de los ítems no son discernibles, mientras que en la TRI las propiedades de los ítems y la habilidad de los sujetos se formulan mediante diferentes parámetros. Tanto las características de los ítems como de los individuos se obtienen en la

misma escala de medición, lo que facilita su interpretación, situación que no ocurre en la TCT, donde se tienen diferentes niveles de medición (Asún y Zúñiga, 2008).

De esa forma, la TRI permite obtener puntuaciones más refinadas del constructo que se analiza porque considera la dificultad y la discriminación de los ítems para obtener una puntuación. Además, las puntuaciones de los sujetos de una población pueden compararse incluso aunque hubiesen respondido a fórmulas diferentes de un test que mide el mismo rasgo (An y Yung, 2014).

La confiabilidad se define como una función condicional a las puntuaciones, ya que la precisión en la medición difiere a lo largo del continuo de valores del constructo analizado. Se puede disponer de medidas de error distintas para cada individuo o nivel de habilidad. La precisión suele representarse mediante curvas de información y error de medición. En contraposición, en la TCT el cálculo de la consistencia interna se realiza para el conjunto de individuos y el instrumento, obteniendo un índice global, que se supone como propiedad del instrumento completo (Asún y Zúñiga, 2008).

Ante la presencia de valores faltantes, en la TCT es necesario imputar las respuestas omitidas, tanto para diseñar el test como para obtener la puntuación de los sujetos. Los modelos de TRI, por otra parte, pueden lidiar con estos valores faltantes si se presentaron de forma aleatoria, mediante la estimación por máxima verosimilitud (An y Yung, 2014).

La TRI, sin embargo, presenta limitaciones por las cuáles incluso no ha tomado auge en ciertas disciplinas (Asún y Zúñiga, 2008). Su uso implica un aumento del tamaño de muestra, determinado por el número de ítems que componen la escala, el número de parámetros que deban estimarse y el tipo de modelo seleccionado. Aunado a esto, la comprensión de los modelos de TRI requiere cierto conocimiento matemático.

Parte de la metodología de la TRI se basa en dos fundamentos (a) la estimación de la curva característica de los ítems<sup>4</sup> y (b) la función de información. La curva característica de un ítem (CCI) es una función matemática que describe la relación entre la habilidad de los sujetos y la probabilidad de acertar a un ítem, o bien, de seleccionar cierta categoría. El propósito es obtener una curva característica para cada ítem con el fin de evaluar su calidad e identificar sus propiedades: dificultad y discriminación (Reise, Ainsworth y Haviland, 2005).

La función de información indica la precisión con que se realiza la estimación de los parámetros, para cada nivel de habilidad. Formalmente se estima como la inversa de la varianza del parámetro de habilidad. Los ítems miden con mayor precisión alrededor del valor de dificultad.

Se puede calcular tanto la función de información de un ítem como la función de información del test. Esta última se obtiene como una medida agregada de las funciones de información de todos los ítems que componen el test y permite describir cómo funciona la medición, como un todo, para cada nivel de habilidad.

En los modelos de TRI se establecen tres supuestos básicos. En primer lugar, se asume la existencia de una variable latente que puede ser explicada por las respuestas brindadas por los examinados y que está siendo medida a través de los ítems y la escala propuesta. El segundo supuesto está ligado al primero, pues establece que la escala es unidimensional, es decir, mide una sola variable latente. El tercero de los supuestos se refiere a la invarianza en la medición.

El principio de invarianza supone dos aspectos. En primer lugar, se establece que las propiedades del test o de los ítems son independientes de las características de los sujetos, salvo por el atributo específico que se está midiendo a través del test. Esto facilita tanto la equiparación horizontal como la equiparación vertical de los puntajes. En segundo lugar, se establece lo que se define como independencia local que asume que la respuesta a un determinado ítem es independiente a la respuesta dada en otro ítem. La probabilidad de

---

<sup>4</sup> En modelos politómicos se habla de Curvas Características de Categoría (CCC).

responder correctamente a un conjunto de ítems es el producto de las probabilidades de contestar correctamente a cada ítem por separado. A raíz de este principio, en la TRI no se definen ítems encadenados.

Los ítems que componen una escala pueden presentarse en múltiples formatos de respuesta. Los ítems de respuesta dicotómica son aquellos en los que solo se presentan dos alternativas de respuesta, mientras que, en los formatos politómicos los examinados pueden elegir una, y solo una, de tres o más alternativas de respuesta.

Los ítems de respuesta dicotómica son utilizados principalmente en la medición de rendimiento o habilidades pues lo que interesa es distinguir la capacidad de las personas de elegir la respuesta correcta. En la práctica, cuando hubiese más de una respuesta incorrecta, todas estas suelen ser tratadas como respuestas incorrectas (Asún y Zúñiga, 2008). Los formatos politómicos tienden a ser utilizados en escalas de personalidad o de actitudes ya que no se supone que existan respuestas correctas o incorrectas, sino que las alternativas de respuesta representan la intensidad que se posee de la característica o constructo, para tener una alta probabilidad de responder a cada alternativa (Asún y Zúñiga, 2008).

A su vez, los ítems politómicos pueden clasificarse como de respuesta ordinal o nominal según el ordenamiento natural que pueda darse a las categorías de respuesta. En los ítems de respuesta ordinal las alternativas pueden ordenarse de menor a mayor, según la intensidad del rasgo que sugieran. En los ítems nominales no existe un orden en las categorías que establezca que alguna opción es más o menos relevante que las otras.

El uso de modelos politómicos, es preferible sobre los modelos dicotómicos, porque un rango más amplio de respuestas permite obtener mayor información acerca del funcionamiento del instrumento, principalmente entre los sujetos con menores niveles de habilidad (De Ayala, 2009). En este tipo de modelos se especifica una función para cada una de las categorías, denominada Curva Característica de la Categoría (CCC) que estima la probabilidad de responder a la categoría  $i$ -ésima del ítem  $j$ -ésimo en función del parámetro de habilidad.

La selección de un modelo de TRI, entre la amplia gama de modelos disponibles, está dada por diferentes aspectos: codificación (dicotómicos o politómicos) y nivel de medición de los ítems (nominales u ordinales), bondad de ajuste del modelo e interpretación que quiera darse a los resultados. Para cada situación alguno de los modelos se ajustará mejor que otros, de modo que la selección de uno u otro dependerá del caso que se desee tratar (Muñiz, 1997).

A continuación se presentan las principales características de los modelos que se proponen en este trabajo y se describen metodológicamente los índices de ajuste empleados en la selección de modelos politómicos de respuesta al ítem. Se describe el concepto de funcionamiento diferencial de los ítems y el método de regresión logística empleado para su detección cuando los ítems tienen más de dos categorías de respuesta.

## **1. Modelos politómicos de respuesta al ítem**

Los ítems de carácter politómico son aquellos que poseen dos o más alternativas de respuesta. Según su naturaleza se les puede clasificar como nominales u ordinales. Las categorías de los primeros no tienen ningún orden en particular. Mientras que, en los ítems de naturaleza ordinal algunas respuestas indican la presencia de más (o menos) del rasgo que se está midiendo, en comparación con las alternativas restantes.

Los modelos politómicos se desarrollan en dos líneas. Existe una división entre aquellos derivados del modelo logístico de un parámetro, también denominado modelo de Rasch, y los que lo son del modelo logístico de dos parámetros (2P). La diferencia entre ellos radica en que los modelos desarrollados en esta segunda línea admiten que los distintos ítems de una escala tengan diferentes parámetros de discriminación.

A continuación se describen solamente los modelos que se desarrollarán en este trabajo: el *Modelo Politómico de Respuesta Nominal* (MPRN), el *Modelo de Crédito Parcial* (MCP) y el

*Modelo de Crédito Parcial Generalizado* (MCPG). Los últimos dos son modelos de carácter ordinal. El MCPG es una extensión del modelo 2P mientras que el MCP lo es del modelo de Rasch.

Existen otros modelos politómicos de uso común como el modelo de Escala de Clasificación (*Rating Scale*; Anderson, 1977 y Andrich, 1978) y el Modelo de Respuesta Graduada (MRG; Samejima, 1969). El primero de estos modelos fue propuesto inicialmente para analizar ítems cuyas respuestas están dadas en una escala tipo Likert pero también se le emplea para datos tipo escala (De Ayala, 2009). Se desarrolló como una extensión del modelo de Rasch para datos politómicos ordinales. Asume que las categorías de respuesta (ordenadas) son equidistantes, separadas una de la otra por una serie de límites. Supone que todos los ítems tienen el mismo número de categorías. Puede considerársele una variación del MCP y un caso particular del MPRN.

El MRG se emplea tanto para ítems con alternativas de respuesta tipo Likert como para ítems con un esquema de clasificación de crédito parcial. Se estima a partir de una serie de categorías ordenadas que provienen de agrupaciones bifurcadas (0 vs. 1, 2; 0,1 vs. 2) por lo que se torna en un modelo de comparaciones acumulativas cuya interpretación se expresa como la probabilidad de que un examinado responda a cierta categoría o a las categorías superiores. Se le considera el modelo sucesivo al modelo 2PL (De Ayala, 2009 y Naumenko, 2014).

### **1.1. Modelos politómicos de respuesta nominal (MPRN)**

El modelo politómico de respuesta nominal (MPRN) fue diseñado con el objetivo de analizar ítems con más de dos categorías de respuesta mutuamente excluyentes, pero en los que dichas categorías no poseen un orden preestablecido. Como consecuencia no es posible suponer de antemano cómo será la relación, directa o inversa, entre las categorías observadas y la magnitud de la habilidad de los sujetos ( $\theta$ ) (De Ayala, 2009).

Este tipo de ítems pueden surgir de un gran número de contextos. El ítem en el recuadro siguiente se presenta a manera de ejemplo. Fue tomado de los ejercicios de práctica para los estudiantes que aplicarán a la Prueba de Aptitud Académica para ingresar a la Universidad de Costa Rica (IIP-UCR, 2017).

*“La paz es tan necesaria como el pan”*

De acuerdo con el texto anterior, sin la paz, el pan es:

- a. dulce
- b. fresco
- c. blando
- d. amargo
- e. insuficiente

Como puede apreciarse, el enfoque tradicional para estimar la habilidad de los sujetos, es decir, generar un puntaje a partir de la suma de códigos asignados a las categorías, perdería sentido cuando se analizan ítems cuyas categorías son nominales (De Ayala, 2009). Las categorías no pueden ordenarse tratando de identificar con certeza cuál de ellas indica mayor presencia de la variable latente que se analiza.

El MRPN fue propuesto por Bock (1972) y se basan en una función logística multinomial para modelar la probabilidad de seleccionar cada una de las alternativas de respuesta, condicional a la estimación de la variable latente o habilidad ( $\theta$ ).

Con este tipo de modelos se obtiene una puntuación para cada sujeto que está basada directamente en las categorías nominales seleccionadas. La probabilidad de que una persona, con un nivel de habilidad de  $\theta$ , responda a la categoría  $k$  para el ítem  $j$  está dada por,

$$P(x = k|\theta, \alpha, \gamma) = \frac{e^{\gamma_{jk} + \alpha_{jk}\theta}}{\sum_{h=1}^{m_j} e^{\gamma_{jh} + \alpha_{jh}\theta}} \quad (1)$$

Donde  $\alpha_{jk}$  y  $\gamma_{jk}$  corresponden a la pendiente y el parámetro de intersección, respectivamente, de la función de respuesta asociada con la categoría  $k$ -ésima del ítem  $j$ -ésimo. El término  $m_j$  representa el número de categorías del ítem  $j$ -ésimo. Para este modelo se estiman un total de  $2j(m - 1)$  parámetros.

El término  $\gamma_{jk}$  refleja la propensión del individuo a responder a la categoría  $k$  mientras que el término  $\alpha_{jk}$  refleja, en parte, la capacidad discriminatoria de la categoría  $k$ -ésima (De Ayala, 2009). Dado que los términos de dificultad  $\gamma_{jk}$  están asociados a la frecuencia de una categoría, aquella que obtuvo el mayor número de respuestas suele presentar el mayor valor positivo de  $\gamma$ , mientras que la categoría con la menor frecuencia está ligado al valor más negativo.

Para los modelos de tipo nominal se tiene una Curva Característica para cada categoría. Una de ellas tiene la máxima pendiente positiva, es decir, es monotónicamente creciente que generalmente representa la categoría correcta o con la frecuencia más alta. Una segunda curva posee la mayor pendiente negativa (monotónicamente decreciente). Las categorías restantes generalmente, aunque no en todas las ocasiones, presentan una curva unimodal.

El MPRN puede emplearse para estimar tanto ítems politómicos como ítems dicotómicos, guardando equivalencia con el modelo 2PL. Los modelos de Crédito Parcial, Crédito Parcial Generalizado y el modelo de Escalamiento de Clasificación son casos especiales del MPRN. Al ser un modelo más general, puede emplearse incluso en ítems de carácter ordinal.

## **1.2. Modelos politómicos de respuesta ordinal**

En los modelos politómicos de respuesta ordinal las categorías de los ítems reciben puntuaciones que denotan una mayor presencia del rasgo que está siendo analizado, Como el que se muestra a manera de ejemplo en el siguiente recuadro.

¿Se siente usted satisfecho con el servicio de la compañía?

1. Nada satisfecho
2. Poco satisfecho
3. Ni insatisfecho ni satisfecho
4. Bastante satisfecho
5. Muy satisfecho

En esta sección se describen los modelos politómicos ordinales que se aplicarán a la escala de conocimiento sobre los medios de transmisión del VIH. En el primer apartado se describe el Modelo de Crédito Parcial y en el segundo se analiza el Modelo de Crédito Parcial Generalizado.

### **1.2.1. Modelo de crédito parcial (MCP)**

El Modelo de Crédito Parcial fue propuesto por Masters (1982). Es una extensión del modelo logístico de un parámetro de Rasch. Al estar en la línea de los modelos de Rasch no admite variaciones en la pendiente de los ítems.

El principio en el que se basa este modelo es la asignación de puntajes con alguna ordinalidad dando "crédito parcial" a respuestas que denotan algún avance en el proceso o el número de pasos completos en un problema, sin que necesariamente se haya llegado a la solución completa o correcta. La puntuación máxima solo se asigna a la categoría que muestra la respuesta correcta o que indique que todos los pasos del problema fueron resueltos adecuadamente.

Las características de este modelo han generado que su uso sea frecuente en test de rendimiento educativo (Martínez, Hernández y Hernández, 2006).

Es posible aplicar el modelo de Rasch a datos politómicos, partiendo del análisis de categorías adyacentes. Suponiendo que se tiene un ítem con tres puntuaciones de crédito parcial: 0,1 y 2, la comparación entre categorías adyacentes está dada por los pares  $x_j=\{0,1\}$  y  $x_j=\{1,2\}$ . Cada par de comparaciones posee lo que se denomina punto de transición.

El primer punto de transición  $\delta_{j1}$  refleja el cambio de obtener ningún crédito 0 a obtener el crédito parcial 1. El siguiente punto  $\delta_{j2}$  refleja el cambio del crédito parcial 1 a la puntuación completa 2. El parámetro  $\delta_{jh}$  refleja la dificultad relativa de seleccionar la categoría  $h$ , en lugar de la categoría  $(h-1)$ .

En ese contexto, la probabilidad de que un examinado obtenga una puntuación de  $x_j$  en el ítem  $j$ , condicional a su nivel de habilidad  $\theta$  está dada por la siguiente fórmula,

$$P(x_j|\theta, \delta_{jh}) = \frac{\exp[\sum_{h=0}^{x_j} (\theta - \delta_{jh})]}{\sum_{k=0}^{m_j} \exp[\sum_{h=0}^k (\theta - \delta_{jh})]} \quad (2)$$

donde el punto de transición se denota por la letra  $h$  y su localización en el continuo de  $\theta$  para el ítem  $j$  se representa como  $\delta_{jh}$ . El término  $m_j$  representa el número de categorías que posee un ítem y puede ser variable de uno a otro. En este modelo se estiman un total de  $(j * 2)+1$  parámetros.

Al igual que en el MPRN, la probabilidad de obtener cierta puntuación puede representarse gráficamente mediante CCC. Existirá para todo ítem, una curva monótonicamente creciente que se asocia a un valor positivo de  $\alpha$  y otra curva monótonicamente decreciente, con un valor de  $\alpha$  negativo y que está asociada a la categoría codificada como 0.

Gráficamente, el parámetro de transición representa el punto donde existe igual probabilidad de seleccionar alguna de dos categorías adyacentes.

### 1.2.2. Modelo de crédito parcial generalizado (MCPG)

El modelo de crédito parcial generalizado fue propuesto por Muraki (1992, 1993). Es una extensión del modelo de Crédito Parcial descrito en la sección anterior, pero permite que los ítems de un test difieran en su pendiente.

La probabilidad de que un entrevistado elija alguna de las categorías  $x_{jk}$  para el ítem  $j$ -ésimo se estima como,

$$P(x_{jk}|\theta, \alpha_j, \delta_{jk}) = \frac{\exp\left[\sum_{h=1}^{k_j} \alpha_j(\theta - \delta_{jh})\right]}{\sum_{c=1}^{m_j} \exp\left[\sum_{h=0}^c \alpha_j(\theta - \delta_{jh})\right]} \quad (3)$$

donde  $\theta$  representa la habilidad del sujeto,  $\alpha_j$  es la discriminación del ítem,  $\delta_{jh}$  es la localización del parámetro de transición entre las categorías  $h$  y  $(h-1)$ ,  $m_j$  representa el número de categorías y  $k=\{1,2,\dots, m_j\}$ . En cada modelo deben estimarse un total de  $j[(m_j-1)+1]$  parámetros.

El parámetro de discriminación  $\alpha_j$  indica el grado en el que las categorías varían entre ítems a medida que cambia  $\theta$ . El parámetro de dificultad puede interpretarse, de acuerdo con Muraki (1992), como la dificultad de seleccionar una categoría en particular.

## 2. Métodos y criterios para la selección de modelos

Esta sección describe algunos elementos metodológicos de análisis que deben desarrollarse vinculados a la estimación de los modelos de TRI. Uno de los supuestos de los modelos de TRI es que el conjunto de ítems permiten medir un único constructo o variable latente, lo que se conoce como unidimensionalidad, por lo que en un primer apartado se describen las principales características del análisis factorial confirmatorio.

Los dos apartados siguientes describen técnicas de post-estimación. Se describen los índices para la selección de modelos que serán utilizados tras ajustar los tres modelos en estudio, con el objetivo de identificar aquel que brinde el mejor ajuste. El último apartado describe el análisis del funcionamiento diferencial de los ítems (DIF), que permite verificar el principio de invarianza en la medida.

## **2.1. Análisis factorial confirmatorio (AFC)**

El análisis factorial, tanto exploratorio (AFE) como confirmatorio (AFC), es una técnica multivariada que permite explorar la dimensionalidad subyacente en los datos (Martínez, Hernández y Hernández, 2014).

Antes de realizar un análisis psicométrico con la TCT o la TRI es necesario verificar que el instrumento mide fundamentalmente solo un rasgo o constructo (unidimensionalidad), para lo que se emplea el análisis factorial. Este es un supuesto que debe cumplirse para que la aplicación de los modelos TRI brinde una medición válida (Jiménez y Montero, 2013).

La diferencia entre el AFE y el AFC radica principalmente en que, antes de realizar los cálculos y los procedimientos, en el AFC debe establecerse de previo la estructura de las variables latentes y su relación con las variables observadas. Antes del desarrollo del análisis, se tiene una hipótesis de cuáles variables componen cada factor, así como la existencia o no de una relación entre los factores (en otro contexto, puede sugerirse la presencia de múltiples factores y correlación entre ellos, en el análisis de la TRI se espera encontrar un único factor).

El AFC forma parte de los modelos de ecuaciones estructurales, dentro de los cuáles el AFC especifica el llamado modelo de medida, que explica cómo variables observadas o medidas reflejan variables latentes (Martínez, Hernández y Hernández, 2014). En su forma más simple, un ítem se definiría como una variable observada y puede representarse como una función lineal de un factor, de la siguiente manera:

$$x = \Lambda_x \xi + \delta \quad (4)$$

donde  $x$  es un vector con las respuestas otorgadas por cada examinado a un ítem,  $\Lambda_x$  es una matriz de saturaciones entre las variables y los factores,  $\xi$  es un vector de variables latentes y  $\delta$  es un vector que representa un término de error.

El objetivo en el AFC es explicar la correlación o covarianza entre las variables observadas. La matriz teórica se puede denotar como:

$$\Sigma = \Lambda_x \Phi \Lambda_x' + \Theta_\delta \quad (5)$$

donde  $\Sigma$  es la matriz de varianzas-covarianzas del modelo teórico,  $\Lambda_x$  es una matriz de saturaciones entre las variables y los factores,  $\Phi$  es la matriz de covarianza entre los factores y  $\Theta_\delta$  es la matriz de residuos.

La estimación parte de la matriz de correlaciones entre las variables observadas,  $S$ , calculada a partir de los datos y será un estimador de  $\Sigma$ . Los parámetros del modelo  $\hat{\Lambda}_x$ ,  $\hat{\Phi}$  y  $\hat{\Theta}_\delta$  son desconocidos y deben estimarse a partir de los datos, generalmente mediante estimadores de máxima verosimilitud.

Comúnmente se emplean algunos de los índices de bondad de ajuste, tras la estimación del AFC, para evaluar la proximidad entre  $S$  y  $\Sigma$ . En este trabajo se contemplan cuatro de los más utilizados: índice comparativo de ajuste (CFI), índice de Tucker-Lewis (TLI), raíz cuadrada del error medio de aproximación (RMSEA) y la raíz cuadrada estandarizada residual (SRMR), todos denotados por sus siglas en inglés.

Los índices CFI y TLI se refieren a medidas de ajuste comparativo pues contrastan el modelo estimado a un modelo hipotético de línea de base, que es el modelo más simple que se le puede aplicar a los datos (Martínez, Hernández y Hernández, 2014). Tanto el CFI como el TLI reflejan la mejora en la bondad de ajuste con respecto al modelo de línea de base (Morgan, Hodger, Wells y Watkins, 2015).

Ambos índices se encuentran en el intervalo [0,1] de tal manera que los modelos próximos a 1 poseen mejor ajuste. Suele considerarse que los modelos cuyo índice es mayor a 0.95 presentan buen ajuste y se califica como de ajuste aceptable a aquellos con cifras superiores a 0.90 (Bentler, 1990 y Hooper, Coughlan y Muller, 2008).

Los dos índices restantes, RMSEA y SRMR, miden el ajuste absoluto de los datos al modelo planteado. El RMSEA indica cuán bien se ajusta el modelo, con parámetros desconocidos pero óptimos, a la matriz de covarianzas poblacional (Hooper, Coughlan y Muller, 2008). El RMSEA puede emplearse para contrastar la hipótesis de que la matriz de covarianza muestral se ajusta a la matriz de covarianza teórica. Este estadístico sigue una distribución  $\chi^2$  y suelen presentarse en conjunto con el RMSEA.

El SRMR es una estimación estandarizada de la raíz cuadrada de los residuos, obtenidos de la matriz de covarianzas muestral y la matriz de covarianzas teórica, que se obtendría con el modelo al que se supone se ajustan los datos.

Aunque no hay gran consenso, tanto para RMSEA y SRMR se considera un indicador de buen ajuste cuando su valor es mejor o igual a 0.05. Valores comprendidos entre 0.05 y 0.08 se consideran de ajuste aceptable (Hooper, Coughlan y Muller, 2008 y Martínez, Hernández y Hernández, 2014).

## **2.2. Criterios para la selección de modelos: índices de ajuste**

A continuación se describen tres índices comúnmente empleados en la selección de modelos politómicos de respuesta al ítem: el test de razón de verosimilitudes (McKinley y Mills, 1985), el Criterio de Información de Akaike (AIC; por sus siglas en inglés; Akaike, 1974) y el Criterio de Información de Bayes (BIC; Schwarz, 1978). La selección de alguno de los modelos propuestos en este trabajo se guiará por la información brindada por estos tres índices.

El test de razón de verosimilitud se emplea con modelos anidados, a lo que debe considerarse que los MPRN, MCP y MGCP están jerárquicamente relacionados. Los índices AIC y BIC pueden emplearse tanto con modelos anidados como no anidados, aunque con estos últimos no es posible realizar pruebas de significancia. En su lugar, los índices AIC y BIC son medidas basadas en la información e incluyen una penalización por la complejidad del modelo.

El cálculo de estos tres índices está basado en la devianza:  $G^2 = -2 \cdot \ln(\text{verosimilitud})$  y son apropiados cuando los parámetros del modelo se obtienen por máxima verosimilitud. En el AIC la complejidad del modelo se penaliza por el número de parámetros mientras que el BIC penaliza a los modelos tanto por su número de parámetros como por el tamaño de muestra.

#### a. Test de razón de verosimilitud ( $G^2$ )

El test de razón de verosimilitud,  $\Delta G^2$ , expresa cuán probable es que los datos sean estimados por un modelo en relación a otro. La razón de verosimilitudes, o equivalentemente su operacionalización logarítmica, puede emplearse para determinar si dos modelos jerárquicamente relacionados difieren significativamente entre ellos.

El cambio en  $G^2$  permite establecer una prueba de significancia para determinar si se rechaza o no la hipótesis nula. Ésta establece que los parámetros adicionales de un modelo completo, pero no de un modelo reducido, son iguales a cero.

El término '*modelo completo*' hace referencia al modelo más complejo y con un mayor número de parámetros, mientras que el '*modelo reducido*' es aquel más simple, con menos parámetros. Este segundo modelo puede verse como un sub-modelo del modelo completo o anidado en él.

El test de razón de verosimilitudes se basa en el estadístico:

$$\Delta G^2 = -2 [\ln(L_R) - \ln(L_F)] = G_R^2 - G_F^2 \quad (6)$$

Donde  $L_R$  hace referencia a la verosimilitud del modelo reducido y  $L_F$  corresponde a la verosimilitud del modelo completo.

El estadístico  $\Delta G^2$  sigue una distribución Chi-cuadrado cuando el tamaño de muestra es grande y el modelo completo se ajusta a los datos. Los grados de libertad para evaluar el test de significancia representan la diferencia entre el número de parámetros de los modelos. Cuando no se encuentra una diferencia estadísticamente significativa, se tiene indicación de que no es necesario emplear un modelo más complejo.

El test de razón de verosimilitudes tiende a seleccionar modelos con más parámetros, en comparación con modelos de menos parámetros, pero se le considera de gran utilidad al contar con una prueba de significancia (Kang, Cohen y Sung, 2005).

#### **b. Criterio de Información de Akaike (AIC)**

El Criterio de Información de Akaike (AIC) incorpora un componente que penaliza a los modelos con el mayor número de parámetros. El índice se estima como,

$$AIC = G^2 + 2p \quad (7)$$

donde  $p$  es el número de parámetros. Si dos modelos se ajustan a partir del mismo conjunto de datos, se considera que aquel con el menor AIC brinda el mejor ajuste.

#### **c. Criterio Bayesiano de información (BIC)**

El Criterio Bayesiano de Información (BIC, por sus siglas en inglés), aplica una penalización a los modelos con mayor número de parámetros y aplica la función logaritmo al número de casos, para penalizar también a los modelos de acuerdo al tamaño de la muestra. El BIC se define como:

$$BIC = G^2 + p \cdot \log(N) \quad (8)$$

donde  $p$  representa el número de parámetros y  $N$  el tamaño de muestra. Se considera que el modelo con el menor BIC, es el que se ajusta mejor a los datos.

El BIC tiende a seleccionar modelos más simples (con menos parámetros) que el índice AIC, precisamente porque da mayor penalización a los modelos con mayor número de parámetros, pero se considera más confiable para la selección del modelo correcto al ajustar modelos politómicos de respuesta al ítem (Kang, Cohen y Sung, 2009). Se considera que los índices basados en la comparación de modelos anidados suelen ser sensibles al tamaño de muestra (Oksana, 2014).

### 2.3. Funcionamiento diferencial de los ítems (DIF)

De acuerdo con Martínez, Hernández y Hernández (2014), el objetivo de la evaluación mediante test es tratar a las personas imparcialmente, independientemente de sus características personales tales como el sexo, la edad, raza o discapacidad. La imparcialidad requiere que las características personales de los sujetos sean irrelevantes para la medición del constructo y que no tengan efectos apreciables en los resultados del test o en su interpretación.

Un ítem puede considerarse sesgado si un individuo tiene distinta probabilidad de responder correctamente, dependiendo de sus características sociodemográficas o de otra índole, aun cuando posea la misma habilidad  $\theta$  que otros sujetos, pero que pertenecen a otra subpoblación (González-Betanzos, Rivera-Heredia y Padrós-Blázquez, 2015). En resumen, se trata de que la probabilidad de seleccionar cierta categoría, condicional a  $\theta$ , difiere para los grupos:  $P(\theta)_{iF} \neq P(\theta)_{iR}$ .

Los métodos de *análisis del funcionamiento diferencial del ítem* pretenden detectar la presencia de este sesgo, comparando como funciona un ítem en grupos distintos, de ahí el uso del término. El fin es determinar si los ítems se comportan de la misma forma (estadísticamente)

para todos los individuos de una población, independientemente de la pertenencia a subgrupos.

Generalmente en el análisis del DIF se analizan las variaciones entre dos subpoblaciones. El grupo de interés, que posee una característica que podría generar variaciones en el desempeño del test, se llamado *grupo focal* (F) y es comparado contra un *grupo de referencia* (R).

Cuando se tienen más de dos subpoblaciones, los métodos de detección del DIF son una extensión de los empleados al analizar a dos grupos. Las comparaciones se realizan, del mismo modo, entre dos grupos a la vez.

En los modelos dicotómicos de TRI, algunos métodos de detección del DIF se basan en la comparación de las curvas características del ítem (CCI) de los grupos focal y de referencia. Las diferencias en estas curvas indican que para un mismo nivel de habilidad, los sujetos no tienen la misma probabilidad de responder correctamente al ítem y pueden deberse a diferencias en el parámetro de dificultad ( $b$ ), de discriminación ( $a$ ) o de ambos.

Cuando el parámetro de dificultad ( $b$ ) varía entre los grupos se presenta lo que se denomina *DIF uniforme*. Cuando solo el parámetro de discriminación ( $a$ ) varía entre los grupos se presenta *DIF no uniforme*. Cada uno de estos tipos de DIF puede visualizarse gráficamente en las CCI, como se muestra en las Figuras 1a y 1b, respectivamente.

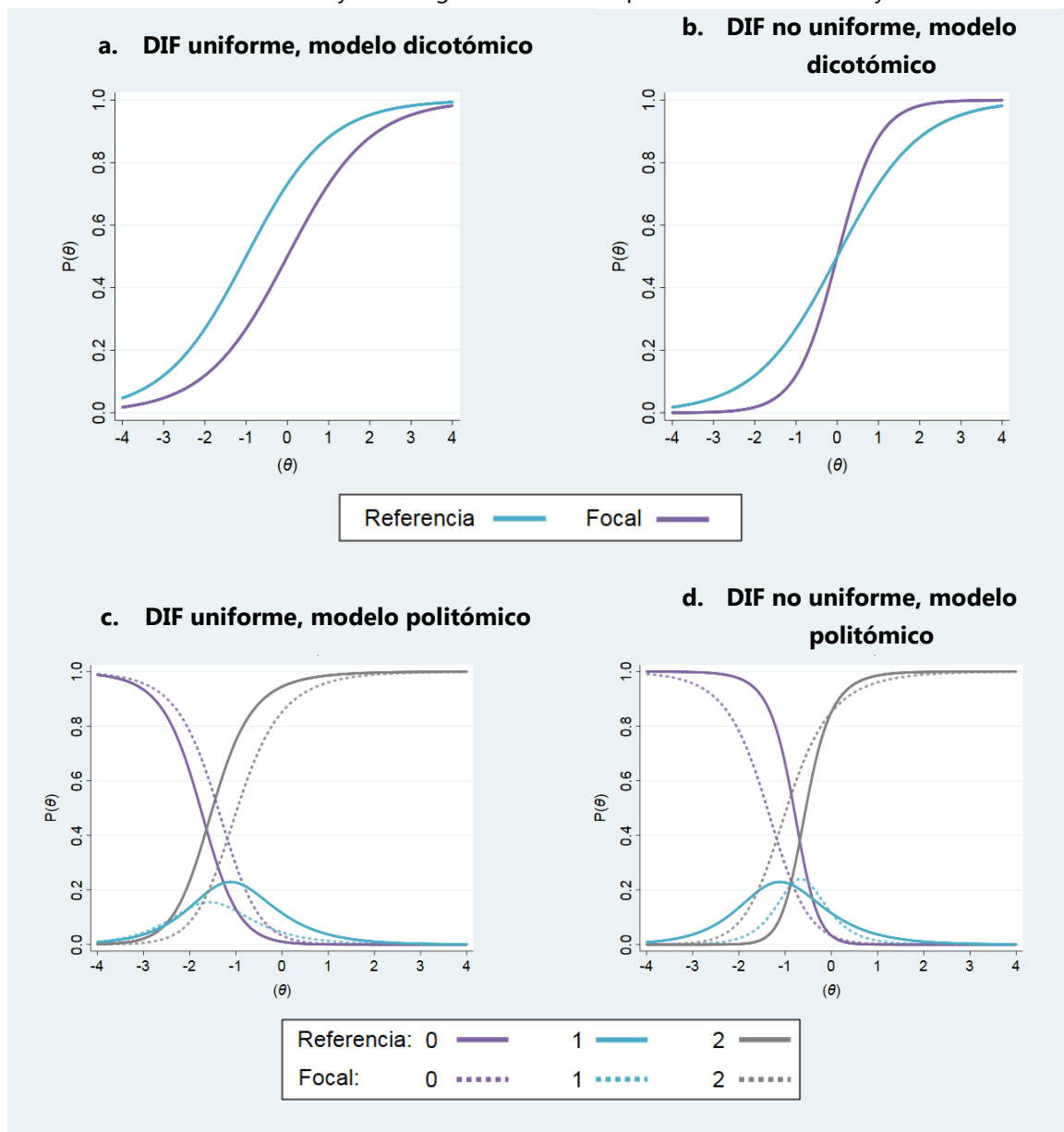
En el DIF uniforme los sujetos que pertenecen a uno de los grupos tienen mayor probabilidad de dar una respuesta correcta en todos los niveles de habilidad, generalmente se trata del grupo de referencia. Nótese que las CCI son diferentes pero no se cruzan, lo que sugiere que la dificultad del ítem es diferente entre los grupos, pero no así la discriminación pues las dos curvas poseen en teoría la misma pendiente.

El DIF no uniforme se presenta cuando la probabilidad de brindar una respuesta correcta es mayor en algunos niveles de  $\theta$ , mientras que en otros niveles esa probabilidad es mayor para el

grupo focal (Figura 1b). En las CCI la presencia de DIF no uniforme se visualiza cuando las dos curvas se intersecan.

**Figura 1**

Curvas características del ítem y de categorías cuando se presenta DIF uniforme y no uniforme



Cuando los ítems son politómicos, la naturaleza del DIF puede volverse más compleja (French y Miller, 1996). A diferencia de los ítems dicotómicos, el DIF puede generarse no solo a causa de

diferencias entre grupos, sino que también puede relacionarse con las categorías de los ítems. El DIF puede presentarse para un conjunto de alternativas de respuesta o entre todas ellas.

Las CCI en los modelos dicotómicos se convierten en curvas características de categorías (CCC) en los modelos politómicos y habrá por ende una curva para cada categoría. En este caso no se grafica la probabilidad de que un individuo responda correctamente a un ítem, sino que cada CCC muestra la probabilidad de que el examinado elija esa categoría de respuesta, condicional a  $\theta$ . La representación gráfica del DIF en modelos politómicos parte del mismo principio de los modelos dicotómicos (Figuras 1c y 1d).

Se documentan en la literatura diversas técnicas para el análisis del DIF, tanto para ítems dicotómicos como politómicos (cf. De Ayala, 2009; Kim, Cohen, Alagoz y Kim, 2007; French y Miller, 1996; Martínez, 1996 y Martínez, Hernández y Hernández, 2014). Entre los más utilizados se encuentran los basados en tablas de contingencia, el test de Mantel-Haenszel (MH) y su extensión para modelos no ordinales (Test Generalizado de Mantel-Haenszel) y los basados en la comparación de modelos de regresión logística.

El test de MH se ha considerado sensible a la dirección del DIF. El DIF no uniforme podría no ser detectado. Si la dirección del DIF cambia en medio de la escala de habilidad, las diferencias entre grupos podrían ser canceladas o contrarrestadas (De Ayala, 2009 y French y Miller, 1996).

El método de regresión logística es uno de los métodos más utilizados. Puede modelar tanto el DIF uniforme como no uniforme con una misma ecuación y pueden realizarse pruebas de significancia para los coeficientes. A este método se le considera tan potente como el test de MH para la detección del DIF uniforme en modelos dicotómicos y más potente al detectar DIF no uniforme (Swaminathan y Rogers, 1990).

Si el método de detección del DIF se basa en procedimientos paramétricos, como los que parten de la comparación de modelos logísticos, el DIF detectado puede deberse a la incorrecta especificación del modelo (Martínez, Hernández y Hernández, 2014). Del mismo

modo, presentan como limitación la necesidad de un mayor tamaño de muestra (Kim, Cohen, Alagoz y Kim, 2007). Hidalgo, López-Martínez, Gómez-Benito y Guilera (2016) identificaron que en test con DIF, las tasas de Error Tipo I dependieron del tamaño del test, de la muestra y el número de categorías del ítem.

Los procedimientos no paramétricos pueden dar lugar a resultados inestables cuando se desarrollan con muestras pequeñas, debido al efecto del error de muestreo (Martínez, Hernández y Hernández, 2016).

Los métodos basados en regresión logística para el análisis del DIF se basan en el desarrollo de una serie de modelos de este tipo. Para cada ítem se estiman tres modelos, comparando dos de ellos a la vez. En cada caso, la respuesta brindada por cada entrevistado se toma como variable dependiente, mientras que el nivel de habilidad, la variable de agrupación y la interacción entre ambas actúan como variables independientes.

Cuando los procedimientos de regresión logística se desarrollan para el análisis de ítems politómicos, aún requieren una variable dependiente dicotómica. Las categorías politómicas deben agruparse en un conjunto de variables dicotómicas que se analizan mediante modelos logísticos, comparando pares de categorías. Para ítems ordinales pueden plantearse esquemas de codificación basados en modelos de razón continua, modelos logísticos acumulativos o modelos logísticos de categorías adyacentes (Agresti, 2011 y French y Miller, 1996).

Para cada ítem, se consideran los siguientes modelos, cada uno es una variación del anterior, por lo que se les considera anidados. Un modelo se ajusta asumiendo la igualdad de los parámetros entre los grupos (modelo compacto) y se compara contra otro modelo donde se supone que los parámetros de los ítems son diferentes (modelo aumentado):

$$z = \tau_0 + \tau_1\Lambda + \tau_2\Gamma + \tau_3(\Lambda * \Gamma) \quad (9)$$

$$z = \tau_0 + \tau_1\Lambda + \tau_2\Gamma \quad (10)$$

$$z = \tau_0 + \tau_1\Lambda \quad (11)$$

donde  $\Lambda$  representa la estimación del individuo para la variable latente que se está analizando (puede tratarse de  $\theta$  o de la suma de las puntuaciones de los ítems).  $\Gamma$  es una variable categórica que indica la pertenencia de cada individuo a uno u otro grupo analizado (miembros del grupo focal o del grupo de referencia). El término  $(\Lambda * \Gamma)$  es la interacción entre el nivel de habilidad de los sujetos y el grupo de pertenencia.

El término  $\tau_1$  indica la relación entre un ítem y la localización de una persona en la escala que mide la variable latente; siendo normal que resulte estadísticamente significativo ya que se espera que los sujetos con distintos niveles de habilidad tiendan a seleccionar distintas alternativas de respuesta, Martínez, Hernández y Hernández, 2016).  $\tau_2$  corresponde a la diferencia promedio entre los subgrupos (*i.e.*:  $\tau_2 = \beta_{0F} - \beta_{0R}$ ) y  $\tau_3$  refleja la interacción del grupo y la localización de la persona (*i.e.*:  $\tau_3 = \beta_{1F} - \beta_{1R}$ ).

Los parámetros de los modelos de regresión logística se estiman por procedimientos de Máxima Verosimilitud lo que permite realizar pruebas de razón de verosimilitudes para determinar cuál modelo explica mejor la estructura de los datos. Las pruebas se realizan entre los modelos (9) y (10) para análisis el DIF no uniforme y (10) y (11) para determinar la presencia de DIF uniforme.

De esa manera puede emplearse el estadístico  $\Delta G^2$ , descrito en la sección 2.2.a., para determinar cuál modelo presenta mejor ajuste:

$$\Delta G^2 = -2 [\ln(L_C) - \ln(L_A)] \quad (12)$$

Este estadístico posee una distribución  $\chi^2$  cuyos grados de libertad están dados por la diferencia en el número de parámetros de los modelos que están siendo comparados. Un resultado estadísticamente significativo en cada comparación brinda indicaciones de que no existe igualdad entre los parámetros. Se dice que no hay invarianza en la medida (González-Betanzos, Rivera-Heredía y Padrós-Blázquez, 2015).

A través  $\Delta G^2$  se verifica la hipótesis de que  $\tau_3 = 0$ . Si para esta prueba no se obtiene un resultado significativo, el término de interacción ( $\Lambda * \Gamma$ ) no es necesario y no se encuentra evidencia de la presencia de DIF no uniforme.

A continuación se comparan los modelos (10) y (11) para analizar la hipótesis  $\tau_2 = 0$ . Si no se encuentran resultados significativos, la variable que define los grupos puede ser omitida, puesto que el término de intercepción es igual en ambos grupos. Si los resultados fueran significativos, se encontraría evidencia de la presencia de DIF uniforme.

## **METODOLOGÍA**

### **1. Objetivos de estudio**

#### **1.1. Objetivo general**

Comparar modelos politómicos de respuesta al ítem para analizar escalas de tipo nominal basándose en una aplicación sobre conocimiento de los medios de transmisión y prevención del VIH.

#### **1.2. Objetivos específicos**

- Analizar el conjunto de ítems sobre conocimiento de los medios de transmisión del VIH mediante modelos politómicos de respuesta al ítem.
- Determinar cuál modelo de respuesta al ítem explica mejor la estructura de los datos.
- Verificar la invarianza de los parámetros de los ítems, por sexo y grupos de edad en el modelo seleccionado.

### **2. Fuentes de información**

Se emplearán los datos de la Encuesta Nacional de Salud Sexual y Reproductiva del año 2015 (ENSSR-15). La encuesta se aplicó a una muestra de todas las personas entre 15 y 64 años de edad, incluyendo tanto a hombres como mujeres, que residían en el territorio costarricense entre noviembre y diciembre del año 2015.

El diseño muestral de esta encuesta fue bietápico y estratificado. Se seleccionaron segmentos censales en una primera etapa, con probabilidad proporcional al tamaño, mientras que en la segunda etapa se seleccionaron viviendas, con igual probabilidad, dentro de cada segmento elegido previamente.

La mayoría de las entrevistas fueron registradas en dispositivos electrónicos (Tablets), con excepción de las realizadas en zonas que se consideraran inseguras, donde la información se completó en cuestionarios impresos.

La base de datos de esta encuesta es de acceso público y puede ser descargada desde la página web del Centro Centroamericano de Población de la Universidad de Costa Rica: [www.ccp.ucr.ac.cr](http://www.ccp.ucr.ac.cr). En este sitio también se puede encontrar otra documentación sobre la encuesta.

### **3. Población de estudio**

Se analizaron un total de 2933 personas participantes en la ENSSR-15. Se incluyeron tanto hombres como mujeres, en un rango etario de entre 15 y 64 años. La edad promedio en la muestra fue de 34.9 años ( $s=14.1$ ). El 52.7% de los entrevistados eran mujeres.

Previo a la presentación de la sección sobre conocimientos acerca del VIH, se planteó a los entrevistados la pregunta: *¿ha oído hablar alguna vez sobre el VIH (sida)?*. Una respuesta negativa a esta interrogante excluía a los sujetos de responder a la sección sobre conocimientos.

Además de esos casos, se ignoraron todos aquellos que se negaron a responder al menos uno de los ítems de la escala de conocimientos. La muestra que se analiza finalmente representa un 91.2% del total de la muestra obtenida originalmente en la encuesta (3216).

### **4. Variables de estudio**

El constructo que se analiza en este trabajo se denominará "conocimiento sobre la prevención y transmisión del VIH". Su medición se establece a partir de una serie de afirmaciones o ítems que permiten analizar el conocimiento sobre las formas de prevenir la transmisión del virus y el rechazo de las principales ideas erróneas sobre su transmisión (Figura 2).

**Figura 2**

Afirmaciones planteadas en la ENSSR-15 relacionadas con el conocimiento de los medios de transmisión del VIH

**1301.** A continuación le voy a leer una serie de afirmaciones. Quisiera que me indicara para cada una de ellas si en su opinión la frase es cierta o falsa.

AFIRMACIÓN	CIERTO	FALSO	NO SABE	NO RESPONDE	RESPUESTA CORRECTA
a. El VIH se puede transmitir al practicar el sexo oral sin protección	1	2	8	9	CIERTO
b. El condón debe colocarse antes de que se produzca la erección	1	2	8	9	FALSO
c. El VIH se puede transmitir al usar el mismo servicio sanitario de una persona infectada	1	2	8	9	FALSO
d. El sobre del condón debe abrirse por una esquina con los dedos	1	2	8	9	CIERTO
e. La única forma de conocer si una persona está infectada con el VIH (sida) es con un examen de sangre específico	1	2	8	9	CIERTO
f. El VIH puede transmitirse por medio de la picadura de un mosquito	1	2	8	9	FALSO
g. Se puede contraer el VIH(sida) por compartir alimentos con una persona infectada	1	2	8	9	FALSO
h. El VIH se puede transmitir por medio del sudor o de la saliva de una persona infectada	1	2	8	9	FALSO
i. El VIH se puede transmitir por medio de la tos y los estornudos de una persona infectada	1	2	8	9	FALSO
j. VIH son las siglas para Virus de Inmunodeficiencia Humana	1	2	8	9	CIERTO
k. El riesgo de transmisión del VIH se reduce si se tienen relaciones sexuales con una única pareja fiel y no infectada	1	2	8	9	CIERTO
l. El VIH puede transmitirse a través de la leche materna	1	2	8	9	CIERTO
m. Puede reducirse el riesgo de transmisión del VIH usando condón cada vez que se mantienen relaciones sexuales	1	2	8	9	CIERTO
n. La persona que tiene el VIH siempre muestra un aspecto enfermizo	1	2	8	9	FALSO

Al momento de la entrevista, se le planteaba al entrevistado cada una de las afirmaciones, a las que debía responder si la idea que se le mencionó era cierta o falsa. Se incorporó además la posibilidad de que el entrevistado respondiera 'no sé'. Esta opción, sin embargo, no era leída a los entrevistados, se generaba tras la respuesta espontánea del entrevistado a la idea planteada. La categoría 'no responde' registró a los entrevistados que se negaron a responder a un ítem en particular, o bien, a toda la escala.

La construcción de esta escala sobre conocimiento de los medios de prevención de VIH no se basó en fundamentos estadístico-matemáticos sino que responde al interés de conocer aspectos puntuales sobre el tema. Sin embargo incorpora algunos de los ítems comúnmente empleados en escalas que miden dicho constructo, incluyendo los ítems propuestos por ONUSIDA (2011, 2014) para desarrollar un indicador que analiza el conocimiento de los medios de prevención del VIH y la identificación de las principales ideas erróneas sobre la transmisión del virus<sup>5</sup>.

Las respuestas otorgadas por el entrevistado fueron recodificadas en las categorías: 0: No sabe, 1: Incorrecto y 2: Correcto, según el patrón que se muestra en la última columna de la Figura 2. En la etapa de análisis se consideraron solamente estas variables recodificadas.

Para poder aplicar los modelos TRI ordinales es necesario que las respuestas de los individuos estén ordenadas de acuerdo al nivel de conocimiento. Este ordenamiento se guía por los resultados del modelo politómico nominal.

Se analizan las variables sexo y edad en el análisis del funcionamiento diferencial de los ítems. En el caso de esta segunda variable, se analizan los grupos etarios: 1) 15 a 24 años, 2) 25 a 34 años y 3) 35 o más. Los primeros dos grupos corresponden a la población joven. En el primero es donde se registra un aumento significativo de las tasas de incidencia, caracterizado por el inicio de las relaciones sexuales entre la población más joven (Ministerio de Salud, 2016 y ONUSIDA, 2014). Al segundo grupo aún se le denomina población joven mientras que el

---

<sup>5</sup> Corresponden a los ítems f, h, k, m y n del listado que se presenta en la Figura 2.

tercero corresponde a la población adulta, donde las tasas de incidencia registradas muestran una disminución con respecto a los otros grupos etarios.

## **5. Análisis estadístico**

El primer paso en el análisis de la escala sobre conocimiento de VIH es verificar la unidimensionalidad de esta escala por medio del análisis factorial confirmatorio.

En segunda instancia se desarrollan los modelos politómicos de TRI. En el contexto del ejemplo que se desarrolla, las alternativas de respuesta corresponden a las categorías correcto, incorrecto y no sabe. El parámetro de habilidad ( $\theta$ ) o la variable latente que se estima, representa el conocimiento sobre los medios de transmisión del VIH y se estimó mediante el procedimiento bayesiano de esperanza a posteriori (EAP, o estimación media de Bayes). La calibración de los parámetros de los ítems se realizó mediante el método de máxima verosimilitud marginal.

Tras esa etapa se comparan los parámetros de dificultad y discriminación, la función de información para cada ítem y para del test de los distintos modelos. Se emplean medidas de bondad de ajuste para identificar el método que brinde una mejor descripción de los datos.

Finalmente se analiza la invarianza de la escala, mediante el funcionamiento diferencial de los ítems en grupos definidos según sexo y edad, pero solamente para el modelo seleccionado en la etapa anterior.

Para el análisis de datos se utilizó el paquete estadístico Stata MP versión 15.1 (Stata Corp. LP, 2017). Los modelos se ajustaron empleando el comando `irt`, especificando las opciones `nrm`, `pcm` y `gpcm` para obtener el MPRN, MCP y MCPG, respectivamente. La función de información del test y las curvas características de las categorías se obtuvieron con el comando `irtgraph`.

En este trabajo se analiza el funcionamiento diferencial de los ítems mediante regresión logística, empleando el paquete difd disponible para usuarios de Stata (Crane, Gibbons, Jolley y van Belle, 2005). Este paquete examina para cada ítem tres modelos de regresión logística ordinal como se especificó en las ecuaciones (9), (10) y (11). Para la detección del DIF uniforme, además del test de razón de verosimilitudes se puede obtener una medida relativa del cambio en el parámetro de habilidad, cuando la variable grupo se incorpora al modelo. Este índice se obtiene al contrastar el coeficiente de regresión de los modelos compacto y aumentado, planteados en las ecuaciones (10) y (11), mediante la siguiente fórmula:

$$\Delta\beta = \frac{\beta_1(2) - \beta_1(3)}{\beta_1(2)} \quad (13)$$

Si la diferencia relativa es grande, la pertenencia a un grupo interfiere con la relación entre la habilidad y la selección de las alternativas de respuesta. Cambios relativos superiores a 5% son indicadores de la presencia de DIF uniforme (Crane, Hart, Gibbons y Cook, 2005).

## RESULTADOS

Esta sección presenta los resultados obtenidos, que corresponden a cuatro líneas de análisis. La primera sección se presenta un análisis descriptivo de la población, las variables sociodemográficas y los ítems analizados. En una segunda fase se desarrolló el AFC como primer paso del análisis de la escala de conocimientos sobre los medios de transmisión del VIH y como requisito para la estimación de modelos TRI. La sección 3 de este capítulo muestra los resultados de la estimación de los tres modelos ajustados: MPRN, MCP y MCPG, presentando en cada caso la estimación de los parámetros de dificultad y discriminación, las CCC, la función de información y las puntuaciones obtenidas. El capítulo finaliza presentando los índices empleados en la selección de modelos y el análisis del DIF para el modelo elegido.

### 1. Descripción de la población y variables de estudio

#### 1.1. Características de la población de estudio

La distribución de los participantes de acuerdo con las principales variables demográficas se presenta en el Cuadro 1. La población de estudio está conformada por 2933 personas, de las cuales 1388 son hombres (50.2%) y 1562 son mujeres (49.8%).

Un 30.1% de la población está conformada por adolescentes y personas menores de 25 años. El resto de la población joven, entre los 25 y 34 años, representa el 23.8%. Los restantes grupos de edad tiene menor importancia relativa (Cuadro 1), principalmente el grupo de 55 a 64 años (12.6%).

La mayor parte de la muestra se concentró en la zona urbana del país (74.1%). En lo que respecta al nivel educativo, la mayoría de los entrevistados tenía algún grado de educación secundaria (47.3%) y una menor proporción había alcanzado los estudios superiores (18.2%).

**Cuadro 1**

Número y porcentaje de entrevistados, según características demográficas

Variable	N° de casos	Porcentaje
<b>Total</b>	<b>2933</b>	<b>100.0</b>
<b>Sexo</b>		
Hombre	1388	47.3
Mujer	1545	52.7
<b>Edad</b>		
15 a 24	882	30.1
25 a 34	698	23.8
35 a 44	559	19.1
45 a 54	425	14.5
55 a 64	369	12.6
<b>Zona de residencia</b>		
Urbana	2174	74.1
Rural	759	25.9
<b>Nivel educativo</b>		
Ninguno	89	3.0
Primaria	923	31.5
Secundaria	1386	47.3
Para/Universitaria	535	18.2

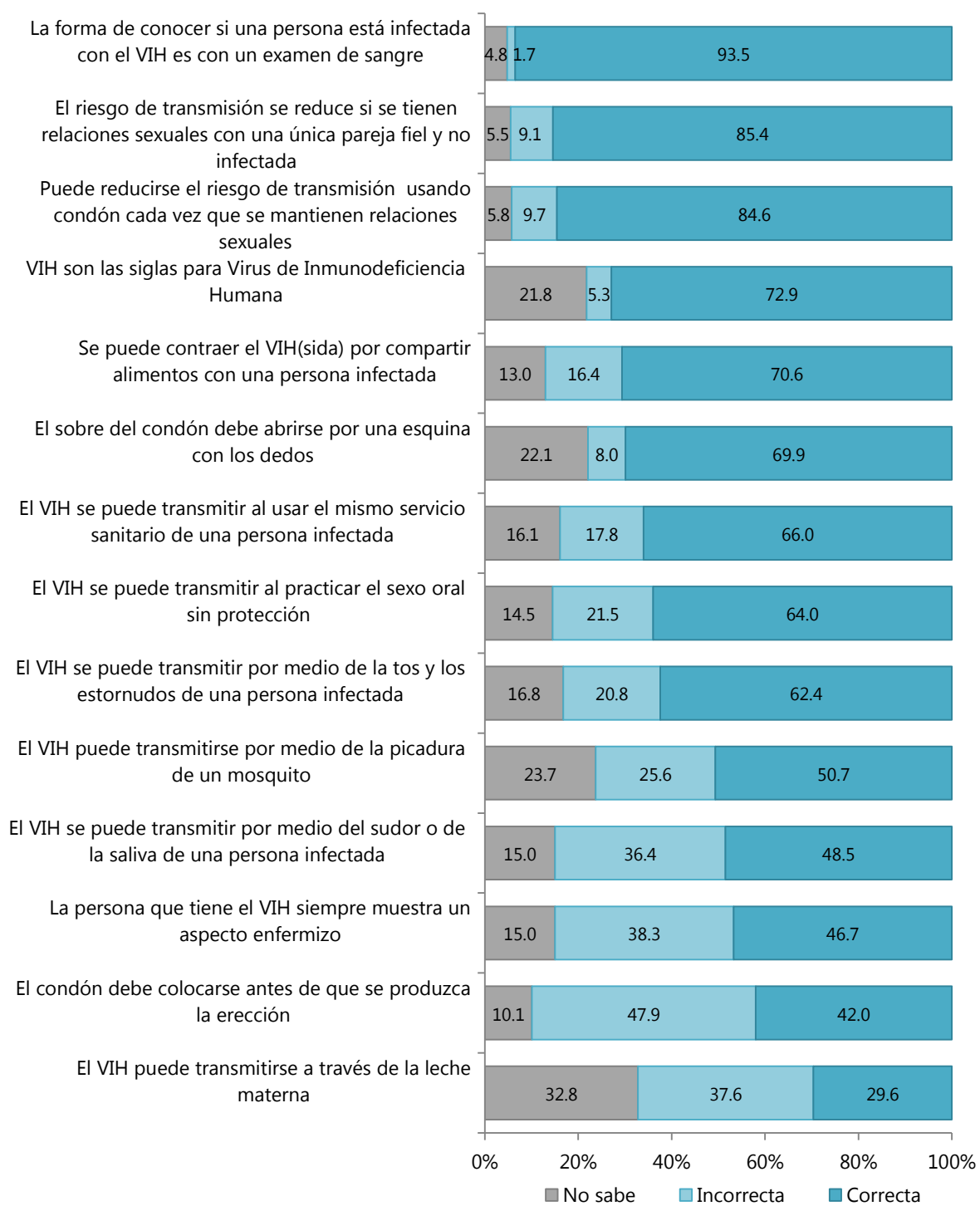
## 1.2. Distribución de respuestas y comportamiento de los ítems

Los ítems con el mayor número de respuestas correctas se muestran encabezando la Figura 3: *"La única forma de conocer si una persona está infectada con el VIH/SIDA es con un examen de sangre específico"* y *"Puede reducirse el riesgo de transmisión del VIH usando condón cada vez que se mantienen relaciones sexuales"*. El porcentaje de entrevistados que respondió correctamente a estos ítems corresponde a 93.5% y 85.4%, respectivamente.

Por otra parte, los ítems con el menor porcentaje de respuestas correctas son *"El VIH puede transmitirse a través de la leche materna"*, *"El condón debe colocarse antes de que se produzca la erección"* y *"La persona que tiene el VIH siempre muestra un aspecto enfermizo"*. Las cifras corresponden a 29.6%, 42% y 46.7%, en ese mismo orden.

**Gráfico 1**

Porcentaje de respuestas correctas, incorrectas y no sabe, para cada ítem de conocimiento sobre las formas de transmisión del VIH



El mayor porcentaje de entrevistados que indicó no saber la respuesta a la idea que se le planteó, se presenta en los ítems "El VIH puede transmitirse a través de la leche materna", "El VIH puede transmitirse a través de la picadura de un mosquito" y "El sobre del condón debe abrirse por una esquina con los dedos" (32.8%, 23.7% y 22.1%).

El Cuadro 2 y el Cuadro 3 muestran el porcentaje de entrevistados que respondieron correcta e incorrectamente a cada uno de los ítems analizados, así como el porcentaje que indicó no saber la respuesta a la afirmación planteada. Estas cifras se presentan según sexo en el Cuadro 2 y según edad de los entrevistados en el Cuadro 3.

**Cuadro 2**

Porcentaje de respuestas correctas, incorrectas y no sabe para cada ítem de conocimiento sobre las formas de transmisión del VIH, según sexo

Ítem	Hombre			Mujer		
	No sabe	Incorrecto	Correcto	No sabe	Incorrecto	Correcto
<b>a</b>	9.5	22.0	68.5	19.0	21.0	59.9
<b>b</b>	4.0	36.8	59.2	15.5	57.9	26.5
<b>c</b>	13.8	22.5	63.7	18.2	13.7	68.2
<b>d</b>	8.5	8.1	83.4	34.2	8.0	57.8
<b>e</b>	4.5	2.4	93.1	5.0	1.2	93.9
<b>f</b>	21.1	31.1	47.8	26.1	20.6	53.3
<b>g</b>	12.0	20.3	67.7	13.8	12.9	73.3
<b>h</b>	11.1	41.1	47.8	18.6	32.2	49.2
<b>i</b>	14.0	25.6	60.4	19.3	16.4	64.3
<b>j</b>	17.1	6.3	76.7	26.0	4.4	69.6
<b>k</b>	4.0	9.3	86.7	6.9	8.9	84.2
<b>l</b>	31.1	39.8	29.1	34.3	35.6	30.1
<b>m</b>	4.5	10.2	85.2	6.9	9.2	83.9
<b>n</b>	14.2	39.5	46.3	15.7	37.2	47.1

Si bien la distribución de las respuestas en las tres categorías presenta variaciones en la mayor parte de los ítems, según sexo las diferencias más notables se observan en los ítems *a*, *b*, *d*, *f*, *g* y *j*. En los primeros tres y también en el último, un mayor porcentaje de hombres que de mujeres responden correctamente a la afirmación planteada, mientras que en los ítems *f* y *g* la mayor proporción de respuestas correctas se presenta entre las mujeres.

**Cuadro 3**

Porcentaje de respuestas correctas, incorrectas y no sabe para cada ítem de conocimiento sobre las formas de transmisión del VIH, según edad

Ítem	15-24			25-34			35 y más		
	No sabe	Incorrecto	Correcto	No sabe	Incorrecto	Correcto	No sabe	Incorrecto	Correcto
<b>a</b>	14.3	20.4	65.3	14.8	21.8	63.5	14.6	22.0	63.4
<b>b</b>	13.4	46.5	40.1	8.0	38.8	53.2	8.9	53.6	37.5
<b>c</b>	21.1	22.4	56.5	13.6	16.0	70.3	14.2	15.7	70.1
<b>d</b>	19.5	6.8	73.7	14.2	6.9	78.9	27.8	9.5	62.7
<b>e</b>	5.9	3.3	90.8	3.9	1.3	94.8	4.5	1.0	94.5
<b>f</b>	27.3	26.8	45.9	24.2	26.9	48.9	21.1	24.2	54.7
<b>g</b>	14.5	20.6	64.9	14.8	13.3	71.9	11.0	15.2	73.8
<b>h</b>	14.3	37.2	48.5	13.9	33.8	52.3	16.1	37.3	46.6
<b>i</b>	17.6	23.0	59.4	17.3	17.9	64.8	16.0	20.8	63.2
<b>j</b>	22.0	6.2	71.8	18.9	5.4	75.6	23.1	4.6	72.3
<b>k</b>	6.5	11.6	82.0	5.0	8.3	86.7	5.2	7.8	87.0
<b>l</b>	37.5	32.9	29.6	36.4	33.8	29.8	27.8	42.6	29.6
<b>m</b>	6.6	10.8	82.7	5.0	8.2	86.8	5.6	9.8	84.6
<b>n</b>	17.8	34.8	47.4	14.6	33.7	51.7	13.4	42.9	43.8

Según grupos de edad (Cuadro 3), las mayores diferencias se presentan en los ítems *b*, *c*, *d* y *f*. En el ítem *b* las personas entre 25 y 34 años son quienes en mayor proporción responden correctamente. En el ítem *c* un mayor porcentaje de personas respondieron incorrectamente o ignoraron la respuesta entre los menores de 25 años, mientras que en el ítem *d* son más bien las personas mayores de 35 años quienes responden en mayor proporción a dichas categorías, si se le compara con los otros grupos etarios. En el ítem *f*, un mayor porcentaje responde correctamente entre las personas de 35 años o más.

En el Cuadro 4 se muestra la matriz de correlaciones policóricas entre los ítems que componen la escala de conocimientos sobre VIH. La mayoría de los coeficientes de correlación resultaron estadísticamente significativos ( $\alpha=0.05$ ).

Los coeficientes de correlación más bajos se presentan al analizar el ítem *a*. Ninguno de los coeficientes de correlación es significativamente distinto de cero al relacionar ese ítems con los ítems *c*, *f*, *g*, *h*, *i* y *m*. Los coeficientes de correlación más altos, por otra parte, se presentan entre el ítem *i* y los ítems *g* y *h* (Cuadro 4).

**Cuadro 4**

Matriz de correlaciones policóricas entre los ítems que componen la escala de conocimientos sobre el VIH

Ítem	a	b	c	d	e	f	g	h	i	j	k	l	m	n
<b>a</b>	-													
<b>b</b>	0.162	-												
<b>c</b>	<b>0.028</b>	0.223	-											
<b>d</b>	0.224	0.460	0.141	-										
<b>e</b>	0.230	0.192	0.340	0.273	-									
<b>f</b>	<b>0.030</b>	0.136	0.347	0.074	0.271	-								
<b>g</b>	<b>-0.004</b>	0.158	0.493	0.122	0.280	0.385	-							
<b>h</b>	<b>-0.015</b>	0.208	0.383	0.127	0.238	0.306	0.485	-						
<b>i</b>	<b>0.016</b>	0.236	0.458	0.157	0.286	0.397	<b>0.577</b>	<b>0.615</b>	-					
<b>j</b>	0.207	0.316	0.304	0.371	0.336	0.217	0.284	0.231	0.301	-				
<b>k</b>	0.089	0.214	0.271	0.145	0.296	0.148	0.208	0.162	0.196	0.310	-			
<b>l</b>	0.197	0.058	0.125	0.066	0.149	0.167	0.087	0.069	0.092	0.166	0.114	-		
<b>m</b>	<b>0.057</b>	0.247	0.285	0.183	0.340	0.127	0.270	0.227	0.269	0.325	0.469	0.082	-	
<b>n</b>	0.119	0.271	0.312	0.173	0.289	0.196	0.298	0.245	0.308	0.342	0.208	0.111	0.319	-

## 2. Análisis factorial confirmatorio

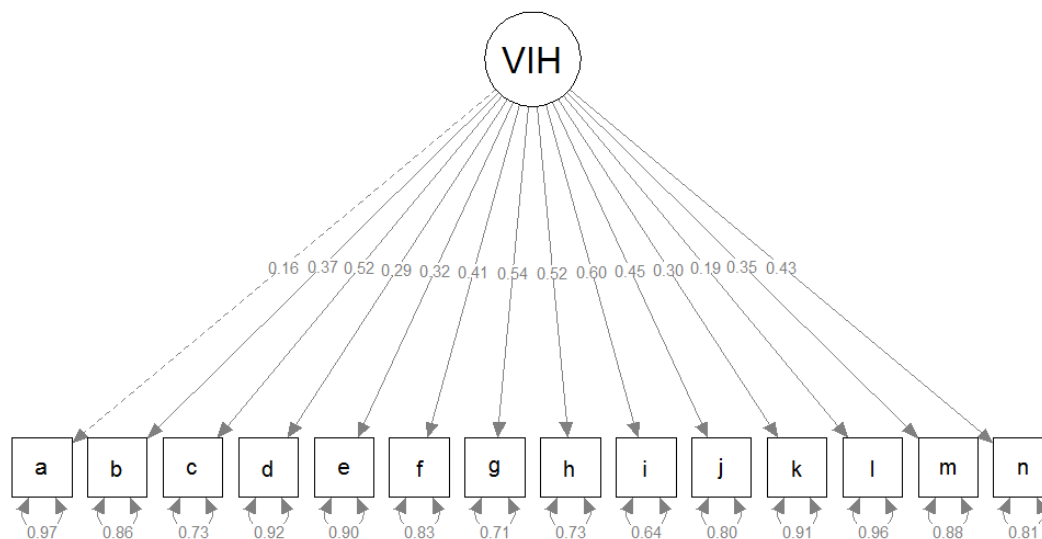
La estructura del análisis factorial confirmatorio se estima partiendo de que todos los ítems miden la misma variable latente. Los resultados revelaron (Gráfico 2) que todos los ítems se relacionan con un único constructo ( $p < 0.05$ , en todos los casos).

Los ítems con las cargas factoriales más altas son los ítems *i*, *g* y *c* con valores de 0.60, 0.54 y 0.52, respectivamente. El ítem *a* presenta la carga factorial más baja (0.16), seguido por el ítem *l* (0.19).

Los índices de ajuste que se obtienen al estimar el AFC (Gráfico 2) permiten determinar que este modelo presenta un ajuste aceptable. El CFI y TLI muestran valores cercanos a 0.90, mientras que los índices RMSEA y SRMR se encuentran cerca de 0.05. Nótese que no se puede rechazar la hipótesis nula planteada, denotando así que la matriz de covarianza muestral se ajusta a la matriz de covarianza teórica ( $p = 0.380$ ).

**Gráfico 2**

Análisis factorial confirmatorio para analizar la unidimensional del conjunto de ítems que estiman el conocimiento sobre las formas de transmisión del VIH



**Nota:** CFI: 0.908; TLI: 0.892; RMSEA: 0.051 ( $p < 0.05 = 0.380$ ); SRMR: 0.053.

### 3. Estimación de modelos

La escala de conocimiento sobre los medios de transmisión del VIH fue analizada mediante los modelos: Politémico de Respuesta Nominal, de Crédito Parcial y de Crédito Parcial Generalizado. En esta sección se describen los resultados obtenidos, comparando las características de esos tres modelos.

En un primer apartado se presentan los parámetros estimados y las curvas características de las categorías para cada ítem y modelo estimado. Posteriormente se analiza la información que brinda cada ítem y el test, nuevamente desglosando los resultados para los tres modelos. En un tercer apartado se comparan las estimaciones de  $\theta$  logradas a partir de estos modelos.

Dado que en los modelos ordinales seleccionados la codificación de las alternativas de respuesta debe asignar algún crédito en función de su relación con el parámetro de habilidad,

para la estimación de los modelos ordinales la codificación de las categorías de respuesta se guio por los resultados del MPRN.

Para esos modelos la codificación se estableció como: 0. No sabe, 1. Incorrecto y 2. Correcto, como se pudo establecer con la información que se aprecia en el Cuadro 5 y el Gráfico 3.

### 3.1. Parámetros de dificultad y discriminación de los modelos

La estimación de los parámetros de discriminación y de dificultad se muestra en el Cuadro 5. Es necesario señalar que en el MCP se estima un único parámetro de discriminación que se supone igual para todos los ítems, en este caso se obtuvo que la dificultad toma un valor de 0.6231.

**Cuadro 5**

Parámetros de dificultad y discriminación de cada ítem, estimados para los modelos analizados

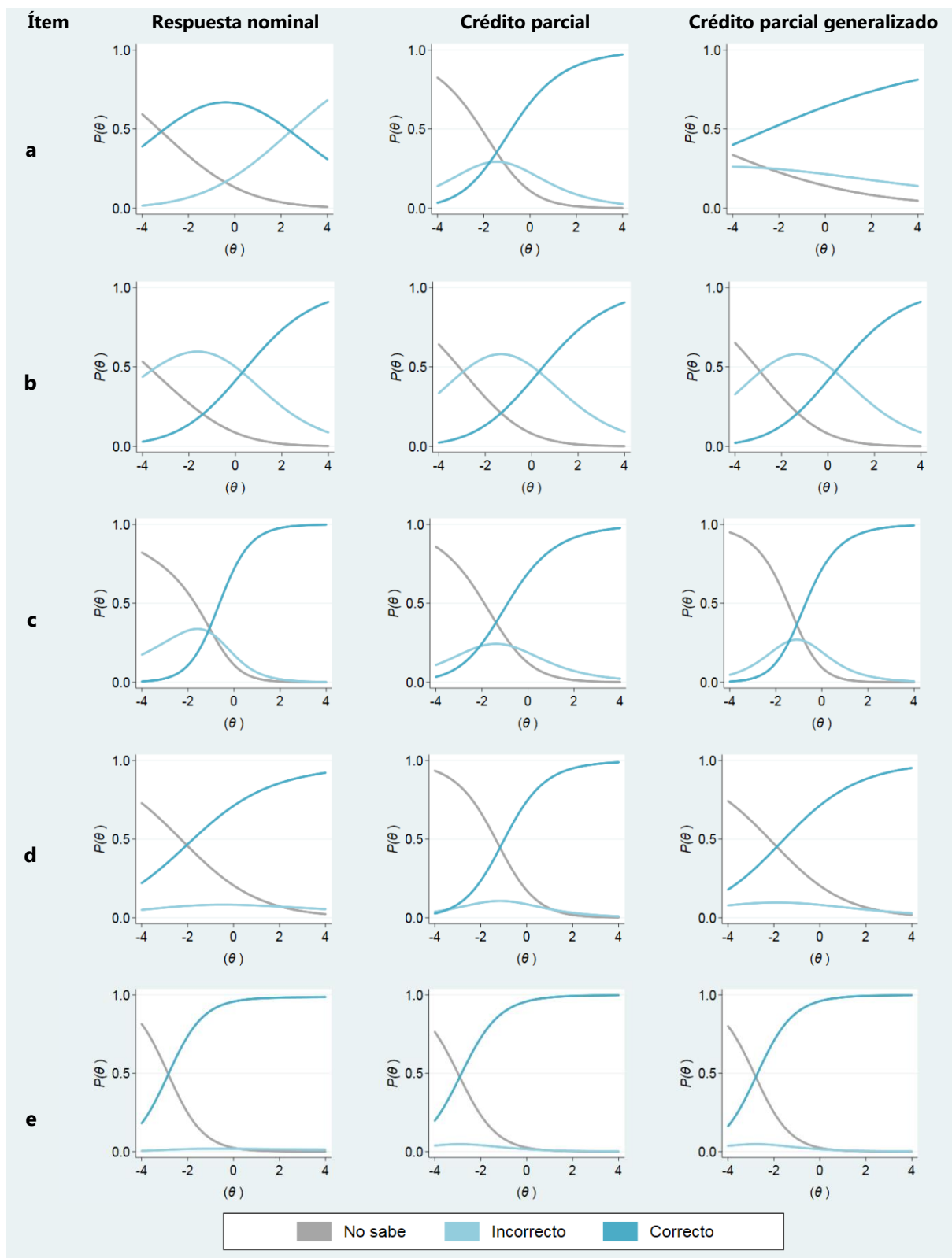
Ítem	MPRN				MCP <sup>1/</sup>		MCPG		
	Discriminación		Dificultad		Dificultad		Discr	Dificultad	
	0 vs 1	0 vs 2	0 vs 1	0 vs 2	0 vs 1	1 vs 2	0 vs 1	0 vs 1	0 vs 2
<b>a</b>	1.00	0.51	-0.43	-3.17	-1.15	-1.74	0.17	-2.50	-6.53
<b>b</b>	0.49	1.12	-3.60	-1.40	-2.95	0.31	0.63	-2.91	0.31
<b>c</b>	0.50	1.77	-0.91	-1.06	-0.68	-2.09	0.93	-0.77	-1.41
<b>d</b>	0.44	0.61	2.06	-2.04	1.13	-3.45	0.33	2.71	-6.46
<b>e</b>	1.21	1.30	0.25	-2.84	0.79	-6.62	0.66	0.69	-6.27
<b>f</b>	0.53	1.20	-0.53	-0.75	-0.51	-0.97	0.62	-0.51	-0.98
<b>g</b>	0.79	2.28	-1.15	-1.14	-0.95	-2.38	1.12	-1.01	-1.38
<b>h</b>	0.66	2.20	-1.97	-0.67	-1.86	-0.37	1.10	-1.42	-0.15
<b>i</b>	0.89	2.83	-1.04	-0.76	-0.83	-1.74	1.38	-0.90	-0.79
<b>j</b>	1.25	1.13	0.92	-1.34	1.77	-4.21	0.58	1.94	-4.50
<b>k</b>	1.22	1.35	-1.00	-2.56	-1.55	-3.76	0.57	-1.55	-4.08
<b>l</b>	0.84	0.35	-0.16	0.07	-0.45	0.66	0.22	-0.73	1.20
<b>m</b>	0.99	1.42	-1.24	-2.42	-1.58	-3.64	0.67	-1.53	-3.40
<b>n</b>	0.64	1.33	-1.91	-1.02	-1.93	-0.22	0.71	-1.78	-0.18

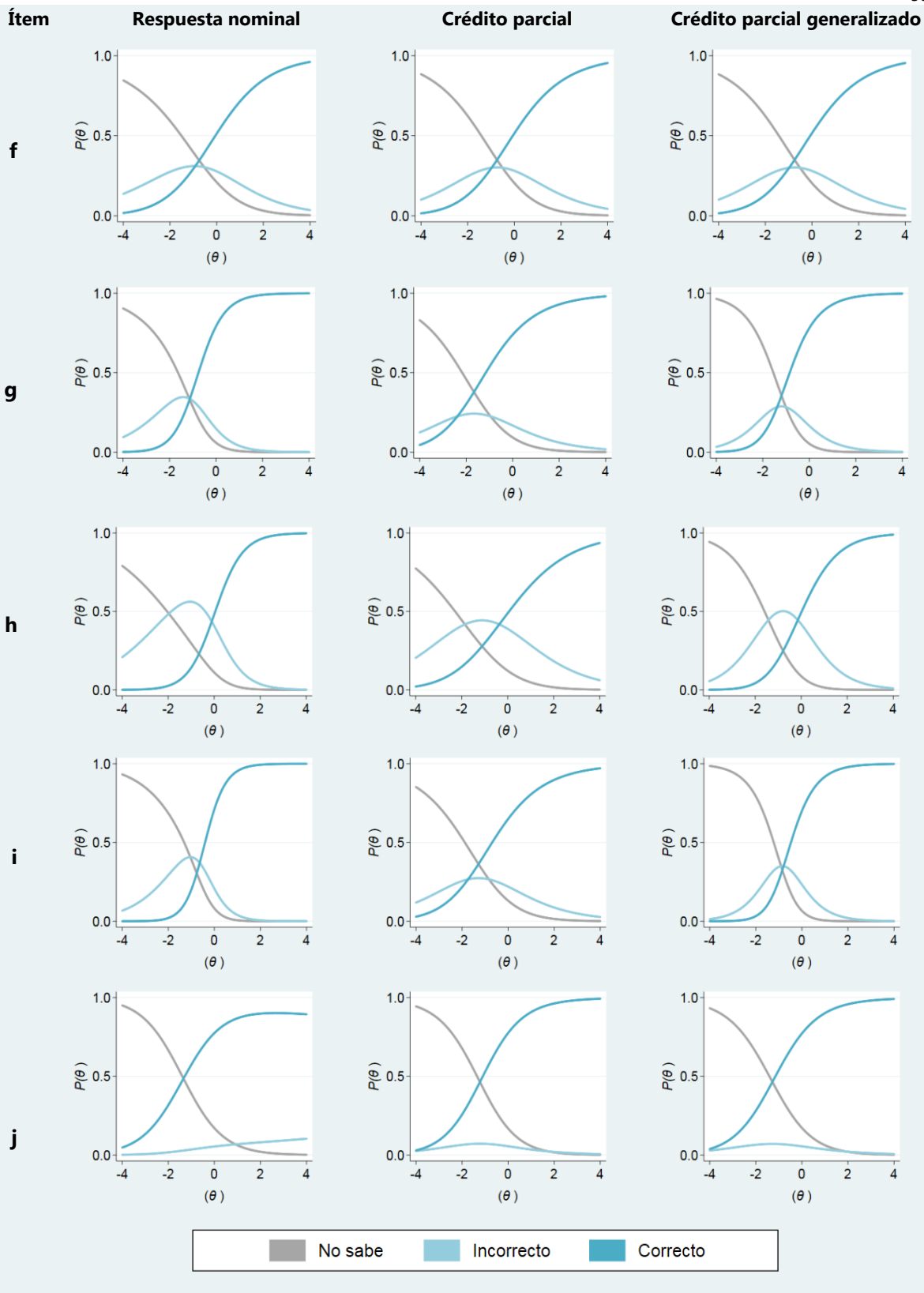
**Nota:** (1) El parámetro de discriminación es igual entre los ítems, se estimó en: 0.6231.

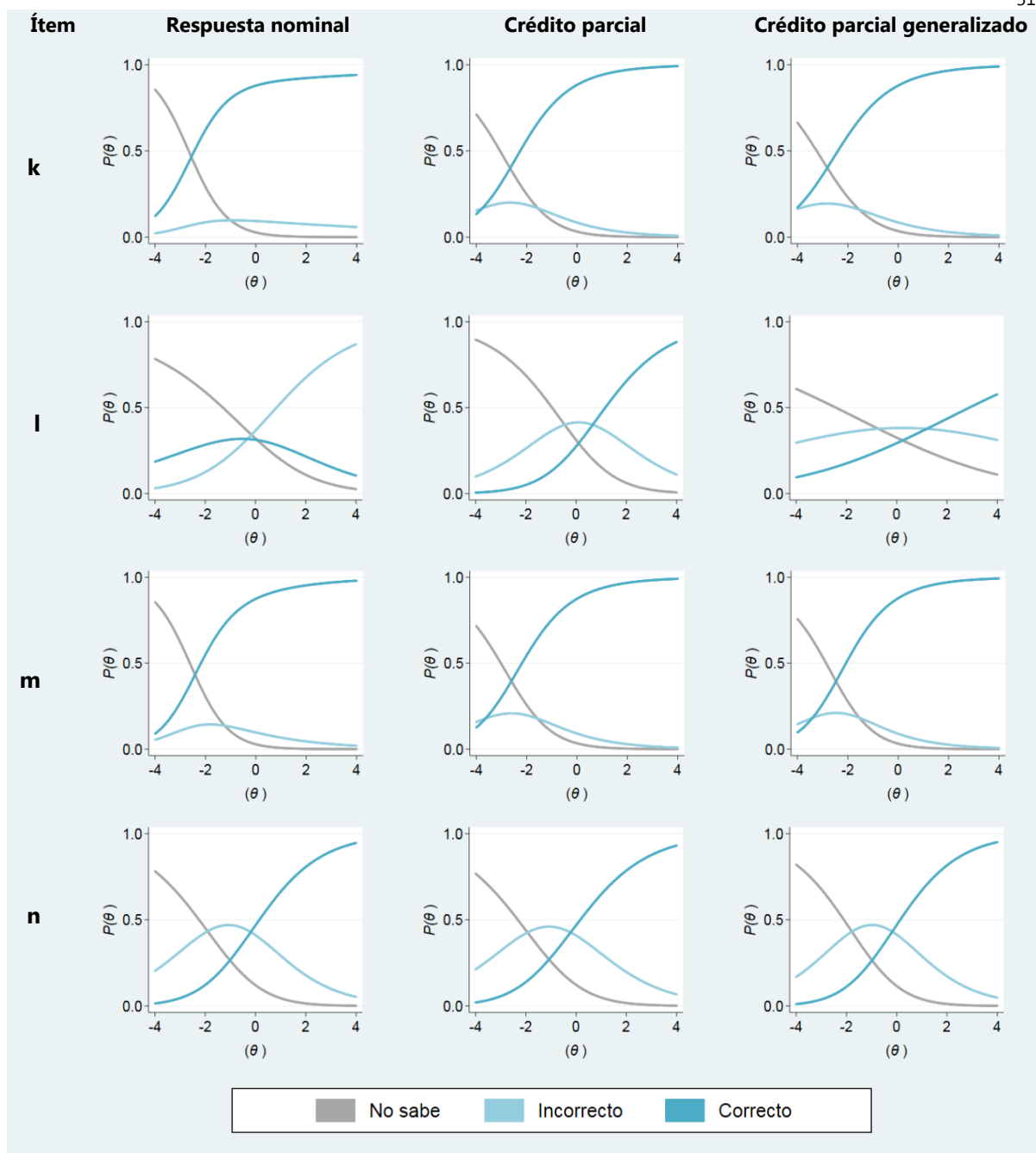
Los parámetros de los ítems pueden analizarse más fácilmente de forma gráfica, al analizar las CCC que se presentan en la siguiente sección.

Gráfico 3

Curvas características de las categorías para cada ítem, según método de análisis







### 3.2. Curvas características de categoría (CCC)

El gráfico 3 muestra las curvas características de las alternativas de respuesta para cada ítem y cada uno de los tres modelos analizados. Las CCC permiten conocer la probabilidad de que un sujeto, que posee cierto nivel de habilidad, seleccione alguna de las categorías de respuesta.

Independientemente del modelo, cuando se hace referencia a un mismo ítem, las CCC muestran un patrón más o menos estándar. Los individuos con mayor conocimiento sobre los medios de transmisión del VIH ( $\theta$ ), poseen una mayor probabilidad de responder correctamente a cada uno de los ítems. Las respuestas incorrectas, por su parte, tienden a presentarse entre los individuos con bajo nivel de conocimiento, pero no necesariamente entre quienes poseen los niveles de habilidad más bajos. La categoría 'no sabe' es más bien la que se asocia con las puntuaciones de  $\theta$  más bajas.

Ese patrón no se cumple para los ítems *a* y *l* al considerar el MPRN, pues los individuos con mayor conocimiento son quienes tienden a responder incorrectamente.

En general, en los modelos politómicos de TRI, habrá una categoría cuya curva será creciente y otra cuya curva será decreciente. En la mayoría de los casos se trata de las categorías 'correcto' y 'no sabe', respectivamente. En los MPRN la forma de estas curvas está dada por la frecuencia, es decir, por el número de sujetos que seleccionó esa alternativa de respuesta, mientras que en los modelos de CP y CPG su forma está dada por el crédito asignado.

Los ítems *d*, *e*, *j*, *k* y *m* se comportan básicamente de manera dicotómica en los tres modelos ajustados, donde las categorías *no sabe* y *correcto* son dominantes. La probabilidad de que un individuo seleccione la categoría 'incorrecto' es baja a lo largo de todo el continuo de  $\theta$ . En su lugar, los examinados en los niveles de habilidad más bajos tienden a desconocer la respuesta y las personas con mayor habilidad tienen mayor probabilidad de responder correctamente a estos ítems.

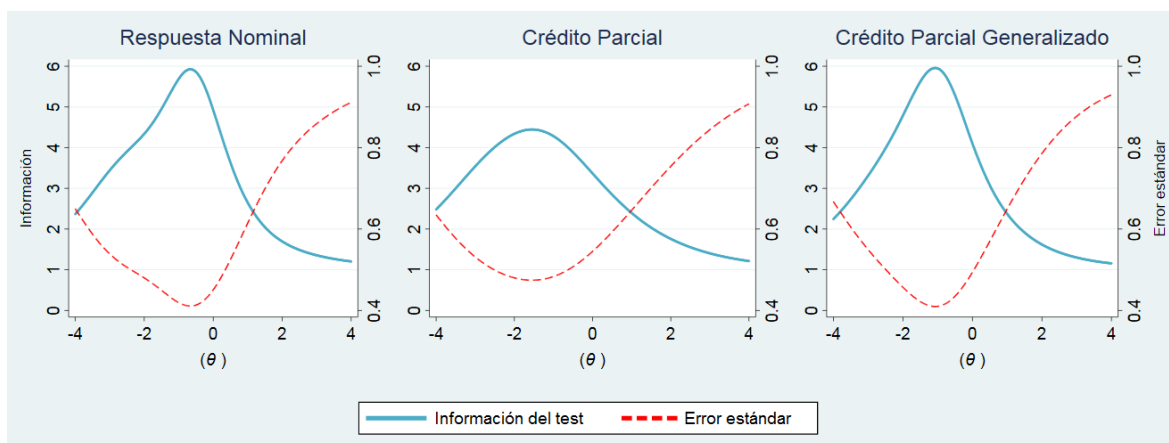
La función de probabilidad asociada a la categoría 'incorrecto' es unimodal en algunos ítems (*b*, *c*, *f*, *g*, *h*, *i* y *n*). La probabilidad de seleccionar esa categoría es baja, a lo largo de los distintos niveles de conocimiento, en el caso de los ítems *d*, *e*, *j*, *k* y *m*. Este patrón es similar entre los modelos.

### 3.3. Información de los ítems y del test

En el Gráfico 4 puede observarse la información del test para los modelos de Respuesta Nominal, Crédito Parcial y Crédito Parcial Generalizado. La función de información estimada mediante el MPRN y MCPG brindan la mayor información, que en general se presenta para los niveles de  $\theta$  de bajo a medio.

**Gráfico 4**

Función de información del test para cada modelo estimado

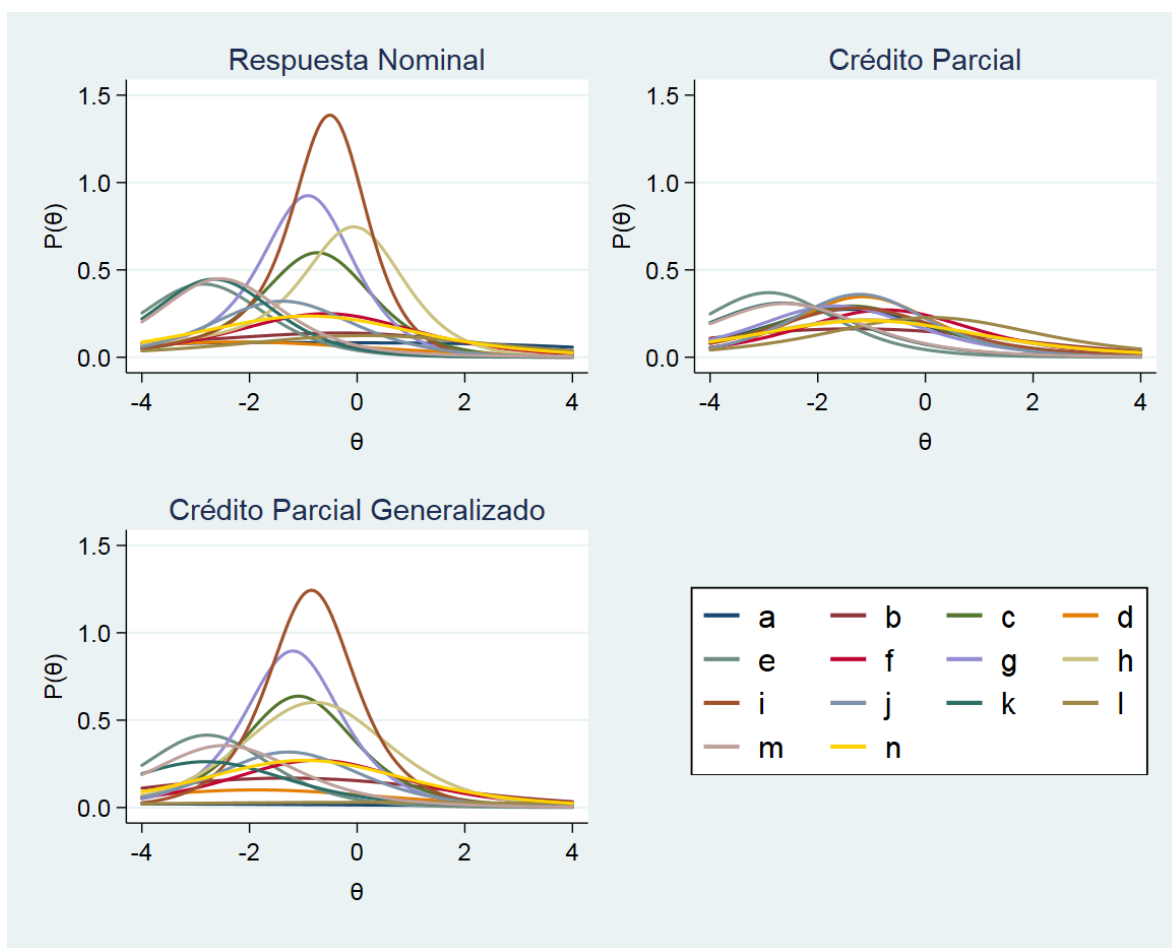


El gráfico 5 muestra la función de información de cada uno de los ítems para los tres métodos analizados. Los ítems brindan mayor información cuando se analizan los modelos de respuesta nominal y crédito parcial generalizado, si se les compara con el modelo de crédito parcial. En general, las funciones de información del MPRN y el MCPG presentan una estructura similar y brindan información en los niveles de habilidad de bajo a medio, y en general, por debajo del promedio de  $\theta$ .

Al considerar esos dos modelos, cuatro de los ítems destacan por presentar la mayor información. El ítem *b* es el que brinda la mayor información en ambos modelos. Le siguen los ítems *g*, *h* y *c*. Todos ellos brindan información en los niveles de habilidad bajo-medio.

**Gráfico 5**

Función de información de los ítems, para cada modelo estimado

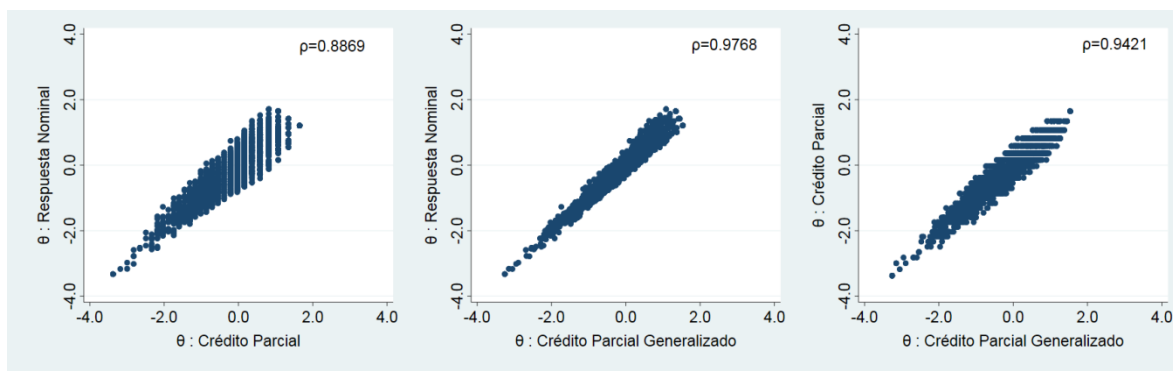


### 3.4. Estimación de las puntuaciones ( $\theta$ )

La puntuación del conocimiento sobre los medios de transmisión del VIH ( $\theta$ ) puede compararse para todos los modelos estimados (Gráfico 6). Esta estimación fue estandarizada de tal forma que la media de la habilidad es igual a 0 y su desviación estándar es 1.

### Gráfico 6

Relación entre la estimación del conocimiento sobre los medios de transmisión del VIH ( $\theta$ ) mediante los tres métodos analizados



De forma general, la estimación del conocimiento es muy similar entre los tres modelos. Se observa mayor dispersión a medida que aumenta el nivel de habilidad. El coeficiente de correlación de Spearman se calculó para establecer una medida numérica de la asociación entre las puntuaciones (Gráfico 6). Todos los coeficientes fueron altos, principalmente al analizar la relación entre las puntuaciones obtenidas mediante el MPRN y el MCPG ( $\rho=0.9768$ ).

## 4. Selección de modelos y el principio de invarianza

### 4.1. Análisis de criterios de selección

La prueba de razón de verosimilitudes para comparar la relación entre los tres modelos se muestra en el Cuadro 6. La hipótesis nula que establece conservar el modelo reducido, se rechaza en todos los casos. Este resultado sugiere que el MPRN brinda un mejor ajuste que los MCP y el MCPG. Al comparar estos últimos modelos, nuevamente el modelo más completo, en este caso el MCPG, brinda mejor ajuste que el MCP.

**Cuadro 6**

Prueba de razón de verosimilitudes entre los modelos anidados

Modelos en comparación	$\Delta G^2$	Grados de libertad	$P(> \chi^2)$
Respuesta Nominal vs. Crédito Parcial*	1057.8	27	0.000
Respuesta Nominal vs. Crédito Parcial Generalizado*	474.7	14	0.000
Crédito Parcial Generalizado vs. Crédito Parcial*	583.1	13	0.000

**Nota:** (\*) Modelo reducido

Los índices AIC y BIC sugieren de la misma manera, la selección del modelo más complejo (Cuadro 7). El MPRN posee las menores cifras para ambos indicadores (63 146.4 y 63 481.5, respectivamente). Si se comparan los modelos ordinales, los índices sugieren la selección de MCPG.

**Cuadro 7**

Estadísticas de bondad de ajuste, AIC y BIC para los modelos estimados

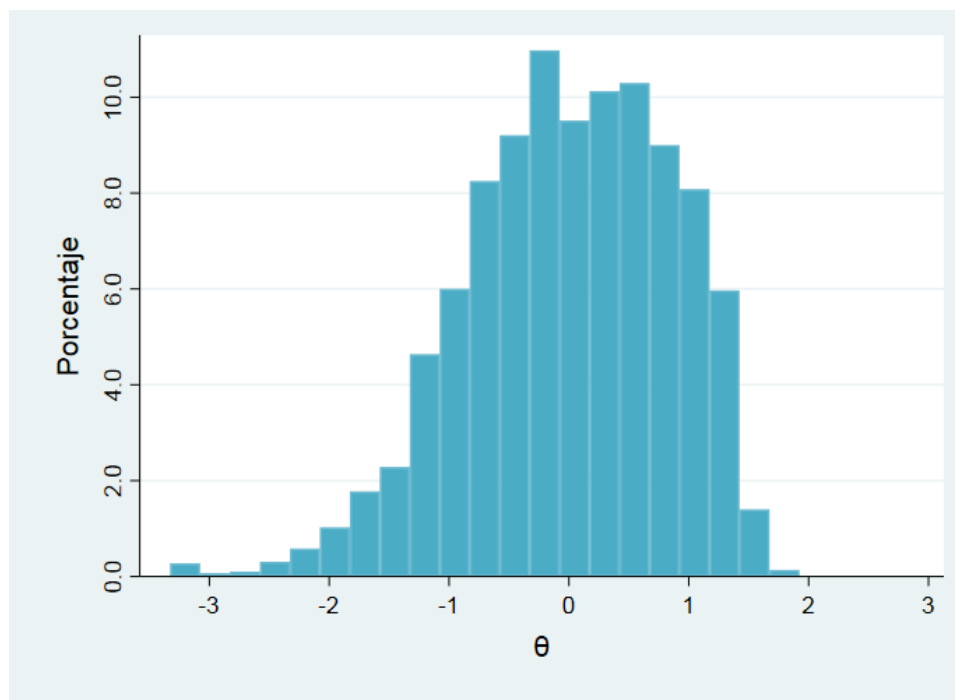
Modelo	Log verosimilitud	Grados de libertad	AIC	BIC
Respuesta Nominal	-31517.2	56	63146.4	63481.5
Crédito Parcial	-32046.1	29	64150.2	64323.8
Crédito Parcial Generalizado	-31754.6	42	63593.1	63844.5

#### 4.2. Distribución de las puntuaciones

La estimación de las puntuaciones obtenidas por los entrevistados se muestra en la Gráfico 7. La mayor parte de los entrevistados obtuvo puntuaciones en torno a los valores medios de  $\theta$ , aunque se presenta cierta asimetría negativa en esta distribución.

**Gráfico 7**

Distribución de las puntuaciones obtenidas en la escala sobre conocimiento en VIH



### 4.3. Funcionamiento diferencial de los ítems

Partiendo de los resultados de la sección anterior, el modelo que presentaba un mejor ajuste fue el MPRN. Como proceso de postestimación se verifica la presencia de funcionamiento diferencial de los ítems, solamente para ese modelo. Los resultados del análisis permitieron determinar que varios ítems presentan funcionamiento diferencial tanto por sexo como por grupos de edad.

Según sexo, se presenta DIF no uniforme en los ítems *a*, *j* y *l*. Al considerar la prueba de razón de verosimilitud para analizar el DIF uniforme, 10 de los 14 ítems presentaron DIF (Cuadro 8). Sin embargo, si se analiza el cambio en el parámetro  $\beta_1$ , solamente el ítem *a* presenta un cambio importante. Los ítems *b*, *d*, *e* y *g* presentan un  $\Delta\beta_1$  que supera el 1%.

**Cuadro 8**

Pruebas de razón de verosimilitudes y cambio en el coeficiente  $\beta$  de la variable latente para analizar el DIF, según sexo

Ítem	NO UNIFORME	UNIFORME	
	P( > $\chi^2$ )	P( > $\chi^2$ )	$\Delta\beta_1$
<b>a</b>	<b>0.000</b>	<b>0.000</b>	<b>-27.1</b>
<b>b</b>	0.721	<b>0.000</b>	2.5
<b>c</b>	0.393	<b>0.000</b>	0.6
<b>d</b>	0.815	<b>0.000</b>	1.1
<b>e</b>	0.062	<b>0.048</b>	1.4
<b>f</b>	0.965	<b>0.005</b>	0.7
<b>g</b>	0.734	<b>0.000</b>	1.5
<b>h</b>	0.688	0.470	0.1
<b>i</b>	0.091	<b>0.001</b>	0.3
<b>j</b>	<b>0.001</b>	<b>0.000</b>	-0.4
<b>k</b>	0.652	0.192	-0.5
<b>l</b>	<b>0.000</b>	0.939	0.1
<b>m</b>	0.299	0.946	0.0
<b>n</b>	0.672	<b>0.045</b>	0.6

**Nota:**  $\Delta\beta_1$  expresado en términos porcentuales

**Cuadro 9**

Pruebas de razón de verosimilitudes y cambio en el coeficiente  $\beta$  de la variable latente para analizar el DIF, según grupos de edad

Ítem	NO UNIFORME P( > $\chi^2$ )			UNIFORME					
	15 a 24		15 a 34	15 a 24 vs. 25 a 34		15 a 24 vs. 35+		25 a 34 vs. 35+	
	25 a 34	35+	vs. 35+	P( > $\chi^2$ )	$\Delta\beta^2$	P( > $\chi^2$ )	$\Delta\beta^2$	P( > $\chi^2$ )	$\Delta\beta^2$
<b>a</b>	0.454	0.347	0.938	0.369	<b>6.5</b>	0.347	<b>5.4</b>	0.972	0.1
<b>b</b>	0.855	0.862	0.789	<b>0.000</b>	-2.3	0.178	0.6	<b>0.000</b>	0.1
<b>c</b>	0.969	0.130	0.163	<b>0.000</b>	-0.8	<b>0.000</b>	0.2	0.303	0.2
<b>d</b>	0.490	<b>0.032</b>	0.214	0.091	-1.7	<b>0.000</b>	<b>5.5</b>	<b>0.000</b>	1.0
<b>e</b>	0.654	0.278	0.631	<b>0.020</b>	-0.6	<b>0.007</b>	-0.1	0.946	0.0
<b>f</b>	0.227	0.052	0.601	0.526	0.4	<b>0.002</b>	-0.5	<b>0.000</b>	1.2
<b>g</b>	0.090	0.487	0.258	0.409	0.3	<b>0.004</b>	-0.1	<b>0.001</b>	1.0
<b>h</b>	0.655	0.816	0.785	<b>0.011</b>	1.3	<b>0.000</b>	1.8	0.103	-0.1
<b>i</b>	0.794	0.384	0.271	<b>0.010</b>	1.3	0.167	0.3	0.096	0.3
<b>j</b>	<b>0.041</b>	0.862	<b>0.039</b>	0.988	0.0	0.148	0.7	0.139	-0.1
<b>k</b>	0.093	<b>0.033</b>	0.862	0.102	-1.5	<b>0.014</b>	-0.9	0.575	0.2
<b>l</b>	0.088	0.457	0.194	0.681	0.8	<b>0.018</b>	-2.7	<b>0.009</b>	1.1
<b>m</b>	0.220	0.786	0.276	0.237	-0.6	0.780	-0.1	0.294	0.0
<b>n</b>	0.591	0.482	<b>0.222</b>	0.988	0.0	0.076	0.6	0.065	-0.2

**Nota:** El  $\Delta\beta_1$  se expresa en términos porcentuales.

De acuerdo con los resultados del Cuadro 9, se presenta DIF no uniforme considerando los grupos etarios en los ítems  $f$ ,  $j$  y  $k$  bajo la estimación del MPRN. Empleando la Prueba de razón de verosimilitudes se presenta DIF uniforme para diez ítems. La identificación de ítems con esa misma clase de DIF varía al considerar el  $\Delta\beta_1$ , pues solamente los ítems  $a$  y  $d$  muestran un cambio mayor a 5% (Cuadro 9).

## CONCLUSIONES Y DISCUSIÓN

Las puntuaciones de los sujetos u otros indicadores estimados tras la aplicación de esta escala sobre medios de transmisión del VIH, son un insumo para múltiples organizaciones cuya labor se enfoca en la prevención de las nuevas infecciones por VIH. El análisis apropiado del conocimiento sobre las formas de prevenir una infección, permite formular distintas estrategias de política pública, con miras a reducir el número de infecciones (ONUSIDA, 2014) a través de la educación sexual y la prevención.

Si existen variaciones en la medición equitativa según subpoblaciones, entonces, eventualmente, las mediciones a través del test y sus resultados no son comparables. A esto debe considerarse adicionalmente que algunos de estos ítems suelen emplearse para generar indicadores que se reportan en informes internacionales, así, debería considerarse incluso el funcionamiento diferencial de los ítems según nacionalidad, principalmente si su redacción se adecúa al contexto de cada nación.

Los ítems planteados en el módulo sobre conocimientos y actitudes hacia el VIH en la ENSSR-15 no surgieron tras el proceso formal recomendado en la elaboración de pruebas estandarizadas (cf. Martínez, 1996). Los resultados de este trabajo pueden ser insumo para la revisión de los ítems que componen la escala, con el fin de mejorar la medición del constructo.

Será necesaria la revisión de los ítems actuales, considerando la redacción e incluso la eliminación de algunas de las ideas planteadas. La selección de nuevos ítems debe realizarse con mayor cuidado en apego a las convenciones establecidas en el desarrollo de ítems y escalas.

Los índices de bondad de ajuste empleados para la selección de modelos permitieron determinar que el modelo politómico de respuesta nominal es el que se ajusta mejor a la estructura de los datos. Este modelo mostró los menores valores para los índices AIC y BIC y al

compararlo con los otros modelos, la prueba de razón de verosimilitudes determinó que el MPRN estima los datos con mejor ajuste.

De acuerdo con esa prueba, el MCPG brinda mejor ajuste que el MCP. El primero de esos modelos es una variación del segundo, donde se libera el supuesto de igualdad del parámetro de discriminación. El mejor ajuste del MCPG sugiere la relevancia de incorporar el parámetro de discriminación al modelar este conjunto de datos.

El test, en general, brinda mayor información para niveles de bajos a medios de habilidad (aproximadamente para  $\theta$  entre -2 y 0). Los MPRN y MCPG poseen funciones de información similares. El MCP brinda la menor información de los tres modelos.

Si se toma como evidencia la forma similar de las funciones de información y la correlación más alta entre las puntuaciones  $\theta$ , en general podría afirmarse que el MPRN y el MCPG son similares. Podría preferirse el modelo nominal, dada su relación más intuitiva con la naturaleza de las categorías de respuesta. Tomando en contexto todos esos elementos, el MCP no parece ser el más adecuado para modelar los datos de esta escala.

Los ítems *b*, *g*, *h* y *c* brindan la mayor información en los MPRN y MCPG, sus funciones de información se destacan del resto. Los ítems restantes brindan poca información. Los mejores ítems no son claramente apreciables en el MCP, pues en general todos ellos brindan menos información si se les compara con los otros dos modelos. Cabe señalar que ningún ítem brinda información para los niveles altos de  $\theta$ .

Los resultados brindan indicios de la presencia de DIF en varios ítems. Considerando el sexo, se identificó la presencia de DIF no uniforme en los ítems *a*, *j* y *l*. El DIF no uniforme se presentó según edad en el ítem *j*, notándose diferencias entre los más jóvenes (15 a 24 años) y el grupo que representa a las personas entre 25 y 34 años.

En la detección del DIF uniforme la prueba de razón de verosimilitudes y el análisis del cambio del coeficiente  $\beta_1$  no guardan coincidencia en todos los casos. Los resultados sugieren que el primero de estos métodos es más sensible en la detección del DIF uniforme y eventualmente podría dar pie a la identificación espuria de ítems que presentan este tipo de sesgo. En algunos ítems para los que se rechaza la hipótesis nula en la prueba de razón de verosimilitudes, se observaron pequeños cambios entre los coeficientes  $\beta_1$ .

La falta de resultados homogéneos entre estos métodos no debería ser distracción del hecho de que algunos ítems deberían revisarse, ante la posibilidad de que presenten funcionamiento diferencial entre las subpoblaciones analizadas.

Los ítems *a* (“El VIH se puede transmitir al practicar el sexo oral sin protección”) y *l* (“El VIH puede transmitirse a través de la leche materna”) presentaron consistentemente resultados desfavorables, pues no solo presentan las cargas factoriales más bajas sino que también muestran DIF. El ítem *a* mostro DIF no uniforme según sexo y DIF uniforme según edad. El ítem *l* según edad, pero solo al considerar la prueba de razón de verosimilitud. A estos aspectos debe sumarse la poca información que brindan, según se observó en las funciones de información.

Los dos ítems mencionados anteriormente mostraron CCC que no coincidían con la lógica esperada en este test cuando se estimó el MPRN. Se espera que los sujetos con más habilidad tengan mayor probabilidad de responder correctamente, lo que no sucede en ese modelo. Esa relación sí se observa con el MCP y el MCPG.

Este hallazgo sugiere la estimación de los denominados modelos híbridos o mixtos, donde los parámetros de estos ítems se estimen con un modelo de tipo ordinal, mientras que los restantes se generen mediante un MPRN.

La eliminación de ambos ítems podría considerarse. Sin embargo, también debe tenerse presente la información sustantiva que las concepciones erróneas sobre estas ideas permiten dilucidar acerca del conocimiento que posee la población.

Solo el 64% de los entrevistados logró identificar correctamente que el VIH puede transmitirse al tener prácticas sexuales orales sin protección (ítem *a*). El condón se utiliza con poca regularidad durante las prácticas sexuales orales (González y Barrientos, 2014) aun cuando este es el tipo de práctica sexual más frecuente entre la población costarricense, según datos de las dos últimas Encuestas de Salud Sexual y Reproductiva realizadas en Costa Rica (2010 y 2015).

Un porcentaje aún más bajo de entrevistados logró responder correctamente que el VIH se transmite a través de la leche materna (ítem *l*; 29.6%). Si bien la transmisión vertical del virus no es la principal vía de infección en Costa Rica (Ministerio de Salud, 2016), los nuevos casos a causa de la transmisión madre-hijo podrían evitarse con información y ampliando el acceso a los servicios de salud. Si la madre recibiera tratamiento podría reducir el riesgo de transmitir el virus a su hijo en menos de 2% (Hernández, León y Varela, 2009). Los resultados sugieren que este es un ítem difícil de responder, pero principalmente porque la población no posee información al respecto.

La escala de conocimientos en materia de VIH descrita por ONUSIDA en sus guías metodológicas (ONUSIDA, 2011, 2014) es ampliamente utilizada para analizar la correcta identificación de las formas de prevenir la transmisión del virus y el rechazo de las principales ideas erróneas sobre su transmisión (Acosta et al, 2015; Arce, 2014 y Espada et al, 2014), bien sea entre la población general como en grupos poblacionales específicos. Algunos de los ítems de esa escala fueron incorporados entre los que aquí se analizaron.

La metodología de ONUSIDA (2011) sugiere que todas las respuestas en las que el entrevistado indica "no sabe" sean consideradas como respuestas incorrectas. A esto debe considerarse que, dado un nivel de conocimiento sobre los medios de transmisión del VIH y suponiendo que las respuestas "no sabe" reflejan un conocimiento erróneo sobre este aspecto, las probabilidades

de responder "no sabe" e "incorrecto" deberían ser equivalentes. Si las respuestas son dicotomizadas en las categorías correcto e incorrecto, se pierde información sobre el proceso de respuesta subyacente, como lo sugieren los resultados de este análisis

En este trabajo, se logra identificar la discrepancia existente entre las curvas características de esas categorías. Dichas categorías presentan una distribución de probabilidad muy disímil. La categoría "no sabe" tiende a ser seleccionada por las personas que en general presentan menor grado de conocimiento sobre los medios de transmisión del VIH. La categoría "incorrecto" tiende, más bien, a ser seleccionada por personas con niveles de conocimiento de medio a bajo.

Esto sugiere que ambas categorías no deberían colapsarse. Las guías metodológicas que agrupan ambas categorías en una sola, sugieren una estimación inapropiada a la estructura de los datos.

A esta última idea, sin embargo, debe considerarse que si bien el uso de modelos de respuesta al ítem brinda una estimación más precisa del constructo analizado, su incorporación en este contexto supone una serie de dificultades. En primera instancia, será necesario que las organizaciones cuenten con profesionales que conozcan de este tipo de modelos, situación que puede ser inviable. La construcción de un indicador a partir de la simple suma de valores puede ser aplicada con relativa facilidad por personal no necesariamente experto. Habrá que considerar también que para el público general y frecuentemente no experto en temas de medición es más fácil comprender la simple suma de los puntajes.

Una posible solución tras las consideraciones anteriores es asignar puntuaciones a las categorías de respuesta, acordes con los resultados generados por el MPRN. Dado que la categoría "no sabe" se asocia con los niveles más bajos de  $\theta$ , seguido de las categorías "incorrecto" y "correcto", siendo esta última la que se asocia con un mayor conocimiento, parece plausible asignar a estas categorías puntajes como 0, 1 y 2, respectivamente. La medida

sobre el conocimiento ( $\theta$ ) se obtendría como la suma de las puntuaciones asignadas a todos los ítems que componen la escala.

Esa estrategia permite en primer lugar asociar las puntuaciones con los niveles de  $\theta$ , y en segundo, permite conservar un esquema para la asignación de puntuaciones y una operacionalización sencilla para el cálculo de un puntaje total, que puede ser aplicada prácticamente por cualquier investigador.

La evidencia que surge tras el análisis aportado en este trabajo, podría guiar estrategias de prevención de nuevas infecciones por VIH. Se revelan vacíos por parte de la población en el conocimiento comprensivo sobre su transmisión y sobre otras ideas relacionadas, sugiriendo que aún deben canalizarse esfuerzos por aumentar el conocimiento en la materia. Aún deben generarse esfuerzos en diferentes ámbitos, inclusive en la medición de este constructo, con el fin alcanzar el objetivo al que se comprometido un gran número de naciones, entre ellas Costa Rica: "Poner fin a la epidemia de VIH y sida".

## BIBLIOGRAFIA

Acosta, S. C., Monteverde, M. P., del Riego, A. P., Barrantes, A. L., Huaranga, E. R., Siancas, J. S. y Segura, E. R. (2015). Conocimientos sobre VIH/SIDA y la disposición a realizar procedimientos médicos: Estudio en una muestra de médicos en formación. *Archivos de medicina*, 11(3), 12.

Agresti, A. y Kateri, M. (2011). Categorical data analysis. In *International encyclopedia of statistical science* (pp. 206-208). Springer Berlin Heidelberg.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.

An, X. y Yung, Y. F. (2014). Item response theory: what it is and how you can use the IRT procedure to apply it. *SAS Institute Inc. SAS364-2014*.

Andersen, E.B. (1977). Sufficient statistics and latent trait models. *Psychometrika* 42, 69–81.

Andrich, D. (1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement* 2, 581–594.

Asún, R. y Zúñiga, C. (2008). Ventajas de los Modelos Politémicos de Teoría de Respuesta al ítem en la Medición de Actitudes Sociales: El Análisis de un Caso. *Psyche (Santiago)*, 17(2), 103-115.

Avert. (2017). History of HIV and AIDS overview. Disponible en: <https://www.avert.org/professionals/history-hiv-aids/overview> [Accesado en octubre, 2017].

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, 107(2), 238.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37(1), 29-51.

Borsboom, D.; Mellenbergh, G.J.; van Heerden, J. (2003). The Theoretical Status of Latent Variables. *Psychological Review* 110 (2): 203-219.

Bunge, M. (1973). La ciencia, su método y filosofía. Buenos Aires: Siglo XX.

Caviezel, V., Bertoli-Barsotti, L. y Lozza, S. O. (2011). Measuring risk profile with a multidimensional Rasch analysis. *Journal of Applied Quantitative Methods*, 6(4).

Carey, M. P., Morrison-Beedy, D. y Johnson, B. T. (1997). The HIV-Knowledge Questionnaire: Development and evaluation of a reliable, valid, and practical self-administered questionnaire. *AIDS and Behavior*, 1(1), 61-74.

Crane, P.; Gibbons, L.E.; Jolley, L. y van Belle G. (2005). DIFd v.1.0. Seattle, WA: University of Washington.

Crane, P. K.; Hart, D. L.; Gibbons, L. E. y Cook, K. F. (2006). A 37-item shoulder functional status item pool had negligible differential item functioning. *Journal of Clinical Epidemiology*, 59(5), 478-484.

De Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Publications.

De Jong, M. G., Steenkamp, J. B. E., Fox, J. P. y Baumgartner, H. (2008). Using item response theory to measure extreme response style in marketing research: A global investigation. *Journal of marketing research*, 45(1), 104-115.

Debera, L. y Nalbarte, L. (2006). Pruebas diagnósticas: una aplicación a la teoría de respuesta al ítem, aproximación clásica y bayesiana. *IESTA–Universidad de la República–Facultad de Ciencias Económicas y de Administración, Serie Documentos de trabajo*.

Edelen, M. O. y Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research, 16*(1), 5.

Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational and psychological measurement, 58*(3), 357-381.

Fraley, R. C., Waller, N. G. y Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*(2), 350-365.

González, A. y Barrientos, A. (2014). Adopción del condón masculino como método de planificación y de prevención de la infección por VIH. Centro Centroamericano de Población, San José, Costa Rica.

González-Betanzos, F.; Rivera-Heredia, M.E. y Padrós-Blásquez, F. 2015. Invarianza por sexo en la Escala de Detección del Transtorno de Ansiedad Generalizada (EDTAG). *Actualidades en Psicología, 29* (119), 141-151.

Gray-Little, B., Williams, V. S. y Hancock, T. D. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 23*(5), 443-451.

Hagman, B. T., Kuerbis, A. N., Morgenstern, J., Bux, D. A., Parsons, J. T. y Heidinger, B. E. (2009). An item response theory (IRT) analysis of the short inventory of problems-alcohol and drugs (SIP-AD) among non-treatment seeking men-who-have-sex-with-men: evidence for a shortened 10-item SIP-AD. *Addictive behaviors, 34*(11), 948-954.

Hays, R. D., Brown, J., Brown, L. U., Spritzer, K. L. y Crall, J. J. (2006). Classical test theory and item response theory analyses of multi-item scales assessing parents' perceptions of their children's dental care. *Medical care*, S60-S68.

Hays, R. D., Morales, L. S. y Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical care*, 38(9 Suppl), II28.

Hernández, G.; León, E. A. y Varela, M. C. (2009). Transmisión vertical del VIH. Por Rodríguez, C. M. y Bárcenas, H. B. (Ed.) *VIH/SIDA y salud pública: Manual para personal de Salud*. (75-80). Censida.

Hidalgo, M.D., López-Martínez, M.D., Gómez-Benito, J. y Guilera, G. (2016). A comparison of discriminant logistic regression and Item Response Theory Likelihood-Ratio Tests for Differential Item Functioning (IRTLRDIF) in polytomous short tests. *Psicothema*, 28(1).

Hooper, D., Coughlan, J. y Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Articles*, 2.

Instituto de Investigaciones Psicológicas-Universidad de Costa Rica (2017). Práctica en línea para la Prueba de Aptitud Académica. Disponible en <http://ori.ucr.ac.cr/sitio/practica-en-linea-paa>. [Accesado en: Agosto, 2017].

Jiménez, K. y Montero, E. (2013). Aplicación del modelo de Rasch, en el análisis psicométrico de una prueba de diagnóstico en matemática. *Revista Digital Matemática*, 13(1), 1-24.

Kang, T., Cohen, A. S. y Sung, H. J. (2009). Model selection indices for polytomous items. *Applied Psychological Measurement*, 33(7), 499-518.

Kim, S. H., Cohen, A. S., Alagoz, C. y Kim, S. (2007). DIF detection and effect size measures for polytomously scored items. *Journal of Educational Measurement*, 44(2), 93-116.

Kohli, N., Koran, J., y Henn, L. (2015). Relationships among classical test theory and item response theory frameworks via factor analytic models. *Educational and Psychological Measurement, 75*(3), 389-405.

Martínez, R. (1995). *Psicometría: Teoría de los test psicológicos y educativos. Madrid: Síntesis.*

Martínez, R: Hernández, M.J. y Hernández, M.V. (2014). *Psicometría: Teoría de los test psicológicos y educativos. Madrid: Editorial Alianza.*

Masters, G.N. (1982). A Rasch model por partial credit scoring. *Psycometrika, 47*, 149-174.

McKinley, R. L. y Mills, C. N. (1985). A comparison of several goodness-of-fit statistics. *Applied Psychological Measurement, 9*(1), 49-57.

Ministerio de Salud (2016). *Plan Estratégico Nacional (PEN) en VIH y sida 2016-2021.* San José, Costa Rica: Dirección de Planificación Estratégica y Evaluación de las Acciones en Salud, 2016.

Ministerio de Salud (2017). *Situación VIH y sida: Costa Rica 2002-2016.*

Morgan, G. B., Hodge, K. J., Wells, K. E. y Watkins, M. W. (2015). Are fit indices biased in favor of bi-factor models in cognitive ability research?: A comparison of fit in correlated factors, higher-order, and bi-factor models via Monte Carlo simulations. *Journal of Intelligence, 3*(1), 2-20.

Moss, A. R. y Bacchetti, P. (1989). Natural history of HIV infection. *Aids, 3*(2), 55-62.

Muñiz, J. (1997). *Introducción a la teoría de respuesta a los ítems. Madrid: Pirámide.*

Naumenko, O. (2014). *Comparison of Various Polytomous Item Response Theory Modeling Approaches for Task-Based Simulation CPA Exam Data.*

ONUSIDA. (2008). Directrices prácticas del ONUSIDA para intensificar la prevención del VIH: Hacia el acceso universal. Ginebra, Suiza: ONUSIDA.

ONUSIDA. (2011). Informe mundial de avances en la lucha contra el Sida. Directrices. Ginebra, Suiza: ONUSIDA.

ONUSIDA. (2013). Informe Mundial: ONUSIDA, informe sobre la epidemia mundial de sida 2013. Ginebra, Suiza: ONUSIDA.

ONUSIDA. (2014). Presentación de informes sobre los avances en la respuesta mundial al sida 2014: Desarrollo de indicadores básicos para el seguimiento de la Declaración política de las Naciones Unidas sobre el VIH y el sida de 2011. Ginebra, Suiza: ONUSIDA.

ONUSIDA. (2015). *El sida en cifras 2015*. Disponible en: [http://www.unaids.org/es/resources/documents/2015/AIDS by the numbers 2015](http://www.unaids.org/es/resources/documents/2015/AIDS%20by%20the%20numbers%202015). [Accesado en octubre 2017].

ONUSIDA. (2016). Informe de brechas en prevención: resumen. Ginebra, Suiza: ONUSIDA.

ONUSIDA. (2017). Hoja informativa: Últimas estadísticas sobre el estado de la epidemia de sida. Disponible en: <http://www.unaids.org/es/resources/fact-sheet> [Accesado en octubre, 2017]

OPS, 2004. La situación del VIH/SIDA en Costa Rica. Serie Análisis de Situación de Salud; n° 11. San José, Costa Rica.

Orlando, M., Sherbourne, C. D. y Thissen, D. (2000). Summed-score linking using item response theory: Application to depression measurement. *Psychological Assessment*, 12(3), 354-359.

Reise, S.P., Ainsworth, A.T. y Haviland, M.G. (2005). Item response theory: Fundamentals, applications, and promise in psychological research. *Current directions in psychological science*, 14(2), 95-101.

Samejima, F. (1969). Estimation of ability using a response pattern of graded scores. *Psychometrika Monograph*, No. 17.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), 461-464.

USAID. (2015). *Questionnaires: Household, woman's and man's. Demographic and Health Surveys Methodology*. Disponible en: [https://dhsprogram.com/publications/publication-dhsq7-dhs-questionnaires -and-manuals.cfm](https://dhsprogram.com/publications/publication-dhsq7-dhs-questionnaires-and-manuals.cfm). [Accesado en octubre, 2017].

WHO (2017). Global Health Observatory (GHO) data. Disponible en: <http://www.who.int/gho/hiv/en/> [Accesado en octubre, 2017].

## **II PARTE**

EFFECTO DE LOS COMPONENTES EDAD, PERIODO Y COHORTE SOBRE  
EL NIVEL DE LA FECUNDIDAD EN COSTA RICA: UN ANÁLISIS DE LAS  
ENCUESTAS DE 1976 A 2015

## INTRODUCCIÓN

Costa Rica cuenta desde hace varias décadas atrás con un registro de estadísticas vitales de gran completitud, que permite conocer el número de nacimiento que ocurren en el país y las principales características del parto y la madre. Además, en el país se han desarrollado en distintos momentos encuestas de fecundidad y sexualidad, aunque su periodicidad ha sido variable.

La trayectoria que tomó la fecundidad en Costa Rica puede ser bien descrita empleando los registros de estadísticas vitales, pero esta fuente provee poca información del patrón que siguió esa trayectoria según las diferencias socioeconómicas de las mujeres. En parte porque variables como el nivel educativo, zona de residencia, ingreso del hogar u ocupación no fueron registradas en todo momento, pero también porque existe un alto porcentaje (superior al 10%) de no respuesta. Ambas circunstancias bastan para preferir las encuestas de fecundidad y salud sexual como fuente de información para analizar los diferenciales socioeconómicos de la transición de la fecundidad en el país.

Las dos primeras encuestas datan de la segunda mitad de los años sesenta; en el área metropolitana en 1964 y el área rural en 1969. Posteriormente, en 1976 se desarrolló otra encuesta pero esta vez con representación a nivel nacional, como también lo fueron las encuestas realizadas posteriormente. La más reciente encuesta se llevó a cabo en el año 2015. La estrategia de muestreo fue similar en todas las encuestas, así como la estructura de los cuestionarios, la redacción de sus preguntas y las opciones de respuesta, lo que permite que estas encuestas sean comparables a lo largo del tiempo. De tal manera, a partir de estas fuentes, es posible seguir el curso que ha tomado la fecundidad en Costa Rica en las últimas décadas y establecer relaciones con otras variables también registradas en las encuestas y en los registros administrativos.

El nivel de la fecundidad en Costa Rica se ha asociado a distintos hitos, entre los que figuran la introducción de los métodos anticonceptivos y su libre comercialización (Gómez, 2009, pág. 3 y Rosero-Bixby y Oberle, 1989), la aplicación de la esterilización femenina como método anticonceptivo y la cesión que se dio a finales de los años noventa a las restricciones que se impusieron a esta práctica médica con fines contraceptivos (Carranza, 2009 y Fernández y Sequeira, 2011), la entrada en vigencia de la ley de paternidad responsable (Ramos, 2009), el deseo de familias más pequeñas entre las cohortes más jóvenes, el aumento de la participación de las mujeres en la fuerza laboral (Mora, 1999 y Rosero, Gómez y Rodríguez, 1982), una mayor escolarización de las mujeres (Behm y Guzmán, 1979; Bongaarts, 2003; Mora, 1999; Rosero, Gómez y Rodríguez, 1982 y Stycos, 1982) y el incremento de la población que reside en las zonas urbanas (Behm y Guzmán, 1979).

La fecundidad en Costa Rica ha venido disminuyendo a lo largo de las últimas décadas y su tendencia a la baja se mantiene hasta la actualidad. La tasa global de fecundidad (TGF), un indicador comúnmente empleado para describir el nivel de la fecundidad de una población, expresa el número de hijos que en promedio tendría una mujer al final de su etapa reproductiva si estuviese expuesta a las tasas de fecundidad por edad que se observan en un momento dado. Mientras que en 1962 la TGF era de 7.1, en el año 2018 esa cifra corresponde a 1.67 hijos por mujer (INEC, 2019).

El descenso de la fecundidad se ha dado en diferentes gradientes para distintos grupos poblacionales. Rosero-Bixby y Casterline (1995, pág. 43) relatan que el descenso de la fecundidad empezó entre las mujeres más educadas y residentes de la región urbana del país. Ese proceso, se extendió posteriormente a otros estratos poblacionales, lo que describen como un proceso de difusión por interacción social. Más recientemente, las últimas encuestas de fecundidad muestran que el número promedio de hijos que han tenido las mujeres varía según zona de residencia y nivel educativo, siendo menor entre las residentes de la zona urbana y las mujeres más educadas (Robles y González, 2012a y Ministerio de Salud, 2016).

Para describir el descenso de la fecundidad a partir de las encuestas que se repiten periódicamente, tres componentes relacionados con el tiempo pueden tener un efecto sobre una variable de interés: edad, periodo y cohorte. De tal manera, la fecundidad no solo se ve determinada por la edad cronológica de la mujer (efecto edad), sino también por las condiciones sociales, políticas y económicas del momento (efecto periodo) y de su experiencia previa a lo largo del curso de la vida (efecto cohorte). Las variables periodo y cohorte, suelen contemplarse como variables indicadoras, operando en el análisis como variables latentes, pues representan múltiples fenómenos no medibles o de los cuales no necesariamente se tiene información.

La incorporación de esos tres componentes en un modelo de regresión podría brindar información sobre las variables involucradas en la compleja dinámica que determina la fecundidad y dar algunas luces de fenómenos macrosociales que pudieron intervenir sobre ella, que están siendo aproximados por las variables periodo y cohorte.

Sin embargo, estas tres variables son linealmente dependientes, ya que, tras conocer dos componentes y combinarlos linealmente es posible encontrar el tercero. Los análisis que se deriven de estudios de este tipo deben lidiar con la relación perfecta entre estas variables, hecho al que se le ha denominado como el *problema de identificación* en los estudios de Edad-Periodo-Cohorte (EPC). Precisamente, el propósito de estas metodologías es distinguir y estimar estadísticamente el efecto puro asociado a cada componente (Yang y Land, 2013).

Se han propuesto distintas estrategias para lidiar con el problema de identificación, que van desde la exclusión de al menos uno de los componentes, la fijación de restricciones a los parámetros en modelos de regresión lineal o valiéndose de estrategias de codificación de las variables EPC que rompa la linealidad entre ellas, hasta los más recientes modelos propuestos por Yang y Land (2006) para analizar los datos mediante modelos lineales mixtos generalizados.

Este trabajo pretende identificar posibles tendencias de los componentes edad, periodo y cohorte sobre el nivel de la fecundidad en Costa Rica, empleando las seis encuestas de fecundidad realizadas entre 1976 y el 2015, en las que se entrevistó a mujeres en edad reproductiva, entre los 15 y 49 años de edad. Se analizó la relación entre el número de hijos tenidos con las variables EPC y la educación, el estado conyugal y la ocupación mediante un modelo lineal mixto generalizado con una distribución de Poisson.

Los resultados de este trabajo sugieren que los efectos de edad, periodo y cohorte están presentes. El número promedio de hijos aumenta con la edad, principalmente entre las mujeres jóvenes pero también disminuye paulatinamente con la edad, una tendencia común en la mayoría de sociedades contemporáneas. Los efectos de periodo y cohorte fueron relevantes solamente en ciertos momentos. El primero muestra una tendencia<sup>6</sup> decreciente, cuyo efecto fue significativo en 1976 y en los años más recientes, 2010 y 2015. El efecto de cohorte, presenta variaciones de magnitud modesta, con efectos estadísticamente significativos entre las cohortes de nacimiento 1926-1934, 1945-1959 y 1970-1974.

---

<sup>6</sup> A lo largo de este trabajo se empleará el término "tendencia" para describir los cambios observados para los efectos edad, cohorte y periodo de un momento al siguiente en el tiempo sin que necesariamente se trate de un análisis de tendencia ni de series temporales. Las variables cohorte y periodo se analizan en categorías, obedecen más bien a factores aleatorios, como se podrá apreciar más adelante en el texto.

## JUSTIFICACIÓN

Las teorías que intentan explicar la transición de la fecundidad, en Costa Rica y en otros países, han identificado un gran número de factores que determinan o actúan en conjunto con otros causales, desencadenando el descenso de la fecundidad. Por mencionar algunos, entre estos factores figuran el aumento del nivel educativo (Behm y Guzmán, 1979; Bongaarts, 2003; Mora, 1999), una mayor urbanización de la población (Leet, 1977; Cutright, 1983), el descenso de la mortalidad infantil (Angeles, 2010; Balakrishnan, 1978), la percepción de los costos-beneficios de tener un (más) hijo(s) (Angeles, 2010); el cambio en el número ideal de hijos, el deseo de limitar el tamaño de la familia y espaciar los nacimientos (van de Walle, 1992; Mason, 1992) y la oferta de métodos anticonceptivos (Cutright, 1983; Rosero, 1983). Diversos trabajos convienen en que el descenso de la fecundidad tiene una explicación multicausal, que ocurre bajo una especie de "sinergia social" (Rosero, 1983).

El patrón que tomó la fecundidad en Costa Rica durante la segunda mitad del siglo XX, se puede considerar atípico, lejos del esquema propuesto por las teorías demográficas que establecen una relación inversa entre la fecundidad y el desarrollo social y económico de las sociedades. De acuerdo con Gómez (2009) y Rosero-Bixby y Oberle (1989) en los años cincuenta la TGF (Gráfico 8) mostró un alza considerable, que no concordaba con la movilización económica y social que experimentaba el país. En ese entonces, Costa Rica alcanzaba un bajo nivel de mortalidad, un nivel relativamente alto de educación y un desarrollo económico y social elevado, pero la fecundidad no mostraba una tendencia a la baja, sino lo contrario (Gómez, 2009).

En la década siguiente, en los años sesenta, la fecundidad muestra un descenso rápido e importante (Gráfico 8), que se alejaba de lo que se observó en otras sociedades, cuando solamente Taiwán y Singapur mostraron un descenso tan rápido en un periodo similar.

A raíz de este patrón atípico de la fecundidad, se cuestionó tanto la naturaleza y la calidad de los datos y del registro de estadísticas vitales, la composición etaria en las mujeres, su

condición conyugal y hasta el impacto que tendría sobre la fecundidad los movimientos de planificación familiar que iniciaban en ese entonces.

Gómez (2009) logra identificar la idoneidad de los datos del registro de nacimientos para el análisis de la fecundidad y concluye que la baja de la fecundidad es real y no se debe a la falla de los registros o a errores en la estimación de la población base, además de identificar que ese descenso no es el resultado de cambios en la composición etaria de las mujeres en edad fértil.

En cuanto a la nupcialidad, Gómez (2009) relata que si bien la tasa de nupcialidad disminuía, lo hacía probablemente a favor de las uniones libres, dada la proporción de niños que nacía fuera del matrimonio. Según sus estimaciones, esta disminución en la proporción de las mujeres casadas explicaría a lo sumo un 25% de la reducción de la fecundidad. Por otra parte, la edad mediana al matrimonio disminuyó levemente durante la década de los sesenta, de tal manera concluye que la edad al casarse no puede ser la causa del descenso de la fecundidad pues un matrimonio más temprano tendría un efecto a la inversa sobre la fecundidad.

Para intentar explicar el descenso de la fecundidad, tras haber descartado la idoneidad de los datos, los cambios en la composición etaria de las mujeres y de la nupcialidad, resta analizar el uso de métodos anticonceptivos y las preferencias de fecundidad. Sin ser concluyentes, los diversos trabajos que relatan el rápido descenso de la fecundidad en Costa Rica señalan la importancia que pudo tener la introducción de la planificación familiar en el descenso de la fecundidad (Carranza, 2009; Gómez, 2009; Reynolds, 1973; Rosero-Bixby y Casterline, 1995 y Rosero-Bixby y Oberle, 1986). Sin embargo, Gómez (2009, pág. 27) señala que existe poca información confiable y organizada sobre las prácticas anticonceptivas a comienzos de la década de 1960.

A las teorías que relacionan la transición de la fecundidad con los cambios económicos y sociales, se suman las denominadas "*teorías de difusión*" que resaltan la posible contribución que pudo tener la "*difusión de innovaciones*" en el descenso de la fecundidad (Casterline, 2001; Rosero-Bixby y Casterline, 1993, 1995). De acuerdo con Casterline (2001; pág. 6) el descenso de

la fecundidad es la consecuencia del incremento en la prevalencia de actitudes y conductas que previamente eran extrañas o estaban ausentes en la población (de ahí el término innovación) pero cuya prevalencia aumenta como consecuencia de la divulgación de esas actitudes y conductas de algunos segmentos de la población hacia otros (de ahí el término difusión).

Desde esta óptica el descenso de la fecundidad no es simplemente la respuesta a cambios demográficos, económicos y sociales que experimentaron las sociedades, sino que refleja la difusión de ciertas actitudes (los costos y beneficios de tener un hijo), comportamientos (uso de métodos anticonceptivos), ideales (número deseado de hijos, matrimonio), conocimientos y valores.

De acuerdo con Palloni (2001) el proceso de difusión de innovaciones requiere de la interacción social, que se refiere a la interdependencia entre las decisiones de los individuos en lo que respecta a su comportamiento reproductivo, pues se postula que los cambios en el conocimiento y conductas de algunos individuos afecta la probabilidad de que otros puedan cambiar su conocimiento o sus propias conductas. Esta interacción social o estos efectos sociales operan en adición o por encima de otros determinantes del cambio reproductivo.

La idea fundamental tras la teoría de la difusión de innovaciones es que la dinámica de la difusión puede alterar el momento y el ritmo al que desciende la fecundidad, clasificándosele incluso como un factor causal (Casterline, 2001) pero la dirección y los medios por los que se difunden estas innovaciones entre los individuos y entre los grupos de la población depende de la interacción o las dinámicas sociales (Palloni, 2001). Por ejemplo, el uso de las pastillas anticonceptivas (innovación) inició entre las mujeres más educadas y de mayores recursos económicos. El uso de este método se difunde a través de la interacción social hacia otras mujeres de ese mismo estrato o hacia mujeres de otras clases sociales, quienes interactuaron con las primeras.

En el caso de Costa Rica, Rosero-Bixby y Casterline (1995) identifican la presencia de efectos de difusión: el descenso de la fecundidad se dio muy rápidamente en comparación con las transformaciones socioeconómicas y alcanzó a todos los estratos sociales. Además, la secuencia espacio-temporal en la transición indica un patrón de ordenamiento espacial que es sugestivo de contagio entre áreas vecinas: inició en el Valle Central y se propagó gradualmente a la periferia del país. Del mismo modo, los autores sugieren que el aumento de la educación y la participación de las mujeres en la fuerza laboral pudieron reducir el aislamiento social de algunas mujeres y expandir la tasa de interacción.

Aun cuando el descenso de la fecundidad en Costa Rica parece ser un fenómeno generalizado, su cronología, orígenes e intensidad varían entre los diferentes estratos de la población (Behm y Guzmán, 1979; Rosero, 1983 y Rosero, Gómez y Rodríguez, 1982). En su más temprana etapa, el descenso se dio principalmente entre las mujeres jóvenes, de 25 a 35 años (Gómez, 2009), de clase media-alta (Rosero, 1983) y más educadas (Behm y Guzmán, 1979). En una segunda etapa se incorporan al proceso los sectores campesinos y de menor condición socioeconómica (Rosero, 1983; Rosero, Gómez y Rodríguez, 1982).

Las encuestas de fecundidad y salud reproductiva que se realizaron en el país (c.f. Robles y González, 2012a) mostraron que la fecundidad ha sido, desde hace años, más elevada en ciertos grupos de la población: entre las mujeres con menor educación, de la zona rural del país, sin ocupación o con una poco calificada. En el presente, aun cuando tiende a observarse una convergencia entre los distintos grupos socioeconómicos, los datos de las dos últimas encuestas (2010 y 2015; c.f. Ministerio de Salud, 2016) todavía muestran diferenciales en la fecundidad y en una serie de indicadores como el uso de métodos anticonceptivos, el número deseado de hijos o el nivel educativo.

El estudio de la transición de la fecundidad ha cobrado importancia en las últimas décadas dadas las implicaciones que tiene su descenso en diversos aspectos macroeconómicos. La fecundidad se considera uno de los componentes demográficos que en el presente y en el futuro determinará, en mayor medida, la estructura etaria y el tamaño de las poblaciones

(Coale, 1986), por lo que el análisis de su trayectoria, la proyección de su nivel y la identificación de los fenómenos asociados a sus variaciones permite entender las consecuencias económicas y sociales vinculadas a la baja de la fecundidad.

El descenso de la fecundidad durante un periodo extenso, conduce a una serie de efectos sociales y económicos que surgen principalmente del cambio de la estructura por edades de la población (Reher, 2011), que también se ve modificada a causa de la reducción de las tasas de mortalidad y el consecuente aumento de la esperanza de vida (Bravo, 1999; CEPAL, 2009).

Los cambios en la estructura por edades tienden a producir un impacto importante sobre el proceso de desarrollo económico, dado que el comportamiento económico de las personas varía según la etapa del ciclo de vida en la que se encuentran (Saad, Miller y Martínez, 2009).

Este cambio en la estructura etaria se refleja en el envejecimiento poblacional, que genera presión sobre los sistemas de salud y de pensiones (Bravo, 1999) al tener un mayor número de personas demandando tales servicios y en el descenso del número de habitantes en edades productivas (Coale, 1986), quienes sostienen económicamente los sistemas de pensiones y se encargan de las tareas que cada sociedad requiere para el desarrollo, sin dejar de lado las tareas de cuidado que los adultos mayores necesitan (Cepal, 2008).

Los países que han alcanzado regímenes de baja o muy baja fecundidad<sup>7</sup> han implementado políticas pronatalistas que van desde las transferencias económicas a las familias que procrean un hijo, apoyo en su cuidado, licencias de maternidad y flexibilización de horarios laborales (McDonald, 2006) u otras que promueven la inmigración (Bouvier, 2001) con el fin de compensar una estructura etaria envejecida.

---

<sup>7</sup> Se habla de regímenes de baja fecundidad cuando la TGF se encuentra por debajo de los 2.1 hijos por mujer y de muy baja fecundidad cuando esta cifra se encuentra por debajo de 1.5 (Billari y Kohler, 2004 y McDonald, 2006). En el segundo caso, el umbral de 1.3 también se menciona en la literatura (Billari y Kohler, 2004).

La estimación de la fecundidad y el conocimiento de su trayectoria también son importantes en otros sentidos. La fecundidad es un elemento fundamental del crecimiento poblacional, de tal manera que las tasas de fecundidad son uno de los componentes necesarios para obtener estimaciones y proyecciones de la población (Lundquist, Anderton y Yaukey, 2014 y Preston, Heuveline y Guillot, 2001).

La estimación de la población por edades específicas es la piedra angular de la epidemiología, la demografía y del análisis económico; es insumo para la planeación de diversas actividades de múltiples sectores de la sociedad. Como denominador de la mayoría de los indicadores, esas estimaciones impregnan todos los aspectos de la comprensión de la salud y el desarrollo. Los errores en las estimaciones poblacionales afectan al seguimiento de los objetivos nacionales e internacionales y a los análisis cronológicos, nacionales y regionales de los indicadores del desarrollo. La importancia de contar con estimaciones demográficas precisas para la planificación gubernamental no debe subestimarse: el tamaño de la población y la composición etaria dictan las necesidades nacionales de infraestructura, vivienda, educación, seguridad, empleo, atención de la salud, atención de los adultos mayores y de la prestación de muchos otros servicios (GBD, 2018).

Del mismo modo, las tasas de fecundidad, tanto por edad materna como en general, son factores clave del crecimiento de la población pero también importantes resultados sociales por sí mismos. Las estimaciones de fecundidad, además de afectar las tendencias demográficas, se utilizan para supervisar la prestación de servicios de salud reproductiva (GBD, 2018).

El descenso de la fecundidad no solo tiene manifestaciones a nivel macrosocial, sino que también lo hace en el plano individual. La disminución del número de hijos dentro de las familias se ha asociado con la postergación de la fecundidad (Gustafsson, 2001 y Timaeus y Moultrie, 2008), en parte a causa de la emancipación femenina que se ha gestado en algunas sociedades durante las últimas décadas (Phan, 2013; Prata, 2017 y Upadhyay et al, 2014).

La edad materna está asociada a la fecundidad: la función ovárica y la calidad de los ovocitos decrece con la edad, mientras que aumenta la frecuencia de anomalías cromosómicas (Inoue et al., 2013). En consecuencia, las tasas de fecundación, desarrollo y embarazo disminuyen, mientras que la tasa de abortos espontáneos aumenta (te Velde y Pearson, 2002).

El aplazamiento de la fecundidad implica que las parejas deciden tener hijos cuando la fertilidad de las mujeres se encuentra en declive (te Velde y Pearson, 2002), lo que deriva en una serie de complicaciones para la madre y su hijo: diabetes gestacional, hemorragia obstétrica, mayores tasas de partos inducidos y cesáreas, sufrimiento fetal, prematuridad, bajo peso al nacer y mortalidad perinatal, mayor presencia de anomalías congénitas (principalmente síndrome de Down), entre otros aspectos relacionados con la salud materna y del recién nacido (Almeida, Almeida y Pedreira, 2015; Koo et al, 2012 y Milner, Barry-Kinsella, Unwin y Harrison, 1992). En la medida en que un mayor número de mujeres tenga hijos a edades tardías y se presente un aumento de las madres e hijos que sufra tales complicaciones, se observaría una presión adicional sobre los servicios de salud.

La postergación de la fecundidad hacia edades reproductivas avanzadas puede tener efectos no deseados entre las mujeres que inician la reproducción tardíamente, pues su tiempo de exposición al embarazo se acorta, lo que condiciona la probabilidad de que las mujeres tengan el número previsto de hijos (Cabella y Nathan, 2018) o que finalicen su ciclo reproductivo sin hijos, pero con el deseo de tenerlos.

No todas estas manifestaciones del descenso de la fecundidad son desfavorables: la disminución del número de hijos suele ser la consecuencia de una mayor equidad de género, del control casi perfecto de la anticoncepción, de la expansión educativa y de mayores oportunidades laborales para las mujeres (Cabella y Nathan, 2018).

En Costa Rica, si bien la tasa de fecundidad ha mostrado una tendencia descendente a lo largo de las últimas cinco décadas, también ha mostrado periodos donde, aun cuando continuó

cayendo, no lo hizo al mismo ritmo que en años previos. Se observan momentos de estancamiento e inclusive aumentos, aunque leves (Gráfico 8).

El curso que tomará en el futuro la fecundidad es desconocido y aún no se dispone de una teoría con validez general, incluso cuando se pone de manifiesto la necesidad de contar con información al respecto y la utilidad práctica de proyectar su curso (Lee, 1974). Ni aún para los países que algunas décadas atrás experimentaron la transición de la fecundidad y alcanzaron una TGF inferior al nivel de reemplazo<sup>8</sup> (2.1 hijos por mujer) se conoce el curso que tomará la fecundidad después de esa transición a regímenes de baja fecundidad (Bongaarts, 1999, 2001; 2002, 2008; Chesnais, 1998).

El curso que tomará la fecundidad en el futuro será predecible y virtualmente modificable en la medida en que se comprendan los factores que la determinan. Es decir, para comprender cabalmente el descenso de la fecundidad en Costa Rica, es necesario contar con información no solo del nivel de la fecundidad en un momento específico, sino también de las diversas variables que la determinan. Además, para estudiar la trayectoria que ha presentado, es necesario contar con información que se haya registrado a lo largo del tiempo.

En estudios de carácter longitudinal, es decir, aquellos que recopilan información de individuos a lo largo del tiempo, cada medición de la que se haga inferencia en un punto específico en el tiempo está influenciada por tres factores: (1) edad (tiempo desde el nacimiento hasta la fecha de la medición), (2) periodo (momento en el que se realiza la medición) y (3) cohorte (grupo de sujetos que nacen en el mismo año) (Twisk, 2013).

El nivel observado de la fecundidad depende no solo de la edad de la mujer (efecto edad), sino también de su experiencia previa dada la cohorte a la que pertenece (efecto cohorte) y de las condiciones sociales y económicas del momento (efecto periodo) (Johnson, 1985) que dan a las parejas incentivos para limitar su fecundidad.

---

<sup>8</sup> El nivel de reemplazo se refiere a la fecundidad mínima necesaria para que una población cerrada (que no tiene una corriente inmigratoria o emigratoria) mantenga su tamaño en el tiempo. Su valor suele establecerse en 2.1 hijos por mujer.

Si bien los componentes Edad-Periodo-Cohorte pueden ser identificados o registrados para un grupo de individuos, el análisis de este tipo de información es complejo, pues existe una dependencia lineal entre estos componentes y no se puede estimar el efecto “crudo” de cada uno, ya que su efecto se confunde con el de los componentes restantes (Twisk, 2013). Por ejemplo, para cada edad y año calendario, existe una única cohorte de nacimiento:  $\text{edad} = \text{periodo} - \text{cohorte}$ .

Esto puede verse como un caso especial de regresores altamente correlacionados, que origina una matriz singular en el proceso de estimación. Dado que la matriz singular produce múltiples estimadores de los tres efectos, es difícil estimar el efecto único de cada componente por separado (Yang y Land, 2013, pág. 61).

La estrategia metodológica que clásicamente han tomado diversos autores para estudiar la fecundidad a lo largo del tiempo ha dejado de lado la confluencia de esos tres componentes. Más bien los diferentes trabajos que estudian el caso de Costa Rica se han limitado a explicar el descenso en función de la edad, aspectos socioeconómicos y a la larga, de ciertos efectos de tipo periodo; al considerar aspectos como la oferta de los métodos anticonceptivos y la esterilización femenina o la entrada en vigencia de la ley de paternidad responsable.

Bell y Jones (2015) señalan que es importante entender cuál o cuáles de los componentes EPC están relacionados con los cambios en la salud, porque pueden tener distintas implicaciones en política pública en salud dependiendo de la interacción de estos componentes. Dado que los distintos datos sobre la fecundidad disponibles para Costa Rica sugieren que el nivel de la fecundidad se vio influenciado no solo por el componente etario, sino por otras variables que podrían ser captadas por los componentes periodo y cohorte, es prácticamente obligatorio desarrollar un análisis del tipo EPC para comprender la trayectoria de la fecundidad en Costa Rica.

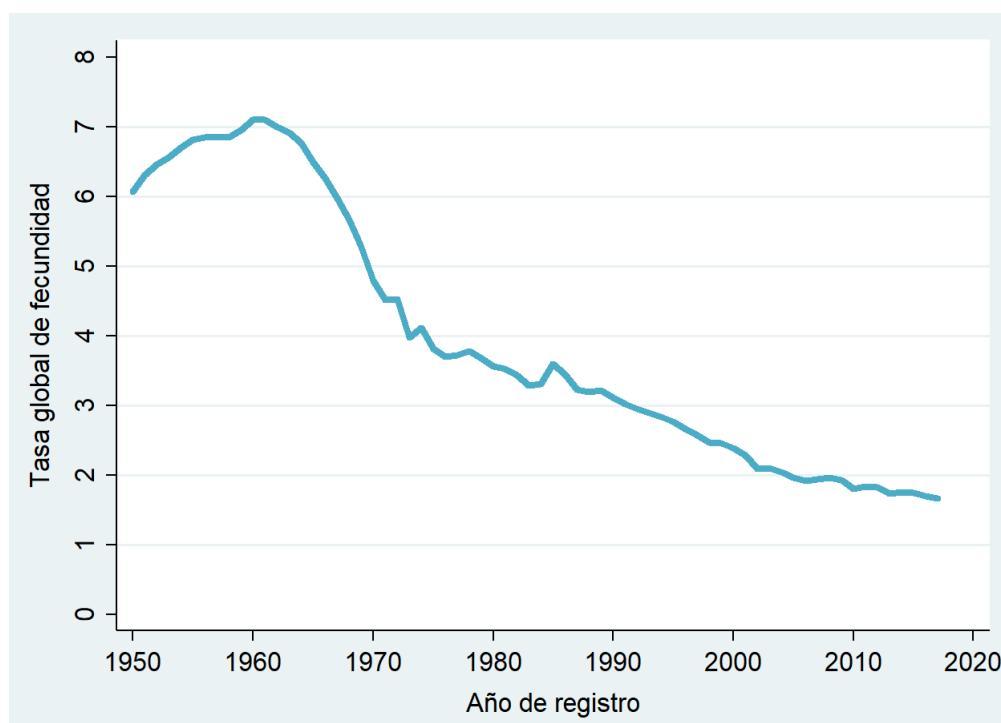
## MARCO DE REFERENCIA

### 1. El descenso de la fecundidad en Costa Rica

Las primeras estimaciones del nivel de la fecundidad en Costa Rica datan de los años cincuenta. Los datos del registro de nacimientos, encuestas y otras fuentes permiten describir la trayectoria de la fecundidad desde varias décadas atrás. El caso de Costa Rica se destaca como hito demográfico por dos razones. La primera, porque a finales de los años cincuenta la tasa de crecimiento natural era una de las más altas en el mundo, a causa de la baja mortalidad y la alta fecundidad: 3.8% entre 1955 y 1960. La segunda, porque a principios de los años sesenta Costa Rica experimentó una reducción abrupta de la tasa de fecundidad, que tampoco registraba precedentes en ningún otro país y era solamente comparable con las tasas observadas para Taiwán y Singapur (Carranza, 2009 y Gómez, 2009).

**Gráfico 8**

Tasa global de fecundidad, Costa Rica 1950-2017. Mujeres de 15-49 años



**Fuente:** INEC, 2018: Registro de Nacimientos de Costa Rica y CCP, 2018: Indicadores demográficos.

La TGF que se puede calcular a partir de los registros de nacimientos, desde los años cincuenta y hasta la actualidad, permite identificar distintos estadios (Gráfico 8) de la fecundidad en Costa Rica.

Durante la década de 1950, la fecundidad aumentó de 6 a 7 hijos por mujer (Gráfico 8). Rosero-Bixby y Oberle (1989) documentaron la trayectoria de la fecundidad desde esa época hasta el año 1989. Hacia 1960 la fecundidad marital rondó los 8 hijos por mujer, la tasa media de lactancia era de 6 meses y la prevalencia anticonceptiva de 16%. La corta duración de la lactancia explica en parte la elevada fecundidad y aunque la prevalencia anticonceptiva es baja, esta es indicativo de que existía un núcleo de mujeres para el que la anticoncepción era común, posiblemente la emergente clase media que empleaba métodos como condón, ritmo o retiro (Rosero-Bixby y Casterline, 1995).

A partir de la década de los años sesenta, aproximadamente a partir de 1962, la fecundidad empezó a descender. Este descenso dio inicio aceleradamente después de que se presentaran las tasas más altas registradas en el país: 7.1 hijos por mujer en 1960. En 1966 la TGF ya había descendido en casi un punto (6.3 hijos por mujer) y a partir de entonces se presentaron las disminuciones más importantes (Gráfico 8). Este rápido descenso se extendió hasta mediados de los años setenta, aproximadamente hasta 1976, cuando la TGF descendió hasta alcanzar los 3.6 hijos por mujer. Desde ese último año el descenso perdió fuerza, aunque su tendencia a la baja continúa incluso en la actualidad.

De acuerdo con Rosero-Bixby y Casterline (1995) la disminución acelerada de la fecundidad a lo largo de ese periodo (1962-1976) coincide con el inicio de la importación de anticonceptivos orales en 1962 y más adelante, en 1968, con la creación del programa de planificación familiar. Aunque antes del año 1962 ya algunas mujeres empleaban el control anticonceptivo, esta práctica no se había difundido al resto de la población. El detonante de este proceso de difusión podría aducirse a la introducción de los anticonceptivos orales en el sector comercial y a la atención prestada por los medios de comunicación a ese nuevo método (Thein y Reynolds, 1972). La píldora tuvo rápida difusión en un periodo corto, siendo adquirida principalmente a

través de médicos privados y sin receta médica en las farmacias (Gómez, 2009; Gómez, Raabe y Bermúdez, 1971).

Carranza (2009) relata que antes de los años sesenta la distribución de métodos anticonceptivos y consejería sobre su uso se brindaba principalmente en la capital del país, únicamente en la consulta médica privada y entre la clientela más adinerada. Documenta, además, que a finales de los años cincuenta el representante de una casa farmacéutica mexicana promocionaba pastillas anticonceptivas entre los médicos, quienes a su vez las distribuían a sus clientes.

La rápida acogida que tuvieron los métodos anticonceptivos pareciera coincidir con el hecho de que, aun cuando tenían familias grandes, las mujeres realmente querían familias de tamaño moderado (Gómez, 1968). La Encuesta Mundial de Fecundidad en la República de Costa Rica realizada en 1964 en la parte urbana del área Metropolitana de San José permitió determinar que entre las mujeres el tamaño deseado de la familia era de 3.9 hijos. En 1969 la Encuesta de Fecundidad realizada en la zona rural del país reveló que el número deseado de hijos era de 4.6 hijos, en promedio (Rosero, Gómez y Rodríguez). Aunado a esto, entre 1962 y 1976, la prevalencia del uso de anticonceptivos aumentó de 16% a 64% (cf. Rosero-Bixby y Oberle, 1989 y Rosero-Bixby et. al., 1984).

Otros hechos importantes que se asocian con el descenso de la fecundidad en este primer periodo son citados por Carranza (2009). A partir de 1962 la Alianza Evangélica Costarricense organizó las Caravanas de Buena Voluntad, un movimiento en el que se ofrecía consejería sobre los medios anticonceptivos, vasectomías y otros servicios médicos en áreas rurales alejadas. En 1963, Henry Tschinkel un ingeniero forestal y profesor del Instituto Interamericano de Ciencias Agrícolas (IICA), preocupado por los efectos negativos que representaba el crecimiento poblacional sobre el ambiente, inició un movimiento de planificación familiar en Turrialba que no solo tuvo impacto a nivel local, sino también a nivel nacional. Tschinkel introdujo entre los médicos turrialbeños y mujeres el anillo vaginal (*zipper ring*) y también proveyó una píldora anticonceptiva (Anovlar) recientemente lanzada al mercado. Más tarde, en

1966, Alberto González, también ingeniero estudiante del IICA, en conjunto con otras personas involucradas en el movimiento de planificación familiar fundaron la Asociación Demográfica Costarricense (ADC).

La ADC inició su labor en 1966 de tal manera que la promoción, consejería y servicios de planificación familiar que puso en marcha tras su fundación, empezó cuando la fecundidad ya venía descendiendo. La actividad privada de planificación familiar ya era fuerte para ese momento (Gómez, 2009). Tras la fundación de la ADC el uso de las pastillas anticonceptivas recibió un fuerte impulso, a través de sus programas de planificación familiar que facilitaron la difusión de la práctica anticonceptiva, brindando acceso a la información y a métodos anticonceptivos entre las mujeres de zonas rurales y semi-urbanas del país.

La primera fase del descenso de la fecundidad en Costa Rica no necesariamente puede achacarse a los programas oficiales de planificación familiar promulgados por la ADC y otras instituciones. De acuerdo con Rosero (1983), entre 1969 y 1976, época que coincide con la rápida expansión del programa de planificación familiar, en la zona urbana del país la prevalencia anticonceptiva se estimó entre un 60 y 65% en 1969 y en un 73% en 1976. En la zona rural, por otra parte, la prevalencia anticonceptiva pasó de 24% a 64% en esos siete años. Estas diferencias según zona de residencia denotan que la anticoncepción ya era una práctica frecuente en la zona urbana antes que en la rural.

En 1968, se empezaron a distribuir métodos anticonceptivos en las Farmacias del Ministerio de Salud. En 1970, prácticamente todas las instancias de este Ministerio contemplaban el suministro de anticonceptivos entre los servicios que brindaba a la población. La Caja Costarricense del Seguro Social incorpora la planificación familiar en sus consultas también en 1970 y en ese mismo año, el Ministerio de Educación incorpora dentro del plan de estudios escolar la planificación familiar como un componente de la educación sexual (Carranza, 2009).

Tras la rápida caída de la fecundidad en ese periodo, el descenso de la fecundidad tuvo distinta intensidad en las décadas posteriores. Durante el periodo siguiente, desde 1977 y hasta cerca

del año 1989, la TGF se mantuvo aproximadamente constante. El descenso de la fecundidad fue poco notable, pues la TGF pasó de 3.7 a 3.2 entre 1977 y 1989. Rosero-Bixby y Casterline (1995) señalan que durante ese periodo la tasa de prevalencia anticonceptiva osciló entre 65 y 69%. El tamaño deseado de la familia permaneció relativamente estable, desde la mitad de los años cincuenta, en torno a 3.1 hijos por mujer en el área urbana y a 4.1 hijos en la zona rural, cosa que se alcanzó a mediados de los años setenta (Robles y González, 2012a, pág. 19).

Sin embargo, a partir del año 1990 la TGF mostró que el nivel de la fecundidad disminuía nuevamente, pasando de 3.2 en ese año hasta alcanzar el nivel de reemplazo en el 2002 (Robles y González, 2012a). La fecundidad se encuentra actualmente en su nivel más bajo registrado: 1.67 hijos por mujer, en el año 2018 (INEC, 2018).

En ese periodo se destacan dos hechos importantes. La promulgación del Decreto Salud Reproductiva N° 27913-S del 9 de junio de 1999 que permitía la esterilización con fines contraceptivos y la entrada en vigencia de la Ley de Paternidad Responsable (N°. 8101) del 16 de abril de 2001.

La práctica de la esterilización femenina, sin embargo, encontró distintos obstáculos en diversos momentos y ha sido probablemente uno de los métodos más debatidos en Costa Rica. Solo hasta 1999, con la promulgación de dicho decreto, su provisión se dio libremente.

Aunque no existe una fecha clara a partir de la cual se empezó a practicar la esterilización femenina, con fines anticonceptivos, algunos estudios señalan que pudo ser a inicios de la década de los años cuarenta (Carranza, 2006), o bien, a principios de los años sesenta (Madrigal, 1987). La esterilización era una práctica relativamente común durante la década de los sesenta y su uso se venía incrementado, en parte a causa de la expansión de los servicios de salud estatales y de la creciente aprobación social a la limitación de la familia (Carranza, 2006). La creación de nuevos establecimientos de salud durante los setenta a lo largo de todo el territorio nacional y la inclusión de la esterilización entre los servicios que brindaba el seguro social permitió que las mujeres tanto de la zona urbana como de la zona rural, de diferentes estratos sociales y niveles de instrucción tuvieran acceso a esta cirugía (Gómez, 2009).

Al inicio, cuando el servicio médico de esterilización femenina se constituyó en los centros hospitalarios, no existían regulaciones de carácter nacional y la aprobación de esta cirugía estaba sujeta al criterio de médicos y comités de profesionales en salud. Las regulaciones que se establecían, generalmente se referían al número de hijos tenidos, cifra que variaba entre los distintos hospitales, pero que rondaba entre los cinco y los siete. En ese entonces, el número de esterilizaciones en las instituciones públicas era relativamente reducido y las usuarias de esta práctica eran mujeres de estratos socioeconómicos bajos, con muchos hijos y en muchas ocasiones con patologías asociadas, por lo cual la esterilización parecía poco cuestionable. Las esterilizaciones en mujeres jóvenes, sanas, con pocos hijos, cuya necesidad hubiera sido difícil de justificar generalmente correspondía a mujeres de mayores recursos económicos, de tal manera que la cirugía se resolvía principalmente en centros médicos privados (Carranza, 2006).

Sin embargo, a partir de la década de los setenta inició un debate sobre la moralidad del uso de la esterilización como método anticonceptivo, ligado principalmente al aumento de su demanda y de la proporción de mujeres que se esterilizaban. Por parte de la iglesia católica y otros grupos religiosos la crítica se centró en el control que se daba a la voluntad divina, pues desde estas agrupaciones se pensaba que las mujeres o las parejas debían tener “los hijos que Dios les enviase”, cosa que se debatía también con otros métodos anticonceptivos, pero en este caso se sumaba la irreversibilidad de esta práctica. Otros detractores mencionaban la posibilidad de que tras ser esterilizadas, las mujeres adoptaran una vida sexual libertina (Fernández y Sequeira, 2011, Rosero-Bixby y Casterline, 1995).

Por parte de la sociedad civil y de la prensa se rumoraba que existía una campaña masiva de esterilizaciones (Carranza, 2006; Gómez, 1983). Estas acusaciones llevaron a la Asamblea Legislativa costarricense a conformar, en junio de 1976, una comisión especial de investigación para indagar en torno al asunto y aunque su conclusión fue que no existía tal campaña, el solo hecho de que se diera una investigación tuvo repercusiones importantes. Sus efectos se vieron gestados principalmente en una reducción considerable del número de esterilizaciones y en la

magnificación del sentimiento de persecución que se generaba entre los médicos por esterilizar mujeres (Carranza, 2006).

Por parte del gremio de médicos y abogados la preocupación estuvo fundamentada en aspectos de carácter legal. El artículo 123 del Código Penal costarricense cita textualmente: *“Se impondrá prisión de tres a diez años, si la lesión causare una enfermedad mental o física que produzca incapacidad permanente para el trabajo, la pérdida de un sentido, de un órgano, de un miembro, del uso de un órgano o miembro, de la palabra, de la capacidad de engendrar o concebir”*. De tal manera, ante el aumento de las esterilizaciones y de la prácticamente nula legislación o reglamentación para la aplicación de esta práctica en Costa Rica, creció la sensación de vulnerabilidad entre los médicos quienes temían que tras el arrepentimiento o insatisfacción de la mujer o su marido se gestara una demanda penal en su contra, en acogimiento al mencionado artículo.

Así las cosas, de acuerdo con Gómez (1983), lo realmente nuevo en Costa Rica no era el uso de la esterilización con fines anticonceptivos, sino la rápida expansión de su práctica en todos los sectores de la población nacional.

Tras este movimiento, el Colegio de Médicos y Cirujanos estableció en agosto de 1976 su propia reglamentación, que posteriormente, en 1988, se convirtió en un Decreto Ejecutivo. Ambos documentos eran prácticamente iguales y normaban las condiciones bajo las cuales podría llevarse a cabo la cirugía detallando una lista finita de patologías, entre las que se mencionan varices, enfermedades genéticas y hereditarias, entre otras, y solo ante las cuales se justificaría una esterilización terapéutica. En el decreto de 1988, sin embargo, desaparecen las condiciones sobre el número de hijos y la edad de la mujer lo que flexibilizó la posibilidad de acción de los médicos. La esterilización anticonceptiva se maquillaba de esterilización terapéutica, buscando una patología que tuviera cabida en el decreto.

Otra de las restricciones que comúnmente se esquivaba era el consentimiento del marido. Además de los requisitos sobre el número de hijos y la edad, a la mujer se le solicitaba una

carta de su esposo o compañero, quien debía firmar dando su consentimiento y autorización a la cirugía. Aunque era un requisito que no estaba escrito expresamente en el decreto, si no que se pedía la firma de un interesado, por razones culturales se solicitaba la firma de la pareja (Fernández y Sequeira, 2011). En ocasiones quienes firmaban este consentimiento eran los mismos médicos, amigos o parientes cercanos de la mujer, principalmente cuando se trataba de mujeres solteras (Carranza, 2009).

Dado el elevado número de esterilizaciones que se realizaban anualmente y la proporción alta de mujeres esterilizadas, no cabe duda que una fracción significativa de ellas habían tenido una motivación básicamente anticonceptiva (Gómez, 1983). Entre 1998 y 1999, antes de la entrada en vigor del Decreto N° 27913-S se realizaron alrededor de 23200 salpingectomías, cinco años después esa cifra se incrementó en un 53%, lo que denota la importancia de ese decreto para dar a muchas mujeres la oportunidad de operarse (Fernández y Sequeira, 2011).

Desde el año 1993 hasta la actualidad, la esterilización es el método anticonceptivo más utilizado entre las mujeres en unión (c.f. Robles y González, 2012b y Ministerio de Salud, 2016). El porcentaje de mujeres en unión que estaban esterilizadas era de 14.7% en 1976, 16.7% en 1986, 19.7% en 1993, 21.3% en 1999, 29.6% en 2010 y 25.5% en 2015, según datos de las encuestas realizadas en esos años.

Como se mencionó anteriormente, la entrada en vigencia de la Ley de Paternidad Responsable (N° 8101, del 16 de abril de 2001), fue otro hecho de relevancia que tuvo lugar en la última fase del descenso de la fecundidad. Esta ley se creó en respuesta al alto número de niños que nacían sin un padre declarado. En el año 2000 el 31.1% de niños nacidos en Costa Rica no tenían un padre declarado (INEC, 2019). Con esta ley se busca garantizar la filiación legal del padre con sus hijos y por ende su participación en su manutención, obligando a los padres a realizarse una prueba de ADN si se rehúsan a reconocer al niño o tienen dudas sobre la paternidad. En el año 2002, después de su entrada en vigencia, el porcentaje de nacimientos sin padre declarado pasó a 8.3% (INEC, 2019).

En Costa Rica, incluso antes de la entrada en rigor de la Ley de Paternidad Responsable, los derechos de los niños estaban protegidos por la legislación costarricense, tanto por la Ley de Pensiones Alimentarias (Nº 7654) como por el Código de Familia (Ley Nº 5476). La Ley de Pensiones Alimentarias establece fuertes mecanismos para que los deudores cumplan con el pago de la pensión. Entre las medidas que se imponen al padre (o deudor alimentario) están el posible embargo salarial, en caso de incumplimiento, a quienes cuenten con una fuente de ingresos regular. En caso de incumplimiento también puede fijarse una pena de prisión si el deudor no es asalariado. Además, figuran restricciones migratorias, salvo que el padre obtenga autorización expresa de la madre o si ha garantizado el pago de, por lo menos, doce mensualidades de cuota alimentaria, el aguinaldo y la totalidad del salario escolar (cf. Asamblea Legislativa, 1996). Si bien la mensualidad por pensión alimentaria puede fijarse a discreción de un juez, este monto ronda entre el 25% y 35% del ingreso del padre y hasta el 50% cuando se trata de varios hijos (Ramos-Chaves, 2010).

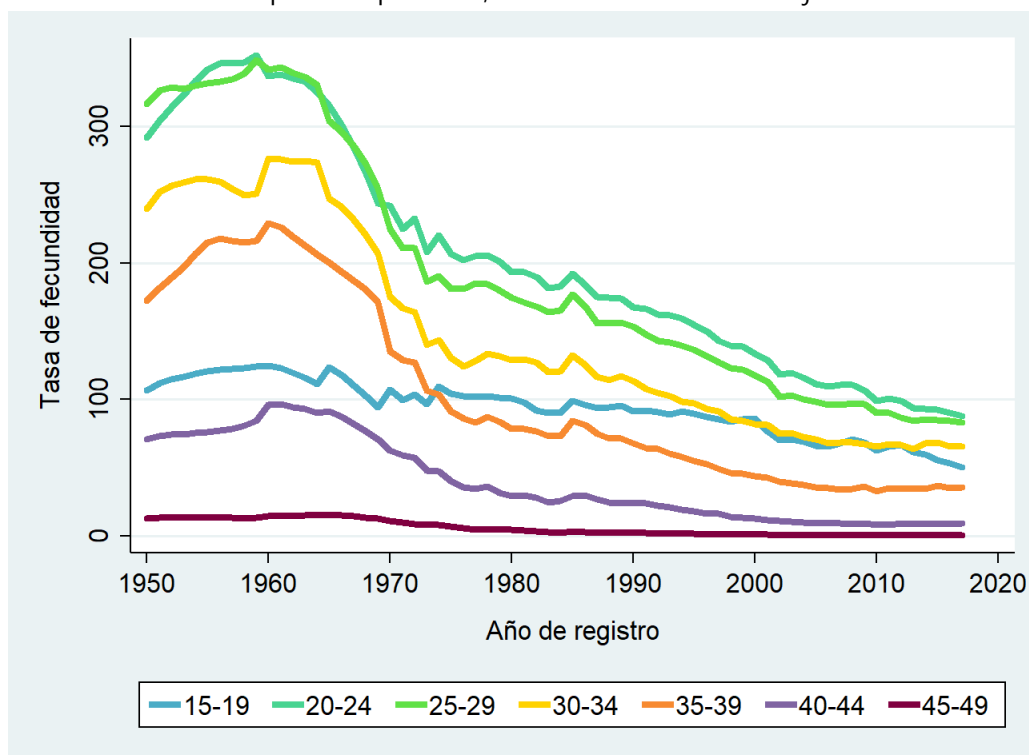
Antes de la entrada en vigencia de la ley, la búsqueda de la filiación parental o la realización de una prueba de ADN implicaban un proceso judicial. Lo segundo solamente se podía obtener mediante una orden judicial, que sería otorgada si la madre presentaba pruebas o testigos de la relación con el presunto padre (cf. Asamblea Legislativa, 1973). Todo este proceso podía tardar hasta tres años, sin garantía de una resolución favorable para la madre, razón por la cual muchas de ellas dimitían del proceso (Camacho, 2006).

Después de la ley, basta con que la madre presente una solicitud ante un Juzgado de Familia para que se solicite la prueba de ADN. Los niños tienen garantía del pago de una pensión alimentaria solo si existe filiación legal con su padre. Es en este sentido que la ley de paternidad responsable por sí misma no garantiza el apoyo económico del padre o una pensión alimentaria, si no, facilidades para conseguir la filiación legal, pero sugiere el fuerte impacto que pudo tener esa nueva legislación sobre la fecundidad, el matrimonio y la conyugalidad.

Ramos-Chaves (2010) determina que la ley tuvo un impacto inmediato en los índices de natalidad y en la TGF. La natalidad se redujo entre un 6% y 10% y la TGF se redujo de 2.4 a 2.2 hijos por mujer a causa de la entrada en vigencia de la ley. La fecundidad se redujo en todos los grupos etarios, aunque no ampliamente entre las mujeres mayores de 33 años. La reducción más notable se presentó en las mujeres de entre 23 y 27 años. Ramos-Chaves también señala que los nacimientos entre las mujeres casadas cayeron tanto o más, con respecto al total de nacimientos, cuando lo esperable es que no se presentaran cambios en ese primer grupo. De tal manera, los resultados sugieren que las relaciones sexuales premaritales eran habituales y ante un embarazo no planificado también se presentaban matrimonios forzados. En ese caso, el número de matrimonios forzados pudo caer en vista de que las mujeres no necesitan contraer matrimonio para recibir apoyo económico para sus hijos.

**Gráfico 9**

Tasas de fecundidad específicas por edad, Costa Rica 1950-2017. Mujeres de 15 a 49 años



**Fuente:** INEC, 2018: Registro de Nacimientos de Costa Rica y CCP, 2018: Indicadores demográficos.

El descenso de la fecundidad en Costa Rica desde 1950, se observó en todos los grupos poblacionales, aunque con diferente periodicidad e intensidad (c.f. Robles y González, 2012a). La fecundidad disminuyó en todos los grupos de edad, aunque la baja de la fecundidad en Costa Rica se caracterizó por concentrarse entre las mujeres jóvenes, entre los 20 y 35 años (Gráfico 9). Esto se contrapone a lo que se observaba en los países desarrollados, en los países occidentales, en Taiwán y Corea, donde la reducción de la fecundidad se daba en las mujeres de edades más avanzadas (Gómez, 2009). De acuerdo con Robles y González (2012a) el único grupo de edad que se aparta del patrón de descenso es el de 15 a 19 años. El descenso de la fecundidad en ese grupo etario se detiene en el año 1969 y no continúa sino hasta la década de los noventas.

## **2. Los estudios de Edad – Periodo – Cohorte**

En distintas disciplinas, aunque comúnmente en epidemiología y demografía, se registra información de individuos a lo largo del tiempo: de distintos sujetos con una característica de interés (medidas transversales) o de los mismos individuos de los que se toman varias mediciones a lo largo de un periodo más amplio de tiempo (estudios de panel o de corte longitudinal).

Cuando se tiene este tipo de información, el efecto de una determinada variable integra el efecto de tres componentes relacionados con el tiempo: edad, periodo y cohorte. La estimación del efecto de estos componentes puede, no solo ser útil para resolver preguntas de investigación sobre estos mismos, sino también para controlar su efecto sobre variables explicativas y afinar la estimación.

El efecto de la edad se refiere a variaciones asociadas a la edad cronológica de los sujetos. Refleja procesos biológicos u otros asociados al curso de la vida. Aunque el efecto de la edad suele ser el de mayor interés en la mayoría de los estudios, suele verse confundido por el de los componentes restantes.

Cuando se estima el efecto de la edad, controlado por otras variaciones de orden temporal o de carácter geográfico, se obtiene una medición que refleja la verdadera naturaleza de los cambios de esta variable. En contraste, los efectos de periodo y cohorte reflejan la influencia de fuerzas externas que operan en distintas direcciones (Yang, 2007).

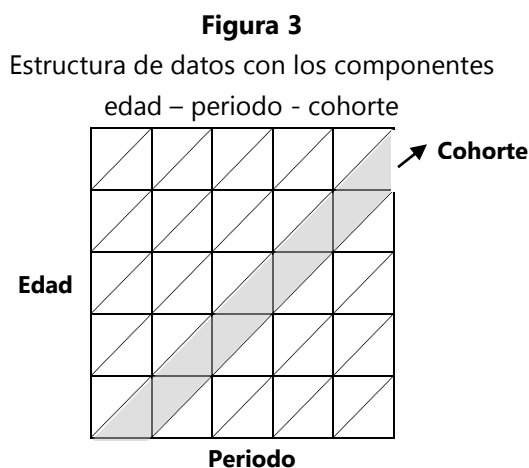
El efecto de periodo, por su parte, se refiere a las variaciones que corresponden a eventos que ocurrieron en un momento o plazo en el tiempo (años calendario), pero que afecta simultáneamente a todos los grupos etarios. Refleja por ende cambios sociales, económicos, políticos, ambientales, intervenciones políticas o de salud pública, entre otras.

Finalmente, el componente de cohorte se refiere al efecto derivado de grupos de individuos que experimentaron simultáneamente un evento inicial o se exponen a la ocurrencia de un evento a partir del mismo momento. Las cohortes por nacimiento son el caso más comúnmente analizado, donde los individuos que nacieron en un mismo periodo (año, quinquenio, etc.), viven los mismos cambios sociales e históricos a la misma edad. El efecto de cohorte refleja la acumulación de experiencias a lo largo de la vida y cada cohorte experimenta un curso de vida distinto.

Se tiene que diferentes cohortes poseen diferentes experiencias de vida, a causa de diferencias sociohistóricas y condiciones epidemiológicas. La sucesión de cohortes, cada una con distintas experiencias de vida, cambia la composición de la población y transforma la sociedad (Ryder, 1965 y Yang, 2007). Desde el punto de vista analítico, la cohorte puede verse estructuralmente como una categoría social que puede reflejar un estrato social (Yang y Land, 2013). Las cohortes que experimentan diferentes condiciones históricas y sociales difieren en su exposición a factores socioeconómicos, conductuales y ambientales. El efecto de cohorte representa el efecto de esos factores ante cuya exposición se inicia en una etapa temprana de la vida y actúan persistentemente a través del tiempo para producir diferentes condiciones de salud (Yang, 2007) y distinta probabilidad de ocurrencia de cierto evento entre cohortes específicas.

En el ámbito de la salud, el efecto de cohorte es particularmente relevante, pues eventualmente refleja diferencias entre las condiciones de vida al nacer, el desarrollo de enfermedades de larga latencia, o largos periodos de exposición a una enfermedad.

Gráficamente la estructura de los datos suele expresarse como se presenta en la Figura 3. En este arreglo rectangular las filas representan información organizada según edad, mientras que las columnas denotan alguna categorización para la variable periodo (momento en que se realiza la medición). Los elementos de la diagonal de esta matriz representan una cohorte<sup>9</sup>. Bajo este tipo de arreglo, la matriz de datos suele registrar cifras agregadas, es decir, indicadores como porcentajes, tasas o razones.



Solamente los estudios de tipo longitudinal pueden proporcionar información de una cohorte verdadera, pues registran información de los mismos individuos a lo largo de un lapso prologando de tiempo. Cuando se tienen datos de corte transversal, tomados en un momento concreto en el tiempo, se confunden los efectos de edad y cohorte. Sin embargo, este último diseño permite estudiar el efecto de dichos componentes mediante cohortes sintéticas (c.f. Mason y Fienberg, 1985).

Yang y Land (2013) describen los siguientes tres esquemas de datos para el análisis EPC:

<sup>9</sup> Se trata del área sombreada delimitada por la línea punteada.

- (1) **Tabla de tasas/proporciones ordenada por edad y periodo:** generalmente se analizan indicadores como tasas o proporciones (tasas de mortalidad, fecundidad, porcentaje de desempleo, entre otros). La información se organiza de manera tabular, siguiendo la estructura que se visualiza en la Figura 3. La edad se despliega en las filas de la tabla y la información sobre periodo en las columnas. Generalmente la información de estas variables se categoriza en grupos quinquenales o decenales y el efecto de cohorte se deriva de ambas.
  
- (2) **Encuestas periódicas por muestreo de corte transversal:** el análisis EPC valiéndose de esta fuente de información requiere que se compilen en un solo archivo los datos de las distintas encuestas, que deben guardar congruencia en el diseño muestral, población meta y redacción de las preguntas. Las mismas cohortes de interés se contemplan en cada ronda, pero normalmente se tienen distintos entrevistados entre encuestas quienes pertenecen a alguna de estas cohortes. De tal manera, el análisis desarrollado con esta fuente de información se basará en el estudio de cohortes sintéticas.

En comparación con el diseño de tipo longitudinal, que generalmente se desarrolla en un periodo de tiempo relativamente abreviado, el enfoque de cohorte sintética facilita el análisis tanto del efecto de edad como de periodo, ya que generalmente los datos provienen de una muestra representativa de todas las edades, desarrollada periódicamente y usualmente durante varias décadas. Además, la dificultad para ubicar a los entrevistados es menor, aunque no está exenta de la selección por mortalidad.

- (3) **Estudios prospectivos de cohorte o diseños longitudinales acelerados (de múltiples cohortes):** este tipo de diseño se basa en el estudio de cohortes verdaderas, es decir, siguiendo prospectivamente a los mismos individuos. A diferencia de los estudios prospectivos de cohorte, que siguen a una sola cohorte, los estudios longitudinales acelerados estudian a múltiples cohortes. Estos últimos permiten diferenciar el efecto de edad con el de cohorte ya que se pueden distinguir las

variaciones intra-individuos y las variaciones entre cohortes. Estos estudios presentan la desventaja de que su costo es más elevado que el de otros diseños y se ven afectados por la pérdida o mortalidad de los sujetos de estudio.

El modelaje del efecto de los componentes edad, periodo y cohorte sobre una variable de interés implica un problema de estimación, ya que los tres componentes son linealmente dependientes. Este *problema de identificación*, al que algunos autores han denominado como el "enigma" de los modelos EPC surge de la imposibilidad de desgranar el efecto "puro" de la edad, el periodo y la cohorte, dada su multicolinealidad perfecta, pues tras conocer dos componentes y combinarlos linealmente es posible encontrar el tercero.

La estimación por mínimos cuadrados o por máxima verosimilitud no puede desarrollarse directamente con datos de EPC, pues el problema de identificación impide que los coeficientes de regresión puedan definirse de manera única ya que existe un número infinito de posibles soluciones. Por lo tanto, no es posible estimar los efectos edad, periodo y cohorte a menos que se establezcan restricciones sobre los coeficientes.

Las primeras propuestas metodológicas para el análisis de datos EPC tuvieron vigor en los años setenta y continuaron durante la década de los ochenta (Mason et al, 1973; Glenn, 1976; Mason, Mason y Winsborough, 1976; Fienberg y Mason, 1979, 1985; Osmond, 1985; Osmond y Gardner, 1989), aunque se reconoce el trabajo de Frost (1939) como una de las primeras contribuciones en el tema. Los primeros modelos o esquemas de análisis se enfocan en técnicas que manejaban datos agrupados.

Las metodologías que se proponen posteriormente se valen del aumento de la capacidad computacional, de la mayor capacidad de almacenamiento y procesamiento de la información en las computadoras más actuales, que pueden incluso analizar datos individuales. En el periodo reciente Yang y Land (2006, 2008) y Yang (2006) proponen un esquema de análisis basado en modelos de efectos mixtos de tipo jerárquico para datos de estudios desarrollados

periódicamente, sin que necesariamente se trate de datos longitudinales y que además se pueden emplear con datos agregados.

Algunos de estos y muchos otros trabajos (Palmore, 1978; Kupper, Janis, Karmous y Greenberg, 1985; Ogata, Katsura, Keiding, Holst y Green, 2000; O'Brien, 2014) han debatido sobre la utilidad y desventajas de ciertos métodos para lidiar con el problema de identificación, sin que exista convención sobre la mejor o más apropiada metodología para solucionarlo y brindar los resultados más confiables.

## 2.1. El problema de identificación

Mason, Mason, Winsborough y Poole (1973) abordaron el uso del modelo de regresión lineal convencional para el análisis EPC, empleando datos agregados, es decir, tablas de tasas o proporciones de datos poblacionales; el primer esquema de datos empleados en el análisis EPC descrito en la sección 2.

El análisis EPC de datos tabulados adolece del problema de identificación, generado por la dependencia lineal exacta entre las variables EPC incorporadas al modelo. La alta correlación entre estas variables explicativas produce una matriz singular, haciendo imposible la estimación del efecto separado de los tres componentes.

Tal modelo de regresión lineal se escribe de la siguiente manera,

$$R_{ij} = \frac{I_{ij}}{P_{ij}} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ij} \quad (14)$$

Donde  $R_{ij}$  denota la tasa de incidencia observada para el  $i$ -ésimo grupo etario donde  $i=1,\dots,a$  grupos etarios, en el  $j$ -ésimo periodo para  $j=1,\dots,p$ ;  $I_{ij}$  denota la incidencia o el número de casos y  $P_{ij}$  representa al tamaño de la población en el grupo  $ij$ -ésimo, es decir, la población expuesta al riesgo,  $\mu$  es el intercepto o la tasa media de incidencia ajustada,  $\alpha_i$  denota el efecto

de la  $i$ -ésima fila o el coeficiente para el  $i$ -ésimo grupo etario,  $\beta_j$  denota el efecto de la  $j$ -ésima columna o el coeficiente del  $j$ -ésimo periodo,  $\gamma_k$  representa el efecto de la diagonal de la tabla o el coeficiente de la cohorte  $k$ -ésima y  $\varepsilon_{ij}$  es el error aleatorio con media  $E(\varepsilon_{ij}) = 0$  y  $var(\varepsilon_{ij}^2) = \sigma^2$ .

El modelo (14) puede expresarse como un modelo lineal generalizado de efectos fijos tras la siguiente reparametrización,

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_k \gamma_k = 0 \quad (15)$$

Con dicha reparametrización, el modelo (14) puede escribirse en forma matricial como una regresión de mínimos cuadrados:

$$Y = Xb + \varepsilon \quad (16)$$

Donde  $Y$  es un vector de la tasa de incidencia,  $X$  es la matriz de diseño que contiene variables dummy para el vector de parámetros  $b$ :

$$b = (\mu, \alpha_1, \dots, \alpha_{a-1}, \beta_1, \dots, \beta_{p-1}, \gamma_1, \dots, \gamma_{a+p-2})^T \quad (17)$$

Los parámetros  $\alpha_a, \beta_p$  y  $\gamma_{a+p-1}$  no se incluyen en el vector de parámetros  $b$  a causa de la restricción definida en (15).  $T$  se refiere a la transpuesta de una matriz o vector.

El estimador por mínimos cuadrados ordinarios de la matriz del modelo de regresión (16), que es también el estimador por máxima verosimilitud cuando se especifica el término de error siguiendo una distribución normal con media cero y varianza constante, es la solución  $b$  de las ecuaciones normales:

$$\hat{b} = (X^T X)^{-1} X^T Y \quad (18)$$

Sin embargo,  $(X^T X)^{-1}$  no existe, debido a que la matriz de diseño  $X$  es singular y por lo tanto existen un número infinito de soluciones a la ecuación (18). Por ende, el estimador por mínimos

cuadrados y máximo verosímil no existe. A este hecho se le conoce como el *problema de identificación* en el análisis EPC e implica que no es posible estimar el efecto separado de las variables edad, periodo y cohorte sin establecer una nueva restricción a los parámetros, adicional a la impuesta en (15).

## 2.2. Métodos para el análisis de datos EPC

En esta sección se describen algunas de las técnicas empleadas en el análisis de datos EPC, intentando mencionar sus ventajas y desventajas para el análisis de datos de este tipo. La primera generación de métodos se caracteriza por (1) analizar datos agrupados, principalmente cifras como razones, proporciones y tasas que se organizaban en tablas, (2) que se basan en establecer restricciones a los parámetros de los efectos periodo y cohorte o (3) intentan romper su dependencia lineal mediante algún esquema de codificación.

La sección 2.3 abordará con mayor detalle otro esquema de análisis: los modelos lineales mixtos generalizados. Estos modelos se describen ampliamente pues serán los empleados para analizar los datos de interés, dadas sus características: (1) se pueden analizar encuestas periódicas de corte transversal, (2) manejan datos individuales y (3) el modelo de regresión admite la inclusión de covariables.

### 2.2.1. Modelo reducido de dos factores

La idea detrás de este método es incluir uno o dos de los componentes edad, periodo o cohorte en el modelo de regresión. Generalmente se incluyen los componentes edad y periodo, excluyendo el de cohorte. Este modelo puede expresarse matemáticamente de la siguiente manera;

$$\log(E_{ij}) = \log\left(\frac{I_{ij}}{P_{ij}}\right) + \mu + \alpha_i + \beta_j \quad (19)$$

Donde  $E_{ij}$  es la tasa de incidencia esperada,  $I_{ij}$  es el número de casos y  $P_{ij}$  es la población expuesta al riesgo en la celda  $ij$ ,  $\mu$  representa el intercepto,  $\alpha_i$  el efecto de la edad y  $\beta_j$  es el efecto de periodo.

Cuando se consideran solo los componentes edad y periodo, se asume que el cambio en la tasa de incidencia es constante entre cohortes a lo largo del tiempo y que el efecto de periodo es independiente de la cohorte y desempeña el papel más importante en la explicación de estos cambios.

Hasta que los tres componentes Edad-Periodo-Cohorte son estimados simultáneamente, aún es incierto si las tendencias observadas se deben al efecto de alguno de ellos. La exclusión de alguno de estos componentes impide establecer inferencias sobre su efecto en el fenómeno estudiado. Por ejemplo, la omisión del efecto cohorte no necesariamente puede generalizarse. Diversos autores han identificado el importante rol que juega el efecto cohorte sobre algunos hechos demográficos (Fogel, 2003; O'Brien, Stockar y Isaacson, 1999 y Ryder, 1985), por lo que en ciertos contextos la omisión de este componente puede derivar en estimaciones ambiguas.

El efecto de cohorte puede entenderse como una forma de interacción entre los componentes edad y periodo. En el modelo planteado anteriormente, se parte del supuesto que tal interacción no existe, por lo que en una tabla en la que se reporten las tasas según edad y periodo, su efecto se obtiene únicamente en base a los efectos marginales de la  $i$ -ésima fila (edad) y la  $j$ -ésima columna (periodo).

La violación a este supuesto puede detectarse gráficamente al presentar las tasas específicas por edad según periodo (en escala lineal o logarítmica). Si las curvas no muestran paralelismo se tendrán indicios de la presencia del efecto de cohorte (Kupper, Janis, Karmous y Greenberg, 1985).

### 2.2.2. Modelos lineales generalizados restringidos

Deviasi (2018) y Yang y Land (2013) citan dos medios para restringir los coeficientes de un modelo de regresión a fin de evitar el problema de identificación: estableciendo restricciones sub-identificadas o restricciones sobre-identificadas. En los modelos del primer tipo se establece al menos una restricción sobre el vector de parámetros (18), siendo generalmente una restricción de igualdad. Es decir, forzando a dos grupos a tener el mismo efecto, como el de los dos primeros grupos de edad  $\alpha_1 = \alpha_2$  o el de los dos primeros periodos  $\beta_1 = \beta_2$ . Tras esta restricción adicional, el modelo (17) es sub-identificado y la matriz  $(X^T X)$  se vuelve no singular y por lo tanto existe un estimador por mínimos cuadrados.

Los modelos sobre-identificados establecen una restricción adicional al modelo cambiando el tamaño de los grupos según edad, periodo o cohorte, lo que generalmente se consigue incluyendo la edad en grupos desagregados y analizando las variables cohorte y periodo en grupos agregados (en pocas categorías).

Los principales problemas asociados a este método se derivan de la necesidad de contar con información externa que permita establecer alguna lógica sobre cuales coeficientes o parámetros restringir. Sin embargo, este tipo de información no suele estar disponible. Además, las estimaciones obtenidas de modelos sub-identificados son sensibles a la elección de los parámetros restringidos. A distintas alternativas en los parámetros restringidos se obtienen diferencias, a veces notables, en los parámetros estimados para las distintas categorías edad, periodo y cohorte. Del mismo, lo hace la categorización establecida para esas variables cuando se emplean los modelos sobre-identificados, pues las estimaciones son sensibles a las longitud del intervalo definido (Yang, Fu y Land, 2004). Todos los modelos sobre-identificados producirán la misma bondad de ajuste a los datos, de tal manera que las medidas de bondad de ajuste no pueden emplearse como criterio para seleccionar el mejor modelo (Yang y Land, 2013).

### 2.2.3. Función de penalización

Este método puede contemplarse como una forma alternativa al modelo de restricción de parámetros, descrito anteriormente. Osmond y Gardner (1982) y Decarli y La Vecchia (1987) sugirieron que se podía imponer una restricción que minimice una función de penalización, escrita en términos del parámetro  $\lambda$ , que permite corregir los parámetros del modelo EPC. Esta función de penalidad es una medida de la distancia entre los modelos de dos factores (edad-periodo, edad-cohorte y periodo-cohorte) y el modelo de tres factores (edad-periodo-cohorte). Robertson, Gandini y Boyle (1999) describen los siguientes pasos para estimar la función de penalización,

- (1) Estimar cada uno de los tres modelos de dos factores y guardar los tres conjuntos de parámetros estimados.
- (2) Estimar el modelo completo, edad-periodo-cohorte usando alguna restricción arbitraria para los parámetros, por ejemplo, suponer que el primer y último periodo tienen el mismo efecto.
- (3) Construir la función de penalización como la suma de cuadrados de las diferencias entre los parámetros de los tres modelos de dos componentes y el modelo completo, ponderado por una medida de bondad de ajuste de los tres modelos de dos factores, como por ejemplo la devianza.
- (4) Esta función de penalidad, es una función de  $\lambda$ , cuyo valor se obtiene tras minimizar la función sobre todos los posibles valores de  $\lambda$ .
- (5) Este valor de  $\lambda$  es utilizado para corregir los parámetros del modelo completo edad-periodo-cohorte, estimado en el paso (2), para conseguir las estimaciones identificadas.

### 2.2.4. Transformaciones paramétricas no lineales

Se puede lidiar con el problema de la identificación incluyendo una función no lineal de al menos uno de los componentes EPC, de tal manera que se rompa la relación lineal que existe entre ellos. Un caso típico suele ser el uso de la función cuadrática para analizar el efecto de edad. Para emplear este método es necesario que estas variables sean continuas para poder establecer funciones polinomiales.

Este método tiene dos desventajas notorias. Primero, no siempre se tiene claridad sobre la función polinomial o algebraica que deba aplicarse y además, incorporar esta función no necesariamente soluciona el problema de identificación. Según Mason y Fienberg (1985) el efecto lineal entre las variables EPC aún permanece indefinido.

### **2.2.5. Empleo de variables proxy**

Bajo este esquema una o más de las variables edad, periodo o cohorte es reemplazada por una o más variables *proxy*. Por ejemplo, la tasa de desempleo o el tamaño de la fuerza laboral pueden emplearse en sustitución del efecto periodo (Pavalko, Gong y Long, 2007) o el porcentaje de población negra (O'Brien, 2000), el tamaño relativo de la cohorte (O'Brien, Stokard y Isaacson, 1999) o los años promedio de consumir cigarrillos (Preston y Wang, 2006) en sustitución del efecto cohorte.

Este es un enfoque popular debido a su atractivo sustantivo, después de todo, las variables EPC se incluyen en el modelo como indicadores de una serie de factores no medidos (Yang y Land, 2013).

Cuando se emplea este enfoque las variables proxy no deben estar relacionadas linealmente con las variables EPC restantes, pues en ese caso la multicolinealidad aún estará presente. Por otra parte, si las variables proxy elegidas no tienen en cuenta la variación total de los factores que pretenden sustituir, dicho enfoque será insuficiente para obtener los efectos de edad, periodo y cohorte. No puede suponerse que toda la variación asociada a los componentes EPC está siendo totalmente contabilizada por las variables proxy seleccionadas.

### **2.3. Modelos lineales mixtos generalizados (MLMG)**

Yang y Land (2006) proponen un modelo jerárquico de efectos mixtos, fijos y aleatorios, para el análisis EPC cuando se emplean datos de encuestas de corte transversal, pero que se realizan

con determinada periodicidad, es decir, abordan el segundo esquema de datos descrito en la sección 2.

En esta sección se describen los modelos lineales mixtos generalizados (MLMG), también de carácter jerárquico. Esta descripción se basa principalmente en la metodología presentada por Yang y Land (2013) pues han adecuado estos modelos al análisis de datos Edad-Periodo-Cohorte. Sin embargo, esta sección se complementa con otras fuentes, como se detalla oportunamente.

Los MLMG son una extensión de los modelos lineales mixtos, pues permiten que la variable dependiente siga alguna de múltiples distribuciones, por ejemplo, binaria. Son también una extensión de los modelos lineales generalizados pues incluyen tanto los efectos fijos como efectos aleatorios. En su versión más simple, se expresan matemáticamente como:

$$y = X\beta + Z\gamma + \varepsilon \quad (20)$$

Donde  $y$  es un vector de dimensión  $N \times 1$  que contiene información sobre la variable dependiente,  $X$  es una matriz  $N \times p$  con la información de las  $p$  variables predictoras,  $\beta$  es un vector  $p \times 1$  con los coeficientes de regresión de los efectos fijos,  $Z$  una matriz de diseño  $N \times q$  para los  $q$  efectos aleatorios, que puede representar tanto variables continuas como dicotómicas al igual que  $X$ ,  $\gamma$  es un vector  $q \times 1$  de parámetros de efectos aleatorios desconocidos y  $\varepsilon$  es un vector  $N \times 1$  de los residuos o errores aleatorios.

Lo que diferencia a los MLMG es el empleo de una función de enlace que relaciona la variable dependiente  $y$  con el predictor lineal. Si se denota este predictor lineal  $\eta$  como la combinación de los efectos fijos y aleatorios, excluyendo los residuales, la ecuación (7) tomará la forma,

$$\eta = X\beta + Z\gamma \quad (21)$$

De tal manera, la esperanza condicional de  $y$  es la siguiente, donde  $g(\cdot)$  representa la función de enlace,

$$g(E(y)) = \eta \quad (22)$$

Los MLMG permiten que la media dependa de una combinación lineal de variables explicativas, denominada predictor lineal, a través de una función de enlace no lineal y donde la distribución probabilística que sigue la variable dependiente pertenece a la familia exponencial. Estos modelos tienen la particularidad de que permiten procesar datos correlacionados al incorporar efectos aleatorios y estimar sus componentes de varianza, adicionales a los del error residual (c.f. Faraway, 2006 y Stroup, 2012).

En el análisis subsecuente se asume que  $\gamma$  y  $\varepsilon$  se distribuyen normalmente con media y varianza como,

$$\begin{aligned} E \begin{bmatrix} \gamma \\ \varepsilon \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \text{Var} \begin{bmatrix} \gamma \\ \varepsilon \end{bmatrix} &= \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \end{aligned} \quad (23)$$

La varianza de  $y$  se calcula mediante  $V = ZGZ' + R$ . Se puede modelar  $V$  estableciendo la matriz de diseño  $Z$  de efectos aleatorios y especificando la estructura de covarianzas para  $G$  y  $R$ .

La construcción de los modelos mixtos implica no solo la estimación de  $\beta$ , si no también de los parámetros desconocidos  $\gamma$  y también de  $G$  y  $R$ . En este caso el método de Mínimos Cuadrados ya no resulta ser el método más eficiente de optimización para minimizar  $(y - X\beta)'V^{-1}(y - X\beta)$  si no la técnica de Mínimos Cuadrados Generalizados Estimados (MCGE).

Ante la falta de información sobre  $V$ ,  $G$  y  $R$ , en el enfoque de MCGE se inserta una estimación razonable de  $V$  en el problema de minimización con el objetivo de obtener estimaciones de  $G$  y  $R$ .

Distintos programas estadísticos, como SAS, construyen la función objetivo asociada a Máxima Verosimilitud (MV) o Máxima Verosimilitud Restringida (MVR) y la maximiza sobre todos los parámetros desconocidos. El problema de maximización se reduce entonces a los parámetros en  $G$  y  $R$ . Los comandos como PROC GLIMMIX de SAS minimizan  $-2$  veces la función de verosimilitud mediante el algoritmo Newton-Raphson (SAS, 2019).

La MV y MVR provee estimaciones de  $G$  y  $R$  que se denotan por  $\hat{G}$  y  $\hat{R}$ . Para obtener estimaciones de  $\beta$  y  $\gamma$  el método estándar es resolver las *ecuaciones del modelo mixto*:

$$\begin{bmatrix} X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z \\ Z'\hat{R}^{-1}X & Z'\hat{R}^{-1}Z + \hat{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} X'\hat{R}^{-1}y \\ Z'\hat{R}^{-1}y \end{bmatrix} \quad (24)$$

La solución puede escribirse como:

$$\hat{\beta} = \left( X'\hat{V}^{-1}X \right)^{-1} X'\hat{V}^{-1}y \quad (25)$$

$$\hat{\gamma} = \hat{G}Z'\hat{V}^{-1}(y - X\hat{\beta}) \quad (26)$$

Bajo el esquema propuesto por Yang y Land (2013) para el análisis de datos EPC se estima un modelo lineal mixto generalizado de dos niveles. En el primer nivel la regresión ajustada se limita a analizar la información individual, propia de cada sujeto de estudio. Tanto la variable dependiente como las explicativas registran información de carácter individual, un intercepto, coeficientes de regresión (pendiente) de efectos fijos y un término de error de nivel individual. El segundo nivel de este modelo recibe la información de los coeficientes de regresión estimados en el primer nivel y ajusta un intercepto y coeficientes de regresión aleatorios para los efectos de las variables periodo y cohorte definidas en el modelo. En este segundo nivel, eventualmente, podrían incorporarse otras variables explicativas que también logren captar efectos de periodo o de cohorte (tasa de participación femenina, tasa de desempleo, etc.).

Ese modelo, como hasta ahora se ha descrito, es un modelo de interceptos aleatorios, que permite que solo el intercepto estimado en el primer nivel varíe aleatoriamente entre cohortes y entre periodos, no así la pendiente definida en ese primer nivel. También es posible

establecer un modelo donde la pendiente sea aleatoria, aunque suelen aplicarse a eventos más puntuales. Para mayor información puede consultarse a Yang y Land (2013).

La estimación de MLMG suele darse principalmente por máxima verosimilitud o máxima verosimilitud restringida. Ambos métodos difieren levemente en cuanto a la estimación de los coeficientes de regresión, pero no así en la estimación de los componentes de varianza. La estimación por MVR toma en consideración la pérdida de grados de libertad a causa de la estimación de los parámetros de regresión. En consecuencia, la estimación por MV presenta un sesgo de subestimación para los componentes de varianza. Este puede ser un problema importante cuando se analizan pequeños grupos, en torno a 30, de tal manera que la estimación por MVR es preferible para modelar datos EPC, pues generalmente se tienen menos de 30 categorías en las variables periodo y cohorte.

Mientras que los coeficientes de regresión en el nivel individual se consideran parámetros a ser estimados en este modelo, los coeficientes del segundo nivel se consideran variables latentes, es decir, captan información que no es directamente observable. De tal manera, los coeficientes de sus efectos aleatorios más que ser estimados son "predichos" (Snijders y Bosker, 2011 y Yang y Land, 2013). Se emplea el método Bayes empírico para predecir estos efectos aleatorios (Casella, 1985; SAS Institute Inc, 2019; Witte, Greenland, Kim y Arab, 2000).

De acuerdo con Yang y Land (2013), para un grupo específico en el modelo, por ejemplo para una cohorte específica, el estimador bayesiano empírico predice su efecto aleatorio como media ponderada de la media general ajustada o media poblacional de la variable respuesta, a través de la muestra combinada de todas las observaciones y el impacto promedio estimado de dicha cohorte en el valor esperado de la variable respuesta a lo largo de todos los períodos. El promedio ponderado adopta la forma de una combinación convexa de estas dos estimaciones, en la que las ponderaciones dependen de las estimaciones de la varianza de los efectos de la cohorte y de los errores del primer nivel. La estimación del efecto de la cohorte específica está sesgada hacia la media poblacional ajustada, pero se considera más confiable

porque se basa en una muestra más grande de observaciones y tiene un menor error cuadrado medio.

Además de su capacidad para tener en cuenta la heterogeneidad en múltiples niveles, el enfoque de efectos mixtos tiene una segunda ventaja importante: es capaz de explicar la variabilidad aleatoria mediante la incorporación de covariables adicionales, tanto a nivel individual (especialmente para el diseño de datos de encuestas) como a nivel superior. Es decir, mediante la especificación e incorporación de atributos individuales y características contextuales de periodo y cohorte en el modelo, este enfoque permite realizar pruebas de hipótesis de los efectos EPC, que no son posibles en el marco del modelo lineal general (Yang y Land, 2013; pág 194).

Estimar el efecto de los componentes EPC mediante modelos lineales convencionales limita la posibilidad de modelar los efectos de cohorte y periodo. Estos se incorporan como si fueran efectos fijos, en lugar de ser efectos aleatorios. Cuando estos efectos se analizan como aleatorios se considera la posibilidad de que las observaciones estén correlacionadas, lo que parece razonable, pues los entrevistados en el mismo año y que pertenecen a una misma cohorte pueden brindar respuestas similares, ya que comparten un error aleatorio no medible y que representa procesos subyacentes, contextuales, originados a causa de la pertenencia de cada individuo a cierta cohorte y al experimentar ciertos eventos en un momento (periodo) dado.

De acuerdo con Yang y Land (2013; pág. 191) el análisis de datos EPC mediante modelos jerárquicos no incurre en el problema de identificación, dado que no se asume que el efecto de estos tres componentes sea lineal o aditivo en el mismo nivel de análisis.

### **2.3.1. Definición de cohorte y periodo como efectos fijos o de aleatorios**

El análisis de tipo jerárquico se puede desarrollar usando modelos con coeficientes tanto fijos como aleatorios para definir la agrupación de datos. La especificación más adecuada depende

de una serie de consideraciones. Yang y Land (2008) mencionan los siguientes criterios aplicados a modelos jerárquicos para datos EPC para definir esas variables como de efectos fijos o de efectos aleatorios, según sea el caso:

- (1) El rango de edad de las poblaciones humanas contemporáneas se puede fijar entre 0 y 125 años, pero la mayoría de los estudios utilizan sólo una parte de este rango fijo. Por ejemplo, cuando la población de estudio corresponde a mujeres en edad reproductiva, de 15 a 49 años. Por lo tanto, las edades individuales o categorías de edad pueden considerarse como entidades únicas y es razonable especificar el efecto de edad como fijo.
- (2) Las categorías de periodo y cohorte disponibles para cualquier análisis empírico son una muestra de todas las categorías que realmente existen, por lo que definir sus efectos como aleatorios es lo más apropiado.
- (3) Un problema clave con la especificación de efectos fijos es el supuesto de que las variables indicadoras o dummy<sup>10</sup> que representan los efectos fijos para cohorte y período explican plenamente todos los efectos de la cohorte y del período. Es decir, el modelo de efectos fijos no admite la posibilidad de una variación aleatoria adicional asociada a los efectos individuales de la cohorte y del periodo. Esto implica que no hay variabilidad sin estar siendo explicada entre cohortes o entre periodos, más allá de lo capturado por sus efectos fijos.

En modelos EPC esta no parece ser la suposición más apropiada. Primero, porque el modelo de efectos fijos tiende a producir errores estándar sustancialmente mayores para el intercepto, lo que implica mayor incertidumbre en la estimación de la media. Además, los datos provenientes de encuestas transversales tienden a ser muy desequilibrados, es decir, cuando los sujetos de estudio en encuestas repetidas se clasifican por cohorte y período, el número de observaciones resulta estar desbalanceado, potencialmente habrá menos sujetos en cohortes más antiguas o en las más recientes, por ejemplo. Bajo esta estructura, los modelos de efectos mixtos utilizan

---

<sup>10</sup> En la práctica los efectos fijos para cohorte y periodo se estiman mediante un conjunto de variables indicadoras o dummy para J-1 cohortes y K-1 periodos.

la información disponible de manera más eficiente y muestran una mejor eficiencia estadística en relación con los modelos de efectos fijos.

- (4) Aunque todo lo anterior sugiera que deben definirse modelos de efectos mixtos en lugar de modelos de efectos fijos, Yang y Land (2008) sugieren una serie de excepciones. Una excepción es cuando la variable periodo tiene pocas categorías, unas dos o tres, caso en el que más bien es razonable especificar los efectos de periodo como fijos y los de cohorte indistintamente, como fijos o aleatorios.
- (5) Una segunda excepción se refiere a situaciones en las que se viola el supuesto de que los efectos aleatorios del segundo nivel se distribuyen independientemente de los regresores del primer nivel. Es decir, no se cumple la condición de que  $corr(r_{0j}, X) = 0$ .

Yang y Land (2008) sugieren un procedimiento de dos pasos para evaluar el supuesto de independencia de los efectos aleatorios y las variables del primer nivel. El primer paso es estimar dos modelos, uno en el que se definan todos los efectos fijos (excluyendo alguna de las categorías de periodo y cohorte para evitar el problema de identificación) y otro en el que las variables periodo y cohorte se definan como de efectos aleatorios. En segundo lugar, se desarrolla el test de Hausman, cuya estadística toma la forma,

$$W = \chi^2[K] = [b - \hat{\beta}]^T \hat{\Sigma}^{-1} [b - \hat{\beta}] \quad (27)$$

Donde  $b$  denota el vector de coeficientes de regresión del primer nivel estimados en el modelo de efectos fijos,  $\hat{\beta}$  denota el vector correspondiente de coeficientes de regresión estimados en el modelo de efectos aleatorios,  $\hat{\Sigma} = Var[\hat{b}] - Var[\hat{\beta}]$  es la diferencia de las matrices de varianza-covariancia de los dos estimadores (el término constante se excluye de todos los vectores y matrices). Bajo la hipótesis nula de que los efectos aleatorios de cohorte y periodo en el modelo de efectos aleatorios son independientes de los regresores de nivel individual,  $W$  se distribuye Chi-cuadrado con grados  $K$  de libertad, donde  $K$  es la dimensión de los vectores  $b$  y  $\beta$ .

### 2.3.2. Significancia de los componentes cohorte y periodo

Yang y Land (2013, pág. 208) describen una serie de procedimientos que se pueden desarrollar para determinar la significancia de los componentes periodo y cohorte. Estos procedimientos descritos por los autores se basan en distintas fuentes de información, pero que posteriormente adaptaron al contexto de los modelos EPC.

Tras estimar un modelo mixto, jerárquico, para datos EPC, se puede conocer la significancia de cada categoría de las variables periodo y cohorte. Sin embargo, tales procedimientos tienen el propósito de determinar la significancia conjunta, de estos componentes y no de sus categorías, exclusivamente. Igualmente, Yang y Land (2013) rescatan que a pesar de que el efecto de estos términos no resulte estadísticamente significativo, dado cierto nivel de significancia, es importante analizar su tendencia o patrón desde lo sustantivo. Los siguientes pasos permiten explorar la significancia de los coeficientes cohorte y periodo.

#### (1) Estudiar el patrón y nivel de significancia de los coeficientes individuales

Examinar el patrón y el nivel de significancia de los coeficientes individuales se puede abordar en dos instancias: (1) graficando la secuencia temporal de los coeficientes estimados, tanto para periodo como para cohorte y (2) examinando la significancia estadística de sus coeficientes individuales.

La inspección gráfica facilita la identificación de patrones y de aquellos periodos o cohortes que parecen tener una importancia sustantiva o que se destaquen.

En modelos mixtos lineales generalizados, puede graficarse la probabilidad predicha o los riesgos relativos, según sea la distribución y función de enlace empleada. En un modelo lineal generalizado cuya variable dependiente sigue una distribución de Poisson con enlace log, el efecto de cohorte puede calcularse como  $\exp(u_{0j})$  y el de periodo como  $\exp(v_{0k})$  para obtener la estimación del conteo, o bien, el riesgo relativo si se ha definido un *offset*.

En cuanto al análisis de la significancia de los coeficientes individuales, se somete a prueba la hipótesis de que los coeficientes de efectos aleatorios de cohorte  $u_{0j}$  y de periodo  $v_{0k}$  son iguales a cero, es decir:

$$\textbf{Cohorte: } H_0: u_{0j} = 0 \text{ vs. } H_1: u_{0j} \neq 0 \quad (28)$$

$$\textbf{Periodo: } H_0: v_{0k} = 0 \text{ vs. } H_1: v_{0k} \neq 0 \quad (29)$$

Si la hipótesis nula no se rechaza, se concluye que el promedio de la variable dependiente para la cohorte  $j$ -ésima o para el periodo  $k$ -ésimo no es diferente del promedio general. Las conclusiones derivadas de estas pruebas son consistentes con la inspección gráfica realizada en el paso anterior.

## (2) Determinar la significancia de los efectos cohorte y periodo como grupo

Dado que algunos coeficientes de periodo y cohorte pueden ser estadísticamente significativos, mientras otros no, el siguiente paso es determinar si el efecto de cohorte y el efecto de periodo son estadísticamente significativos, como conjunto, y por ende contribuyen a explicar la varianza en el modelo. Para ello, Yang y Land (2013) sugieren (1) analizar la devianza y los componentes de varianza y (2) desarrollar una prueba F para determinar la presencia de efectos aleatorios.

La devianza del modelo estimado se define como  $-2\log(\text{verosimilitud})$  y puede calificarse como una medida de bondad de ajuste entre el modelo y los datos. De tal manera, si se compara la devianza con los grados de libertad y esa primera medida es mucho más grande que la segunda, se tendrá indicación de que existe una asociación significativa entre las variables independientes con la variable respuesta.

Para probar si los efectos de cohorte y periodo contribuyen significativamente a explicar la variabilidad de la variable dependiente es posible emplear una prueba F para probar la hipótesis de la presencia de efectos aleatorios. La distribución de muestreo del estadístico F es

exacta cuando los efectos aleatorios se distribuyen independientemente como variables aleatorias normales.

Como los modelos generados bajo REML no producen una estimación máximo verosímil si no que se basan en la estimación pseudo máximo verosímil, no se recomienda seleccionar el mejor modelo a partir de criterios que se basan en el valor de la máxima-verosimilitud. Bajo la estimación REML los criterios de información no son comparables (Littell, 2016) aun cuando se comparen modelos en competencia, incluso si están anidados con respecto a los efectos fijos o los efectos aleatorios. En algunos casos podría ser posible que al comparar dos modelos anidados, el que posee el mayor número de parámetros tenga una pseudo-probabilidad más pequeña (SAS Institute Inc, 2019; pág. 3382).

Cuando el modelo propuesto incorpora como términos aleatorios únicamente los componentes de periodo y cohorte, se tendrá un conjunto de residuos aleatorios para la cohorte  $j$ -ésima,  $u_{0j}$  y un conjunto de residuos aleatorios para el efecto de periodo  $k$ -ésimo,  $v_{0k}$ . Se asume que cada conjunto de coeficientes es independiente y estos se distribuyen normalmente con media cero y varianza  $\tau_u$  para la cohorte y  $\tau_v$  para el periodo.

Para un modelo de interceptos aleatorios, la prueba exacta F contrasta las siguientes hipótesis a fin de identificar la relevancia de los coeficientes de efectos aleatorios de cohorte y de periodo:

$$\mathbf{Cohorte:} \quad H_0: \tau_u = 0 \text{ vs. } H_1: \tau_u > 0 \quad (30)$$

$$\mathbf{Periodo:} \quad H_0: \tau_v = 0 \text{ vs. } H_1: \tau_v > 0 \quad (31)$$

De igual manera, puede probarse la hipótesis para determinar conjuntamente la significancia de ambos componentes:

$$H_0: \tau_u = \tau_v = 0 \text{ vs. } H_1: \tau_u > 0 \text{ ó } \tau_v > 0 \quad (32)$$

La hipótesis nula representa aquellas situaciones en las que las variaciones de los efectos de cohorte, los de periodo o los de cohorte y periodo simultáneamente no difieren significativamente de cero.

En el caso de modelos lineales, la idea tras la prueba F es que cuando la varianza de un efecto aleatorio es igual a cero (por ejemplo, cuando el efecto de cohorte  $\tau_u = 0$ ), la diferencia entre la suma de cuadrados del modelo que incluye los componentes aleatorios cohorte y periodo  $S_{min}$  y el mínimo de la suma de cuadrados sin estos efectos aleatorios, estimado mediante una regresión por mínimos cuadrados ordinarios  $S_{OLS}$ , deben ser cercanas. Partiendo de este principio, se puede estimar la suma de cuadrados de residuales como,

$$S_{OLS} = \sum_{i=1}^N \|y_i - X_i \hat{\beta}_{OLS}\|^2 \quad (33)$$

Posteriormente, el mínimo de la suma de cuadrados en un modelo de efectos aleatorios, se calcula como,

$$S_{min_\delta} = \min_{\delta} \|y - W\delta\|^2 \quad (34)$$

donde  $W$  consiste en la matriz  $X$  de observaciones de las variables explicativas a nivel individual en conjunto con la matriz de diseño  $Z$  para los efectos aleatorios de la cohorte, es decir,  $W = [X, Z]$  y  $\delta = (\beta', u'_0)'$ . Bajo el supuesto de que las variables explicativas en la matriz  $X$  son linealmente independientes, el rango de  $X$  es  $m$ .

Bajo la hipótesis nula planteada, la razón de ambas sumas de cuadrados sigue una distribución F,

$$\frac{(S_{OLS} - S_{min})/(r - m)}{S_{min}/(N_T - r)} \sim F(r - m, N_T - r) \quad (35)$$

donde  $N_T$  denota el tamaño de la muestra,  $r$  es el rango de la matriz  $W$  y  $m$  es el número de variables explicativas que se incluyen en la regresión por mínimos cuadrados ordinarios.

Cuando el efecto aleatorio está presente en un modelo lineal mixto, es decir, cuando  $\tau_\mu$  no es cero,  $S_{min}$  debería ser relativamente pequeño y la razón (35) tenderá a ser grande y se rechazaría la hipótesis nula.

Al rechazar la hipótesis nula, se concluye que la varianza de los efectos de periodo o cohorte, según sea el caso, no es cero, y por ende, incluir este conjunto de efectos aleatorios es relevante para explicar la variabilidad de la variable dependiente.

Esta prueba F puede desarrollarse para ambos efectos aleatorios (periodo y cohorte), en conjunto, pero en ese caso el rango de la matriz  $Z$  es mayor, pues incluye los términos  $\mu'_0$  y  $\nu'_0$ . Sin embargo, la hipótesis nula asociada a esta prueba se rechazará cuando el efecto de al menos uno de los componentes, periodo o cohorte, sea estadísticamente significativo.

Cuando la variable dependiente analizada no es cuantitativa sino más bien categórica, o cuando no se puede suponer que tenga una distribución normal, se emplean modelos lineales generalizados para analizar datos EPC. También se asume una distribución normal multivariante para los efectos aleatorios.

En ese tipo de modelos, la prueba F para determinar la significancia de los efectos aleatorios debe modificarse. Esta modificación a las fórmulas mostradas anteriormente parte de que la devianza estadística se comporta asintóticamente como una suma de cuadrados (Demidenko, 2004 y McCullagh y Nelder, 1989). De tal manera, el estadístico F corresponde a,

$$\frac{(l_0 - l_{max})/(r - m)}{l_{max}/(N_T - r)} \cong F(r - m, N_T - r) \quad (36)$$

Donde  $l_0$  es el máximo de la log-verosimilitud que se obtiene en un modelo de regresión en el que se incluyen las mismas variables de nivel 1 pero sin controlar por el efecto de cohorte o de periodo.  $l_{max}$  es el máximo de la log-verosimilitud de un modelo en el que el efecto de periodo o cohorte se incorporan como efectos fijos. Se puede demostrar que un modelo de efectos fijos que trata los efectos aleatorios como fijos corresponde a un modelo de efectos

aleatorios con matriz de covariancia infinita (Yang y Land, 2013; pág. 220), de tal manera que  $l_{max}$  es una estimación del límite superior para la suma de cuadrados de un modelo mixto.  $N_T$  denota el tamaño de la muestra,  $r$  es el rango de la matriz  $W$  y  $m$  es el número de variables explicativas en el nivel 1 del modelo.

## **METODOLOGÍA**

### **1. Objetivos de estudio**

#### **1.1. Objetivo general**

Describir la trayectoria de la fecundidad en Costa Rica mediante un modelo lineal generalizado de efectos mixtos que permita controlar por el efecto de los componentes Edad-Periodo-Cohorte, empleando las encuestas de fecundidad, salud sexual y reproductiva realizadas en el periodo 1976-2015.

#### **1.2. Objetivos específicos**

- Estimar un modelo de regresión lineal generalizado de efectos mixtos para describir el número de hijos tenidos y el efecto de las covariables: zona de residencia, estado conyugal y ocupación, controlando por el efecto de los componentes Edad, Periodo y Cohorte.
- Describir el efecto de los componentes Edad, Periodo y Cohorte sobre el nivel de la fecundidad entre los años 1976-2015.

### **2. Fuentes de información**

Costa Rica cuenta con información de encuestas en materia de salud sexual y reproductiva desde el año 1964, la más reciente encuesta fue llevada a cabo en el año 2015. Su formato es análogo a las conocidas Demographic Health Surveys y cada nueva encuesta guarda coincidencia con las anteriores, tanto en la redacción de las preguntas como en las opciones de respuesta y los métodos de selección de los entrevistados, para garantizar la comparabilidad de los datos a lo largo del tiempo.

En cada una de las encuestas se empleó un diseño muestral probabilístico, con representatividad a nivel nacional y regional. Tanto el cuestionario como las bases de datos son de acceso público y se pueden descargar desde la página del Centro Centroamericano de Población de la Universidad de Costa Rica: [www.ccp.ucr.ac.cr](http://www.ccp.ucr.ac.cr)

En este trabajo se analizan las encuestas realizadas en 1976, 1986, 1993, 1999, 2010 y 2015. La mayoría recopilaron información de mujeres en edad reproductiva: 15-49 años. Sin embargo, la encuesta del año 1976 incluye solamente a mujeres mayores de 20 años y la realizada en 1999 registra información solamente de mujeres entre 18 y 44 años. Las encuestas de 2010 y 2015 incluyeron tanto a hombres como mujeres. En esa primera, se entrevistó a personas entre los 15 y 80 años de edad y se realizó un sobremuestreo de la población adolescente y adulta mayor. En la encuesta del año 2015 se incluyó a la población desde los 15 hasta los 64 años.

En las encuestas de 2010 y 2015 se aplicó un factor de ponderación para corregir por los sesgos de selección inherentes al esquema muestral empleado. Las encuestas de los años anteriores, desde 1976 hasta 1999 son autoponderadas y por lo tanto la variable de ponderación toma el valor de uno. El factor de ponderación de los casos registrados en las encuestas de 2010 y 2015 fue reescalado para que su media se aproxime a uno, evitando que los casos registrados en esas dos últimas encuestas tengan mayor peso que otros.

### **3. Población de estudio**

La población objetivo de este trabajo es la población femenina en edad reproductiva, concretamente entre los 15 y los 49 años de edad. El Cuadro 10 muestra el tamaño de muestra y el rango etario de las mujeres entrevistadas en cada una de las encuestas. En 1976 y 1999 se redujo el rango etario de la población de mujeres entrevistadas.

**Cuadro 10**

Tamaño de muestra y rango etario de las mujeres entrevistadas en las encuestas de 1976 al 2015

Año de encuesta	Tamaño de muestra en análisis	Edades analizadas
<b>Total</b>	<b>14596</b>	<b>15-49</b>
<b>1976</b>	3935	20-49
<b>1986</b>	3527	15-49
<b>1993</b>	3618	15-49
<b>1999</b>	1029	18-44
<b>2010</b>	1206	15-49
<b>2015</b>	1281	15-49

**4. Variables de estudio**

La variable dependiente en este estudio corresponde al número de hijos que ha tenido cada mujer hasta el momento de la entrevista. Esta variable se construye a partir de una serie de preguntas que se plantean a las entrevistadas, generalmente en un módulo que trata sobre la historia de embarazos y nacimientos. A cada mujer se le pregunta: *¿Tiene usted hijas o hijos vivos?, ¿Cuántas hijas e hijos suyos viven con usted?, ¿Cuántos de estos hijos (as) son adoptados?, ¿Cuántas hijas e hijos suyos no viven con usted? y ¿Cuántas hijas e hijos suyos nacieron vivos y luego han muerto?* El número de hijos tenidos se obtiene al sumar los hijos que viven con la mujer, los que viven fuera del hogar y los que nacieron vivos aunque hayan muerto algunas horas después. Los hijos adoptados se excluyen de esta suma.

El planteamiento de esta serie de preguntas se hace necesario para estimar adecuadamente el número de hijos nacidos vivos (c.f. El-Badry, 1961; ONU, 1983; ONU, 2005; Rodríguez y Oliva, 2011), indagando tanto por hijos como por hijas para evitar la omisión de las mujeres por efectos del lenguaje de las preguntas, evitando el olvido de hijos que no viven con la madre al momento de la entrevista y omitiendo la fecundidad que no corresponde a la mujer entrevistada tras restar a los hijos adoptados.

Las variables independientes y su esquema de codificación se describen en la Figura 4. Las variables edad, zona de residencia, estado conyugal y nivel educativo se incorporan como variables de efectos fijos en el primer nivel del modelo jerárquico. Las variables periodo y cohorte se incorporan en el segundo nivel del modelo como variables de efectos aleatorios.

**Figura 4**  
Descripción y codificación de las variables analizadas

Variable	Codificación
Edad	Continua, edades simples: 15 – 49 años. Se incorpora en los modelos de regresión calculando para cada caso la diferencia entre la edad y la edad promedio de toda la muestra.
Zona de residencia	0: Rural ; 1: Urbano
Estado conyugal	0: Soltera, nunca en unión ; 1: Ha vivido alguna vez en unión
Años de educación	Años de educación Se incorpora en los modelos de regresión como la diferencia entre los años de educación de cada caso con respecto a los años de educación promedio de la cohorte a la pertenece que cada mujer.
Ocupación	0: No trabaja ; 1: Trabaja al momento de la encuesta
Periodo	6 categorías que corresponden al año en que se desarrolló la encuesta: 1976, 1986, 1993, 1999, 2010 y 2015.
Cohorte	Se analizan en total 15 cohortes, que corresponden a la agrupación por años quinquenales del año de nacimiento de las mujeres entrevistadas en las distintas encuestas. Este grupo de mujeres nació entre 1926-1929 y 1995-2000.

Al analizar modelos jerárquicos o de efectos mixtos debe elegirse el tipo de métrica que tomarán las variables explicativas que se incorporen en el primer nivel del modelo (Enders y Tofighi, 2007; Paccagnella, 2006; Raudenbush y Bryk, 2002). Además de la métrica que toma naturalmente la variable, el centrado con respecto a la media general y el centrado con respecto a la media de subgrupos son los más habituales para los predictores de tipo continuo.

Para centrar una variable con respecto a la media general se resta a cada observación la media de la variable, estimada con la información de toda la muestra. Bajo este método de centrado cada coeficiente representa el cambio esperado en la variable respuesta, dado el cambio de una unidad en la variable explicativa en cuestión, utilizando la escala de medición original. El

centrado con respecto a la media de grupos se obtiene al restarle a cada observación la media de cierto grupo al que pertenezca, por ejemplo, la media de la cohorte. Se emplea cuando se quiere obtener el efecto puro del primer nivel sin considerar las variables del segundo nivel. La variable centrada por este método no estará correlacionada con las variables del segundo nivel, es decir, la estimación de su efecto no estará mediado por las variables de nivel 2 (Enders y Tofighi, 2007).

En modelos jerárquicos en los que solo se estima como aleatorio el intercepto (no la pendiente de las variables del primer nivel), Raudenbush y Bryk (2002) sugieren que los tres métodos son equivalentes en términos de la estimación de los parámetros, pero no así en lo que se refiere a la estimación de los componentes de varianza y las estadísticas de bondad de ajuste.

En este trabajo, no se espera que el efecto de la edad, varíe de una cohorte a otra, por lo que esta variable se centra con respecto a la edad promedio de toda la muestra. Se incluyó además el cuadrado de esta variable (centrada con respecto a la media) como un término más en el modelo de regresión, pues se considera que el efecto de la edad no es lineal. El modelo también se ajustó incluyendo edad como una variable categórica, pero se excluyó porque la edad y su componente cuadrático ajustan mejor el modelo.

El nivel educativo de la población costarricense aumentó en las últimas décadas (c.f. PEN, 2019, Mora, 1999 y Román, Murillo y León, 2014) y así como la asistencia al sistema educativo formal (c.f. PEN, 2019). El nivel educativo promedio se habría incrementado de una cohorte de mujeres a la siguiente. Se considera que el incremento en el nivel educativo se confunde, en alguna medida, con el efecto cohorte, por lo que, la variable "años de educación" fue centrada con respecto a los años promedio de educación según cohorte, para mitigar su correlación con esa variable de segundo nivel. También se evaluó la pertinencia de analizar el nivel educativo en categorías, probando diferentes agrupaciones, pero se excluyó este esquema por mostrar un ajuste más pobre.

## 5. Análisis de datos

El análisis de los datos se divide en dos fases. En una primera fase se desarrollan análisis de carácter descriptivo que permiten comprender los patrones generales de los datos, principalmente el de los componentes edad, periodo y cohorte, así como del número de hijos nacidos vivos. En la segunda fase de análisis, se desarrollan modelos de efectos fijos y mixtos para describir la relación entre el número de hijos tenidos nacidos vivos y las variables independientes seleccionadas.

Sea  $Y_i$ , donde  $i = 1, \dots, n$  es un vector que registra el número de hijos (nacidos vivos) que había tenido cada mujer hasta el momento de la entrevista. Se asume que  $Y_i$  sigue una distribución de Poisson:  $Y_i \sim \text{Poisson}(\mu_i)$ . Se ajusta un modelo lineal mixto generalizado con una distribución de Poisson que asocia el promedio de la variable  $Y_i$  con las covariables  $X_i$  individuales, la cohorte y el periodo mediante la relación que se expresa en las ecuaciones (37), (38) y (39):

El modelo de nivel 1, o de efectos fijos, se escribe como:

$$\begin{aligned} \log(\mu_i) = & \beta_{0jk} + \beta_1 EDAD_{ijk} + \beta_2 EDAD_{ijk}^2 + \beta_3 ZONA_{ijk} \\ & + \beta_4 ECONYUGAL_{ijk} + \beta_5 EDUCACION_{ijk} \\ & + \beta_6 OCUPACION_{ijk} \end{aligned} \quad (37)$$

El modelo de nivel 2, o de efectos aleatorios, se define como:

$$\beta_{0jk} = \gamma_0 + u_{0j} + v_{0k} \quad ; \quad u_{0j} \sim N(0, \tau_u), v_{0k} \sim N(0, \tau_v) \quad (38)$$

El modelo combinado se escribe como;

$$\begin{aligned} \log(\mu_i) = & \gamma_0 + \beta_1 EDAD_{ijk} + \beta_2 EDAD_{ijk}^2 + \beta_3 ZONA_{ijk} \\ & + \beta_4 ECONYUGAL_{ijk} + \beta_5 EDUCACION_{ijk} \\ & + \beta_6 OCUPACION_{ijk} + u_{0j} + v_{0k} \end{aligned} \quad (39)$$

Donde en todos los casos,

$i = 1, 2, \dots, n_{jk}$  individuos en la cohorte  $j$  y en el periodo  $k$ ,

$j = 1, 2, \dots, 15$  cohortes de nacimiento,

$k = 1, 2, \dots, 6$  periodos, años en los que se realizaron las encuestas.

El desarrollo del análisis bajo este esquema obedece a la metodología sugerida por Yang y Land (2013) para el análisis de modelos EPC provenientes de encuestas periódicas de corte transversal.

El modelo descrito en (39), define que el intercepto y la media esperada tienen efectos aleatorios para periodo y cohorte y no define efectos aleatorios para los parámetros de pendiente, en aras de mantener un modelo simple.

Cuando se ajustan modelos en los que la variable respuesta sigue una distribución de Poisson es usual que se incluya un *offset*, una variable estrictamente positiva que opera como denominador del conteo analizado. La incorporación del *offset* al modelo es relevante si se espera que el número de ocurrencias del evento sea proporcional a la exposición y los sujetos de estudio han estado expuestos al evento durante un intervalo de tiempo que no es constante. Bajo estas condiciones, se estarían modelando tasas o razones de cambio, en lugar de un conteo, lo que facilita la interpretación de los coeficientes. Cuando se omite el *offset* y los sujetos tienen diferente tiempo de exposición el modelo de regresión de Poisson podría estar sesgado.

El modelo ajustado, tal cual se detalla en las ecuaciones anteriores (37 y 39), no incorpora un *offset* a pesar de que las mujeres entrevistadas no han vivido el mismo tiempo de exposición al embarazo. Los coeficientes del modelo expresan variaciones en el número promedio de hijos tenidos y no variaciones en términos de una tasa o razón. En el contexto de este estudio, el tiempo de exposición a la fecundidad (*offset*) está definido por la edad, de tal manera que la inclusión del *offset* generaría conflicto con dicha variable, que ya se incorpora como variable explicativa y es virtualmente, una de las más importantes en el modelo a ajustar.

La estimación de la fecundidad en este trabajo emplea como variable respuesta el número de hijos nacidos vivos y no tasas de fecundidad o alguna otra medida que permita asociar ese conteo con el tiempo de exposición. La omisión de este factor de exposición se debe a dos motivos.

El primero es que, aunque se logra conocer el número total de hijos nacidos vivos que ha tenido cada mujer, en las encuestas de 1986, 2010 y 2015 se registró la fecha de nacimiento solamente de los hijos nacidos durante los cinco años previos al momento de la entrevista. Esto impide determinar a cuál grupo etario contribuye el nacimiento de cada uno de los hijos que haya tenido una mujer a lo largo de su periodo reproductivo.

La forma en la que se calculan habitualmente las tasas de fecundidad (c.f. Rutstein y Rojas, 2006) define como numerador el número de hijos nacidos vivos asociado al grupo de edad que correspondía a la madre al momento de ese nacimiento. El denominador de dicha tasa serán los años persona de exposición al embarazo, que haya vivido cada mujer en cada grupo etario. Así las cosas, una mujer puede brindar información de nacimientos en uno o más grupos etarios (e incluso en ninguno si no tuviera hijos), según sea su edad al momento de cada nacimiento. Además, el tiempo de exposición que se estima para cada mujer también debe asignarse a uno o más grupos etarios, según su edad.

Por ejemplo, si en una determinada encuesta se entrevista a una mujer de 36 años quien ha tenido tres hijos, nacidos hipotéticamente cuando ella tenía 24, 29 y 32 años, cualquier medición que enlace este conteo con el tiempo de exposición debe anotar la "contribución" de un hijo a los grupos etarios 20-24, 25-29 y 30-34, pero asociándolo con un periodo de exposición (denominador) en los grupos etarios (quinquenales) que van desde los 15 hasta los 36 años.

No contar con la fecha de nacimiento de todos los hijos no imposibilita la estimación de una medida de fecundidad, pero esta haría referencia a un periodo de cinco años, en vista de que las encuestas registran a los hijos nacidos en ese plazo. Siguiendo el caso presentado

anteriormente, para el numerador de una tasa solamente se tendría información del último hijo, que se relacionaría con un denominador que contabilice el tiempo de exposición a un embarazo delimitado al rango de 31 a 36 años.

Para efectos de este trabajo, se descartó ese último esquema porque se pretendía describir la fecundidad que habían alcanzado las mujeres a lo largo de todo su periodo reproductivo pues este es el que refleja el efecto producido por las variables edad, periodo y cohorte, presentes a lo largo de todo el curso de la vida y no solamente durante un lapso de cinco años.

Además, si se contempla la información del momento de nacimiento de los hijos, independientemente del periodo al que se delimite su registro, el análisis de este tipo de información se torna más complejo. Partiendo del principio de parsimonia, el análisis se desarrolló manteniendo un esquema más simple, que en todo caso produce resultados aceptables.

De tal manera, los modelos lineales generalizados, tanto el de efectos fijos (37) como el de efectos mixtos (39), se estimaron sin definir un *offset*. Aun conociendo su relevancia, la omisión del *offset* en estos modelos se debe a la complejidad que suma el cálculo de tasas de fecundidad, como se ha justificado anteriormente. Además, el tiempo de exposición guarda correspondencia con la edad, variable que ya se incluyó en el modelo. El *offset* por definir en este modelo estaría dado en función de la edad.

La varianza marginal del modelo se puede expresar como  $Var(Y_i) = \Phi\mu_i$  donde  $\Phi$  representa el parámetro de dispersión. La distribución de Poisson asume que  $\Phi=1$ . Cuando  $\Phi>1$  existe sobredispersión, es decir, la varianza es superior a la esperada y es señal de que el modelo ajustado no es el adecuado. En ese caso, pueden emplearse otras distribuciones de probabilidad como la Binomial Negativo o Poisson inflada con cero para modelar los datos.

El MLMG presentado en el Cuadro 19 surgió de una serie de consideraciones y pruebas que contribuyeron a establecer ese como el modelo final. Las variables edad y nivel educativo

fueron analizadas como categóricas. El valor que tomaban los coeficientes sugirió que la variable edad tenía un efecto cuadrático mientras que la educación tenía una relación lineal con el número de hijos. Los índices de bondad de ajuste permitieron establecer bajo cual esquema de codificación se obtenía un modelo que ajustara mejor a los datos.

Para el modelo (39) el parámetro de sobredispersión  $\Phi=1.01$ , lo que sugiere que no hay sobredispersión y la distribución de Poisson es apropiada para modelar la variable dependiente. Si se ajusta una distribución binomial negativa, comúnmente empleada para modelar datos de conteo, el modelo estimado no brinda un mejor ajuste o resultados que sean sustancialmente diferentes que los del modelo (39).

El análisis de datos se realizó empleando el programa SAS 9.4 (SAS Institute Inc, 2014). Para la generación de los modelos fijos y mixtos se empleó el comando PROC GLIMMIX de dicho programa, que permite ajustar modelos con efectos aleatorios normales en los que la distribución condicional de los datos pertenece a la familia exponencial, también conocidos como modelos lineales mixtos generalizados (MLMG). Debido a que la distribución normal también es miembro de la familia exponencial y ante la ausencia de efectos aleatorios, GLIMMIX permite ajustar modelos lineales mixtos (Littell et al, 2006). GLIMMIX no ajusta modelos jerárquicos con efectos aleatorios no normales, con este procedimiento se selecciona la distribución de la variable respuesta condicionada a los efectos aleatorios normalmente distribuidos (SAS Institute Inc, 2014).

La estimación de los parámetros aleatorios que se obtiene con el comando GLIMMIX de SAS representan la desviación estimada con respecto al intercepto  $\gamma_0$  (SAS, 2019). En este trabajo, la estimación de los parámetros de efectos aleatorios que corresponde al componente de cohorte se denota por el término  $u$  y los de periodo se denotan por  $v$ .

## **RESULTADOS**

Este capítulo presenta los resultados obtenidos, divididos en dos secciones. La primera sección permite caracterizar a la población en estudio mediante una serie de cuadros con información descriptiva, mostrando primero su distribución de acuerdo con las principales variables en análisis e incluyendo las variables EPC. En esa misma sección se analiza la distribución del número de hijos tenidos (la variable dependiente en el estudio). En la segunda sección del capítulo se presentan los resultados derivados del MLMG seleccionado y ahonda en el efecto estimado para las variables cohorte y periodo.

### **1. Características de la población y variables de estudio**

#### **1.1. Descripción de las variables edad, periodo y cohorte**

La muestra global de este estudio está compuesta por 14 605 mujeres entrevistadas en alguna de las seis encuestas de fecundidad y salud sexual. Esta muestra se distribuye según edad en el momento de la encuesta, periodo y cohorte como se muestra en el Cuadro 11.

Se analizan mujeres entre los 15 y 49 años de edad. La muestra está bien distribuida según edad, aunque en los últimos grupos etarios el porcentaje de mujeres disminuye.

La proporción de casos es menor en las tres últimas encuestas, pues el tamaño de muestra disminuyó en ese periodo. El porcentaje de casos tiende a ser menor en las primeras cohortes (1926-1944) y en las más recientes (1975-2000), ya que tan solo una o dos encuestas reúnen información de mujeres nacidas en esos años. Sin embargo, el número de casos en las distintas categorías a analizar por periodo y cohorte es suficiente para contemplar esta categorización.

**Cuadro 11**

Número y porcentaje de mujeres que se analizan en las categorías de las variables edad, periodo y cohorte

Variable	N° de casos	Porcentaje <sup>/a</sup>
<b>Total</b>	<b>14605</b>	<b>100.0</b>
<b>Edad</b>		
15-19	1,946	12.4
20-24	2,943	19.8
25-29	2,761	19.1
30-34	2,273	15.8
35-39	1,987	13.8
40-44	1,552	10.8
45-49	1,143	8.3
<b>Periodo</b>		
1976	3,935	26.9
1986	3,527	24.2
1992	3,618	24.8
1999	1,029	7.1
2010	1,206	8.3
2015	1,290	8.8
<b>Cohorte</b>		
1926-1929	278	1.9
1930-1934	438	3.0
1935-1939	729	5.0
1940-1944	1,001	6.9
1945-1949	1,430	9.8
1950-1954	1,870	12.8
1955-1959	1,688	11.6
1960-1964	1,581	10.9
1965-1969	1,774	12.6
1970-1974	1,234	8.7
1975-1979	872	6.1
1980-1984	429	3.2
1985-1989	465	2.9
1990-1994	514	2.9
1995-2000	302	1.7

**Nota:** (a) Las cifras relativas han sido ponderadas

La distribución porcentual por los distintos grupos etarios es similar en cada una de las encuestas (Cuadro 12), aunque el número de casos disminuye nuevamente en las últimas tres encuestas, en todos los grupos de edad, a causa del tamaño de muestra más pequeño.

**Cuadro 12**

Número y porcentaje de mujeres entrevistadas, según grupos de edad y periodo. Costa Rica, 1976-2015.

Grupo de edad	Periodo						Total
	1976	1986	1993	1999	2010	2015	
<b>Total</b>	3935	3527	3618	1029	1206	1290	14605
<b>Porcentaje <sup>/a</sup></b>							
15-19	-	19.9	16.9	8.9 <sup>b</sup>	25.2	18.2	13.3
20-24	25.1	19.2	16.5	21.1	20.2	17.2	20.2
25-29	21.3	18.8	18.7	18.9	15.3	15.9	18.9
30-34	16.6	15.2	16.6	18.4	9.7	13.6	15.6
35-39	14.8	11.4	13.6	18.5	11.2	14.2	13.6
40-44	11.4	8.4	10.4	14.3	10.5	12.4	10.6
45-49	10.8	7.1	7.2	-	8.0	8.5	7.8
<b>Número de casos</b>							
15-19	-	702	613	92 <sup>b</sup>	304	235	1946
20-24	986	677	598	217	243	222	2943
25-29	839	662	676	194	185	205	2761
30-34	653	537	601	189	117	176	2273
35-39	583	403	493	190	135	183	1987
40-44	448	296	375	147	126	160	1552
45-49	426	250	262	-	96	109	1143

**Nota:** (a) Las cifras relativas han sido ponderadas (b) Incluye solamente mujeres entre 18 y 19 años de edad.

El Cuadro 13 muestra el número y el porcentaje de casos según cohorte de nacimiento y periodo. Las mujeres entrevistadas a lo largo del periodo de estudio nacieron entre los años 1926 y 2000. Los quinquenios cercanos a ambos extremos cuentan con un menor número de casos, mientras que las cohortes de nacimiento de 1945 hasta 1974 suman el mayor número de casos, pues agrupan información proveniente de al menos tres encuestas.

A lo largo del periodo de estudio, aumentó el porcentaje de mujeres que residía en la zona urbana del país (Cuadro 14). En los primeros años, aproximadamente la mitad de las mujeres habitaba en la zona urbana del país, en la actualidad esa proporción corresponde prácticamente a tres cuartas partes de la población.

**Cuadro 13**

Número y porcentaje de mujeres entrevistadas, según cohorte de nacimiento y periodo. Costa Rica, 1976-2015

Cohorte	Periodo						Total
	1976	1986	1993	1999	2010	2015	
<b>Total</b>	3935	3527	3618	1029	1206	1290	14605
<b>Porcentaje <sup>/a</sup></b>							
1926-1929	7.1	-	-	-	-	-	1.9
1930-1934	11.1	-	-	-	-	-	3.0
1935-1939	13.8	5.3	-	-	-	-	5.0
1940-1944	16.0	8.1	2.4	-	-	-	6.9
1945-1949	19.1	10.4	8.6	-	-	-	9.8
1950-1954	24.2	14.4	11.3	0.1	-	-	12.8
1955-1959	8.7	17.6	15.9	14.7	-	-	11.6
1960-1964	-	19.5	17.2	18.5	8.4	-	10.9
1965-1969	-	20.6	17.7	18.4	13.5	10.0	12.6
1970-1974	-	4.3	17.1	19.1	13.0	11.7	8.7
1975-1979	-	-	9.9	20.9	12.4	13.3	6.1
1980-1984	-	-	-	8.5	16.5	14.4	3.2
1985-1989	-	-	-	-	17.4	16.4	2.9
1990-1994	-	-	-	-	17.5	15.9	2.9
1995-2000	-	-	-	-	1.4	18.4	1.7
<b>Número de casos</b>							
1926-1929	278	-	-	-	-	-	278
1930-1934	438	-	-	-	-	-	438
1935-1939	544	185	-	-	-	-	729
1940-1944	630	284	87	-	-	-	1001
1945-1949	752	367	311	-	-	-	1430
1950-1954	952	509	408	1	-	-	1870
1955-1959	341	621	575	151	-	-	1688
1960-1964	-	686	621	190	84	-	1581
1965-1969	-	725	641	189	133	86	1774
1970-1974	-	150	617	196	125	146	1234
1975-1979	-	-	358	215	119	180	872
1980-1984	-	-	-	87	164	178	429
1985-1989	-	-	-	-	250	215	465
1990-1994	-	-	-	-	302	212	514
1995-2000	-	-	-	-	29	273	302

**Nota:** (a) Las cifras relativas han sido ponderadas

La población femenina que se contempla en este estudio tiene una edad promedio cercana a los 30 años, cifra que se mantiene a lo largo del periodo de estudio. La edad mediana es, en términos generales, similar a lo largo de las encuestas. Es mayor en 1976 y 1999, cuando se entrevistó a mujeres de 20 años o más, en el primer caso, y entre 18 y 44 años, en el segundo.

**Cuadro 14**

Características de las mujeres por periodo y principales variables analizadas. Costa Rica, 1975-2015.

Variable	Año de la encuesta						Total
	1976	1986	1993	1999	2010	2015	
<b>Total</b>	3935	3527	3618	1029	1206	1290	14605
<b>Zona de residencia</b>							
Rural	44.7	43.6	54.6	52.0	39.3	25.7	45.3
Urbano	55.3	56.4	45.4	48.0	60.7	74.3	54.7
<b>Edad</b>							
- Promedio	31.9	28.6	29.6	30.1	30.3	31.0	30.2
- Desv. estándar	8.5	9.3	9.2	7.7	9.8	10.1	9.1
- Mediana	31	27	29	30	29	30	29
<b>Nivel educativo</b>							
Primaria o menos	70.6	53.9	50.0	45.5	32.9	29.6	53.0
Secundaria	19.2	35.5	35.1	37.9	45.3	45.8	32.9
Para/Universitaria	10.2	10.6	15.0	16.6	21.8	24.7	14.2
- Promedio	6.0	7.4	8.0	8.4	9.5	10.1	7.6
- Desv. estándar	4.0	3.7	3.8	3.9	4.5	4.7	4.2
<b>Estado conyugal</b>							
Nunca en unión	22.8	30.6	29.1	23.0	30.2	32.1	27.7
Alguna vez en unión	77.2	69.4	71.0	77.0	69.8	67.9	72.3
<b>Ocupación</b>							
Trabaja actualmente	31.5	31.6	35.2	41.0	33.8	32.6	33.4
No trabaja	68.5	68.4	64.8	59.0	66.2	67.4	66.6

El aumento de la escolaridad entre las mujeres puede notarse en las cifras presentadas en el Cuadro 14. Mientras que en el año 1976 prácticamente el 70% de las mujeres no había completado más allá de la educación primaria, una década después la proporción de mujeres

con ese nivel educativo había disminuido a cerca del 50%. En las encuestas de 2010 y 2015, la proporción de mujeres poco escolarizadas no supera la tercera parte de la población. A lo largo del periodo, aumentó progresivamente el porcentaje de mujeres con educación superior. El aumento de la escolaridad, también puede notarse en la estimación de los años promedio de escolaridad (Cuadro 14), que pasó de 6 años en 1976, a 7.4 en 1986, hasta aproximarse a los 10 años en 2010 y 2015.

Cerca de una tercera parte de las mujeres no había vivido en unión, proporción que varió relativamente poco a lo largo del periodo en estudio. Esta cifra fue más baja en 1976 y 1999, pero en estas encuestas el rango etario de las mujeres entrevistadas fue mayor, excluyendo a las más jóvenes quienes en muchos casos no habían tenido la oportunidad de iniciar una vida en pareja.

El porcentaje de mujeres que se encontraba trabajando tampoco varió sustancialmente y corresponde a poco más del 30%. Estas cifras, sin embargo, se refieren a la actividad en un momento concreto, pero no a la experiencia laboral que haya tenido previamente.

Las características de las mujeres entrevistadas según edad se muestran en el Cuadro 15. El porcentaje de mujeres que vive en la zona urbana del país es similar en todos los grupos etarios, poco más del 50%.

El nivel educativo tiende a ser mayor entre las mujeres más jóvenes. Se observa un mayor porcentaje de mujeres con educación superior en los primeros grupos de edad, salvo entre las mujeres de 15 a 19 años. La menor escolaridad en ese grupo etario en particular, obedece en muchas ocasiones, a que por su edad estas mujeres aún no han completado los niveles educativos previos, pero como consecuencia precisamente de su edad. Los años promedio de educación también denotan mayor escolaridad entre las mujeres jóvenes, particularmente entre las menores de 30 años.

El porcentaje de mujeres que ha vivido alguna vez en unión aumenta progresivamente entre los grupos de edad. Esa cifra corresponde a 19% entre las adolescentes y 90.8% entre las

mujeres mayores de 45 años. Esta relación es esperable, puesto que las mujeres más jóvenes han vivido menor tiempo de exposición a una unión conyugal. Este hecho ocurre también, en alguna medida, al referirse a la ocupación que se ubica por encima del 30% entre las mujeres de 25 años o más y los grupos etarios subsecuentes.

**Cuadro 15**

Características de las mujeres por grupos de edad y principales variables analizadas. Costa Rica, 1975-2015

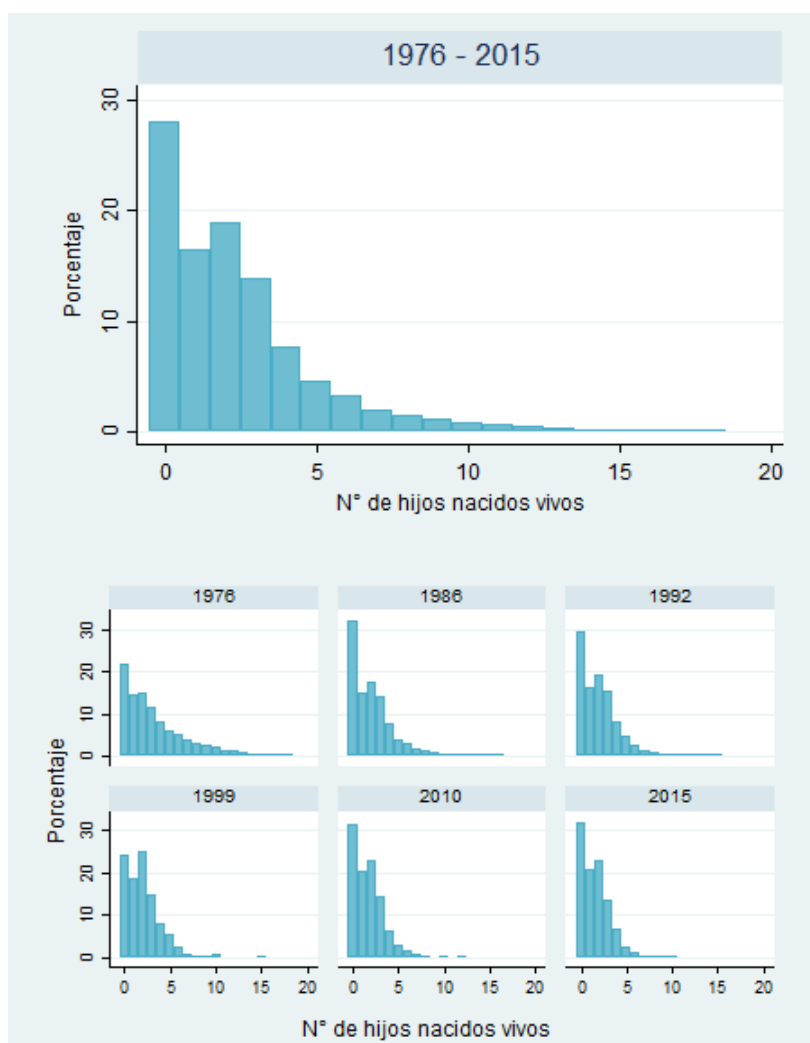
Variable	Edad							Total
	15-19	20-24	25-29	30-34	35-39	40-44	45-49	
<b>Total</b>	1946	2943	2761	2273	1987	1552	1143	14605
<b>Zona de residencia</b>								
Rural	44.5	45.9	44.4	46.3	46.8	44.9	42.8	45.3
Urbano	55.5	54.1	55.6	53.7	53.2	55.2	57.2	54.7
<b>Nivel educativo</b>								
Primaria o menos	36.2	46.4	49.1	54.2	59.3	67.6	70.6	53.0
Secundaria	57.0	37.3	34.6	29.2	25.9	19.7	18.0	32.9
Para/Universitaria	6.7	16.3	16.3	16.6	14.9	12.7	11.4	14.2
- Promedio	8.1	8.3	8.1	7.8	7.3	6.5	6.1	7.6
- Desv. estándar	2.7	3.7	4.2	4.5	4.5	4.6	4.6	4.2
<b>Estado conyugal</b>								
Nunca en unión	81.0	43.0	20.8	12.1	11.0	8.7	9.3	27.7
Alguna vez en unión	19.0	57.0	79.2	87.9	89.1	91.3	90.8	72.3
<b>Ocupación</b>								
Trabaja actualmente	21.3	26.0	31.6	33.9	36.8	33.9	30.5	30.4
No trabaja	78.7	74.1	68.4	66.2	63.2	66.1	69.5	69.6

## 1.2. Número de hijos tenidos

El Gráfico 10 muestra la distribución del número de hijos nacidos vivos. En el panel superior de este gráfico se muestra la distribución referente a toda la muestra, mientras que en la sección inferior se presenta la distribución en cada encuesta.

**Gráfico 10**

Distribución del número de hijos nacidos vivos para el total de la muestra y por periodo. Costa Rica, 1976-2015.



En todos los casos la distribución posee asimetría positiva, propia del tipo de información analizada. Sin embargo, en las encuestas más recientes, hay un mayor porcentaje de mujeres sin hijos o con pocos hijos<sup>11</sup>. Mientras en 1976 el 51.6% de las mujeres tenía 2 hijos o menos, en 1992 esa cifra corresponde a 65.3% y a 75.5% en 2015.

<sup>11</sup> Hay que recordar que en 1976 se entrevistaron mujeres entre los 20 y 49 años de edad y en la encuesta de 1999 a mujeres en el rango de 18 a 44 años. Debido a que no se capta información de las mujeres más jóvenes, la distribución del número de hijos en esos años podría concentrarse más a la derecha, en vista de que estas mujeres han tenido mayor tiempo de exposición al embarazo.

**Cuadro 16**

Número promedio de hijos nacidos vivos que han tenido las mujeres según edad, periodo y cohorte. Costa Rica, 1976-2015

<b>Variable</b>	<b>Promedio</b>	<b>Desv. Estándar</b>
<b>Total</b>	<b>2.4</b>	<b>2.6</b>
<b>Edad</b>		
15-19	0.2	0.5
20-24	0.9	1.1
25-29	1.9	1.5
30-34	2.8	2.0
35-39	3.5	2.6
40-44	4.3	3.2
45-49	5.0	3.7
<b>Periodo</b>		
1976	3.4	3.4
1986	2.2	2.4
1992	2.1	2.1
1999	2.0	1.8
2010	1.6	1.6
2015	1.6	1.6
<b>Cohorte</b>		
1926-1929	6.6	4.4
1930-1934	6.7	4.2
1935-1939	5.1	3.5
1940-1944	4.3	3.0
1945-1949	3.1	2.4
1950-1954	2.2	2.0
1955-1959	2.3	1.9
1960-1964	1.9	1.6
1965-1969	1.4	1.5
1970-1974	1.1	1.4
1975-1979	1.0	1.4
1980-1984	1.5	1.3
1985-1989	1.1	1.1
1990-1994	0.5	0.9
1995-2000	0.2	0.7

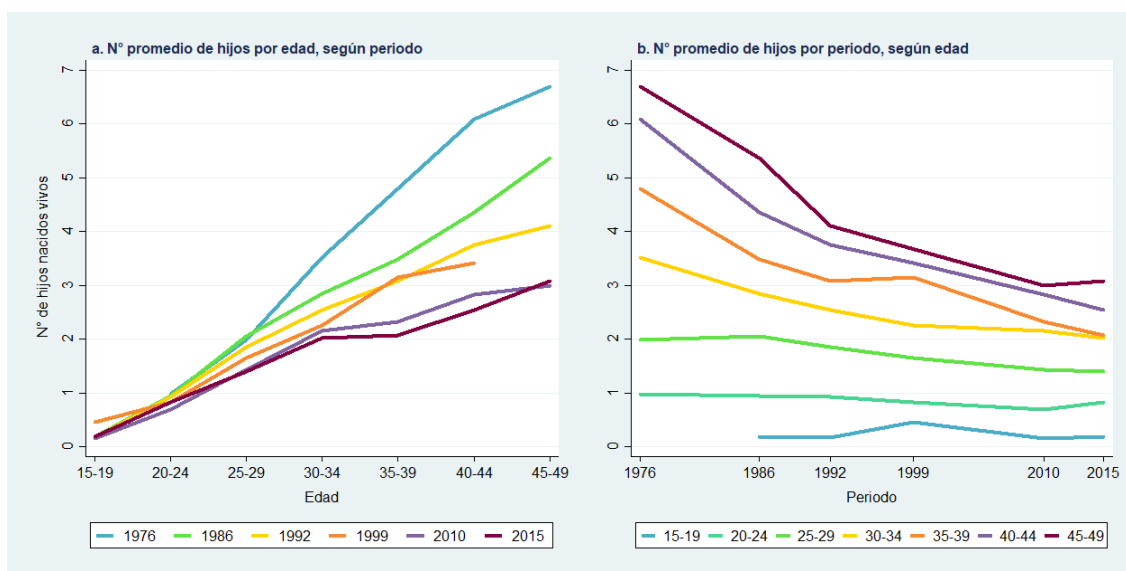
El Cuadro 16 muestra el número promedio de hijos que han tenido las mujeres según edad, periodo y cohorte. Entre las mujeres más jóvenes, el número promedio de hijos no es superior a uno. Se observa un incremento del número promedio de hijos a lo largo de los grupos etarios y una reducción del número de hijos al considerar el periodo y la cohorte. Sin embargo, es necesario resaltar que las mujeres de los primeros grupos etarios y las cohortes más recientes, han tenido menor tiempo de exposición al embarazo y no necesariamente han alcanzado el número de hijos que finalmente tendrán a lo largo de su ciclo reproductivo.

El Gráfico 11a, en el panel de la izquierda, muestra el número promedio de hijos para cada periodo y cómo varía en función de la edad. Esa cifra se incrementa a medida que también lo hace la edad, tendencia que se mantiene en todas las encuestas. Sin embargo, si se compara el número medio de hijos a una edad dada, esta imagen sugiere que progresivamente las mujeres de cohortes recientes tenían menos hijos, en cuanto que, la fecundidad fue más alta en las mujeres entrevistadas en los primeros años y disminuyó en los más recientes. Además, es posible notar que las mujeres tenían menos hijos al final de su ciclo reproductivo. En 1976, las mujeres de 45 a 49 años tenían en promedio 6.7 hijos mientras que en 2015 las mujeres de ese mismo grupo etario tenían en promedio 3.1 hijos.

El gráfico anterior, guarda correspondencia con el que se observa en el panel de la derecha. En el Gráfico 11b se presentan las mismas variables, pero se ha invertido el ordenamiento. En este caso se estimó el número promedio de hijos tenidos por cada grupo de edad y como varía entre las distintas encuestas. Esta nueva visualización permite identificar que para un mismo grupo etario el número promedio de hijos termina siendo diferente, con tendencia a la baja. Las diferencias más notables se presentan en los últimos tres grupos etarios, en los que esa medida disminuyó cerca de 3 hijos, en el periodo 1976-2015. El número promedio de hijos prácticamente no varió en los grupos de 15-19 y 20-24. La información de ambas figuras permite derivar que cada cohorte tiene una experiencia distinta a cada edad y periodo.

### Gráfico 11

Número promedio de hijos nacidos vivos según periodo y grupos de edad.  
Costa Rica, 1976 – 2015.



El descenso de la fecundidad en las últimas décadas tuvo diferentes gradientes según distintos grupos, como lo sugiere la información presentada en el Cuadro 17. En general, se tiene como patrón que el número medio de hijos es menor en la zona urbana, en mujeres de mayor nivel educativo, que nunca han vivido en unión y entre aquellas que trabajaban.

Al observar el número medio de hijos por periodo, es posible notar que la fecundidad empezó a descender desde cifras muy diferentes si se comparan los distintos grupos sociodemográficos. El número medio de hijos disminuyó de manera importante en el periodo 1976-1986, lo que coincide con la estimación obtenida de la tasa global de fecundidad y su patrón histórico a partir de los registros de nacimientos. Sin embargo, la razón a la que decreció el número medio de hijos fue mayor en la zona rural que en la urbana, entre las mujeres con educación primaria (el número medio de hijos entre las mujeres más educadas prácticamente no cambió) y entre las que trabajaban. En el último periodo las diferencias entre la zona residencial urbana y la rural, entre los grupos educativos y la ocupación tienden a estrecharse.

**Cuadro 17**

Número promedio de hijos nacidos vivos por periodo, edad y cohorte,  
según variables de estudio. Costa Rica, 1976-2015

Variable	Zona de residencia		Nivel educativo			Estado conyugal		Ocupación	
	Rural	Urbano	Primaria	Secund	Para/Univ	Nunca en unión	Alguna vez	Trabaja actualm	No trabaja
<b>Total</b>	<b>2.8</b>	<b>2.0</b>	<b>3.2</b>	<b>1.5</b>	<b>1.3</b>	<b>0.3</b>	<b>3.2</b>	<b>2.0</b>	<b>2.5</b>
<b>Periodo</b>									
1976	4.3	2.6	4.1	1.8	1.3	0.5	4.2	3.7	3.3
1986	2.5	1.9	2.9	1.3	1.4	0.2	3.1	1.7	2.4
1992	2.3	1.7	2.7	1.5	1.4	0.3	2.8	1.7	2.3
1999	2.2	1.8	2.6	1.7	1.1	0.3	2.5	1.7	2.2
2010	1.8	1.5	2.5	1.3	1.1	0.3	2.2	1.5	1.7
2015	1.8	1.5	2.4	1.3	1.1	0.4	2.2	1.5	1.6
<b>Edad</b>									
15-19	0.3	0.1	0.3	0.1	0.0	0.0	0.9	0.1	0.2
20-24	1.1	0.7	1.3	0.7	0.2	0.2	1.5	0.6	1.0
25-29	2.2	1.6	2.4	1.5	1.0	0.4	2.2	1.4	2.1
30-34	3.3	2.4	3.4	2.3	1.8	0.8	3.1	2.2	3.1
35-39	4.1	3.0	4.2	2.7	2.0	0.9	3.8	3.0	3.8
40-44	5.3	3.5	5.0	3.0	2.4	1.2	4.6	3.4	4.7
45-49	6.2	4.2	5.9	3.3	2.5	1.4	5.4	4.1	5.4
<b>Cohorte</b>									
1926	8.2	5.3	7.2	3.8	2.1	2.1	7.1	6.3	6.6
1930	8.5	5.1	7.2	4.5	3.5	1.3	7.2	6.0	6.9
1935	6.3	4.0	5.5	3.2	2.8	0.9	5.7	4.7	5.2
1940	5.3	3.5	4.9	2.7	2.3	0.9	4.7	3.7	4.5
1945	3.9	2.5	3.5	2.2	1.9	0.6	3.5	2.8	3.2
1950	2.7	1.8	2.6	1.8	1.4	0.4	2.8	2.4	2.2
1955	2.7	2.0	2.7	2.0	1.6	0.4	2.8	2.0	2.4
1960	2.2	1.7	2.5	1.6	1.1	0.4	2.4	1.5	2.2
1965	1.5	1.2	1.6	1.2	1.1	0.2	2.1	1.1	1.5
1970	1.1	1.1	1.4	0.9	0.7	0.1	1.9	1.0	1.1
1975	0.8	1.2	1.3	0.7	1.1	0.1	2.0	0.9	1.1
1980	1.5	1.5	1.9	1.4	1.2	0.4	1.8	1.2	1.7
1985	1.3	1.0	1.8	1.2	0.5	0.4	1.6	0.8	1.3
1990	0.6	0.5	1.1	0.5	0.3	0.1	1.1	0.4	0.5
1995	0.2	0.2	0.6	0.2	0.0	0.1	1.0	0.1	0.2

La relación de estas variables socioeconómicas con el número de hijos se mantiene al analizar los componentes de edad y cohorte. En el primer caso, cabe destacar que el número medio de hijos que alcanzan las mujeres de 45-49 años, grupo etario en el que se supone que la fecundidad es completa, es mayor entre las mujeres de la zona rural, de menor nivel educativo, en unión y las no ocupadas. Al referirse al componente cohorte, la fecundidad es mayor en las primeras cohortes y bastante menor en las cohortes más jóvenes, sin embargo cabe destacar que éstas últimas han tenido menos tiempo de exposición al embarazo.

## 2. MLMG para datos de fecundidad

El Cuadro 18 muestra los resultados obtenidos tras la estimación de un modelo de efectos fijos, que incorpora solo las variables del primer nivel (Ec. 37) con el propósito de determinar la relevancia de incluir las variables de cohorte y periodo como componentes aleatorios.

**Cuadro 18**

Resultados de la estimación del modelo de efectos fijos,  
sin incluir las variables cohorte y periodo. Costa Rica, 1976-2015

Variable	Coefficiente	exp(Coef)	Error estándar	Valor t	Pr >  t
<b>Intercepto</b>	-0.677		0.030	-22.2	<.0001
<b>Edad</b>	0.076	1.079	0.001	68.9	<.0001
<b>Edad<sup>2</sup></b>	-0.002	0.998	0.000	-25.2	<.0001
<b>Años de educación</b>	-0.054	0.948	0.002	-34.5	<.0001
<b>Estado conyugal</b>					
Alguna vez en unión	1.576	4.837	0.030	53.2	<.0001
Nunca en unión		- Categoría de referencia -			
<b>Zona de residencia</b>					
Rural	0.157	1.169	0.011	13.7	<.0001
Urbano		- Categoría de referencia -			
<b>Ocupación</b>					
Trabaja actualmente	-0.128	0.880	0.013	-10.1	<.0001
No trabaja		- Categoría de referencia -			

Tanto en el modelo cuyos resultados se presentan en el Cuadro 18 como en el que se presenta en el Cuadro 19, la variable dependiente es el número de hijos que han tenido las mujeres. El modelo se ajustó mediante una regresión de Poisson que modela el logaritmo del conteo esperado como una función de las variables explicativas pero en el que no se incluyó el tiempo de exposición. Los coeficientes de regresión expresan la diferencia en el logaritmo del conteo esperado cuando la variable explicativa en cuestión cambia en una unidad, bajo la condición de que las variables explicativas restantes se mantienen constantes. En la columna  $\exp(\text{Coef})$  se muestra el valor del coeficiente de regresión tras aplicar la función exponencial y su valor se interpreta como un cambio relativo en el número de hijos.

El modelo presentado en el Cuadro 18, estimado para propósitos comparativos, muestra que todas las variables en análisis tienen coeficientes significativos y se relacionan con la fecundidad de la forma esperada. La edad, vivir en unión conyugal y residir en la zona rural están asociadas positivamente con el número de hijos tenidos. Mientras que un mayor nivel educativo y tener una ocupación, se asocian con un menor número de hijos.

El efecto cuadrático de la variable edad resultó estadísticamente significativo, denotando que el número medio de hijos aumenta con la edad pero este conteo se desacelera a mayores edades, lo que es esperable al referirse a la fecundidad.

El Cuadro 19 muestra los resultados de la estimación del MLMG, de efectos fijos y aleatorios, para explicar los cambios en el número de hijos nacidos vivos. A diferencia del modelo presentado en el Cuadro 18, este último incorpora los efectos de periodo y cohorte como efectos aleatorios, conservando las mismas variables de efectos fijos.

Los resultados muestran que el efecto neto de la edad, controlado por los efectos aleatorios de periodo y cohorte, es significativo pero no lineal, en vista de que el efecto cuadrático de la edad también es significativo. Este coeficiente negativo denota una relación cóncava (con respecto al origen) con el número de hijos, que en términos sustantivos, implica que el número medio de hijos aumenta con la edad, pero este conteo se desacelera a mayores edades.

**Cuadro 19**

Resultados de la estimación del MLMG. Costa Rica, 1976-2015

<b>Efectos fijos</b>	<b>Coefficiente</b>	<b>exp(Coef)</b>	<b>Error estándar</b>	<b>Valor t</b>	<b>Pr &gt;  t </b>
<b>Intercepto</b>	-0.727		0.098	-7.4	0.0007
<b>Edad</b>	0.074	1.077	0.002	43.0	<.0001
<b>Edad<sup>2</sup></b>	-0.002	0.998	0.000	-24.8	<.0001
<b>Años de educación</b>	-0.055	0.947	0.002	-34.3	<.0001
<b>Estado conyugal</b>					
Alguna vez en unión	1.587	4.890	0.030	53.5	<.0001
Nunca en unión					- Categoría de referencia -
<b>Zona de residencia</b>					
Rural	0.138	1.148	0.012	11.9	<.0001
Urbano					- Categoría de referencia -
<b>Ocupación</b>					
Trabaja actualmente	-0.093	0.911	0.013	-7.3	<.0001
No trabaja					- Categoría de referencia -
<b>Efectos aleatorios</b>	<b>Coefficiente</b>	<b>exp(Coef)</b>	<b>Error estándar</b>	<b>Valor t</b>	<b>Pr &gt;  t </b>
<b>Periodo</b>					
1976	0.311	1.365	0.093	3.36	0.001
1986	0.155	1.168	0.090	1.72	0.086
1992	0.045	1.046	0.089	0.50	0.616
1999	-0.056	0.946	0.091	-0.61	0.540
2010	-0.218	0.804	0.092	-2.37	0.018
2015	-0.238	0.788	0.094	-2.54	0.011
<b>Cohorte</b>					
1926-1929	0.156	1.169	0.056	2.80	0.005
1930-1934	0.157	1.170	0.050	3.15	0.002
1935-1939	0.042	1.042	0.044	0.94	0.348
1940-1944	0.002	1.002	0.040	0.05	0.963
1945-1949	-0.106	0.899	0.037	-2.87	0.004
1950-1954	-0.138	0.871	0.035	-3.92	<.0001
1955-1959	-0.068	0.934	0.035	-1.96	0.050
1960-1964	-0.060	0.941	0.035	-1.71	0.088
1965-1969	-0.046	0.955	0.037	-1.27	0.205
1970-1974	-0.127	0.880	0.041	-3.08	0.002
1975-1979	-0.067	0.935	0.047	-1.42	0.155
1980-1984	-0.022	0.978	0.052	-0.42	0.672
1985-1989	0.076	1.079	0.059	1.29	0.197
1990-1994	0.120	1.127	0.070	1.70	0.088
1995-2000	0.084	1.088	0.089	0.95	0.342
<b>Componentes de varianza</b>	<b>Varianza</b>		<b>Error estándar</b>	<b>Valor z</b>	<b>Pr &gt;  z </b>
Periodo	0.047		0.031	1.51	0.066
Cohorte	0.012		0.006	2.13	0.017

El efecto de la edad se ilustra en el Gráfico 12. Para la edad  $i$ -ésima se calculó  $\hat{\beta}_i^* = \hat{\beta}_1(\text{Edad centrada}) + \hat{\beta}_2(\text{Edad centrada}^2)$ , donde  $\hat{\beta}_1$  y  $\hat{\beta}_2$  son los coeficientes de regresión de la edad y la edad al cuadrado, respectivamente. Las variables edad y la edad al cuadrado fueron centradas con respecto a la edad promedio de toda la muestra, por lo que esta estimación de su efecto refleja variaciones con respecto a ese valor. La edad promedio de todas las mujeres en la muestra fue 30.18 años. El gráfico muestra los valores de  $\exp(\hat{\beta}_i^*)$ .

En el Gráfico 12 se visualiza fácilmente el patrón descrito anteriormente. También se observa que el incremento en el número de hijos es menos marcado entre las mujeres menores de 20 años. El conteo se incrementa más rápidamente entre los 20 y 40 años, aproximadamente. Para las mujeres menores de 30 años esta estimación toma valores por debajo de 1, lo que indica que estas mujeres tienen menos hijos que una mujer con una edad igual a la edad promedio de la muestra. Así por ejemplo, el efecto de edad estimado a los 15 años es  $\exp(\hat{\beta}_{15}^*)=0.20$ , lo que quiere decir que una mujer de 15 años tiene 80% menos hijos que una mujer de 30.18 años. Por encima de los 30 años, estos coeficientes son positivos, lo que denota un incremento en el número de hijos con respecto a esa edad promedio. Para una mujer de 49 años  $\exp(\hat{\beta}_{49}^*)=1.98$ , lo que expresa que tras controlar por el efecto de otras variables explicativas, las mujeres llegan al final de su periodo reproductivo con casi el doble de hijos que una mujer de poco más de 30 años.

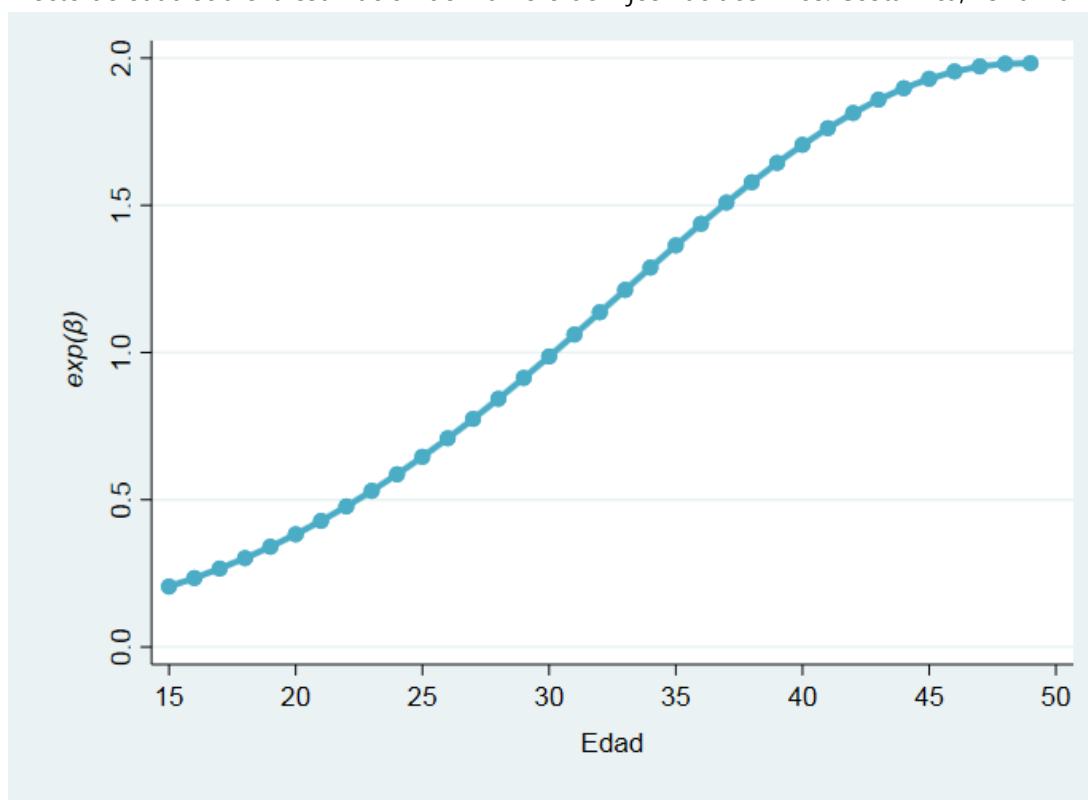
La información de los componentes fijos restantes muestra que el número promedio de hijos varía al considerar estas covariables. Poseer mayor nivel educativo y tener una ocupación, en términos generales, están relacionados con un menor número promedio de hijos. El aumento de un año en el nivel educativo de una mujer, por encima del nivel educativo promedio de su cohorte de nacimiento, conlleva a una reducción del número de hijos de 5.2%. Mientras que, entre las mujeres que han tenido una ocupación el número de hijos se reduce en un 8.9%, si se les compara con las mujeres nunca ocupadas.

Haber vivido alguna vez en unión y ser residente de la zona rural del país son condiciones que se relacionan, en promedio, con un mayor número de hijos. Las mujeres que vivieron alguna

vez en unión tuvieron un número de hijos 4.9 veces mayor que el que tuvieron las mujeres que nunca habían vivido en unión conyugal. Las mujeres de la zona rural tuvieron en promedio un número de hijos 14.8% mayor que las mujeres residentes de la zona urbana.

**Gráfico 12**

Efecto de edad sobre la estimación del número de hijos nacidos vivos. Costa Rica, 1976-2015



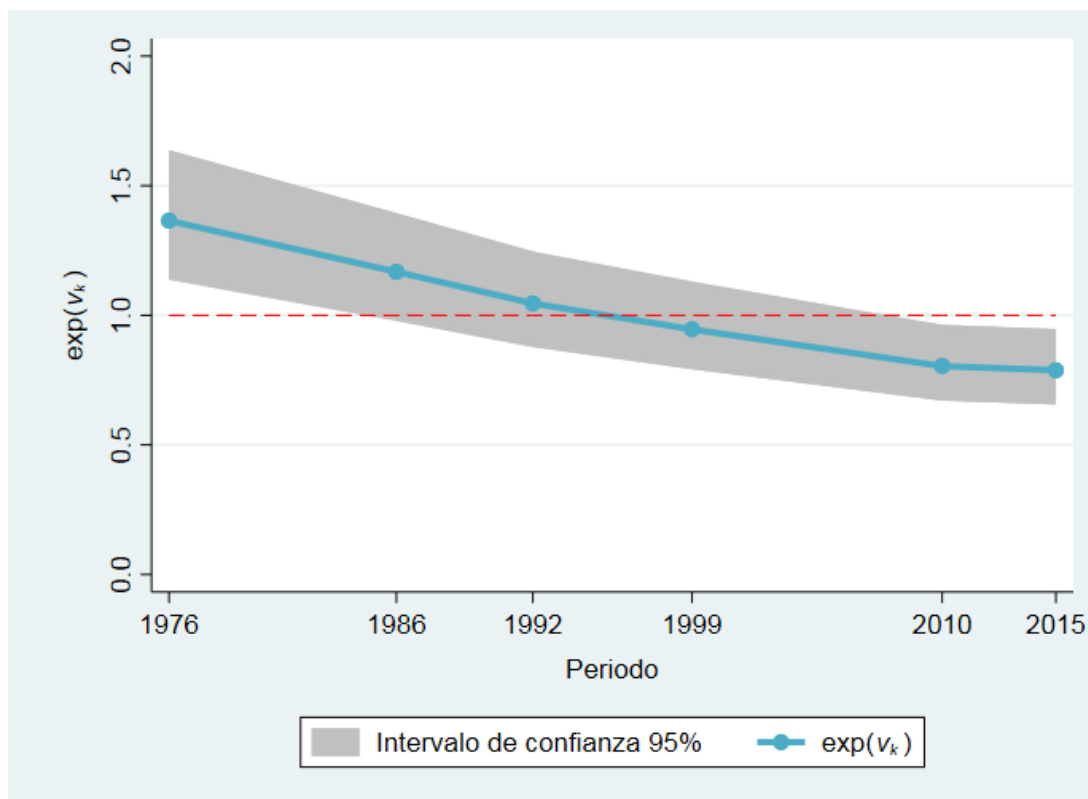
El comando GLIMMIX de SAS estima los parámetros de efectos aleatorios como desviaciones con respecto al intercepto global  $\hat{\gamma}_0$ , su estimación se aprecia en la parte inferior del Cuadro 19.

La inspección visual de estos efectos de periodo y cohorte facilita la comprensión de la tendencia temporal que han tomado, teniendo presente que este no es un análisis de tendencias ni de series temporales, en este caso se habla del patrón de cambio de un periodo a otro o de una cohorte a otra. En el Gráfico 13 se presenta el efecto de periodo, aplicando la función exponencial, i.e.  $exp(\nu_{0k})$ , como inversa de la función de enlace empleada en la

estimación del modelo. El Gráfico 14 presenta información equivalente para el efecto cohorte:  $exp(u_{0k})$ .

**Gráfico 13**

Efecto de periodo sobre la estimación del número de hijos nacidos vivos. Costa Rica, 1976-2015



La tendencia del efecto de periodo es decreciente y solo los efectos de 1976, 2010 y 2015 se desvían sustancialmente de la media general (Gráfico 13). El efecto de periodo decrece en forma prácticamente lineal, esto quiere decir que una mujer que experimentó su vida reproductiva en 1976 tiene en promedio un mayor número de hijos que una mujer que lo hizo más cerca del presente, particularmente en 2010 y 2015. A pesar de esta tendencia, los resultados del modelo de regresión (Cuadro 19) indican que son estadísticamente significativos ( $\alpha=0.05$ ) solamente los efectos de periodo 1976, 2010 y 2015.

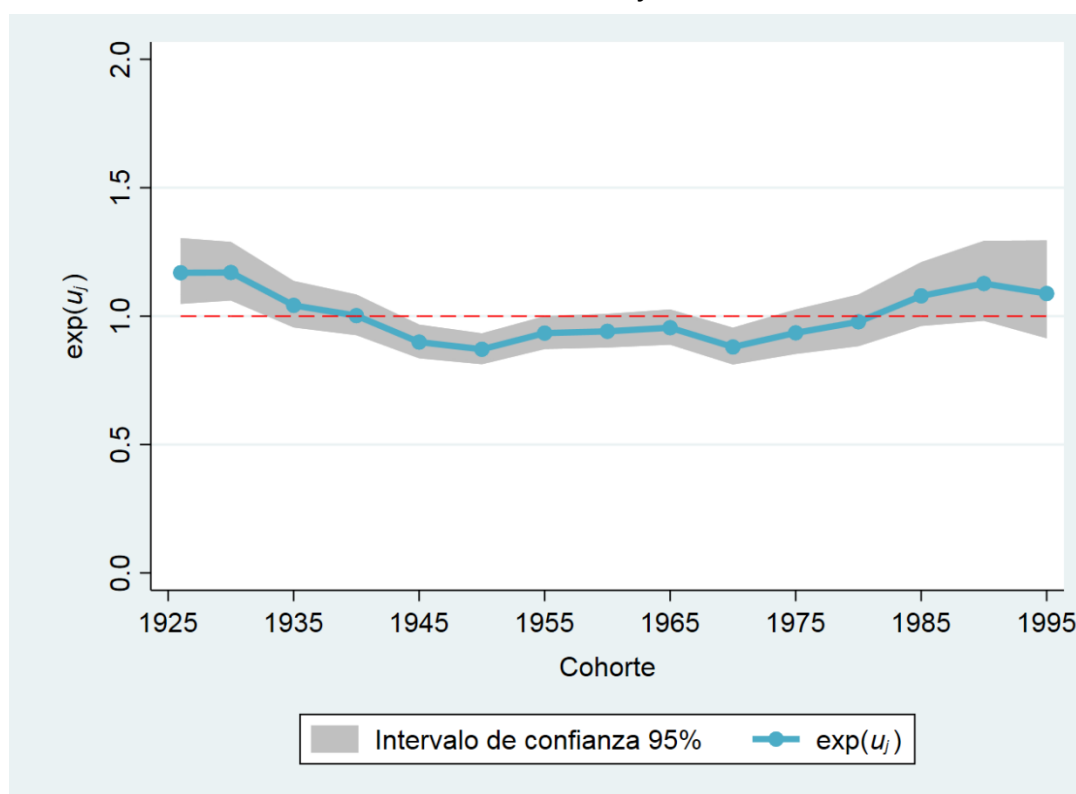
Los efectos de periodo que resultaron estadísticamente significativos constituyen los puntos más extremos de la distribución. En 1976 aun cuando la fecundidad ya había caído drásticamente, si se le compara con la tasa observada en los años sesenta, en ese momento mostraba las mayores cifras de todo el periodo estudiado en este trabajo (1976-2015), por lo que es esperable que el coeficiente de 1976 sugiera que ese periodo contribuye a un aumento de la fecundidad. Tras controlar por el efecto de las variables explicativas incorporadas al modelo y también por el efecto de cohorte, el efecto de periodo indica que en ese momento las mujeres tenían 36.5% más hijos, por encima del número promedio de hijos que tuvieron las mujeres a lo largo de todo el periodo de estudio 1976-2015. En los años 2010 y 2015 la relación se invierte, en ese último año, por ejemplo, el número de hijos que tuvieron las mujeres se redujo en 21.2%.

Los resultados obtenidos para el componente de periodo no permiten desentrañar la influencia que pudieron tener diferentes acontecimientos sociales y económicos sobre el nivel de la fecundidad, más bien muestra la tendencia que tuvo la fecundidad durante el periodo de estudio (Gráfico 8): desde 1976 hasta el año 2015, aún con diferenciales socioeconómicos y a diferente ritmo en diferentes periodos, la fecundidad no ha hecho otra cosa que descender. Esto puede deberse a la periodicidad irregular y distante entre una encuesta y la siguiente, que provee información pobre sobre los fenómenos que pretenden ser captados por el componente periodo. Por ejemplo, Costa Rica experimentó una crisis económica entre 1980 y 1982 momento en el que el país experimentó hiperinflación y un aumento de las tasas de desempleo y subempleo (Barahona, 1999). Se ha documentado que existe una relación negativa entre eventos que generan periodos de crisis económica o de convulsión social con la nupcialidad y la fecundidad (Caldwell, 2006 y Lee, 1990). Así que de contarse con información para esos años, se esperaría observar un efecto de periodo a la baja en la fecundidad. La encuesta de fecundidad que recaba información más cercana a esa fecha es la de 1986, de manera tal que el efecto de periodo que pudo tener la crisis económica de esos años probablemente se confunde con otros eventos que ya de por sí conllevan al descenso de la fecundidad.

La inspección gráfica del efecto de cohorte sugiere un patrón cíclico (Gráfico 14). Se observa una tendencia a la baja hasta las cohortes de mujeres nacidas en 1945-1949. El efecto de las cohortes desde 1945-1949 y hasta 1970-1974, en términos generales, muestra un estancamiento por debajo de la media general. Pero el efecto de cohorte vuelve a aumentar a partir de la cohorte de mujeres nacidas en 1975-1979. Sin embargo, los resultados presentados en el Cuadro 19 señalan que solo fueron estadísticamente significativos los coeficientes de las cohortes de nacimiento: 1926-1929, 1930-1934, 1945-1949, 1950-1954, 1955-1959 y 1970-1974.

**Gráfico 14**

Efecto de cohorte sobre la estimación del número de hijos nacidos vivos. Costa Rica, 1976-2015



El efecto de las primeras dos cohortes (1926-1929 y 1930-1934) se encuentra por encima del promedio global (intercepto) lo que apunta a que las mujeres nacidas en esos años tuvieron en promedio más hijos que todo el grupo de mujeres en estudio. Las mujeres nacidas entre 1926-1929 y 1930-1934 tuvieron un número de hijos 16.9% y 17% mayor, respectivamente.

Ese grupo de mujeres vivió buena parte de su edad reproductiva antes de que iniciara el descenso de la fecundidad y antes de que se introdujeran los métodos anticonceptivos en Costa Rica. A 1962, cuando inician esos acontecimientos, estas mujeres habían cumplido entre 28 y 37 años. Siendo así, estas cohortes empezaron su vida reproductiva cuando no había acceso a métodos anticonceptivos (modernos) y además experimentaron el alza de la fecundidad observada antes de los años sesenta (Gráfico 8). De acuerdo con los datos de las encuestas este grupo de mujeres, en comparación con las cohortes restantes, se caracterizan por ser las que, en promedio, alcanzaron el menor nivel educativo (4.3 años,  $s=3.7$ ) y presentaban el menor porcentaje de ocupación (24.8%) pero también el número deseado de hijos fue el más alto registrado en todas las cohortes ( $\bar{x}=6.2$ ,  $s=4.2$ ).

El siguiente lapso en el que los coeficientes del componente cohorte fueron significativos corresponde a las mujeres nacidas entre 1945-1959. Los coeficientes de regresión fueron negativos (Cuadro 19) lo que sugiere un efecto a la baja sobre la fecundidad (Gráfico 14). Estas cohortes de nacimiento experimentaron la mayor parte de su periodo reproductivo durante el descenso más marcado de la fecundidad en Costa Rica: 1962-1976. En 1962, estas mujeres eran niñas o adolescentes, con edades entre los 3 y 17 años, mientras que en 1976 cuando finalizaba el descenso de la fecundidad, este grupo de mujeres había cumplido entre 17 y 31 años. Esto quiere decir que empezaron su vida reproductiva en plena ampliación de los programas de planificación familiar. En este grupo de mujeres el número promedio de hijos deseados fue de 3.5 ( $s=1.7$ ), casi la mitad de la cifra registrada para las primeras cohortes estudiadas, el porcentaje de mujeres que había tenido una ocupación fue de 33% y los años promedio de educación aumentaron a 7.2 ( $s=4.0$ ). Este grupo de cohortes de nacimiento también transcurre parte de su edad reproductiva durante la crisis económica que tomó lugar en Costa Rica entre 1980 y 1982.

La cohorte de nacimiento 1970-1974 también se asocia con un coeficiente negativo y estadísticamente significativo. El efecto en esta cohorte muestra una contribución a la baja en 11.9%, sobre el número medio de hijos que tendrían las mujeres, controlando ese efecto por el

resto de variables explicativas y también por el efecto de periodo. El efecto a la baja en la fecundidad de esta cohorte podría vincularse con la entrada en vigencia de la Ley de Paternidad Responsable, en el año 2001. Este grupo de mujeres tenía a lo sumo 32 años en ese momento. De acuerdo con Ramos (2010, pág. 17), la tasa global de fecundidad cayó entre un 6 y 10% como impacto inmediato de la regencia de esta ley. Sin embargo, destaca que la fecundidad de las mujeres mayores de 33 años se vio poco afectada, la mayor caída en la fecundidad se observó en las mujeres de 23 a 27 años, precisamente parte del rango etario que contempla la cohorte de nacimiento 1970-1974.

Aunque la influencia de dicha ley se vislumbra como el principal evento vinculante relacionado con la baja fecundidad en esa cohorte, también cabe resaltar que estas mujeres iniciaban su edad reproductiva cuando empezó la segunda fase del descenso de la fecundidad, en los años noventa. En 1990 se encontraban entre los 16 y 20 años, por lo que probablemente iniciaron su vida reproductiva expuestas al cambio de las diversas normas que implicasen el descenso de la fecundidad. Entre 1990 y el 2002, lapso que abarca ese segundo descenso, el número deseado de hijos disminuyó (Rosero-Bixby y Casterline, 1995). Los datos de las encuestas revelan que para esa cohorte en concreto el número deseado de hijos fue de 2.6, en promedio ( $s=1.3$ ).

Para determinar la relevancia de incorporar los efectos de periodo y cohorte se analiza la devianza y los componentes de varianza del modelo. La devianza (37555.36) es bastante mayor que los grados de libertad (14 577) del modelo, lo que denota una asociación importante y significativa entre las variables explicativas y el número de hijos tenidos. Por su parte, los componentes de varianza (Cuadro 19) indican que las variaciones por cohorte son estadísticamente significativas ( $p=0.0166$ ) mientras que las de periodo no ( $p=0.0660$ ).

En este trabajo no se emplean los criterios más usuales para comparar la bondad de ajuste y seleccionar el mejor modelo, como los son el Criterio de información de Akaike (AIC; por sus siglas en inglés; Akaike, 1974), el Criterio de información Bayesiano (BIC; Schwarz, 1978) y el test de razón de verosimilitudes (McKinley y Mills, 1985). Estos tres métodos se estiman a partir de la log-verosimilitud obtenida al estimar cada modelo, sin embargo, los modelos (37) y (39)

se obtuvieron a partir del método de estimación por *máxima verosimilitud restringida* (REML, por sus siglas en inglés) que produce una pseudo-máxima verosimilitud.

En lugar de los métodos basados en la verosimilitud en este trabajo se recurre a la prueba F aproximada, descrita en la sección 2.3.2 y cuyos resultados se presentan en el Cuadro 19, para determinar la relevancia de incorporar los efectos aleatorios como componentes del modelo de regresión.

Esta prueba de hipótesis permite determinar si los efectos de cohorte y periodo, como conjuntos, contribuyen a explicar la varianza en el modelo, en vista de que algunos coeficientes son estadísticamente significativos y otros no. Se prueban las hipótesis planteadas en (30) y (31).

**Cuadro 20**

Resultados de la prueba F para determinar la presencia de efectos aleatorios. Costa Rica, 1976-2015

	<b>Efecto Periodo</b>	<b>Efecto cohorte</b>	<b>Efecto periodo y cohorte</b>
<b>F</b>	<b>87.2</b>	<b>38.1</b>	<b>27.6</b>
<b>Pr &gt; F</b>	<b>&lt; 0.001</b>	<b>&lt; 0.001</b>	<b>&lt; 0.001</b>

Los resultados de la prueba F muestran que la incorporación de los efectos aleatorios de periodo y cohorte producen una variación estadísticamente significativa en el intercepto del nivel 1. Este hecho sugiere, por ende, la relevancia de incluir ambos términos para explicar la trayectoria de la fecundidad.

Al referirse al efecto de periodo, como conjunto, la conclusión que se deriva bajo el estadístico z presentado al analizar los componentes de varianza y la que se observa con la prueba F para probar la presencia de efectos aleatorios no guardan correspondencia, pues con el primero se concluye que el efecto de periodo no es significativo mientras que con la segunda prueba se llega a lo contrario. En cuanto al estadístico F es importante señalar que SAS emplea un algoritmo de estimación que genera una pseudo-máxima verosimilitud, bajo la cual ciertos parámetros se estiman normal, consistente y asintóticamente, pero no se basan en una

estimación máximo-verosímil. En ese caso, la función de log-verosimilitud necesaria para estimar la prueba F asintótica puede solo ser aproximada. El estadístico z, sin embargo, suele brindar resultados más conservadores (Yang y Land, 2013, pág. 221-222).

Ante esta disyuntiva cabe apelar a lógica que se derive de este conjunto de resultados, mismos que sugieren la relevancia de incorporar este componente. En primer lugar, permiten contabilizar y mostrar los importantes cambios que se produjeron en la fecundidad por los efectos del primer periodo y los dos últimos. Los coeficientes de periodo incluso presentan un mayor rango de variación que los observados para el efecto cohorte, es decir, los cambios porcentuales en el número de hijos debidos a los efectos de periodo son mayores en algunos casos, que los efectos de cohorte u otras variables explicativas presentes en el modelo.

## CONCLUSIONES Y DISCUSIÓN

Los resultados de este trabajo sugieren que el descenso de la fecundidad se ha dado sostenidamente a través de los periodos pero con limitaciones en el número de hijos tenidos a través de las cohortes. Las variaciones estimadas para el componente edad muestran que esta variable tiene una contribución importante en la determinación del número de hijos, el efecto identificado está muy atado a lo que se esperaría para esta variable: una contribución de pequeña a modesta durante la adolescencia y por encima de los 40 años y una probabilidad creciente de tener un hijo durante las edades intermedias.

Al menos durante el inicio del descenso de la fecundidad, el cambio está más asociado a los efectos de periodo que de cohorte. Las variaciones más modestas que presenta el efecto de cohorte se contraponen al descenso sostenido del efecto periodo, que pasa de tener una contribución positiva sobre el número promedio de hijos en 1976 a una negativa en los más recientes estadios. El descenso de la fecundidad que se expresa como un efecto de periodo a partir de 1976 y su sintonía con la baja relacionada a las cohortes de nacimiento de 1945 a 1959 (las mujeres en edad fértil entre 1962-1976) concuerda con dos eventos también vinculados con el comienzo del descenso de la fecundidad, discutidos previamente por otros autores: 1) el aumento de la oferta de anticonceptivos modernos así como una fuerte y eficiente ejecución de los programas de planificación familiar y 2) el rápido desarrollo económico que se gestaba desde años atrás (Gómez, 2009 y Rosero, 1983) en conjunto con las mejoras en materia de salud y la creciente cobertura de la seguridad social (Miranda, 1988). La presencia de efectos más pronunciados de periodo que de cohorte durante la primera fase del descenso de la fecundidad son esperables si se coteja con la teoría de la transición demográfica, que establece como precondiciones del cambio el descenso de la mortalidad, la mortalidad infantil, las mejoras en salud y el creciente bienestar socioeconómico de las sociedades (Lundquist, Anderton y Yaukey, 2014; van de Walle, 1992).

Ese hallazgo concuerda con los descritos por otros autores (Hsueh y Anderton, 1990; Kye, 2012 y Pullum, 1980) quienes también estudiaron el efecto de los componentes edad, periodo y

cohorte sobre los cambios de la fecundidad. En estos trabajos se identificó una mayor contribución del efecto periodo que del efecto cohorte (al menos durante el inicio de la transición de la fecundidad). El descenso inicial de la fecundidad fue asociado a la introducción de métodos anticonceptivos, campañas que promovían la planificación familiar y facilitaban el acceso a estos métodos, mejoras en materia de salud y calidad de vida.

Sin embargo, también resalta como una precondition del descenso de la fecundidad, el deseo de las mujeres de limitar su fecundidad o de espaciar los nacimientos (Rosero, 1983 y van de Walle, 1992). La conceptualización sobre el tamaño de la familia es una condición necesaria para adoptar medios que permitan limitar el número de hijos (van de Walle, 1992). Este deseo de familias más pequeñas (de no menos de dos ni de más de cuatro hijos) ya se manifestaba entre las mujeres costarricenses desde la década de 1940 y los datos de encuestas también revelaban la preferencia de familias de tamaño moderado al inicio del descenso de la fecundidad, en la década de 1960 (Rosero, 1983; pág. 93). Esto sugiere que por décadas las mujeres deseaban una familia de tamaño moderado, pero tenían familias grandes, tal cual resaltan Rosero-Bixby y Casterline (1995), quienes además destacan una prevalencia anticonceptiva relativamente baja en la década de los sesenta.

Tras la oferta de métodos efectivos de planificación familiar, de pastillas y DIUs a partir de 1962, de la importación de un mayor volumen de ambos métodos a partir de 1965 y del inicio del Programa Nacional de Planificación Familiar en 1968 desde el que se brindaron anticonceptivos gratuitamente (Rosero, 1983; 1984), el control del tamaño de la familia empezó a ser plausible para las mujeres costarricenses, quienes pasaron de tener los hijos que la biología les permitía al número de hijos que deseaban o incluso a espaciar los nacimientos.

La introducción de métodos anticonceptivos se visualiza entonces, como el detonante de la rápida caída de la fecundidad en la década de los sesenta, que respondía al deseo lejano de controlar la fecundidad. Pero a esta ampliación de la oferta de anticonceptivos, la acompañaban los cambios sociales y económicos que experimentaba el país, que evolucionaba hacia otro estilo de desarrollo, uno llamado "Estado benefactor" o "Estado gestor" en el que el

Estado desempeñaba un rol de patrono, asumiendo el papel de la actividad económica, promoviendo el bienestar social, la salud, educación, vivienda e infraestructura.

Si bien las disminuciones de la fecundidad entre periodos implican adaptaciones a las circunstancias específicas de cada periodo, la limitación que las mujeres y sus parejas establecen al número de hijos en cohortes específicas sugiere un aumento difuso de las innovaciones anticonceptivas y los estilos de vida que limitan la fecundidad. La transición de la fecundidad por interacción social no tiene un efecto inmediato, empieza siendo pequeño y aumenta cuando también lo hace la proporción de mujeres que emplea el control natal (Casterline, 2001).

El aumento del porcentaje de mujeres en unión que había utilizado algún método anticonceptivo se dio rápidamente de un periodo a otro, pero con variaciones por grupos etarios (la combinación del cambio en ambas variables implica variaciones de cohorte). Según datos de las encuestas, en 1976 el 81.6% de las mujeres que habían vivido alguna vez en unión había utilizado un método anticonceptivo y el 72.6% había utilizado alguno moderno (incluyendo la esterilización, en ambos casos). Estas cifras, como se mencionaba, fueron variables por grupos etarios: el porcentaje de mujeres en unión que había empleado algún método rondó el 85% entre las mujeres de 20 y 39 años, pero fue del 80% entre las de 40 a 44 años y de 59% entre las de 45 y 49 años. En las encuestas siguientes, estas proporciones aumentan en todos los grupos etarios, aunque aún se reportaba una menor proporción de usuarias en los dos últimos grupos, con respecto al resto. Esta diferencia si bien se estrecha con el tiempo, no alcanza las mismas cifras que las observadas en las mujeres jóvenes hasta en la encuesta de 1999.

El descenso que se describe por periodo en combinación con el efecto a la baja en las cohortes de nacimiento de 1945-1959, descrita en los resultados de este trabajo, y la mayor proporción de usuarias de métodos entre las jóvenes, es indicativo de la difusión de las innovaciones anticonceptivas, que empezaría entre las mujeres más jóvenes y se extendería a las de mayor edad, subsecuentemente. Esto coincide con lo que Gómez (2009) relata, con relación al patrón

de disminución de la fecundidad que se produjo principalmente entre las mujeres de 20 a 35 años. Rosero-Bixby y Casterline (1995; pág. 49) afirman que “las tasas de fecundidad por edad muestran que las cohortes nacidas entre 1930 y 1945 aproximadamente, fueron responsables de la mayor parte del descenso de la fecundidad en Costa Rica. Las cohortes anteriores y posteriores añadieron poco al descenso de la fecundidad ocurrido en las cohortes 1930-1945”.

Una vez resuelta la necesidad insatisfecha de métodos de planificación familiar, en el presente, el nivel de la fecundidad se vincula más bien con cambios en las dinámicas de género y de la formación de las familias. Los cambios en las dinámicas de género, hacia la igualdad de género, habrían permitido el incremento del nivel educativo de las mujeres, su mayor participación en la fuerza laboral y potenciado el empoderamiento femenino (Goldscheider, Bernhardt y Lappegård, 2015; Sobotka, 2017). Mientras que los cambios en las dinámicas de formación de las familias determinarían cómo las mujeres y hombres perciben la idea del matrimonio o la unión conyugal, la llegada de los hijos, el nexo entre el tiempo destinado a las tareas de cuidado de los hijos y la vida laboral y el uso del tiempo. Todos estos hechos derivarían en la disminución paulatina del número deseado de hijos o incluso en el deseo de no tenerlos (Crimmins, Easterlin y Saito, 1991; Sobotka, 2017).

El cambio en el número deseado de hijos se puede asociar más bien con efectos de cohorte, que de periodo. Entre las primeras cohortes de nacimiento estudiadas, 1926-1934, el tamaño deseado de la familia superaba los 6 hijos. Pero esta cifra cae rápidamente entre las siguientes cohortes, hasta encontrarse por debajo de 4 hijos entre las mujeres nacidas en la década de 1940 y el primer quinquenio de los años cincuenta. Después de ese momento el número deseado de hijos sigue disminuyendo paulatinamente hasta alcanzar los 2 hijos entre las mujeres de la última cohorte de nacimiento en estudio, 1995-2000.

Distintos trabajos, desarrollados en distintas regiones y poblaciones, han posicionado el empoderamiento o emancipación femenina que se ha gestado durante las últimas décadas como una revolución que ha permitido a las mujeres alcanzar logros en diferentes ámbitos (Phan, 2013; Prata, 2017 y Upadhyay et al., 2014). El empoderamiento femenino, definido como

la ampliación de la capacidad de una persona para tomar decisiones estratégicas en la vida en un contexto en el que antes se les negaba esa capacidad (Kabeer, 2001), se considera un factor clave en la planificación familiar y en la salud reproductiva de las mujeres. A medida que las mujeres estén en mejores condiciones de tomar decisiones estratégicas en su vida, deseen planear su vida a futuro y su rol como mujer, más allá de ser esposa y madre, la planificación familiar les permitiría retrasar, espaciar o limitar sus embarazos, liberando su tiempo para otras actividades (Prata et al., 2017). Esta emancipación femenina se asocia con una menor fecundidad y menores tasas de embarazos no deseados (Upadhyay et al., 2014), con el uso constante de la anticoncepción, la intención de utilizar la anticoncepción en el futuro y con la comunicación conyugal en relación con la planificación familiar (Prata et al., 2017).

Rosero-Bixby y Casterline (1995) documentan a partir de un estudio de tipo cualitativo, la experiencia de mujeres costarricenses acerca de su percepción de porqué antes las familias eran grandes y ahora tienden a ser más pequeñas. Entre los múltiples factores resaltan el machismo y la subordinación de la mujer en el hogar. Los hombres en muchos casos fueron descritos como deseosos de tener muchos hijos, como signo de hombría, imponiendo su voluntad a las mujeres. Las mismas entrevistadas dieron a entender que en los últimos años se experimentaban cambios en la relación entre los sexos, habiendo mayor dialogo entre la pareja y que se venía dando una mayor participación de los hombres en el cuidado de los hijos y las tareas del hogar e incluso tomaban la iniciativa para planificar la familia.

De acuerdo con Crimmins et al. (1991) las preferencias por el número deseado de hijos, el uso del tiempo y los bienes no se supone que vayan cambiando con el tiempo, sino que se espera que cambien entre cohortes con las diferentes condiciones sociales y económicas experimentadas durante cada período. A medida que las cohortes experimenten cambios en las condiciones económicas y sociales durante su crianza, llegarán a la edad adulta con diferentes aspiraciones relativas a la vida familiar y matrimonial, la vida laboral y el bienestar material.

El cambio de una serie de ideales y valores se visualiza como el generador de esos cambios de comportamiento relevantes para la formación de la familia y el matrimonio (Crimmins, Easterlin y Saito, 1991; Brinton et al., 2018, Goldscheider, Bernhardt y Lappegård, 2015; Lesthaeghe, 2010; 2011). La creciente tendencia entre los jóvenes de priorizar la realización personal, la satisfacción de las necesidades materiales básicas y el disfrute de tiempo de ocio, sobre el matrimonio y la familia, implica que las personas evalúen con mayor exigencia los costos de tener hijos (Angeles, 2010; Cabella y Nathan, 2018; Crimmins et al., 1991).

Durante la primera fase del descenso de la fecundidad, el nacimiento y crianza de los hijos estaba prácticamente confinado al matrimonio y el número de parejas sin hijos permaneció siendo bajo (Lesthaeghe, 2010), el matrimonio y la llegada de los hijos se consideraban hechos esperables, parte del curso vital, y cada género tenía un rol definido. Sin embargo, en el periodo más reciente, aproximadamente cuando la fecundidad se acercó al nivel de reemplazo (etapa post-transición) la dinámica familiar cambió: las relaciones de pareja se consideran débilmente comprometidas y transitorias, se ha producido una renuencia al matrimonio: un número cada vez mayor de adultos interrumpe, retrasa o evita los vínculos formales, ya sea al entrar en relaciones de cohabitación o vivir fuera de una pareja. Además, la procreación y la crianza de los hijos se han ido separando cada vez más del matrimonio, con un gran aumento del número de hijos nacidos fuera del matrimonio, ya sea en cohabitación o madres solteras y la paternidad se suele minimizar o evitar por completo entre los jóvenes, lo que conduce aún más a la disminución de la fecundidad (Goldscheider et al., 2015).

Para el caso de Costa Rica, en el estudio de Rosero y Casterline (1995) descrito anteriormente, algunas participantes establecen un vínculo del control natal con los niveles de vida, así como con aspiraciones de consumo de ciertos bienes materiales y de una vida mejor. Se rescatan aspectos como el deseo de procurar una vida mejor para los hijos a través de la educación y llevar una vida económica más holgada como familia, al tener que distribuir los recursos económicos entre un número menor de integrantes.

Esta última idea, permite poner de manifiesto la relación entre la situación económica de los individuos con sus ideales de fecundidad. De acuerdo con Sobotka (2017), en los países desarrollados la situación económica de los adultos jóvenes se ha deteriorado en las últimas décadas, las tasas de empleo han declinado, los trabajos se han vuelto más inestables y peor pagados, en comparación con las personas de edades intermedias, las ganancias relativas han disminuido y la adquisición de viviendas se ha vuelto inaccesible para muchos jóvenes. Estos cambios estructurales en la económica han vuelto cada vez más incompatible el trabajo en el mercado laboral con las tareas de crianza (Cabella y Nathan, 2018).

El efecto a la baja en el número de hijos que se observa en los dos últimos periodos de estudio (Gráfico 6) podría denotar, en alguna medida, la incertidumbre económica que se experimentó en el país en años recientes y su influencia a la baja sobre la fecundidad. La tasa de desempleo abierto (TDA), calculada con base en la información de las Encuestas de Hogares (INEC, 2019), se ha incrementado para Costa Rica paulatinamente en las últimas décadas: en 1990 la TDA fue de 4.5%, de 5.1% en el 2000 y de 7.7% en 2009, momento a partir del cual empezó a aumentar hasta alcanzar cifras en torno a 9% y 10%. Sin embargo, se observan grandes diferencias por grupos etarios, en detrimento de los más jóvenes. Durante el periodo 1990-2018, la TDA de las personas de 15 a 24 años, prácticamente duplicaba las cifras observadas para el total de la población. Entre 2008 y 2009, cuando Costa Rica experimentó los efectos de la crisis económica internacional (c.f. Ulate et al., 2013), el desempleo aumentó en todos los grupos de la población, pero lo hizo principalmente entre los jóvenes. Entre esos dos últimos años, la TDA pasó de 10.8% a 17.6% y en la actualidad se encuentra en 24.7%, entre los menores de 25 años.

Jiménez-Fontana (2018) estudiando datos para Costa Rica, coincide en que históricamente las personas más jóvenes siempre tienen tasas de desempleo más altas pero señala que han aumentado de forma acelerada en el tiempo. Menciona, además, que las generaciones más jóvenes enfrentan condiciones más adversas que las que enfrentaron sus antecesores a la misma edad. La denominada generación de los *millennials*, es decir, los nacidos entre 1986 y 1999, tiene un nivel de escolaridad promedio superior al resto de generaciones y una mayor proporción domina un segundo idioma, pero enfrenta condiciones más adversas en materia de

empleo: casi una tercera parte no está asegurado y no cotiza para el régimen de pensiones y su salario promedio es más bajo.

En cuanto al curso futuro que tomará la fecundidad, sería de esperarse que la TGF no aumente por encima de dos hijos por mujer. El número deseado de hijos en la actualidad ya ronda esa cifra y el número promedio de hijos tenidos en la población es también menor a dos en algunos grupos poblacionales (Cuadro 17).

Los datos de la última encuesta de fecundidad (Ministerio de Salud, 2016) indican que en 2015, las mujeres tenían en promedio 1.6 hijos mientras que el número promedio deseado es 2.2 hijos. Aunque el desfase entre ambas cifras podría ser producto del menor tiempo de exposición al embarazo entre las mujeres jóvenes, también sugiere que algunas mujeres habían tenido menos hijos de los que realmente deseaban. Este hecho ya se ha documentado entre los países que experimentaron el descenso de la fecundidad desde hace décadas y que experimentaron desde años atrás las bajas tasas de fecundidad que América Latina empezó a mostrar desde hace relativamente poco tiempo (United Nations, 2015).

Algunos de esos países han establecido políticas pronatalistas (McDonald, 2006; Pardo y Varela, 2013; Thévenon, 2011) cuya experiencia puede ser un referente para formular políticas en pro de la fecundidad en Costa Rica. Aunque se han instaurado diversidad de políticas, no tuvieron gran éxito aquellas que otorgaban estímulos económicos al nacimiento de un hijo ni las que estimulan la inmigración. Las más exitosas fueron aquellas que promovían la equidad de género y las que permiten consensuar el tiempo destinado a la vida laboral y la vida familiar mediante provisión pública de cuidados infantiles, licencias parentales, transferencias a los hogares y horarios de trabajo flexibles y promoción de la división de las tareas de cuidado de los hijos entre hombres y mujeres (Cabella y Nathan, 2018).

La incertidumbre económica que se ha gestado entre la población en los últimos años a causa de la crisis fiscal que experimenta el país (Herrera, 2018 y IDESPO, 2018), posiblemente tenga un efecto (de periodo) a la baja sobre la fecundidad. La contante disminución que ha mostrado

el número deseado de hijos de una cohorte a la siguiente, sugiere que esta cifra podría disminuir entre las nuevas cohortes. Por lo tanto, se podría conjeturar, que la TGF disminuirá en el futuro próximo.

El aumento de la esperanza de vida y la disminución de las tasas de mortalidad y fecundidad, prácticamente definen el destino de la estructura etaria de la población, asegurando que ésta seguirá envejeciendo. Siendo el caso, Costa Rica debería prepararse para afrontar esta etapa, no necesariamente buscando que la fecundidad aumente pues como se mencionó en líneas anteriores, las personas no desean tener más de dos hijos. Si no, Costa Rica debería mirar a necesidades más inminentes en materia de infraestructura y servicios dirigidos a la creciente proporción de adultos mayores, crear una sociedad más inclusiva, garantizar la sostenibilidad de los sistemas de pensiones, el ahorro y la eficiencia económica y aprovechando los bonos demográficos.

## BIBLIOGRAFIA

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Allison, P. (1988). Event history analysis. Regression for longitudinal event data. 3rd edition. Sage Publications, Inc. United States.
- Almeida, N. K. O., Almeida, R. M. y Pedreira, C. E. (2015). Adverse perinatal outcomes for advanced maternal age: a cross-sectional study of Brazilian births. *Jornal de pediatria*, 91(5), 493-498.
- Angeles, L. (2010). Demographic transitions: analyzing the effects of mortality on fertility. *Journal of Population Economics*, 23(1), 99-120.
- Asamblea Legislativa. (1973). Ley 5476: Código de Familia. La Gaceta, 21 de diciembre 1973.
- Asamblea Legislativa. (1996). Ley 7654: Ley de Pensiones Alimentarias. La Gaceta, 19 de diciembre 1996.
- Asamblea Legislativa. (2001). Ley 8101: Ley de Paternidad Responsable. La Gaceta, 16 de abril 2001.
- Balakrishnan, T. R. (1978). Effects of child mortality on subsequent fertility of women in some rural and semi-urban areas of certain Latin American countries. *Population Studies*, 32(1), 135-145.
- Behm, H. y Guzmán, J.M. (1979). Diferencias socioeconómicas del descenso de la fecundidad en Costa Rica, 1960-1970.
- Billari, F. y Kohler, H. P. (2004). Patterns of low and lowest-low fertility in Europe. *Population studies*, 58(2), 161-176.
- Bongaarts, J. (1978). A Framework for Analyzing the Proximate Determinants of Fertility. *Population and Development Review*. 4(1):105-132.
- Bongaarts, J. (1999). Fertility decline in the developed world: Where will it end?. *American Economic Review*, 89(2), 256-260.

Bongaarts, J. (2001). Fertility and Reproductive Preferences in Post-Transitional Societies. *Population and Development Review*, 27, 260-281.

Bongaarts, J. (2002). The end of the fertility transition in the developed world. *Population and development review*, 419-443.

Bongaarts, J. (2008). Fertility transitions in developing countries: Progress or stagnation?. *Studies in family planning*, 39(2), 105-110.

Bongaarts, J. (2015). Modeling the fertility impact of the proximate determinants: Time for a tune-up. *Demographic Research*, 33, 535-560.

Bravo, J. (1999). Envejecimiento de la población y sistemas de pensiones en América Latina. En: *Encuentro Latinoamericano y Caribeño sobre las Personas de Edad; documentos de seminario técnico-Santiago: CELADE, 1999-v. 1, p. IV. 1.1-12.*

Brinton, M.C., Bueno, X., Oláh, L. y Hellum, M. (2018). Postindustrial fertility ideals, intentions, and gender inequality: A comparative qualitative analysis. *Population and development review*, 44(2), 281-309.

Bouvier, L. F. (2001). Replacement migration: Is it a solution to declining and aging populations?. *Population and Environment*, 22(4), 377-381.

Cabella, W. y Nathan, M. (2018). Los desafíos de la baja fecundidad en América Latina y el Caribe.

Caldwell, J. C. (2006). Social upheaval and fertility decline. In *Demographic Transition Theory* (pp. 273-299). Springer, Dordrecht.

Camacho, E. La filiación matrimonial y la presunción de paternidad del marido. *Revista Sala Segunda*, (1), 2006.

Casella, G. (1985). An introduction to empirical Bayes data analysis. *The American Statistician*, 39(2), 83-87.

Carranza, M. (2009). A brief account of the history of family planning in Costa Rica.

Centro Centroamericano de Población (CCP). (2012). Proyecciones Nacionales de Población de Costa Rica 2000-2050. Recuperado en febrero, 2019. Disponible en <https://censos.ccp.ucr.ac.cr/>

CEPAL. (2008). Transformaciones demográficas y su influencia en el desarrollo en América Latina y el Caribe.

CEPAL. (2009). El envejecimiento y las personas de edad: indicadores sociodemográficos para América Latina y El Caribe.

Chesnais, J. C. (1998). Below-replacement fertility in the European Union (EU-15): Facts and policies, 1960-1997. *Review of population and social policy*, 7(101), 83-101.

Coale, A. (1986). Demographic Effects of Below-Replacement Fertility and Their Social Implications. *Population and Development Review*, 12, 203-216.

Crimmins, E., Easterlin, R., y Saito, Y. (1991). Preference Changes Among American Youth: Family, Work, and Goods Aspirations, 1976-86.

Cutright, P. (1983). The ingredients of recent fertility decline in developing countries. *International Family Planning Perspectives*, 101-109.

Decarli, A. y La Vecchia, C. (1987). Age, period and cohort models: review of knowledge and implementation in GLIM.

Debiasi, E. (2018). Age-Period-Cohort Analysis: a Summary of Analytical Approaches and Results.

Demidenko, E. 2004. *Mixed models: Theory and applications*. Hoboken, NJ: Wiley.

El-Badry, M. A. (1961). Failure of enumerators to make entries of zero: errors in recording childless cases in population censuses. *Journal of the American Statistical Association*, 56(296), 909-924.

Enders, C. K. y Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods*, 12(2), 121.

Faraway, J. J. (2006). *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models*. Chapman and Hall/CRC.

Fernández, D. y Sequeira, P. (2011). Diez años después de aprobada la ley del derecho a la esterilización en Costa Rica. *Rev. Costarric. Salud Pública*, 20(2), 107-118.

Fienberg, S. E. y Mason, W. M. (1979). Identification and estimation of age-period-cohort models in the analysis of discrete archival data. *Sociological methodology*, 10, 1-67.

Fienberg, S. E. y Mason, W. M. (1985). Specification and implementation of age, period and cohort models. In *Cohort analysis in social research* (pp. 45-88). Springer, New York, NY.

Fogel, R. W. (2003). *Changes in the process of aging during the twentieth century: findings and procedures of the early indicators project* (Working paper N° 9941). National Bureau of Economic Research.

Frost, W. H. (1939). The age selection of mortality from tuberculosis in successive decades. *American Journal of Epidemiology*, 30(3), 91-96.

GBD 2017 Population and Fertility Collaborators. (2018). Population and fertility by age and sex for 195 countries and territories, 1950–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 392(10159), 1995-2051.

Glenn, N. D. (1976). Cohort Analysis' Futile Quest: Statistical Attempts to Separate age, period, and cohort effects. *American Sociological Review* 41: 900-904.

Goldscheider, F., Bernhardt, E., y Lappegård, T. (2015). The Gender Revolution: A Framework for Understanding Changing Family and Demographic Behavior. *Population and Development Review*, 41(2), 207-239.

Gómez, M. (1968). Informe de la Encuesta de Fecundidad en el Área Metropolitana. *Universidad de Costa Rica, Instituto Centroamericano de Estadística. San José*, 155-220.

Gómez Barrantes, M. (1983). Aspectos demográficos de la esterilización femenina en Costa Rica.

Gómez, M. (2009). El rápido descenso de la fecundidad en Costa Rica. *Población y Salud en Mesoamérica*.

Gómez, M., Raabe, C. y Bermúdez, V. (1971). *Práctica anticonceptiva y uso de los servicios de planificación familiar en Costa Rica según encuestas recientes*. Departamento de Estadística, Escuela de Ciencias Económicas, Universidad de Costa Rica.

Gustafsson, S. (2001). Optimal age at motherhood. Theoretical and empirical considerations on postponement of maternity in Europe. *Journal of population economics*, 14(2), 225-247.

Herrera, B. (2018). Panorama sociopolítico en Costa Rica (2018): una lectura ciudadana. *Revista PRAXIS*, (78), 1-24.

Hsueh, Y. C. y Anderton, D. L. (1990). Temporal Dimensions of the Fertility Transition: An Age-Period-Cohort Analysis of Frontier Fertility. *Sociological Perspectives*, 33(4), 447-464.

IDESPO, 2018. Encuesta Percepción sobre aspectos de la coyuntura y culturas políticas en Costa Rica 2018.

INEC & CCP (2008). *Tasas de fecundidad por edad. Costa Rica 1950-*. San José, Costa Rica. Recuperado en febrero, 2019. Disponible en <http://ccp.ucr.ac.cr/>

INEC. (s.f.). Sistema de indicadores de seguimiento a los Objetivos de Desarrollo Sostenible. recuperado en Noviembre, 2019. Disponible en: <http://inec.cr/objetivos-de-desarrollo-sostenible>

Instituto Nacional de Estadística y Censos (INEC). (2019). Registro Nacional de Nacimientos de Costa Rica, 2000-2018. Recuperado en febrero, 2019. Disponible en <http://redatam.inec.go.cr/>

Inoue, T., Ono, Y., Yonezawa, Y., Oi, M., Kobayashi, N., Kishi, J. y Emi, N. (2013). Oocyte quality improvement using a herbal medicine comprising 7 crude drugs. *Open Journal of Obstetrics and Gynecology*, 3(01), 195.

Johnson, R. 1985. Analysis of age, period and cohort effects in marital fertility. En: Mason, W. M. y Fienberg, S. E. (Eds). *Cohort analysis in social research: Beyond the identification problem*. Springer, New York, NY.

Kabeer, N. (2001). Reflections on the Measurement of Women's Empowerment: Theory and Practice. Discussing women's empowerment: Theory and Practice. Novum Grafiska, AB, Stockholm.

Klein, J. y Moeschberger, M. (2003). Survival analysis: techniques for censored and truncated data. 2da Edición. Springer.

Koo, Y. J., Ryu, H. M., Yang, J. H., Lim, J. H., Lee, J. E., Kim, M. Y. y Chung, J. H. (2012). Pregnancy outcomes according to increasing maternal age. *Taiwanese Journal of Obstetrics and Gynecology*, 51(1), 60-65.

Kupper, L. L., Janis, J. M., Karmous, A. y Greenberg, B. G. (1985). Statistical age-period-cohort analysis: a review and critique. *Journal of chronic diseases*, 38(10), 811-830.

Lee, R. D. (1974). Forecasting births in post-transition populations: Stochastic renewal with serially correlated fertility. *Journal of the American Statistical Association*, 69(347), 607-617.

Lee, R. (1990). The demographic response to economic crisis in historical and contemporary populations. *Population bulletin of the United Nations*, (29), 1.

Leet, D. R. (1977). Interrelations of population density, urbanization, literacy, and fertility. *Explorations in Economic History*, 14(4), 388.

Lesthaeghe, R. (2010). The Unfolding Story of the Second Demographic Transition. *Population and Development Review*, 36(2), 211-251.

Lesthaeghe, R. (2011). The "Second Demographic Transition": A Conceptual Map for the Understanding of Late Modern Demographic Developments in Fertility and Family Formation. *Historical Social Research / Historische Sozialforschung*, 36(2 (136)), 179-218.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D. y Schabenberger, O. (2006). *SAS for mixed models* (Vol. 2). Cary, NC: SAS Institute.

Lundquist, J. H., Anderton, D. L. y Yaukey, D. (2014). *Demography: the study of human population*. Waveland Press.

Madrigal, J. (1987) Evolución de la esterilización femenina en Costa Rica (Análisis de tres encuestas). IX Seminario Nacional de demografía. Asociación Demográfica Costarricense. Disponible en: [http://ccp.ucr.ac.cr/bvp/pdf/seminarios/seminario\\_9/noveno\\_15.pdf](http://ccp.ucr.ac.cr/bvp/pdf/seminarios/seminario_9/noveno_15.pdf), accesado en febrero, 2019.

Mason, K. O. (1992). Culture and the fertility transition: Thoughts on theories of fertility decline. *Genus*, 1-14.

Mason, W. M. y Fienberg, S. E. (1985). Cohort analysis in social research: Beyond the identification problem. Springer, New York, NY.

Mason, K. O.; Mason, W. M.; Winsborough, H. H. y Poole, W. K. (1973). Some methodological issues in cohort analysis of archival data. *American sociological review*, 242-258.

Mason, W. M.; Mason, K. O. y Winsborough, H. H. (1976). Reply to Glenn. *American Sociological Review*, 41(5), 904-905.

McCullagh, P. y Nelder, J. A. 1989. *Generalized linear models*. New York: Chapman and Hall.

McDonald, P. (2006). Low fertility and the state: The efficacy of policy. *Population and development review*, 32(3), 485-510.

McKinley, R. L. y Mills, C. N. (1985). A comparison of several goodness-of-fit statistics. *Applied Psychological Measurement*, 9(1), 49-57.

Ministerio de Salud. (2016). II Encuesta nacional de salud sexual y salud reproductiva: Costa Rica 2015.

Milner, M., Barry-Kinsella, C., Unwin, A. y Harrison, R. F. (1992). The impact of maternal age on pregnancy and its outcome. *International Journal of Gynecology & Obstetrics*, 38(4), 281-286.

Mora, R. (1999). Efecto de la escolaridad sobre la fecundidad en Costa Rica, 1971-1992 (Trabajo Final de Graduación). Universidad de Costa Rica, San Pedro de Montes de Oca, Costa Rica.

United Nations (1983). Manual X: Indirect Techniques for Demographic Estimation. United Nations publication.

O'Brien, R. M. (2000). Age period cohort characteristic models. *Social science research*, 29(1), 123-139.

O'Brien, R.M. (2014). *Age-period-cohort models: approaches and analyses with aggregate data*. Chapman and Hall/CRC.

O'Brien, R. M., Stockard, J. y Isaacson, L. (1999). The enduring effects of cohort characteristics on age-specific homicide rates, 1960-1995. *American Journal of Sociology*, 104(4), 1061-95.

Ogata, Y., Katsura, K., Keiding, N., Holst, C., & Green, A. (2000). Empirical Bayes age-period-cohort analysis of retrospective incidence data. *Scandinavian Journal of Statistics*, 27(3), 415-432.

ONU, (2005). Manual sobre la recolección de datos de fecundidad y mortalidad.

Osmond, C. (1985). Using age, period and cohort models to estimate future mortality rates. *International Journal of Epidemiology*, 14(1), 124-129.

Osmond, C. y Gardner, M. J. (1982). Age, period and cohort models applied to cancer mortality rates. *Statistics in Medicine*, 1(3), 245-259.

Osmond, C. y Gardner, M. J. (1989). Age, period, and cohort models. Non-overlapping cohorts don't resolve the identification problem. *American Journal of Epidemiology*, 129(1), 31-35.

Paccagnella, O. (2006). Centering or not centering in multilevel models? The role of the group mean and the assessment of group effects. *Evaluation review*, 30(1), 66-85.

Palmore, E. (1978). When can age, period, and cohort be separated?. *Social Forces*, 57(1), 282-295.

Pavalko, E. K., Gong, F. y Long, J. S. (2007). Women's work, cohort change, and health. *Journal of Health and Social Behavior*, 48(4), 352-368.

PEN. (2019). Séptimo Informe Estado de la Educación. *San José, Programa Estado de la Nación*.

Phan, L. (2013). Women's empowerment and fertility changes. *International Journal of Sociology of the Family*, 39(1/2), 49-75

Prata, N., Fraser, A., Huchko, M. J., Gipson, J. D., Withers, M., Lewis, S., ... y Upadhyay, U. D. (2017). Women's empowerment and family planning: A review of the literature. *Journal of biosocial science*, 49(6), 713-743.

Preston, S., Heuveline, P. y Guillot, M. (2001). Demography: measuring and modeling population processes.

Preston, S. H. y Wang, H. (2006). Sex mortality differences in the United States: The role of cohort smoking patterns. *Demography*, 43(4), 631-646.

Pullum, T. W. (1980). Separating age, period, and cohort effects in white US fertility, 1920–1970. *Social Science Research*, 9(3), 225-244.

Ramos-Chaves, Á. (2010). Essays on Economic Development in Costa Rica. Tesis para optar al grado de PhD. en Economía, Universidad de California, Berkeley.

Raudenbush, S. W. y Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Reher, D. (2011). Economic and Social Implications of the Demographic Transition. *Population and Development Review*, 37, 11-33.

Reynolds, J. (1973). Costa Rica: measuring the demographic impact of family planning programs. *Studies in Family Planning*, (4), 310-316.

Robertson, C., Gandini, S. y Boyle, P. (1999). Age-period-cohort models: a comparative study of available methodologies. *Journal of clinical epidemiology*, 52(6), 569-583.

Robles, A. y González, A. (2012a). La fecundidad en Costa Rica: 1992-2010. En: Ministerio de Salud, *Visualizando la salud reproductiva y la sexualidad desde diversas perspectivas: un análisis a partir de la Encuesta de Salud Sexual y Reproductiva Costa Rica 2010*. San José, C.R.

Robles, A. y González, A. (2012b). Planificación familiar. En: Ministerio de Salud, *Visualizando la salud reproductiva y la sexualidad desde diversas perspectivas: un análisis a partir de la Encuesta de Salud Sexual y Reproductiva, Costa Rica 2010*. San José, C.R.

Rodríguez, J. (2008). Reproducción adolescente y desigualdades en América Latina y el Caribe: un llamado a la reflexión y a la acción. Santiago de Chile.

Román, I., Murillo, D. y León, J. (2014). Educación y mercado laboral en Costa Rica: tendencias, cambios y prospectiva desde una perspectiva comparada.

Rosero Bixby, L. (1983). Determinantes de la fecundidad en Costa Rica. *Notas de población*.

Rosero-Bixby, L., y Casterline, J. B. (1993). Modelling diffusion effects in fertility transition. *Population studies*, 47(1), 147-167.

Rosero-Bixby, L. y Casterline, J.B. (1995). Difusión por interacción social y transición de la fecundidad: evidencia cuantitativa y cualitativa de Costa Rica. *Notas de Población*.

Rosero-Bixby, L., Gómez Barrantes, M. y Rodríguez, V. (1982). Determinantes de la fecundidad en Costa Rica: análisis longitudinal de tres encuestas.

Rosero-Bixby, L. y Oberle, M.W. (1989). Fertility change in Costa Rica 1960–84: analysis of retrospective lifetime reproductive histories. *Journal of biosocial science*, 21(4), 419-432.

Rutstein, S.O. y Rojas, G. (2006). *Guide to DHS Statistics: Demographic and Health Surveys Methodology*. Calverton, NY: US Agency for International Development, ORC Macro.

Ryder, N. B. (1985). The cohort as a concept in the study of social change. *American Sociological Review*. Vol. 30, No. 6 (Dec., 1965), pp. 843-861

Saad, P. M.; Miller, T. y Martínez, C. (2013). Impacto de los cambios demográficos en las demandas sectoriales en América Latina. *Revista brasileira de estudos de população*, 26(2), 237-261.

Samandari, G. y Speizer, I. (2010). Adolescent sexual behavior and reproductive outcomes in Central America: trends over the past two decades. *International Perspectives on Sexual and Reproductive Health*. 36 (1): 26-35.

SAS Institute Inc. (2019). SAS/STAT(R) 15.1 User's Guide: The GLIMMIX Procedure.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), 461-464.

Simmons, G.B. (1991). Research on the determinants of fertility. En: Farooq, G.M. y Simmons, G.B. (Eds.), *Fertility in developing countries*, Macmillan series.

Snijders, T. A. B y Bosker, R. J. (2011). *Multilevel analysis: an introduction to basic and advanced multilevel modelling*.

Stycos, J. M. (1982). The decline of fertility in Costa Rica: Literacy, modernization and family planning. *Population Studies*, 36(1), 15-30.

Stroup, W. W. (2012). *Generalized linear mixed models: modern concepts, methods and applications*. Chapman & Hall/CRC.

Te Velde, E. R. y Pearson, P. L. (2002). The variability of female reproductive ageing. *Human reproduction update*, 8(2), 141-154.

Timæus, I. y Moultrie, T. (2008). On Postponement and Birth Intervals. *Population and Development Review*, 34(3), 483-510.

Thein, T. M. y Reynolds, J. (1972). Contraception in Costa Rica. The Role of the Private Sector, 1959-1969. *San Jose: Demographic Association of Costa Rica*.

Twisk, J. W. (2013). *Applied longitudinal data analysis for epidemiology: a practical guide*. Cambridge University Press.

Ulate, A., Soto, M., Jiménez, M., y Miranda, V. (2013). Costa Rica frente a la gran contracción. En: *Costa Rica: Impactos y lecciones de la crisis internacional*, 14-30.

United Nations. (2015). World fertility patterns 2015.

Upadhyay, U. D.; Gipson, J. D.; Withers, M.; Lewis, S.; Ciaraldi, E. J.; Fraser, A., ... y Prata, N. (2014). Women's empowerment and fertility: a review of the literature. *Social Science & Medicine*, 115, 111-120.

Van de Walle, E. (1992). Fertility transition, conscious choice, and numeracy. *Demography*, 29(4), 487-502.

Witte, J. S., Greenland, S., Kim, L. L. y Arab, L. (2000). Multilevel modeling in epidemiology with GLIMMIX. *Epidemiology*, 11(6), 684-688.

Yang, Y. y Land, K. C. (2006). A mixed models approach to the age-period-cohort analysis of repeated cross-section surveys, with an application to data on trends in verbal test scores. *Sociological methodology*, 36(1), 75-97.

Yang, Y. y Land, K. C. (2008). Age-period-cohort analysis of repeated cross-section surveys: fixed or random effects?. *Sociological methods & research*, 36(3), 297-326.

Yang, Y., Fu, W. J. y Land, K. C. (2004). A methodological comparison of age-period-cohort models: the intrinsic estimator and conventional generalized linear models. *Sociological methodology*, 34(1), 75-110.

Yang, Y. y Land, K. C. (2013). *Age-Period-Cohort Analysis: New Models, Methods, and Empirical Applications*. CRC Press.