

Estimación bayesiana de la función de consumo en la Eurozona.

Agustín ALONSO RODRÍGUEZ

Real Centro Universitario
“Escorial-María Cristina”
San Lorenzo del Escorial

Resumen: La Zona Euro, al momento de escribir este artículo, está formada por 17 países, miembros de la Unión Europea. A pesar de que esta unión monetaria impone unas normas de actuación comunes, el caso es que estos 17 países son muy diferentes por su lengua, cultura, situación geográfica, extensión territorial e idiosincrasia. En este contexto el presente trabajo estudia la función de consumo keynesiana para comprobar las semejanzas o desemejanzas que presenta en este conjunto de países.

Abstract: The Eurozone, at the moment of writing this paper, is integrated by 17 countries of the European Union. Despite the fact that this monetary union establishes norms of common action, however these 17 states are rather different in many aspects: language, culture, geographical situation, territory and idiosyncrasy. In this context, this paper considers the Keynesian consumption function to discover the similarities and the dissimilarities of this function in this group of countries.

Palabras clave: Eurostat, función de consumo keynesiana, consumo en términos reales, renta disponible en términos reales, modelos de datos de panel, modelos multinivel/jerárquicos, estimación bayesiana, cadenas de Markov Monte Carlo, plataforma estadística de cálculo R, paquetes WinBUGS, JAGS.

Keywords: Eurostat, Keynesian consumption function, real consumption, real disposable income, panel data models, multilevel/hierarchical models, Bayesian models, Markov chain Monte Carlo, R statistical system, WinBUGS, JAGS.

Sumario:

- I. La Zona Euro.**
- II. Razón de este estudio.**
- III. Software utilizado.**
- IV. Los datos.**
- V. Posibles modelos.**
- VI. Primeros pasos.**
- VII. Modelos con pendiente constante y ordenada en el origen cambiante.**
- VIII. Modelos con pendiente y ordenada en el origen cambiantes.**
- IX. Estimación bayesiana.**
- X. Conclusiones.**
- XI. Bibliografía.**

Recido: enero de 2011.

Aceptado: marzo de 2011.

I. LA ZONA EURO

La Zona Euro o Eurozona es el conjunto de estados miembros de la Unión Europea que tienen como moneda oficial el euro. Su existencia data del 1 de enero de 1999, y en la actualidad la integran 17 países, con Estonia incorporada a la zona el 1 de enero de 2011.

Estos 17 países son: Alemania, Bélgica, Austria, Chipre, Eslovaquia, Eslovenia, Estonia, España, Finlandia, Francia, Grecia, Irlanda, Italia, Luxemburgo, Malta, Países Bajos y Portugal.

Al margen de estos países, el euro es también utilizado como moneda oficial por otros países; unos con acuerdos oficiales con la Unión Europea, y otros sin tales acuerdos. Entre los primeros hay que mencionar a Mónaco, San Marino y el Vaticano. También lo utiliza Andorra, si bien el acuerdo con la Unión Europea entrará en vigor el 1 de enero de 2012. Fuera de Europa, hay que mencionar a los territorios franceses de Mayotte y San Pedro y Miguelón.

Entre los que no tienen acuerdo con la Unión Europea pero utilizan el euro, están Montenegro y la auto proclamada República de Kosovo.

Para completar este resumen sobre la Zona Euro, cabe mencionar a los 10 países europeos que no pertenecen a la misma. A saber: Bulgaria, Dinamarca, Reino Unido, Suecia, Hungría, Letonia, Lituania, Polonia, República Checa y Rumania.

Bosnia-Herzegovina, por su vinculación pasada con el marco alemán, mantiene una tasa de cambio fija con relación al euro.

Por último, cabe mencionar el hecho de que 14 países africanos, por su vinculación a Francia y al franco francés, mantienen una paridad fija con respecto al euro.

II. RAZÓN DE ESTE ESTUDIO

Dentro de la unidad formal resultante de la pertenencia a la Eurozona, los países integrantes presentan diferencias estructurales que no se pueden obviar, y entre las que se pueden destacar la extensión territorial, la población, la lengua, el clima, la riqueza y la propia idiosincrasia.

Por este motivo, se analiza en este artículo la función keynesiana de consumo partiendo con un modelo de datos de panel o modelo multinivel/jerárquico, y terminando con la estimación bayesiana del modelo multinivel más favorecido, estadísticamente, en la primera fase.

La razón para acudir al enfoque bayesiano estriba en el hecho de que, en este enfoque, tanto las variables como los parámetros del modelo son considerados variables aleatorias, a las que se asocia una distribución de probabilidad. Esta es la diferencia más fundamental entre el enfoque bayesiano y el más extendido, el enfoque *frecuentista*, en el que los parámetros del modelo, que se tratan de estimar, son desconocidos pero fijos.

Con el enfoque bayesiano se pretende dar cabida a las diferencias estructurales, antes aludidas, existentes entre los países de la Eurozona. De hecho, el enfoque bayesiano permite obtener simulaciones que representan la incertidumbre inferencial en torno a los parámetros de un modelo.

III. SOFTWARE UTILIZADO

En este trabajo se utiliza la plataforma estadística R, en la que corren programas específicos para el estudio de datos de panel, como son *plm*, y *arm*, y paquetes como WinBUGS y JAGS, con sus diversas adaptaciones, que permiten la estimación bayesiana de modelos. Cabe destacar que WinBUGS es el fruto de esfuerzos colectivos, (cf. Alonso Rodríguez, 2008), que ponen al alcance de los estadísticos y econométricos la realización del enfoque bayesiano, que antes del advenimiento de los ordenadores, era impensable. El lector interesado puede encontrar en CRAN (*The Comprehensive R Archive Network*) los más de 2700 paquetes, que al momento de escribir estas líneas, están disponibles para R, y que agrupados por materias se pueden consultar en el apartado *Task Views* de CRAN.

IV. LOS DATOS

Los datos sobre el consumo y la renta disponible están tomados de Eurostat, y donde, aunque parezca increíble, no se encuentran los datos referidos a Bélgica y Chipre. Por su reciente incorporación, tampoco se utilizan los datos referentes a Estonia.

Los datos anuales aquí considerados van desde el año 2001 al año 2009, actualizados al 30-11-2010. Un total de 9 observaciones por estado, que al tener en cuenta los 14 estados considerados, dan un total de 126 observaciones referidas al consumo y a la renta.

Los datos del consumo son transformados en datos de consumo real al dividirlos por el IPC de cada país. Los datos de la renta disponible, ya vienen en valores reales, y son centrados para su representación gráfica. Los datos de ambas variables son utilizados en logaritmos. Los países analizados, con sus identificadores respectivos son los siguientes:

Alemania, identificador: D
Austria, identificador: A
Eslovaquia, identificador: SLK
Eslovenia, identificador: SLV
España, identificador: E
Finlandia, identificador: FIN
Francia, identificador: F
Grecia, identificador: GR
Irlanda, identificador: IR
Italia, identificador: I
Luxemburgo, identificador: LUX
Malta, identificador: MA
Países Bajos, identificador: NL
Portugal, identificador: P.

La representación gráfica de las variables consumo real y renta disponible, en logaritmos: $lconsR$ y $lrentaD$, respectivamente, para nuestro conjunto de países viene recogida en la figura 1, tras centrar el logaritmo de la variable renta disponible.

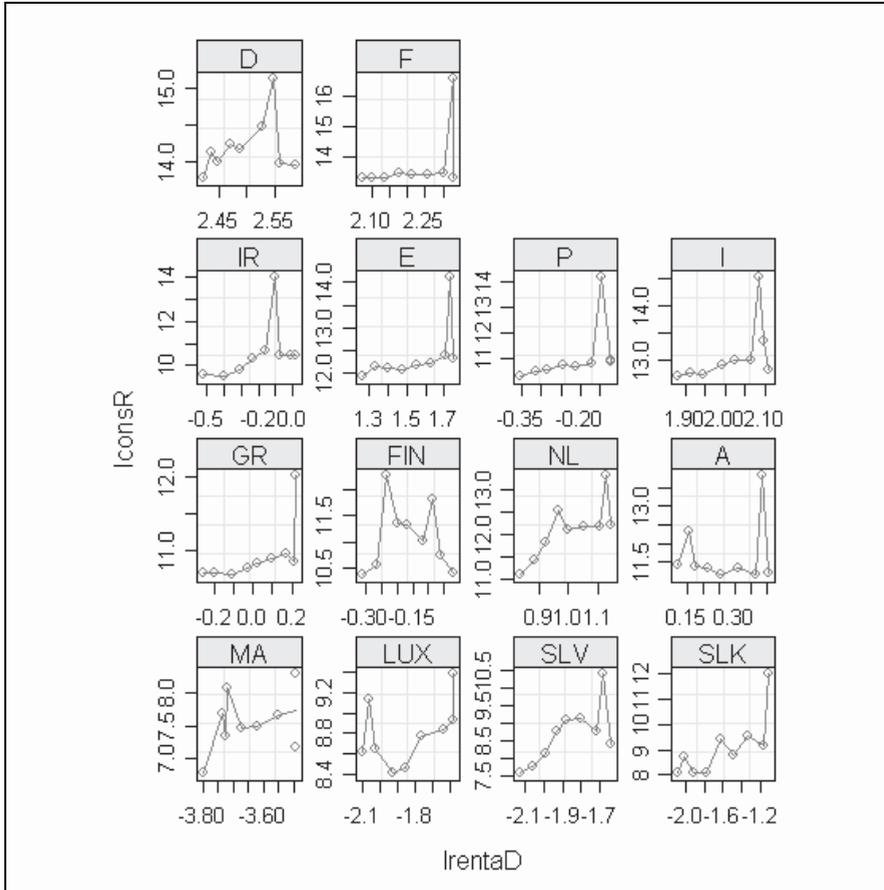


Figura 1

V. POSIBLES MODELOS

En una primera inspección de la figura 1, la función de consumo keynesiana para estos países, entre 2001 y 2009, parece requerir un modelo con ordenadas en el origen y pendientes diferentes para cada caso. Es decir un modelo multinivel

$$C_{ij} = \beta_{1j} + \beta_{2j} x_{ij} + u_{ij}; \text{ para } i=1 \dots n; j=1 \dots J$$

siendo i el número de observaciones por país y J el número de países. Por ende, C_{ij} representa el consumo en el año i -ésimo, del país j -ésimo. La renta disponible viene indicada por la variable x .

Los modelos multinivel y de datos de panel permiten establecer modelos en los que cambie únicamente la ordenada en el origen o *intercept*, pero no la pendiente; la pendiente pero no el *intercept*, o el *intercept* y pendiente simultáneamente.

Siguiendo a Gelman-Hill (2007, pp. 251-252 y pp. 262-265), es preferible escribir el modelo multinivel descomponiendo sus distintos elementos de la siguiente manera:

Para modelos con variaciones en sólo la ordenada en el origen:

$$C_i = \beta_{1j(i)} + \beta x_i + u_i; i = 1 \dots 9$$

$$\beta_{1j} = a + b u_j + w_j; j = 1 \dots 14$$

Y para modelos con ordenada en el origen y pendiente cambiantes:

$$C_i = \beta_{1j[i]} + \beta_{2j[i]} x_i + u_i; i = 1 \dots 9$$

$$\beta_{1j} = a_0 + b_0 u_j + w_{j1}; j = 1 \dots 14$$

$$\beta_{2j} = a_1 + b_1 u_j + w_{j2}; j = 1 \dots 14$$

Ahora tenemos dos veces más vectores de coeficientes cambiantes (β_1, β_2) con más vectores de coeficientes de segundo nivel (a, b), con posible correlación entre los errores (w_1, w_2).

La diferencia con el modelo clásico de regresión radica en la posibilidad de establecer un modelo para la posible variación entre los grupos, aquí: estados.

VI. PRIMEROS PASOS

Dado que la representación de la figura 1 se ha obtenido organizando los datos según el paquete *nlme*, de Pinheiro y Bates, hagamos los primeros pasos de nuestro estudio con este paquete.

Si no consideramos el carácter de datos agrupados de la muestra, el modelo a estimar sería (para aligerar la notación suprimimos subíndices fácilmente comprensibles):

$$\ln \text{consR} = \beta_1 + \beta_2 \ln \text{income} + u$$

siendo

$\ln \text{consR}$: logaritmo de la variable consumo, en términos reales;

$\ln \text{income}$: logaritmo de la variable renta disponible, en términos reales.

El resultado de la estimación es:

```
lm(formula = lconsR ~ lincome)
      coef.est coef.se
(Intercept) -2.43    0.49
lincome      1.11    0.04
---
n = 126, k = 2
residual sd = 0.78, R-Squared = 0.86
```

Es decir, el modelo estimado lo podemos escribir:

$$\text{lconsR} = -2.43 + 1.11 \text{ lincome} + \hat{u}; R^2 = 0.86; s(\text{residual}) = 0.78$$

(0.49) (0.04)

(entre paréntesis, los errores estándar de los coeficientes estimados).

Si tenemos presente el reagrupamiento en estados de esta muestra, obtenemos el modelo estimado con los resultados siguientes:

```
Model: lconsR ~ lincome | estados
Data: datos.org

Coefficients:
      (Intercept) lincome
MA      -4.999461  1.4802166
LUX      1.887333  0.6716982
SLV     -26.794552  3.4585429
SLK     -17.374733  2.5171789
GR       -6.811874  1.4617719
FIN      12.656818 -0.1307771
NL     -45.470119  4.3819597
A       -8.089107  1.5960998
IR     -30.571634  3.4496173
E      -16.408049  2.1058156
P     -65.351195  6.4155781
I     -38.435308  3.6439965
D     -20.644351  2.3830662
F     -58.643990  5.0484593

Degrees of freedom: 126 total; 98 residual
Residual standard error: 0.7472256
>
```

Antes de entrar en detalles, hagamos una primera diagnosis de este modelo. Al representar, con la función *pairs(modelo)*, los coeficientes estimados, tenemos la figura 2:

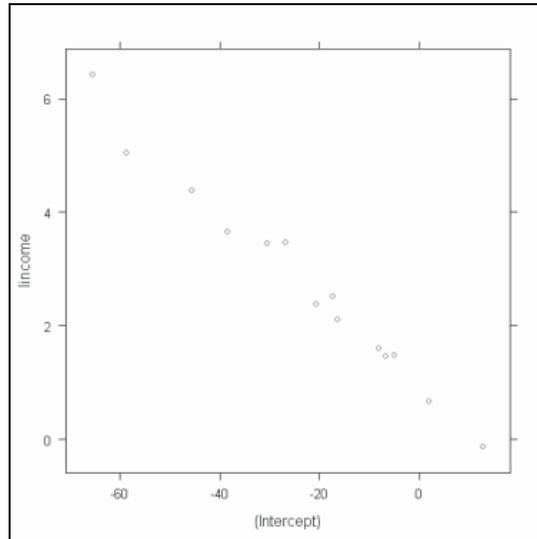


Figura 2

En esta figura se evidencia la muy alta correlación negativa entre la ordenada en el origen o *intercept* y la pendiente. Para corregir la situación, centramos la variable *lincome* en torno a su media, obteniendo la nueva variable: *lincome2*.

Con esta variable, y sin tener presente el agrupamiento por estados, ahora el modelo estimado por Mínimos Cuadrados Ordinarios es:

```
lm(formula = lconsR ~ lincome2)
      coef.est coef.se
(Intercept)  11.03   0.06
lincome2     1.12   0.03
---
n = 117, k = 2
residual sd = 0.67, R-Squared = 0.90
```

o sea:

$$lconsR = 11.03 + 1.12lincome2 + \hat{u}; R^2 = 0.90; s(\text{residual}) = 0.67$$

(0.06) (0.03)

(entre paréntesis los errores estándar de los coeficientes).

Con esta variable como variable explicativa, el modelo estimado, teniendo presente el agrupamiento en estados, muestra el siguiente resultado:

```

Model: lconsR ~ lincome2 | estados
Data: datos2

Coefficients:
      (Intercept)  lincome2
MA      12.952989  1.4802166
LUX     10.033863  0.6716982
SLV     15.151552  3.4585429
SLK     13.154265  2.5171789
GR       10.916875  1.4617719
FIN     11.070719 -0.1307771
NL       7.675424  4.3819597
A       11.268805  1.5960998
IR     11.266219  3.4496173
E        9.131828  2.1058156
P       12.458599  6.4155781
I        5.760027  3.6439965
D        8.258094  2.3830662
F        2.585033  5.0484593

Degrees of freedom: 126 total; 98 residual
Residual standard error: 0.7472256
>

```

Ahora la alta correlación entre *intercept* y pendiente ha desaparecido como muestra la figura 3.

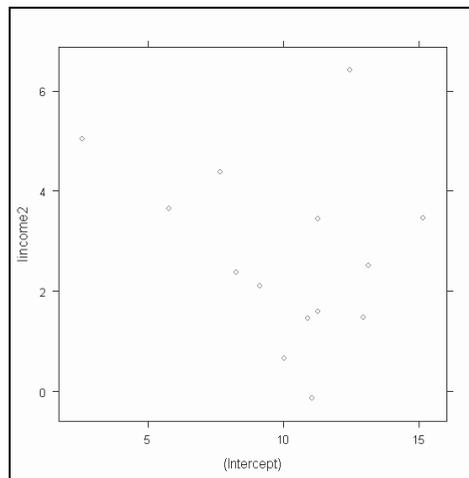


Figura 3

Como una guía inicial para el establecimiento de los elementos aleatorios, la función `plot(intervals(modelo))`, muestra el siguiente aspecto, figura 4:

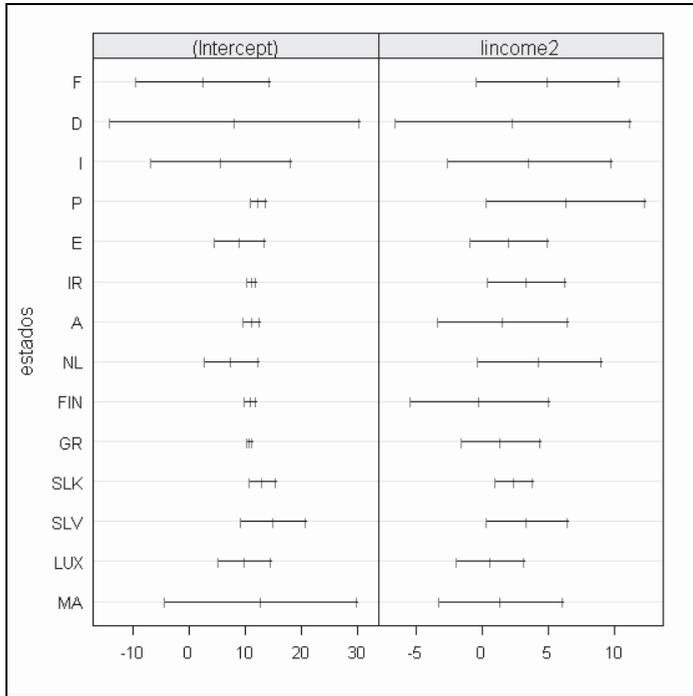


Figura 4

Mientras en el ámbito de la pendiente hay solapamiento, indicando que la pendiente puede ser considerado como un efecto fijo, en el modelo de efectos mixtos, en el de la ordenada en el origen se puede afirmar que es mucho menor tal solapamiento, indicando la conveniencia de incluir un efecto aleatorio para tomar en cuenta la variabilidad, estado a estado, en el *intercept* o término constante.

Para continuar, utilizaremos la función `lmer()` que permite estimar modelos multinivel lineales con coeficientes variables. El nombre de la función hace referencia a *linear mixed effects in R*, si bien esta función permite estimar también modelos lineales generalizados. Con el término “mixed effects” se hace referencia a los efectos *random*, coeficientes que varían por grupo, y a los efectos “fixed” o coeficientes que no varían. Esta terminología muy utilizada es, sin embargo, discutible, y en lugar de “fixed” y “random” utilizamos, siguiendo a Gelman y Hill, los términos de coeficientes *modelizados*, es decir, agrupados, y/o *no modelizados*. (Cf. Gelman-Hill, 2007, p. 259)

La función *lmer()* es parte del paquete *Matrix* y del paquete *lme4*, actualmente en desarrollo.

VII. MODELO CON PENDIENTE CONSTANTE Y ORDENADA EN EL ORIGEN CAMBIANTE

El modelo más simple para nuestro caso puede escribirse:

$$C_i \sim N(\beta_{1[j]} + \beta x_i; \sigma_y^2); i=1 \dots n; j=1 \dots J$$

que escribiremos, para simplificar la notación como

$$Y_i \sim N(\alpha_{j[i]} + \beta x_i; \sigma_y^2); i=1 \dots n; j=1 \dots J$$

Este modelo permite una ordenada en el origen cambiante y una pendiente constante; al *intercept* cambiante se asigna una distribución de probabilidad:

$$\alpha_j \sim N(\mu_\alpha, \sigma_\alpha^2); j=1 \dots J$$

siendo μ_α y σ_α^2 estimables a partir de los datos. La distribución de probabilidad asignada tiene el efecto de “empujar” la estimación de α hacia la media μ_α , pero no totalmente. Es decir, en el límite, si σ_α tiende a infinito, no tiene efecto notable el agrupamiento; pero si σ_α tiende a cero, entonces se termina en una estimación agrupada.

La línea media de regresión y las varianzas individuales y de grupos.

Los modelos multinivel tienen tantos parámetros que no es útil su consideración pormenorizada. En su lugar, las representaciones gráficas y los resúmenes estadísticos globales merecen mayor consideración. En estos modelos tenemos *hyperparámetros*: los parámetros sin subíndices de grupo.

En nuestro caso, la estimación del modelo, que denominamos M2, con *lmer*

$$Y_i \sim N(\alpha_{j[i]} + \beta x_i; \sigma_y^2); i=1 \dots n; j=1 \dots J$$

genera los resultados siguientes:

```

lmer(formula = lconsR ~ lincome2 + (1 | estados))
      coef.est coef.se
(Intercept) 11.07    0.08
lincome2     1.12    0.05

Error terms:
Groups   Name          Std.Dev.
estados (Intercept) 0.19
Residual                    0.76
---
number of obs: 126, groups: estados, 14
AIC = 307.2, DIC = 284.4
deviance = 291.8
    
```

o con más detalle, con *summary(M2)*:

```

Linear mixed model fit by REML
Formula: lconsR ~ lincome2 + (1 | estados)
      AIC   BIC logLik deviance REMLdev
307.2 318.5 -149.6   291.8   299.2
Random effects:
Groups   Name          Variance Std.Dev.
estados (Intercept) 0.034528 0.18582
Residual                    0.574634 0.75805
Number of obs: 126, groups: estados, 14

Fixed effects:
              Estimate Std. Error t value
(Intercept) 11.07169    0.08382  132.09
lincome2     1.11880    0.04838   23.12
Correlation of Fixed Effects:
              (Intr)
lincome2 0.000
    
```

Es decir, $\hat{\mu}_\alpha = 11.07$ y $\hat{\sigma}_y = 0.76$ y el modelo que promedia sobre los estados queda como

$$lconsR = 11.07 + 1.12 \text{ lincome2} + \hat{u}$$

(0.08) (0.05)

con un error estándar residual igual a 0.76 y un error estándar por grupo igual a

$$\hat{\sigma}_\alpha = 0.19$$

Cabe resaltar que la desviación estándar de la media de los niveles entre estados, casi coincide con la desviación estándar de 17 mediaciones dentro de cada estado, ya que

$$\frac{0.76}{\sqrt{17}} = 0.18386$$

Una manera de interpretar la variación entre estados σ_{α} consiste en ponerla en relación con la varianza residual:

$$\frac{\hat{\sigma}_{\alpha}^2}{\hat{\sigma}_y^2} = \frac{0.18582^2}{0.75805^2} = 0.06009$$

Esta relación entre varianzas también se suele expresar mediante la correlación inter-grupos:

$$\frac{\sigma_{\alpha}^2}{\sigma_{\alpha}^2 + \sigma_y^2}$$

Expresión que varía entre cero, si el agrupamiento no aporta información relevante, y uno, en caso contrario. En nuestro caso:

$$\frac{\sigma_{\alpha}^2}{\sigma_{\alpha}^2 + \sigma_y^2} = \frac{0.18582^2}{0.18582^2 + 0.75805^2} = 0.05668$$

En resumen, la estimación con grupos está próxima a la estimación sin tener en cuenta los grupos. Al aumentar el tamaño muestral, la estimación multinivel por grupos se acerca a la estimación sin tener en cuenta los grupos.

Otros resultados derivados de la estimación del modelo M2.

La instrucción *coef(M2)* devuelve los resultados siguientes:

```

coef (M2)
$estados
  (Intercept) lincome2
A      11.18527  1.118800
D      11.19275  1.118800
E      10.92342  1.118800
F      11.14461  1.118800
FIN    11.14278  1.118800
GR     11.01905  1.118800
I      10.99112  1.118800
IR     10.97756  1.118800
LUX    10.99566  1.118800
MA     11.26957  1.118800
NL     11.03137  1.118800
P      11.15191  1.118800
SLK    11.01110  1.118800
SLV    10.96750  1.118800

```

Es decir, para Austria, el modelo estimado es:

$$y = 11.18527 + 1.118800x + \hat{u}$$

para Alemania:

$$y = 11.19275 + 1.118800x + \hat{u}$$

Y así sucesivamente, para el resto de países.

pendientes son idénticas, pues así se ha especificado en el modelo mediante: (*l|estados*).

Y la instrucción *fixef(M2)* devuelve:

```

fixef (M2)
(Intercept) lincome2
  11.071692   1.118800

```

Es la regresión que promedia sobre los estados que nos permite escribir el modelo estimado como:

$$y = 11.071692 + 1.118800x + \hat{u}$$

Con *se.fixef(M2)*, los errores estándar correspondientes son:

```
se.fixef(M2)
(Intercept)    lincome2
0.08382206    0.04838081
```

indicando que los coeficientes de la regresión *media* son significativos.

Por último, *ranef(M2)*

```
ranef(M2)
$estados
(Intercept)
A    0.11357913
D    0.12106166
E   -0.14826842
F    0.07291516
FIN  0.07109088
GR   -0.05263773
I   -0.08057521
IR   -0.09413029
LUX -0.07602829
MA   0.19788133
NL   -0.04032004
P    0.08021928
SLK -0.06059206
SLV -0.10419539
```

valores que nos indican la magnitud del desplazamiento en la ordenada en el origen, para cada estado particular.

Para estos coeficientes, los errores estándar, estimador por *se.ranef(M2)* son:

```
se.ranef(M2)
$country
[,1]
[1,] 0.1497061
[2,] 0.1497061
[3,] 0.1497061
[4,] 0.1497061
[5,] 0.1497061
[6,] 0.1497061
[7,] 0.1497061
[8,] 0.1497061
```

```

[9, ] 0.1497061
[10, ] 0.1497061
[11, ] 0.1497061
[12, ] 0.1497061
[13, ] 0.1497061
[14, ] 0.1497061

```

De esta última información, y tomando como referencia la regresión *media*, resultan los modelos estimados. Así, para Austria,

$$y = (11.071692 + 0.11357913) + 1.118800x + \hat{u}$$

o sea

$$y = 11.18527 + 1.118800x + \hat{u}$$

Resultado que coincide con el dado por *coef(M2)*. Y así se continúa.

VIII. MODELO CON ORDENADA EN EL ORIGEN Y PENDIENTE CAMBIANTES

La siguiente instrucción estima el modelo con ambos coeficientes cambiantes. Sea el modelo M3:

$$M3 = \text{lmer}(\text{lconsR} \sim \text{lincome2} + (1 + \text{lincome2} | \text{estados}))$$

Con *display(M3)* obtenemos la siguiente información:

```

lmer(formula = lconsR ~ lincome2 + (1 + lincome2 | estados))
      coef.est coef.se
(Intercept) 11.05    0.08
lincome2     1.14    0.06

Error terms:
  Groups   Name      Std.Dev. Corr
estados  (Intercept) 0.00
          lincome2  0.11    0.00
Residual                    0.75
---
number of obs: 126, groups: estados, 14
AIC = 309.1, DIC = 282.8
deviance = 289.9.

```

o más en detalle, con *summary(M3)*

```

Linear mixed model fit by REML
Formula: lconsR ~ lincome2 + (1 + lincome2 | estados)
   AIC   BIC logLik deviance REMLdev
 309.1 326.1 -148.5   289.9   297.1
Random effects:
Groups   Name             Variance  Std.Dev.  Corr
estados  (Intercept)  8.1882e-16  2.8615e-08
          lincome2    1.3165e-02  1.1474e-01  0.000
Residual                    5.6789e-01  7.5359e-01
Number of obs: 126, groups: estados, 14

Fixed effects:
              Estimate Std. Error t value
(Intercept) 11.05474    0.07945   139.1
lincome2     1.13871    0.05722    19.9

Correlation of Fixed Effects:
          (Intr)
lincome2 -0.049

```

Ahora

$$\hat{\sigma}_\alpha = 2.8615e - 0.8$$

Y la desviación estándar residual: $\hat{\sigma}_y = 0,75$

Con *coef(M3)* obtenemos:

```

coef(M3)
$estados
  (Intercept) lincome2
A      11.05474  1.157887
D      11.05474  1.209707
E      11.05474  1.049321
F      11.05474  1.183339
FIN    11.05474  1.128874
GR     11.05474  1.140138
I      11.05474  1.083034
IR     11.05474  1.162121
LUX    11.05474  1.173024
MA     11.05474  1.007279
NL     11.05474  1.124551

```

P	11.05474	1.134754
SLK	11.05474	1.187304
SLV	11.05474	1.200543

En un primer momento no parece que se haya conseguido el objetivo, pues parece que sólo es cambiante la pendiente. La aparente constancia de la ordenada en el origen es debido a los valores insignificantes que nos muestra *ranef(M3)* para el *intercept*:

```
> ranef(M3)
$estados
      (Intercept)      lincome2
A      3.697372e-15    0.019181406
D     -4.265458e-16    0.071001571
E     -1.137206e-15   -0.089384371
F     -3.045856e-16    0.044633585
FIN    3.171100e-15   -0.009831387
GR    -1.774039e-15    0.001432921
I     -1.193102e-16   -0.055671419
IR    -3.866255e-15    0.023415478
LUX   -2.350217e-15    0.034318548
MA     6.288185e-15   -0.131426516
NL    -9.121179e-16   -0.014154770
P      3.352439e-15   -0.003951665
SLK   -2.075784e-15    0.048599090
SLV   -3.543035e-15    0.061837529
```

Sumar estos valores a la regresión *media* dada por *fixef(M3)*

```
> fixef(M3)
(Intercept)      lincome2
 11.054737      1.138705
```

es decir:

$$y = 11.054737 + 1.138705x + \hat{u}$$

deja sin cambios a la ordenada en el origen.

Los errores estándar para los coeficientes de esta regresión son:

```
se.fixef(M3)
(Intercept)      lincome2
 0.07945467      0.05722226
```

que validan los coeficientes del modelo M3.

Llegados aquí parece oportuno pronunciarse sobre ambos modelos. Con la instrucción *anova(M3,M2)*. Obtenemos:

Data:

Models:

M2: $lconsR \sim lincome2 + (1 | estados)$

M3: $lconsR \sim lincome2 + (1 + lincome2 | estados)$

	Df	AIC	BIC	logLik	Chisq	Chi	Df	Pr(>Chisq)
M2	4	299.80	311.15	-145.90				
M3	6	301.94	318.95	-144.97	1.8671		2	0.3932

que muestra que ambos modelos son, estadísticamente equivalentes, si bien por el *principio de parsimonia*, resulta favorecido el más sencillo, el M2.

IX. ESTIMACIÓN BAYESIANA

Llegados aquí, quizás sea oportuna la pregunta ¿por qué continuar con el enfoque bayesiano?

Varias son las posibles razones.

Cuando el número de grupos es pequeño, o el modelo multinivel es complicado, puede que no exista suficiente información para estimar las varianzas con precisión. En estos casos el enfoque bayesiano promedia la incertidumbre sobre todos los parámetros del modelo, y permite expandir más los modelos.

Además la aparente heterogeneidad, al menos teórica, de los países de la Zona Euro se puede recoger con las distribuciones de probabilidad atribuidas a los parámetros, en lugar de considerarlos simplemente como desconocidos pero fijos.

La estrategia aconsejable en estos análisis conlleva los siguientes pasos:

- 1.- Comenzar con un modelo global, sin reconocimiento de grupos, utilizando la función *lm()*.
- 2.- Establecer modelos multinivel. Es decir, permitir que el intercept y/o la pendiente varíen, utilizando la función *lmer()*.
- 3.- Con BUGS (Bayesian analysis Using Gibbs Sampler), simular la incertidumbre sobre todos los parámetros del modelo, utilizando estas simulaciones para resumir la incertidumbre en torno a los coeficientes, predicciones, etc.

Tras haber llevado a cabo los anteriores pasos, se estima un modelo teniendo presente el punto de vista inicial, sugerido por la figura 1: a saber, un modelo con ordenada en el origen y pendiente cambiantes.

Siguiendo a Gelman y Hill, para modelizar la matriz de varianzas y covarianzas de este modelo se utiliza la distribución inversa de Wishart, con sus ventajas e inconvenientes; cf. Gelman y Hill, (2007, pp. 286-287 y 376-377).

La estimación con R2WinBUGS del modelo M3, anteriormente presentado, muestra el resumen de resultados siguiente:

```
M3 <-
bugs(zonaue.data, zonaue.inits, zonaue.parameters, "wishart1.b
ug",
+ n.chains=3, n.iter=700, debug=T)

> print(M3)
Inference for Bugs model at "wishart1.bug", fit using
WinBUGS,
  3 chains, each with 700 iterations (first 350 discarded)
  n.sims = 1050 iterations saved
```

	mean	sd	2.5%	25%	50%	75%	97.5%	Rhat	n.eff
a[1]	8.7	5.5	-2.5	5.3	9.5	12.1	17.7	1.1	59
a[2]	9.0	5.5	-2.2	5.6	10.0	12.6	18.6	1.1	82
a[3]	8.9	5.7	-2.3	5.2	9.8	12.5	17.6	1.1	57
a[4]	9.0	5.7	-2.3	5.5	9.9	12.6	18.5	1.1	50
a[5]	8.9	5.5	-1.8	5.2	9.9	12.4	17.9	1.1	59
a[6]	9.1	5.7	-1.8	5.4	10.1	12.8	18.2	1.1	53
a[7]	8.9	5.6	-2.3	5.7	9.6	12.3	18.4	1.1	65
a[8]	9.0	5.6	-2.2	5.6	9.9	12.5	18.0	1.1	56
a[9]	9.4	5.8	-1.6	5.5	10.4	13.2	19.2	1.1	50
a[10]	9.1	6.0	-2.4	5.4	9.8	12.6	20.0	1.1	100
a[11]	8.9	6.1	-3.0	5.3	9.7	12.4	19.5	1.1	68
a[12]	9.1	6.1	-2.2	5.4	9.8	12.5	19.8	1.1	82
a[13]	8.8	6.0	-2.7	5.2	9.7	12.3	19.2	1.1	78
a[14]	9.0	5.9	-2.2	5.2	9.6	12.5	20.0	1.1	62
b[1]	0.4	0.4	-0.3	0.1	0.3	0.6	1.1	1.1	64
b[2]	0.4	0.4	-0.3	0.1	0.3	0.6	1.1	1.1	82
b[3]	0.4	0.4	-0.2	0.1	0.3	0.6	1.1	1.1	59
b[4]	0.4	0.4	-0.3	0.1	0.3	0.6	1.1	1.1	56
b[5]	0.4	0.4	-0.2	0.1	0.3	0.6	1.1	1.1	67
b[6]	0.4	0.4	-0.2	0.1	0.3	0.6	1.2	1.1	54
b[7]	0.4	0.4	-0.3	0.1	0.3	0.6	1.2	1.1	63
b[8]	0.3	0.4	-0.3	0.1	0.3	0.6	1.1	1.1	62
b[9]	0.4	0.4	-0.3	0.1	0.3	0.6	1.2	1.1	49
b[10]	0.3	0.4	-0.3	0.1	0.3	0.6	1.2	1.1	100
b[11]	0.4	0.4	-0.3	0.1	0.3	0.6	1.3	1.1	66

b[12]	0.3	0.4	-0.3	0.1	0.3	0.6	1.3	1.1	67
b[13]	0.4	0.4	-0.3	0.1	0.3	0.6	1.3	1.1	69
b[14]	0.4	0.4	-0.3	0.1	0.3	0.6	1.2	1.1	74
mu.a	9.0	5.2	-1.2	5.7	9.6	12.2	17.5	1.1	53
mu.b	0.4	0.4	-0.2	0.1	0.3	0.6	1.1	1.1	48
sigma.y	0.8	0.8	0.1	0.3	0.5	0.9	2.9	1.3	10
sigma.a	2.3	2.1	0.4	0.8	1.8	3.1	8.2	1.8	6
sigma.b	0.2	0.1	0.0	0.0	0.1	0.2	0.6	1.8	6
rho	-0.4	0.4	-0.9	-0.7	-0.4	-0.1	0.4	1.0	130
deviance	12.6	16.7	-20.7	1.4	12.8	23.2	42.9	1.3	10

For each parameter, *n.eff* is a crude measure of effective sample size, and *Rhat* is the potential scale reduction factor (at convergence, *Rhat*=1).

DIC info (using the rule, $pD = \bar{D} - \hat{D}$)

$pD = -662.9$ and $DIC = -650.3$

DIC is an estimate of expected predictive error (lower deviance is better).

Como la columna **Rhat** muestra valores en torno a 1, se cumple uno de los criterios para asegurar la convergencia de las tres cadenas de Markov generadas.

En la primera columna, aparecen las medias o valores esperados de las distribuciones de probabilidad de la ordenada en el origen y de la pendiente para cada uno de los 14 estados, y en las siguientes columnas los cuantiles correspondientes. En función de los mismos es posible establecer intervalos de confianza para las medias calculadas.

Así, para el caso de España, el intervalo del 95% de confianza para el valor calculado para la ordenada en el origen, que es 8.9, viene dado por (-2.3, 17.6). En este intervalo se encajan los valores estimados para dicho coeficiente por los modelos M2 y M3. Y para la pendiente, la propensión marginal a consumir, cuyo valor medio calculado es 0.4, el intervalo del 95% de confianza viene dado por (-0.2, 1.1), en cuyo entorno encaja el valor estimado por M3, y de venir con más dígitos la tabla de resultados bayesiana, podríamos afirmar que también encajaría el valor estimado por M2.

X. CONCLUSIONES

Dos son las posibles conclusiones derivadas de nuestro estudio, teniendo en vista el objetivo del mismo.

La primera conclusión es que la propensión marginal al consumo es semejante en los 14 países considerados, y está en torno al punto porcentual.

La segunda conclusión es que la ordenada en el origen es también bastante semejante en los 14 países.

Se puede, por tanto, avanzar como conclusión general que existe una gran homogeneidad en la función de consumo keynesiana para los países de la Eurozona, lo que obliga a reconsiderar la primera impresión tomada tras la inspección de la figura 1.

XI. BIBLIOGRAFÍA

- ALBERT, J., *Bayesian Computation with R*, Springer, New York 2007.
- ALONSO RODRÍGUEZ, A., El resurgir de Thomas Bayes, en *Anuario Jurídico y Económico Escurialense*, 41(2008) 327-360.
- arm, URL: <http://cran.r-project.org/web/packages/arm>, v(ersion).1.3-08 2010.
- BATES, D. M., *lme4: Mixed-effects modeling with R*, Springer, de próxima aparición.
- BATES, D. M., Y MARTIN MAECHLER, *lme4*, URL:<http://cran.rproject.org/web/packages/lme4package/>, v.0.999375-37, 2010.
- BUGS Project, Bayesian inference Using Gibbs Sampling, URL: <http://www.mrc-bsu.cam.ac.uk/bugs/>, MRC Biostatistics Unit, Cambridge, UK.
- BRUGS, URL: <http://cran.r-project.org/web/packages/BRugs/>, actualmente retirado del CRAN repository, Viena.
- CROISSANT, Y. Y MILLO, G., Panel Data Econometrics in R: The plm Package, URL:<http://cran.r-project.org/web/packages/plm/>, v.1.2-6, 2010.
- GELMAN, A. Y HILL, J., *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, New York 2007.
- GELMAN ANDREW, YU-SUNG SU, MASANAO YAJIMA, JENNIFER HILL, MARIA GRAZIA PITTAU, JOUNI KERMAN Y TIAN ZHENG, arm package, URL: <http://cran.r-project.org/web/packages/arm>, v.1.3-08, 2010.

- jags, URL: <http://calvin.iarc.fr/~martyn/software/jags/>, v. 2.2.0, 2010.
- LANCASTER, T., *Introduction to Modern Bayesian Econometrics*, Blackwell Publishing, Oxford 2004.
- lme4, URL:<http://cran.rproject.org/web/packages/lme4package/>, v.0.999375-37, 2010.
- MARTIN, A.D., QUINN, K. M. Y PARK, J. H., *MCMCpack Packge*, R version 1.0-8, 2010,
- MCMCpack,URL:<http://cran.r-project.org/web/packages/MCMCpack/>,1.0-8, 2010.
- nlme, URL:<http://cran.r-project.org/web/packages/nlme/>, v.3.1-97, 2010.
- NTZOUFRAS, I., *Bayesian Modeling Using WinBUGS*, John Wiley & Sons, Hoboken, New Jersey 2009.
- PINHEIRO, J. C.Y BATES, D. M., *Mixed-Effects Models in S and S-PLUS*, Springer, New York 2000.
- plm, URL:<http://cran.r-project.org/web/packages/plm/>, v.1.2-6, 2010.
- PLUMMER, A., JAGS: Jet Another Gibbs Sampler, URL: <http://calvin.iarc.fr/~martyn/software/jags/>, v. 2.2.0, 2010.
- PLUMMER, M., BEST, N., COWLES, K. y VINES, K., coda: Output Analysis and Diagnostics for MCMC, URL:<http://cran.r-project.org/web/packages/coda/>, v.0.14-2, 2010.
- R Development Core Team, R: *A Language and Environment for Statistical Computing*, Viena, Austria, URL: <http://www.r-project.org>. Versión actual: 2.12.1, 2010. Cada seis meses, aproximadamente, se actualiza la versión.
- R2jags, URL:<http://cran.r-project.org/web/packages/r2jags/>, v. 0.02-11, 2010.
- rjags, URL:<http://cran.r-project.org/web/packages/rjags/>, v. 2.2.0-2, 2010.
- R2WinBUGS, URL:<http://cran.r-project.org/web/packages/R2WinBUGS/>, v. 2.1-16, 2009.

- SMITH, B. J., BOA: Bayesian Output Analysis program for MCMC, R package, URL:<http://cran.r-project.org/web/packages/boa/>, v. 1.1-7-2, 2008.
- SPIEGELHALTER, D. J., THOMAS, A., BEST, N.G., Y LUNN, D., *WinBUGS Version 1.4 Users Manual*, MRC Biostatistics Unit, Cambridge, 2003.
- STURTZ, S., LIGGES, U., AND GELMAN, A., *R2WinBUGS: A Package for Running WinBUGS from R*. *Journal of Statistical Software*, 12(3) (2005) 1-16.
- SU, Y-S. Y YAJIMA, M., R2jags, URL:<http://cran.r-project.org/web/packages/R2jags/>, v. 0.02-11, 2010.

