

Un Modelo Estadístico Flexible para la Estructura Intertemporal de Tasas en Chile.*

Dante Jara**

Agosto 2004
(Versión Preliminar)

Resumen

El presente trabajo plantea un modelo estadístico flexible que captura paramétricamente la compleja condicionalidad y desvíos de normalidad que caracteriza a la estructura intertemporal de tasas de interés en la economía chilena entre los años 1992 y 2003. El modelo general consiste en una aproximación SemiNoParamétrica a la densidad condicional del proceso conjunto que siguen las tasas a 6 meses y 5 años. Este trabajo realiza el ejercicio de elección del mejor modelo, escogiendo el grado de flexibilidad necesario para capturar los hechos estilizados, por lo que son los datos los que indican la forma funcional específica para su densidad conjunta. Se implementa un algoritmo de simulación del modelo estadístico estimado, lo cual permite simular series artificiales de estructura de tasas a fin de verificar sus principales regularidades empíricas. El modelo SNP estimado puede ser utilizado como métrica para discriminar entre modelos alternativos de equilibrio general que pretendan dar cuenta de los hechos estilizados de estructura de tasas en la economía chilena.

* Agradezco a Rómulo Chumacero por su asesoría para la realización de la presente Tesis y el financiamiento del proyecto FONDECYT Nro. 1030681. Los descargos usuales se aplican.

** Universidad de Chile. Magister en Economía. djara@facea.uchile.cl

1. Introducción

El estudio de las características del proceso que está detrás de las realizaciones de la estructura intertemporal de tasas de interés ha motivado el desarrollo de numerosas metodologías, tanto para su recuperación a partir de la información financiera, para su modelación estadística y su explicación teórica. Tener modelos que den cuenta de la estructura intertemporal de tasas es importante por varios motivos. Una motivación particular consiste en la elaboración de predicciones. Las tasas de largo plazo son valores esperados de futuras tasas de corto plazo ajustadas por riesgo, lo cual implica que la curva de rendimiento contiene información sobre el desempeño futuro de la economía. Por tanto, el diferencial entre las tasas largas y cortas no sólo es útil para predecir futuras tasas cortas, sino también variables reales e inflación. Dichas predicciones son necesarias para decisiones de inversión por parte de firmas y para decisiones de ahorro por parte de consumidores. Otra motivación gira en torno a la necesidad de los Bancos Centrales de contar con modelos estructurales que caractericen el mecanismo de transmisión de tasas de interés, haciendo explícitos los determinantes de la relación entre la Tasa de Política Monetaria y las tasas relevantes para las decisiones de consumo e inversión, y asimismo identificar las expectativas sobre la posición futura de política monetaria.

El presente trabajo aporta a esta agenda de investigación proponiendo una forma funcional flexible para aproximar estadísticamente el proceso que rige las realizaciones de tasas de interés. Es bien sabido que las formas paramétricas estándar tienen problemas para ajustar las características que poseen las realizaciones de tasas y formas no paramétricas presentan dificultad para capturar la compleja condicionalidad de los datos, más aún en procesos multivariados. Es por eso que este trabajo propone el uso de métodos SemiNoParamétricos (SNP) para modelar el proceso conjunto que determina las realizaciones de tasas. Una ventaja de estos métodos es que no sufren de la maldición de la multidimensionalidad como los kernels, permiten modelar en forma flexible la condicionalidad de todos los momentos estadísticos de la estructura conjunta, permitiendo a su vez que sean los datos los que brinden información respecto de la forma específica que debe tener la representación estadística que mejor los ajuste.

Este trabajo implementa la metodología señalada con datos de la estructura intertemporal de tasas en UF para la economía chilena entre enero 1992 y diciembre del 2003. El ejercicio consistió en encontrar el mejor modelo estadístico que caracterice la mencionada estructura, para luego reportar los principales hechos estilizados. Entre ellos se señala la marcada persistencia temporal en la estructura y desvíos de normalidad en las innovaciones del proceso, específicamente excesos de kurtosis y asimetría. Adicionalmente, se encuentra evidencia de una relación negativa entre las realizaciones contemporáneas de los shocks de las tasas cortas y largas.

En Gallant y Tauchen (2001) se plantea la forma de utilizar la información estadística para recuperar la estructura económica que genera los datos, señalándose que el

gradiente de la estimación por máxima verosimilitud del mejor modelo estadístico que represente el Proceso Generador de Datos (PGD), es un estadístico que resume todas las propiedades estadísticas de los datos y que un candidato a modelo estructural debe ser capaz de generar realizaciones artificiales que satisfagan la esperanza de dicho gradiente. En ese sentido, el modelo estimado en este trabajo constituye el modelo estadístico que mejor ajusta a los datos de tasas de interés en UF para la economía chilena para el periodo en cuestión. Por tanto el modelo estimado sirve como métrica para discriminar entre modelos estructurales alternativos que intenten dar cuenta de la estructura de tasas en la economía chilena y su gradiente esta disponible para futuras investigaciones.

El trabajo se organiza del siguiente modo. En la segunda parte se efectúa una breve descripción de la base de datos utilizada para las estimaciones, en la tercera parte se efectúa una descripción de la metodología SNP implementada, en la cuarta parte se reportan los principales hechos estilizados encontrados, en la quinta parte se efectúa una discusión del uso del modelo SNP como métrica para la estimación de modelos estructurales y en la sexta parte se presentan las conclusiones.

2. Los Datos

El objetivo del trabajo consistió en caracterizar estadísticamente la estructura intertemporal de la tasa de interés para la economía chilena entre los años 1992 y 2003 y a fin de reportar sus principales regularidades empíricas. En particular, se caracterizó la estructura de la tasa de interés que está implícita en la información de precios y Tasa Interna de Retorno (TIR) al vencimiento de los instrumentos en UF emitidos por el Banco Central (BC) como parte de la implementación de su política monetaria.

Se define como estructura intertemporal de la tasa de interés en UF, o curva cero cupón en UF, a las tasas de interés utilizadas para la valorización de bonos cero cupón cuyo principal esta indizado a la UF. Esta información no se encuentra explícita en los datos de mercado puesto que no existen bonos cero cupón para todas las madureces requeridas. Sin embargo, existen varias metodologías alternativas para filtrar la curva cero cupón a partir de la información disponible de bonos con cupones. La base de datos utilizada en este trabajo fue estimada en Jara y Ochoa (2004), quienes implementaron las metodologías propuestas por Nelson y Siegel (1997) y por Svensson (2001), las cuales imponen una estructura paramétrica para la tasa cero cupón, que se define en función de la madurez y un conjunto de parametros.

Para estimar la curva cero cupón, Nelson y Siegel proponen la siguiente forma funcional paramétrica para la tasa cero cupón y_{ti} :

$$y_{ti} = \beta_0 + \frac{(\beta_1 + \beta_2)(1 - \exp(-\frac{i}{\tau}))}{\frac{i}{\tau}} - \beta_2 \exp(-\frac{i}{\tau}) \quad (1)$$

donde y_{ti} en el momento t es una función de la madurez i y del vector de parámetros $\theta' = (\beta_0, \beta_1, \beta_2, \tau)$. Esta forma funcional es lo suficientemente flexible como para capturar los perfiles que típicamente presenta la curva. En el periodo t , la secuencia y_{ti} con $i = 1, \dots, n$, (siendo n la madurez máxima) puede ser utilizada para valorizar un bono con cupones al que le resta n cupones por entregar :

$$P_{tn}^C = \sum_{i=1}^n C \exp(-iy_{ti}) + \exp(-ny_{tn}) \quad (2)$$

Utilizando datos para el periodo t , correspondientes a precios de mercado (P_{tn}^C), valor del cupón (C), madurez (n) o TIR al vencimiento de los bonos, y reemplazando (1) en (2), se estima $\theta' = (\beta_0, \beta_1, \beta_2, \tau)$ por mínimos cuadrados no lineales. De esta forma se tiene una aproximación paramétrica de la curva cero cupón implícita en los datos para el periodo t utilizando $\hat{\theta}'$ y la forma funcional (1). La metodología de Svensson es análoga a la de Nelson y Siegel, siendo su forma paramétrica propuesta aún más flexible a fin de capturar formas particulares que pudiese presentar la curva.

En la estimación efectuada de los datos de tasas cero cupón, los periodos fueron definidos en forma mensual desde enero de 1992 hasta diciembre de 1993. Para cada mes se construyó una base de datos de precios de mercado y licitación de los bonos PRC y PRBC emitidos por el BC y cuyo principal se encuentra indizado a la UF. Se adjuntó a dicha información la madurez al vencimiento y valor de los cupones correspondientes. A partir de la base de cada mes se estimó la curva cero cupón en UF para cada mes. Las tasas fueron estimadas directamente en base semestral, puesto que los PRC entregan cupones semestralmente. Se escogió alternativamente en cada caso los estimadores obtenidos con la metodología de Nelson y Siegel o con la de Svensson, según el grado de ajuste de ambos modelos respecto de los datos.

La Tabla 1 reporta algunos estadísticos para madureces escogidas de la estructura de tasas estimada. En promedio durante el periodo en cuestión la curva ha tenido una pendiente ligeramente positiva, dado que la media en promedio ha crecido con la madurez. Asimismo se aprecia que la volatilidad de las realizaciones de tasas ha ido disminuyendo con la madurez. Se aprecia también asimetría en las tasas y exceso de kurtosis. Este último fue disminuyendo con la madurez.

Tabla 1: Estadísticos de la Estructura de Tasas (base semestral)

	3m	6m	9m	1a	3a	5a	10a	20a
Media	0.026	0.026	0.027	0.027	0.028	0.028	0.029	0.029
Desv St.	0.022	0.016	0.014	0.012	0.008	0.006	0.004	0.003
Asim.	-1.437	-1.261	-1.218	-1.209	-1.060	-1.045	-0.764	-0,240
Kurt.	8.476	5.712	5.394	5.193	3.579	3.560	3.303	3.035

Una vez disponible la base de datos de estructura intertemporal de tasas, este trabajo la caracterizó estadísticamente a fin de reportar sus principales hechos estilizados. Para ello se tomó de la estructura completa, la tasa cero cupón a 6 meses como representativa de tasa corta y a la tasa a 5 años como representativa de tasa larga y se procedió a modelar estadísticamente este proceso bivariado¹.

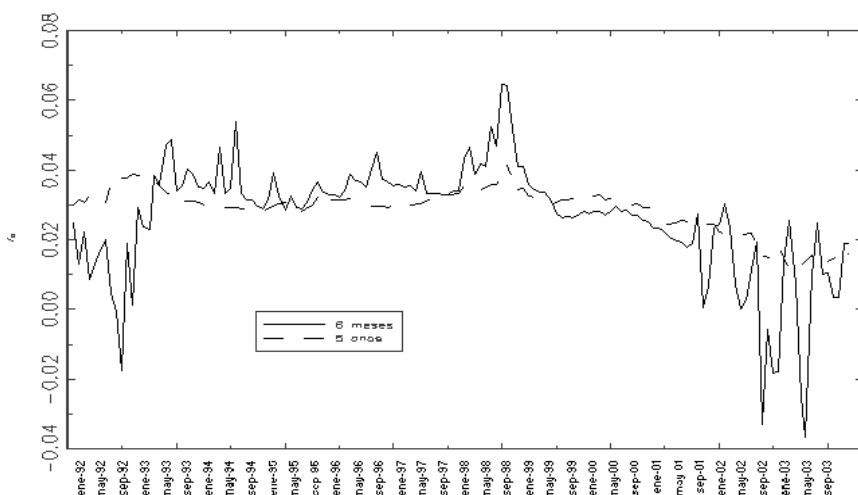


Figura 1. Estructura Intertemporal de Tasas de Interes en UF.

3. Densidad condicional de la estructura de tasas

En esta sección se efectúa la estimación de la densidad conjunta de la tasa cero cupón en UF libre de riesgo a 6 meses y 5 años para la economía chilena. Para ello se utiliza la forma funcional flexible SemiNoParamétrica (SNP), la cual constituye un

¹Se eligió la tasa a 6 meses por la mayor confiabilidad que existe en su estimación, al existir pocos datos de mercado para madureces menores, principalmente al inicio del periodo de análisis. Se eligió la tasa a 5 años por ser la duración de mayor liquidez en promedio para los instrumentos del BC durante el periodo en cuestión.

aproximador universal a funciones de densidad. El objetivo consistió en proponer una forma funcional que permita modelar la condicionalidad que rige la forma de la densidad del proceso conjunto en un instante del tiempo. Ello porque las características del PGD de un sistema multivariado estacionario se encuentran en su densidad condicional, dado que esta incorpora toda la información sobre heterocedasticidad condicional, desvíos de normalidad, irreversibilidad temporal, etc. en general cualquier forma de no linealidad. Estas características se aceptan hoy como comunes en series de tiempo financieras y son mejor pensadas en términos de la densidad condicional del proceso. Mediante la utilización de la forma funcional flexible SNP se evita cometer errores de especificación del modelo estadístico, lo cual es muy útil, dado que la imposición de modelos paramétricos estándar, que no ajustan a los datos, lleva a que los errores de especificación sesguen en forma seria las estimaciones y lleven a resultados espurios (White 1980).

3.1. Densidad SemiNoParamétrica (SNP)

El modelo estadístico SNP emplea expansiones de Hermite para aproximar la densidad condicional de procesos multivariados. Este modelo se estima por máxima verosimilitud utilizando métodos convencionales junto con una estrategia de selección que determina el apropiado orden de expansión. Una descripción del fundamento teórico de esta metodología se encuentra en Gallant et al. (1987), mientras que una descripción de la construcción y estimación de la densidad SNP se encuentra en Gallant et al. (1990b). Seguidamente se efectúa una breve descripción de esta metodología.

Sea z_t un vector de dimensión M . Una particular expansión de Hermite de la densidad de z_t tiene la forma: $h(z_t) \propto [P(z_t)]^2 \phi(z_t)$, donde $P(z_t)$ denota un polinomio multivariado en z_t y $\phi[\cdot]$ es la densidad normal estándar. La constante de proporcionalidad $1/\int [P(s_t)]^2 \phi(s_t) ds$ hace que $h(z_t)$ integre a 1 y sea densidad²:

$$h(z_t) = \frac{[P(z_t)]^2 \phi(z_t)}{\int [P(s_t)]^2 \phi(s_t) ds} \quad (3)$$

La densidad $h(z_t)$ es una función homogénea de los coeficientes de $P(z_t)$, por lo que para tener una única representación para la aproximación de la densidad $h(z_t)$, el término constante del polinomio se normaliza a 1.

Considérese ahora el proceso $y_t = Rz_t + u$, donde R es una matriz triangular superior y u un vector de dimensión M . Una particular expansión de Hermite de la densidad condicional del proceso y_t tiene la forma $f(y_t/\theta) \propto [P\{R^{-1}(y_t - u)\}]^2 \phi[R^{-1}(y_t - u)]/|det(R)|$. Como $\phi[\cdot]/|det(R)|$ es la densidad normal multivariada con media u y matriz de varianzas y covarianzas $\Sigma = RR'$, y dado que el primer término del polinomio es 1, la expansión entera $f(y_t/\theta)$ es proporcional a la densidad normal multivariada:

²El anexo A del presente trabajo efectúa una descripción de las expansiones de Hermite y su utilización para la construcción de la densidad (3).

$$f(y_t/\theta) = \frac{[P(z_t)]^2 \eta(y_t/u, \Sigma)}{\int [P(s_t)]^2 \phi(s_t) ds} \quad (4)$$

donde $\eta(y_t/u, \Sigma) = \phi [R^{-1}(y_t - u)] / |\det(R)|$ y $z_t = R^{-1}(y_t - u)$. Se define Kz como el orden máximo de los polinomios de Hermite con que se construye $P(z_t)$. Si $Kz = 0$, entonces $P(z_t) = 1$ y $f(y_t/\theta) = \eta(y_t/u, R)$, lo que lleva a que el modelo sea una normal multivariada. Si $Kz > 0$ la densidad normal es modificada por la multiplicación por el polinomio $P(z_t)$. La forma así alcanzada para $f(y_t/\theta)$ es lo suficientemente flexible para aproximar en principio cualquier tipo de densidad suave.

Adicionalmente, si el primer momento de la densidad condicional de y_t depende de Lu rezagos, se define $u_{x_{t-1}} = b_0 + Bx_{t-1}$, donde $x_{t-1} = (y_{t-Lu}, \dots, y_{t-1})$. Por tanto, la expansión sería en este caso de la forma $f(y_t/x_{t-1}, \theta) \propto [P(z_t)]^2 \eta(y_t/u_{x_{t-1}}, \Sigma)$. Si $Kz = 0$ el modelo propuesto sería un VAR con innovaciones normales. Si $Kz > 0$ el modelo sería un VAR semiparamétrico, el cual permite aproximar densidades cuyos primeros momentos dependen linealmente de x_{t-1} , pero cuya forma es independiente de x_{t-1} .

Para aproximar procesos con heterogeneidad condicional (dependencia temporal de momentos superiores al primero) se hace que los coeficientes de $P(z_t)$ sean a su vez polinomios de grado Kx en x_{t-1} , donde en este caso $x_{t-1} = (y_{t-Lp}, \dots, y_{t-1})$. Un polinomio de grado Kz en z_t cuyos coeficientes son polinomios de grado Kx en x_{t-1} es un polinomio en z_t y x_{t-1} de grado $Kz + Kx$, lo cual lleva a que la expansión de la densidad de y_t sea en este caso $f(y_t/x_{t-1}, \theta) \propto [P(z_t, x_{t-1})]^2 \eta(y_t/u_{x_{t-1}}, \Sigma)$, siendo ahora la constante de proporcionalidad igual a $1 / \int [P(s_t, x_{t-1})]^2 \phi(s_t) ds$. Cuando $Kx > 0$ la forma de la densidad dependerá de x_{t-1} , y la densidad puede en principio, aproximar la heterogeneidad condicional presente en cualquiera de los momentos de y_t (Gallant y Tauchen 1989).

En datos de mercados financieros, los segundos momentos exhiben marcada dependencia sobre x_{t-1} . En un intento por ajustar el segundo momento, Kx puede requerir ser muy grande. Para evitar esto se hace R una función lineal de Lr rezagos del valor absoluto de innovaciones de la media condicional, tal que se define $R_{x_{t-1}}$ como:

$$vech(R_{x_{t-1}}) = \rho_0 + \sum_{i=1}^{Lr} P_{(i)} |y_{t-1-Lr+i} - u_{x_{t-2-Lr+i}}| \quad (5)$$

de modo que $f(y_t/x_{t-1}, \theta) \propto [P(z_t, x_{t-1})]^2 \eta(y_t/u_{x_{t-1}}, \Sigma_{x_{t-1}})$, donde $\Sigma_{x_{t-1}} = R_{x_{t-1}} R'_{x_{t-1}}$.

Por tanto, la forma general de la densidad SNP es la siguiente:

$$f(y_t/x_{t-1}, \theta) = \frac{[P(z_t, x_{t-1})]^2 \eta(y_t/u_{x_{t-1}}, \Sigma_{x_{t-1}})}{\int [P(s_t, x_{t-1})]^2 \phi(s_t) ds} \quad (6)$$

donde $\eta(y_t/u_{x_{t-1}}, \Sigma_{x_{t-1}}) = \phi [R_{x_{t-1}}^{-1}(y_t - u_{x_{t-1}})] / |\det(R_{x_{t-1}})|$ y $z_t = R_{x_{t-1}}^{-1}(y_t - u_{x_{t-1}})$.

Dado que la función valor absoluto no es diferenciable, esta es aproximada por la función Andrew, la cual es doblemente diferenciable y que se define a como:

$$a(u) = \begin{cases} (|100u| - \pi/2 + 1)/100 & \text{si } |100u| \geq \pi/2 \\ (1 - \cos(100u))/100 & \text{si } |100u| < \pi/2 \end{cases} \quad (7)$$

En procesos donde $M > 1$ (más de una variable), la construcción de $P(z_t)$ implica la inclusión de un gran número de interacciones entre las variables del modelo, aún para Kz pequeños. De modo análogo si $Kx > 0$, en caso se quiera capturar la condicionalidad de momentos superiores al segundo. Por tanto, se define como Iz al mayor orden de interacción a ser suprimido de $P(z_t)$. Si $Iz = 0$ ninguna interacción es suprimida. Si $Iz > 0$, entonces sólo las interacciones menores o iguales a $Kz - Iz$ son incluidas en $P(z_t)$. De modo análogo para Kx .

De este modo, una forma particular de (6) dependerá de los valores que se asignen a los parámetros Lu, Lr, Lp, Kz, Iz, Kx e Ix , los cuales se detallan en la Tabla 2:

Tabla 2: Parámetros que definen la densidad SNP

<i>Pars.</i>	<i>Definición</i>
Lu	Rezagos de y_t en la media condicional en t
Lr	Rezagos de $ y_t - u_{x-1} $ en la varianza condicional en t
Lp	Rezagos de y_t en el polinomio que corresponde a los coeficientes de $P(\cdot)$
Kz	Orden máximo de los polinomios de Hermite en z_t en $P(\cdot)$
Iz	Interacciones suprimidas en los polinomios de Hermite en z_t en $P(\cdot)$
Kx	Orden máximo de los polinomios de Hermite en x_{t-1} en $P(\cdot)$
Ix	Interacciones suprimidas en los polinomios de Hermite en x_{t-1} en $P(\cdot)$

Se define una Configuración de la densidad SNP a un conjunto de valores específicos asignados a los parámetros de la Tabla 2, los cuales determinan la parametrización asignada a la densidad en (6) y en consecuencia determinan el modelo estadístico a utilizar para imponer estructura al proceso y_t . La Tabla 3 reporta algunas densidades a las que se llega a partir del rango de valores que pueden tomar los parámetros de la Tabla 2.

Tabla 3: Modelos Alternativos

Lu	Lr	Lp	Kz	Iz	Kx	Ix	Modelos
> 0	$= 0$	≥ 0	0	0	0	0	VAR
> 0	> 0	≥ 0	0	0	0	0	VAR-ARCH
> 0	> 0	≥ 0	> 0	0	0	0	VAR-ARCH-SNP
> 0	> 0	> 0	> 0	0	> 0	0	No linear, No parametrico

Cada linea de la Tabla 3 implica una dimensión distinta en la cual se extiende el grado de flexibilidad de la densidad SNP que se propone para el proceso que sigue y_t . Finalmente, con la configuración seleccionada se construye $f(y_t/x_{t-1}, \theta)$ y se estima el vector θ , el cual incluye a los coeficientes de $P(z_t, x_{t-1})$, los coeficientes de la media condicional, $b_0|B$ y de la varianza condicional $\rho_0|P_{(1)}..P_{(Lr)}$. Los parámetros se estiman por máxima verosimilitud de modo que se maximice:

$$S_n(\theta) = \frac{1}{n} \sum_{t=1}^n \log [f(y_t/x_{t-1}, \theta)] \quad (8)$$

El modelo SNP presentado es potencialmente capaz de capturar toda la complejidad, en terminos de condicionalidad de todos los momentos, de cualquier estructura de generadora de datos.

3.2. Estimación SNP de la estructura de tasas

Sea y_t con $M = 2$, las realizaciones en el periodo t del proceso bivariado que siguen las tasas cero cupón en UF a 6 meses y 5 años en la economía chilena, cuya gráfica se reporta en la Figura 1. Las realizaciones de tasas utilizadas en esta sección están expresadas en base semestral. El objetivo consistió en imponer estructura al proceso estocástico que siguen ambas realizaciones conjuntas en el momento t y para ello se asumió la forma funcional flexible (6), se estimaron los modelos de la Tabla 3 y se escogió el que mejor ajusta a los datos.

La estrategia de estimación utilizada partió de la configuración más simple para luego ir avanzando en configuraciones superiores utilizando como semilla a los estimadores de θ de la configuración inmediatamente anterior. Por ejemplo, si se empieza estimando un modelo VAR-ARCH con $Lu = 1$ y $Lr = 1$, los parámetros estimados para esta configuración se utilizaron como semillas para la estimación del modelo VAR-ARCH con $Lu = 1$ y $Lr = 2$. La extrema no linealidad en los parámetros de la función (6) lleva a la presencia de varios mínimos locales, por lo que es muy importante

asegurar un óptimo global en una configuración antes de utilizar sus estimadores para pasar a una configuración superior. Ello se consigue partiendo en la optimización de varios puntos distintos y verificando llegar al mismo vector de estimadores³.

La Tabla 4 presenta los resultados del ejercicio de elección de mejor modelo estadístico. Gran parte del esfuerzo de este trabajo consistió en el cálculo de los valores que conforman dicha Tabla, principalmente por la necesidad de construir un código flexible para la estimación de cada uno de los modelos presentados y por la necesidad de encontrar estimadores que sean óptimos globales, lo cual se dificultaba en forma exponencial a medida que crecía la complejidad de cada modelo. El encabezado de la mencionada Tabla corresponde a los parámetros definidos en la Tabla 2, seguidos del total de coeficientes estimados (P_θ), el valor del negativo de la loglikelihood⁴ y los criterios de información de Schwarz, Hannan-Quinn y Akaike. Cada una de las líneas de la Tabla 4 corresponde a la estimación por máxima verosimilitud de la configuración de la densidad (6) respectiva⁵.

³En términos prácticos, se implementó la estrategia de búsqueda de óptimos globales señalada por Gallant et al (1990b). Para la estimación de una configuración determinada, se utilizó como semilla el óptimo global de la configuración inmediatamente anterior. Esta semilla era perturbada aleatoriamente varias veces, junto con los nuevos coeficientes que aparecían para la configuración más grande, y con cada perturbación se efectuaba una cantidad fija de pasos en la optimización. Luego se tomaba los coeficientes que alcanzaban el valor más bajo para el negativo de (8) y con ellos se optimizaba hasta converger.

⁴Se reporta el negativo dado que la rutina de optimización utilizada minimiza el negativo de (8) para obtener mayor estabilidad numérica.

⁵Se entregó al optimizador una base de datos normalizada (media incondicional cero y matriz de varianzas y covarianzas incondicional identidad) a fin de ganar estabilidad en términos computacionales, lo cual no modifica las propiedades asintóticas del estimador.

Tabla 4: Resultados de Estimación

Lu	Lr	Lp	Kz	Iz	Kx	Ix	P_θ	$-LogLik$	BIC	HQN	AIC
Modelos VAR											
1	0	1	0	0	0	0	9	0.998	1.157	1.101	1.063
2	0	1	0	0	0	0	13	0.994	1.224	1.143	1.087
Modelos VAR-ARCH											
1	1	1	0	0	0	0	15	0.600	0.865	0.771	0.707
1	2	1	0	0	0	0	21	0.441	0.812	0.681	0.591
1	3	1	0	0	0	0	27	0.445	0.921	0.753	0.638
Modelos VAR-ARCH-SNP											
1	2	1	1	0	0	0	23	0.384	0.790	0.647	0.549
1	2	1	2	1	0	0	25	0.345	0.786	0.630	0.523
1	2	1	2	0	0	0	26	0.298	0.757	0.595	0.484
1	2	1	3	2	0	0	27	0.277	0.754	0.579	0.470
1	2	1	3	1	0	0	28	0.259	0.753	0.603	0.459
1	2	1	3	0	0	0	30	0.141	0.671	0.484	0.356
⇒	1	2	1	4	3	0	29	0.080	0.592	0.411	0.287
1	2	1	4	2	0	0	30	0.096	0.626	0.439	0.311
1	2	1	4	1	0	0	32	0.180	0.745	0.545	0.409

La Tabla 4 esta separada en tres bloques, según la dimensión por la cual se aumenta el grado de flexibilidad de la densidad (6). La estrategia consistió primero en obtener la mejor configuración para el modelo VAR, luego pasar a buscar la mejor configuración para el modelo VAR-ARCH, para luego terminar buscando la mejor configuración VAR-ARCH-SNP. La estrategia buscó que sean los datos los que señalen la forma de la densidad que mejor ajuste a sus realizaciones .

El ejercicio llevado a cabo para la búsqueda de la mejor configuración VAR se presenta en el primer bloque de la Tabla 3. Se paró de buscar en $Lu = 2$ puesto que los criterios de elección de mejor modelo preferían $Lu = 1$. El valor de la loglikelihood del mejor modelo en esta dimensión fue de 1.219 y el BIC llegó a 1.378.

Una vez escogida la estructura de rezagos para la media condicional, se pasó a buscar la mejor configuración que incluyese estructura para la varianza condicional, para lo cual se fijó $Lu = 1$, dado que corresponde a la mejor especificación para la media condicional. Los resultados de este ejercicio se presentan en el segundo bloque de las Tabla 4. En este caso se paró de buscar en $Lr = 3$, puesto que los criterios de información preferían $Lr = 2$. La estructura de tasas estimada presenta evidencia de condicionalidad en sus segundos momentos, puesto que la mejor configuración del segundo bloque, donde se modeló los dos primeros momentos, permite alcanzar un menor valor para los criterios de información en relación a los modelos del primer bloque, donde sólo se modeló la media.

Sin embargo, los dos primeros bloques de la Tabla 4 tienen en común que sus modelos estimados poseen la representación $\eta(y_t/u_{x_{t-1}}, \Sigma_{x_{t-1}})$, es decir, se esta asumiendo normalidad conjunta en las realizaciones de tasas de interés. En el tercer bloque de la Tabla 4 se verifica si es conveniente o no, en términos de los criterios de información, modelar paramétricamente desvíos de normalidad, lo cual se implementa proponiendo la densidad general $f(y_t/x_{t-1}, \theta)$ presentada en (6). Una vez escogida la mejor estructura para capturar la condicionalidad de los primeros y segundos momentos de la densidad conjunta de las tasas de interés, se pasó a estimar las configuraciones asociadas a modelos VAR-ARCH-SNP.

El tercer bloque es el más extenso de la Tabla 4, dado que son varias las configuraciones posibles que se pueden proponer. Se estimaron cada uno de los modelos presentados y se computaron los respectivos criterios de información a fin de hacerlos comparables.

La flecha en la primera columna de la derecha en la Tabla 4 señala la configuración finalmente escogida, que corresponde a la configuración 1214300. Según el modelo escogido, la densidad condicional en t de la estructura de tasas tiene una media que depende del primer rezago de las variables ($Lu = 1$), tiene una varianza que depende de los dos primeros rezagos del valor absoluto de las innovaciones de la media condicional ($Lr = 2$) y presenta desvíos de normalidad, lo cual hizo necesario escoger un $Kz > 0$ a fin de capturar dicho desvío en forma paramétrica. Específicamente fue necesario utilizar un polinomio en z_t de grado $Kz = 4$. No fue necesario incluir interacciones entre las innovaciones estandarizadas en $P(z_t)$, por lo que en la configuración escogida $Iz = 3$.

La representación particular escogida de la forma general (6) se hace explícita en las siguiente ecuaciones:

$$\eta(y_t/u_{x_{t-1}}, \Sigma_{x_{t-1}}) = \frac{(2\pi)^{-1} \exp(-z_t' z_t / 2)}{|\det(R_{x_{t-1}})|}$$

donde,

$$z_t = R_{x_{t-1}}^{-1}(y_t - u_{x_{t-1}}) = \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix}$$

$$u_{x_{t-1}} = \begin{bmatrix} u_{1x_{t-1}} \\ u_{2x_{t-1}} \end{bmatrix} = \begin{bmatrix} b_0(1) \\ b_0(2) \end{bmatrix} + \begin{bmatrix} B(1,1) & B(1,2) \\ B(2,1) & B(2,2) \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix}$$

$$vech(R_{x_{t-1}}) = \begin{bmatrix} R_{x_{t-1}}(1,1) \\ R_{x_{t-1}}(1,2) \\ R_{x_{t-1}}(2,2) \end{bmatrix} = \begin{bmatrix} p_0(1) \\ p_0(2) \\ p_0(3) \end{bmatrix} + \begin{bmatrix} P^1(1,1) & P^1(1,2) \\ P^1(2,1) & P^1(2,2) \\ P^1(3,1) & P^1(3,2) \end{bmatrix} \begin{bmatrix} |y_{1t-1} - u_{1x-2}| \\ |y_{2t-1} - u_{2x-2}| \end{bmatrix}$$

$$+ \begin{bmatrix} P^2(1,1) & P^2(1,2) \\ P^2(2,1) & P^2(2,2) \\ P^2(3,1) & P^2(3,2) \end{bmatrix} \begin{bmatrix} | y_{1t-2} - u_{1x-3} | \\ | y_{2t-2} - u_{2x-3} | \end{bmatrix}$$

y

$$P(z_t) = P(z_{1t}, z_{2t}) = 1 + A(0,1)z_{2t} + A(0,2)z_{2t}^2 + A(0,3)z_{2t}^3 + A(0,4)z_{2t}^4 \\ + A(1,0)z_{1t} + A(2,0)z_{1t}^2 + A(3,0)z_{1t}^3 + A(4,0)z_{1t}^4$$

Los términos Kx e Ix se mantuvieron constantes en todas las configuraciones estimadas e iguales a cero⁶. Ello porque la configuración óptima escogida implica la estimación de 29 coeficientes con un total de 141 datos⁷, es decir, 4.8 observaciones por parámetro. Hacer $Kx > 0$ para capturar condicionalidad en momentos superiores al segundo, implicaría un crecimiento exponencial en la cantidad de coeficientes a estimar lo cual reducía gravemente su grado de eficiencia.

Puede notarse las grandes ganancias en términos de ajuste en la estimación debido a la flexibilidad de la forma paramétrica propuesta. El valor más bajo para el BIC asumiendo normalidad conjunta es de 0.817 (configuración 1210000), el cual es mayor respecto al valor del BIC para el mejor modelo SNP escogido, el cual es 0.581 (configuración 1214300). Se encuentra entonces evidencia de desvíos de normalidad conjunta en la densidad condicional de tasas de interés a 6 meses y 5 años, por lo cual es necesario alterar las formas paramétricas usuales a fin de poder efectuar una mejor especificación del modelo estadístico para las tasas de interés en cuestión.

La Tabla 5.1 presenta los valores estimados de los parámetros para la media y varianza condicional del mejor modelo así como las desviaciones estándar respectivas. La matriz de varianzas y covarianzas fue construida utilizando el OPG (outer product of gradients) de la estimación, para lo cual fue necesario computar analíticamente y en forma flexible para cada uno de los modelos, el gradiente de la densidad SNP⁸.

⁶Si $Kx = 0$ el término Lp no se activa en la parametrización de $f(y_t/x_{t-1}, \theta)$.

⁷La muestra incluye 144 meses entre enero 1992 y diciembre del 2003. La cantidad máxima de rezagos utilizada fue de 3 ($Lu + Lr$ de la configuración más grande) por lo que a fin de hacer comparables los modelos, todos fueron estimados con 141 datos.

⁸Es decir, se computó analíticamente la derivada de la loglikelihood de cada modelo estimado y con esta se computó V , donde:

$$V = \left[\sum_{t=1}^T \left(\frac{\partial \log f [(y_t/x_{t-1}, \theta)]}{\partial \theta} \right)' \left(\frac{\partial \log f [(y_t/x_{t-1}, \theta)]}{\partial \theta} \right) \right]^{-1}$$

Tabla 5.1: Estimaciones SNP
de $u = b_0|B$ y $R = p|(P_1 P_2)$

Configuración: 1214300			
Coefficientes	Estimador	Error	Standar
b_0	$b_0(1)$	-0.020	(0.014)
	$b_0(2)$	-0.022	(0.012)
B	$B(1, 1)$	0.971	(0.020)
	$B(2, 1)$	0.016	(0.013)
	$B(1, 2)$	0.082	(0.020)
	$B(2, 2)$	0.908	(0.018)
p_0	$p_0(1)$	0.091	(0.010)
	$p_0(2)$	0.098	(0.022)
	$p_0(3)$	0.133	(0.027)
P_1	$P_1(1, 1)$	0.021	(0.061)
	$P_1(2, 1)$	-0.557	(0.177)
	$P_1(3, 1)$	0.522	(0.205)
	$P_1(1, 2)$	0.298	(0.088)
	$P_1(2, 2)$	-0.630	(0.154)
	$P_1(3, 2)$	0.622	(0.129)
P_2	$P_2(1, 1)$	0.138	(0.053)
	$P_2(2, 1)$	0.150	(0.056)
	$P_2(3, 1)$	-0.342	(0.053)
	$P_2(1, 2)$	-0.111	(0.046)
	$P_2(2, 2)$	-0.464	(0.106)
	$P_2(3, 2)$	0.377	(0.082)

La Tabla 5.2 presenta los coeficientes estimados para el polinomio que conforma la densidad SNP, así como las respectivas desviaciones estándar.

Tabla 5.2: Estimaciones SNP
de coeficientes de $P(z_t)$

Configuración: 1214300			
	Coefficientes	Estimador	Error Standar
A	$A(0, 0)$	1	(0.000)
	$A(0, 1)$	0.090	(0.067)
	$A(0, 2)$	-0.305	(0.057)
	$A(0, 3)$	-0.035	(0.026)
	$A(0, 4)$	0.038	(0.007)
	$A(1, 0)$	-0.091	(0.078)
	$A(2, 0)$	-0.342	(0.042)
	$A(3, 0)$	0.027	(0.017)
	$A(4, 0)$	0.056	(0.006)

Se efectuaron tests simples de normalidad a los residuos estandarizados del modelo, computando el coeficiente de asimetría, kurtosis y el pvalue del test Jarque Bera (JB). También se verificó la capacidad del modelo escogido de capturar la condicionalidad de los primeros y segundos momentos de los datos de tasas de interés, mediante la implementación del test Ljung Box a las series de residuos en niveles y al cuadrado. Los resultados de este ejercicio se presentan en las Tabla 5:

Tabla 5: Análisis de residuos estandarizados del modelo escogido

	Asimetria	Kurtosis	JB(pv)	LB $Z(10, pv)$	LB $Z^2(10, pv)$
Z1 (6 meses)	0.47	4.73	0.00	0.620	0.987
Z2 (5 años)	0.24	8.65	0.00	0.626	0.267

Los residuos estandarizados del modelo escogido presentan algún grado de asimetría y exceso de kurtosis, lo cual es característico en las innovaciones de las series financieras. El Test JB rechaza en ambos casos la hipótesis nula de normalidad en los residuos en forma individual, lo cual era de esperarse dado que la densidad condicional que mejor captura el PGD de las tasas en cada t es una que presenta desvíos de normalidad ($Kz > 0$). El modelo escogido es capaz de capturar la persistencia de la media y la varianza que presentan las series, puesto que los residuos estandarizados no presentan correlación serial hasta el orden 10, ni en niveles ni al cuadrado, dado que los pvalues de los test no rechazan la hipótesis nula de ausencia de correlación serial de las series en niveles y al cuadrado.

4. Hechos Estilizados

4.1. Desvíos de Normalidad

Este trabajo ha encontrado evidencia de desvíos de normalidad en la densidad condicional del proceso conjunto que siguen las tasas de interés en cuestión y ha propuesto una forma funcional flexible que captura paramétricamente dicha regularidad empírica. Un aspecto interesante consiste en la posibilidad de simular realizaciones artificiales de la densidad conjunta estimada, con el objetivo de apreciar visualmente la forma en que se produce este desvío de normalidad. Gallant et al. (1990) propusieron un algoritmo para encontrar una envolvente de la densidad SNP, la cual corresponde a una mixture de variables distribuidas CHI y desarrolla un algoritmo para simular de dicha mixture como paso intermedio para simular de la densidad SNP utilizando el método de rechazo⁹.

En el instante t la densidad conjunta de las tasas $f(y_t/x_{t-1}, \theta)$ esta centrada en la media condicional en t , denominada $u_{x_{t-1}}$. La aleatoriedad de las realizaciones efectivas de tasas depende de las innovaciones estandarizadas en t , z_t , cuya estructura es la estimada por SNP, y estan ajustadas por la matriz de varianzas y covarianzas del periodo t , denominada $\Sigma_{x_{t-1}}$. El algoritmo de simulación de la densidad SNP permite muestrear de la densidad bivariada de las innovaciones estandarizadas z_t pues estas son las que presentan los desvíos de normalidad¹⁰. Es decir se efectúan simulaciones artificiales del proceso (3), siendo los valores de los coeficientes estimados en la Tabla 5.2 los que indican el desvío de normalidad en particular. Los elementos del vector z_t se simulan en forma conjunta.

La Figura 2 muestra la densidad marginal de las simulaciones de las innovaciones estandarizadas para la serie de tasas a 6 meses junto con la densidad de una normal estándar. Se aprecia gráficamente el exceso de kurtosis y las colas más delgadas que poseen los shocks estandarizados a la tasa a 6 meses en relación a la normal estándar.

⁹El anexo B del presente trabajo describe en detalle el algoritmo de simulación de la densidad SNP utilizada en este trabajo.

¹⁰Por tanto, las realizaciones simuladas del proceso bivariado estimado corresponden a innovaciones con media cero y matriz de varianzas y covarianzas identidad.

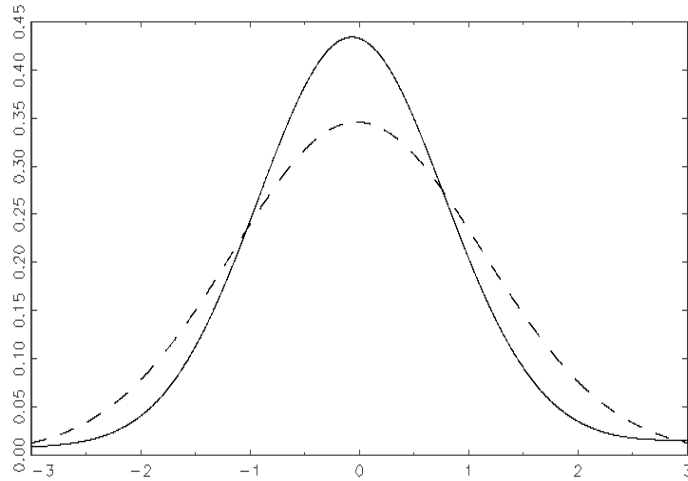


Figura 2: Densidad Marginal innovaciones de tasa corta

La Figura 3 presenta la densidad marginal de las simulaciones de las innovaciones estandarizadas para la serie de tasas a 5 años junto con la densidad de una normal estándar. Gráficamente se ve el exceso de kurtosis, el cual sin embargo es menor que en el caso de las tasa a 6 meses. Adicionalmente se aprecia asimetría positiva, lo que implica una mayor masa de probabilidad de ocurrencia para valores positivos de los shocks.

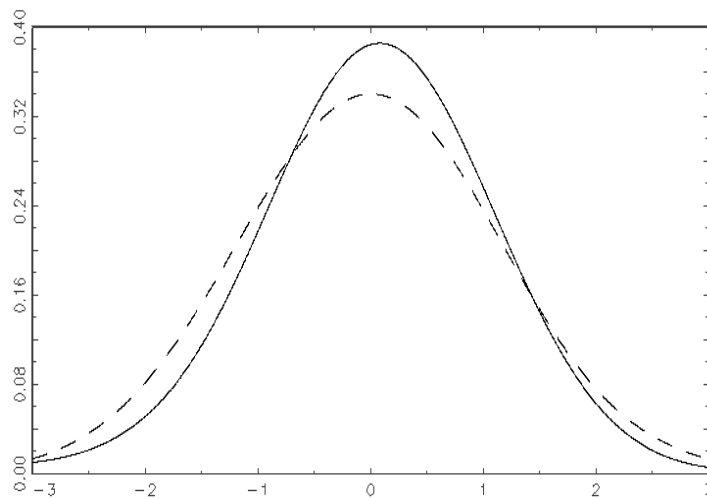


Figura 3: Densidad Marginal innovaciones de tasa larga

La estimación de la forma funcional flexible SNP y la implementación del algoritmo de simulación señalado permiten identificar este hecho estilizado de la estructura de

tasas. La Figura 4 presenta la densidad bivariada de las innovaciones con varianza igual a la matriz de varianzas y covarianzas incondicional de los datos de tasas con que se realizó la estimación y evaluado en su respectiva media incondicional. Existe un exceso de kurtosis respecto de la densidad normal bivariada estándar. Se aprecia también que se asigna una probabilidad de ocurrencia significativa a realizaciones extremas de la densidad conjunta.

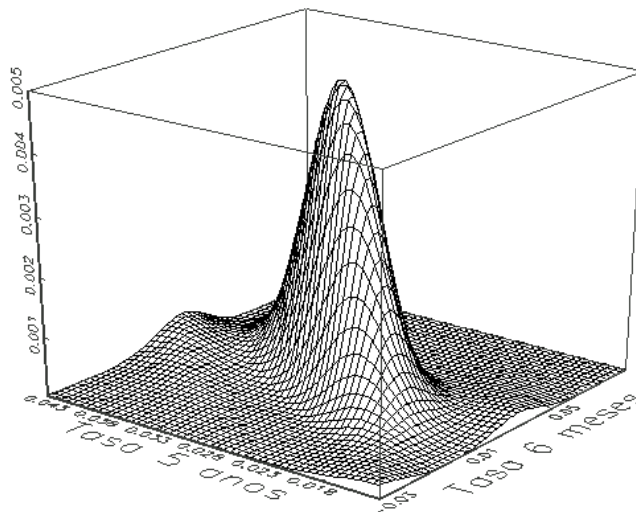


Figura 4: Densidad Conjunta de tasas de interes

4.2. Pendiente de la curva y persistencia de la Estructura de Tasas

Tal como lo muestra la Figura 1, existieron durante el periodo de análisis momentos donde la pendiente de la curva cero cupón fue negativa (tasa a 6 meses mayor que tasa a 5 años), principalmente al inicio de la muestra, y existieron periodos donde ocurría lo contrario, es decir, la curva presentaba una pendiente positiva, principalmente hacia el final de la muestra. Las simulaciones de series de tiempo de la estructura de tasas obtenidas a partir del modelo escogido debieran ser capaces de replicar este hecho estilizado. Para simular series de tiempo se utilizó el algoritmo de simulación de innovaciones estandarizadas, las cuales se multiplicaron por la matriz de varianzas y covarianzas predicha de cada periodo y se les sumó la media condicional predicha para cada periodo. De este modo se utilizó el modelo estadístico estimado para proyectar series de estructura de tasas de interés en UF. Una proyección en particular se presenta en la Figura 5. En ella se aprecia la marcada persistencia de las realizaciones de tasas, pero sin embargo estas presentan reversión a la media. Esto lleva que el modelo sea capaz de replicar periodos con pendiente positiva en la curva y periodos con pendiente negativa.

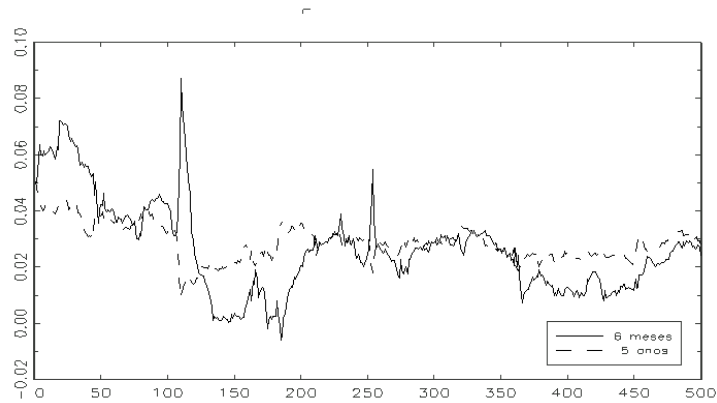


Figura 5: Simulacion Estructura de Tasas

Las simulaciones de series de tiempo implican estacionariedad en las series estimadas. Si bien la estructura estimada es no lineal, una manera de ver el grado de persistencia a partir de los coeficientes estimados para la media condicional es viendo los valores propios de la matriz B . Su mayor valor propio 0,98, lo que implica gran persistencia en las realizaciones conjuntas de la media condicional de las series de tasas. Esto puede identificarse como otro hecho estilizado de la estructura de tasas. Los signos de los coeficientes estimados en B son todos positivos, por lo que existiría una relación positiva entre las realizaciones contemporáneas y los rezagos de ambas tasas, lo cual además es consistente con la marcada persistencia señalada.

4.3. Estructura de la volatilidad de las innovaciones

A partir de los datos de tasas y la estructura estimada para la matriz de varianzas y covarianzas, se proyectó la estructura de volatilidad de las innovaciones al proceso estimado para las tasas durante el periodo bajo análisis. La Figura 6 presenta los resultados de este ejercicio. Se aprecia que existen episodios de mayor volatilidad hacia el inicio y hacia el final del periodo de análisis. Esto se aprecia también en la Figura 1, donde las tasas fueron muy volátiles hacia el principio y final de la muestra, por lo que el modelo estimado captura adecuadamente esta hecho.

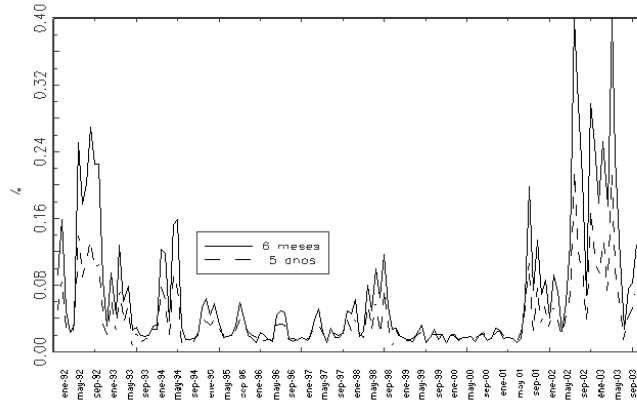


Figura 6: Volatilidad condicional de innovaciones

El modelo también captura la mayor volatilidad condicional de la tasa a 6 meses respecto de la tasa a 5 años, dado que en Figura 6, la línea continua que representa la secuencia de volatilidad de la tasa a 6 meses es siempre mayor a la línea discontinua que corresponde a la tasa a 5 años.

Un aspecto interesante del modelo propuesto es que permite estimar la secuencia de covarianzas contemporáneas que tuvieron las innovaciones a las tasas durante el periodo en cuestión, lo que permite computar el coeficiente de correlación condicional contemporáneo existente entre las tasas para cada periodo de la muestra, lo cual otorga información sobre el signo de la relación contemporánea de las realizaciones conjuntas. El resultado se presenta en la figura 7.

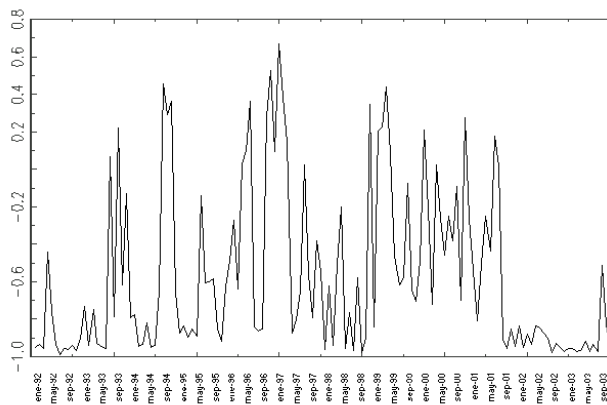


Figura 7: Correlacion innovaciones a Tasa corta y larga.

Puede notarse que la correlación condicional entre las innovaciones de las tasas es mayormente negativa, y lo es en más grado cuando es mayor la volatilidad de los

shocks a las tasas de interés. La correlación condicional de las innovaciones en promedio durante el periodo bajo análisis fue de -0.55. Una manera simple de ver la relación que existe entre la varianza de las innovaciones de las tasas y la correlación condicional entre las mismas es viendo la matriz de correlaciones computada a partir de las series de varianzas y correlaciones presentadas en la Figuras 6 y Figura 7. Esto se reporta en la Tabla 6:

Tabla 6: Relación entre volatilidad y correlación condicional de innovaciones

Matriz de Correlaciones			
	$var_t(e_{1t})$	$var_t(e_{2t})$	$corr_t(e_{1t}, e_{2t})$
$var_t(e_{1t})$	1.00	0.97	-0.53
$var_t(e_{2t})$	-	1.00	-0.56
$corr_t(e_{1t}, e_{2t})$	-	-	1.00

donde $e_t = \{e_{1t}, e_{2t}\}$ corresponden a las innovaciones mencionadas. Este es otro hecho estilizado encontrado en el presente trabajo. A mayor volatilidad de las innovaciones a las tasas de interés, mayor es la relación negativa contemporánea entre las realizaciones de efectivas de las tasas, lo cual se ve porque la correlación es negativa en ambos casos.

5. El Modelo SNP como métrica

El ejercicio llevado a cabo hasta ahora es puramente estadístico, pues a pesar de tener un modelo que de cuenta de las propiedades estadísticas de los datos, este no sirve para entender la economía que esta detrás de las realizaciones de tasas. Es decir, el modelo estadístico no es capaz de explicar la forma como los agentes utilizan la información disponible en los mercados financieros, forman sus expectativas y deciden las tasas que utilizarán para descontar los bonos cero cupón disponibles, en este caso bonos en UF.

Sin embargo, el modelo estadístico hallado en este trabajo, al dar cuenta de las propiedades estadísticas de las realizaciones de tasas, puede ser utilizado como métrica para evaluar modelos alternativos basados en primeros principios que pretendan explicar el comportamiento de la estructura de tasas en Chile. Gallant et al (1997) señalan que el gradiente de la estimación por máxima verosimilitud del mejor modelo estadístico que represente una estructura en particular, es un estadístico suficiente pues resume todas las propiedades estadísticas de los datos. Este gradiente es una función de los datos y del vector de parámetros estimados por máxima verosimilitud y depende de la forma particular del modelo estadístico propuesto. Para el caso del presente trabajo, el modelo escogido es el modelo SemiNoParamétrico, el cual tiene un vector gradiente que posee la siguiente forma general:

$$\frac{1}{T} \sum_{i=1}^T \frac{\partial f(y_t/x_{t-1}, \theta)}{\partial \theta} \Big|_{\hat{\theta}_{MV}} = \vec{0} \quad (9)$$

tal que se cumple que el gradiente analítico de la densidad SNP evaluada en los parámetros estimados y en los datos de tasas es cero en sentido vectorial. Cualquier modelo teórico propuesto para la estructura de tasas de interés en Chile debe ser capaz de generar realizaciones artificiales de \tilde{y}_t que junto a los $\hat{\theta}_{MV}$ satisfagan dicho gradiente. En términos generales deben satisfacer:

$$E \left[\frac{\partial f(\tilde{y}_t/x_{t-1}, \theta)}{\partial \theta} \right] \Big|_{\hat{\theta}_{MV}} = \vec{0} \quad (10)$$

De este modo es posible discriminar entre modelos teóricos alternativos que pretendan dar cuenta de las realizaciones de la estructura de tasas de interés en Chile en UF. La utilidad de la forma funcional flexible estimada es doble, puesto que por un lado permite que sean las realizaciones efectivas las que brinden información respecto de la forma funcional para la densidad que mejor caracteriza sus propiedades estadísticas, y por otro lado, representa una métrica objetiva para discriminar entre modelos alternativos, dado que esta métrica representada en el vector gradiente del mejor modelo estadístico, resume todas las propiedades estadísticas de los datos, las cuales debieran ser encompasadas por el modelo teórico que mejor se aproxime a la estructura económica verdadera que genera los datos.

Por ello se construyó el procedimiento que computa el mencionado gradiente a fin de ser utilizado para posteriores investigaciones que intenten proveer un marco analítico sobre la estructura económica detrás de los datos de tasas de interés, lo cual queda pendiente para la agenda futura de investigación.

6. Conclusiones

El principal objetivo de este trabajo consistió en investigar las características de la relación entre las tasas de corto plazo y de largo plazo dentro de la estructura intertemporal de tasas de interés en UF en Chile entre los años 1992 y 2003. La motivación consistió en reportar un conjunto de regularidades empíricas que modelos alternativos de equilibrio sean capaces de confrontar. Se organizó el esfuerzo sobre la tarea de estimar la densidad condicional de la estructura de tasas, puesto que para un proceso estacionario la densidad condicional es un estadístico condicionalmente invariante y encompasa toda la información probabilística del proceso. En particular predictibilidad, volatilidad y otros momentos condicionales. Tales elementos se piensan mejor en términos de la densidad condicional del proceso y no en signos y magnitudes de parámetros específicos.

Este trabajo encontró evidencia acerca del requerimiento de densidades más complejas que las estándar, lo cual en términos del modelo implica ordenes de expansión de hermite mayores a uno, para capturar la compleja estructura de los datos. Estas complejidades incluyen entre otras cosas, la estructura de la función de varianza condicional bivariada, las colas delgadas de la densidad de tasas, la interacción no lineal de tasas dentro de la estructura y la dependencia temporal.

La tarea pendiente consiste en proponer modelos de equilibrio basados en el comportamiento del mercado, que incluya la especificación de los problemas de optimización que enfrentan los agentes y la estructura de información que utilizan en sus decisiones. El desafío de estos modelos económicos consiste en confrontar los datos y dar cuenta de sus mayores hechos estilizados: volatilidades correlacionadas, correlación contemporánea entre las tasas dentro de la estructura, excesos de kurtosis, etc. El presente trabajo contribuye en esta línea de investigación, reportando dichos hechos estilizados y construyendo una métrica a utilizar para discriminar entre modelos de equilibrio alternativos.

Referencias

- [1] Castro, Ruben. (2003): "Polinomios de Hermite: su aplicación en SNP," Manuscrito, Universidad de Chile.
- [2] Fenton, Victor M., y Ronald Gallant (1996), "Qualitative and Asymptotic Performance of SNP Density Estimators," *Journal of Econometrics* 74, 77-118.
- [3] Gallant, A. Ronald, y Douglas W. Nychka (1987), "Seminonparametric Maximum Likelihood Estimation," *Econometrica* 55, 363-390.
- [4] Gallant, A. Ronald y George Tauchen (1989), "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications," *Econometrica* 57, 1091-1120.
- [5] Gallant, R. y G. Tauchen (1990a), "A Nonparametric Approach to Nonlinear Time Series Analysis: Estimation and Simulation" en David Brilliguer, Peter Caines, John Geweke, Enmanuel Parzen, Murray Rosenblatt, and Murad S. Taqqu eds. *New Directions in Time Series Analysis*, Part II. New York: Springer-Verlag, 71-92.
- [6] Gallant, R. y G. Tauchen (1990b) "SNP: A Program for Nonparametric Time Series Analysis," Manuscrito, Departamento de Economía, Duke University.
- [7] Gallant, A. Ronald, David A. Hsieh, y George E. Tauchen (1991), "On Fitting Recalcitrant Series: The pound/Dollar Exchange Rate, 1974-83," en William A. Barnett, James Powell, George E. Tauchen, eds. *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Proceedings of the Fifth International

- Symposium in Economic Theory and Econometrics, Cambridge: Cambridge University Press, Capítulo 8, 199-240.
- [8] Gallant, R. P. Rossi y Tauchen G. (1992): "Stock Prices and Volume" *Review of Financial Studies* v 5 n 2, pp. 199-242.
 - [9] Gallant, A. Ronald, Peter E. Rossi y George Tauchen (1993), "Nonlinear Dynamic Structures" *Econometrica* 61, 871-907.
 - [10] Gallant, R. y G. Tauchen (1999), "The Relative Efficiency of Method of Moment Estimators," *Journal of Econometrics* 92, 149-172.
 - [11] Gallant, R. y G. Tauchen (2001) "Efficient Method of Moments," Manuscript, Department of Economics, Duke University.
 - [12] Jara, D. y M. Ochoa (2004), "Estimación de la Estructura de la Tasa de Interés en Chile 1992-2003," Manuscrito, Universidad de Chile.
 - [13] Monahan, John F. (1987), An Algorithm for Generating Chi Random Variables, Algorithm 651, ACM Transactions on Mathematical Software 13, pp 168-172.
 - [14] White, H. (1980), Using Least Squares to Approximate Unknown Regression Functions," *International Economic Review*, 21, 149-170

Anexo A: Expansiones de Hermite.

Se describirá la manera en que se utilizan las expansiones de Hermite para aproximar funciones de densidad multivariadas.

Un conjunto particular de polinomios de Hermite $H_n(z)$ es ortogonal de grado n en el intervalo $]-\infty, +\infty[$ con respecto a la función de peso $\exp(-z^2/2)$ si se cumple que:

$$\int_{-\infty}^{+\infty} \exp(z^2/2) H_n(z) H_m(z) dz = \sqrt{2\pi n!} \delta_{nm} \quad (\text{A1})$$

siendo $\delta_{nm} = 1$ si $n = m$ y cero en otro caso. Para el caso de la función de peso señalada, los polinomios de Hermite poseen la siguiente representación general:

$$H_n(z) = (-1)^n \exp(z^2/2) \frac{\partial^n}{\partial z^n} \exp(-z^2/2) \quad (\text{A2})$$

siendo los primeros polinomios de Hermite:

$$\begin{aligned} H_0(z) &= 1 \\ H_1(z) &= z \\ H_2(z) &= z^2 - 1 \\ H_3(z) &= z^3 - 3z \\ H_4(z) &= z^4 - 6z^2 + 3 \\ H_5(z) &= z^5 - 10z^3 + 15z \\ H_6(z) &= z^6 - 15z^4 + 45z^2 - 15 \end{aligned}$$

Se define $\psi(z, t) = \exp(zt - t^2/2)$ como una función generadora de $H_n(z)$, dado que su desarrollo de Taylor en t tiene como coeficientes los dichos polinomios:

$$\psi(z, t) = \sum_{n=0}^{\infty} \frac{H_n(z)}{n!} t^n \quad (\text{11})$$

dado que $\partial^n \psi(z, t) / \partial t^n |_{t=0}$ es igual a A2. La suma infinita de polinomios representa en forma exacta la función $\psi(\cdot)$. A partir de A1 se define un nuevo conjunto de polinomios de Hermite:

$$\tilde{H}_n = \frac{H_n \exp(-z^2/4)}{\sqrt{\sqrt{2\pi n!}}} \quad (\text{A4})$$

tal que satisfacen A1 con función de peso 1. Esta forma particular de polinomios es la que se utilizó para la aproximación polinomial de Hermite efectuada en el presente

trabajo. El objetivo es aproximar la densidad $h(z)$. Por conveniencia, la aproximación se efectúa sobre $\sqrt{h(z)}$. Dado que la suma infinita de polinomios es una representación exacta de la función $\sqrt{h(z)}$, la suma finita de polinomios sólo hará una aproximación. Entonces una aproximación de orden Kz será:

$$\sqrt{h(z)} \cong \sum_{n=0}^{Kz} \frac{H_n(z)}{n!} c_n \quad (\text{A5})$$

donde c_n se define en forma arbitraria en la aproximación. Redefinimos los c_n tal que se reexpresen A5 como:

$$\sqrt{h(z)} \cong \sum_{n=0}^{Kz} \tilde{H}_n(z) b_n \quad (\text{A6})$$

donde $\tilde{H}_n(z)$ se define en A4. Elevando al cuadrado ambos términos de A6, reemplazando A4 en A6 y recordando que $\phi(z) = (2\pi)^{-1/2} \exp(-z^2/2)$, se llega a la expansión polinomial de Hermite de la densidad $h(z)$:

$$h(z) \propto \left[\sum_{n=0}^{Kz} \hat{H}_n(z) b_n \right]^2 \phi(z) \quad (\text{A7})$$

donde $\hat{H}_n(z) = H_n(z)/n!$. Una vez escogido orden de la expansión de Hermite Kz es posible reordenar los polinomios en la sumatoria, efectuar factorizaciones según las potencias de z y redefinir los parámetros correspondientes a fin de reexpresar A7 y llegar a :

$$h(z) \propto [P(z)]^2 \phi(z) \quad (\text{A8})$$

Anexo B: Algoritmo de Simulación de densidad SNP.

En esta parte se describe el algoritmo de simulación de la densidad SNP utilizada para el presente trabajo y que está descrito en Gallant et al (1990).

El método de rechazo

El método de rechazo para muestrear de una densidad multivariada $h(z)$ depende de encontrar una función $b(z)$ positiva e integrable que sea envolvente o domine a $h(z)$. Es decir $0 \leq h(z) \leq b(z)$. Se deriva la densidad $g(v)$ a partir de $b(v)$ haciendo $g(v) = b(v) / \int b(s) ds$. La forma de utilizar $g(v)$ y $b(z)$ para simular de $h(z)$ es la siguiente: generar un par (u, v) tal que v pertenezca a $g(v)$ y u tenga una distribución uniforme $[0, 1]$. Si $u > h(v)/b(v)$ se rechaza el par (u, v) y se intenta de nuevo. Si $u \leq h(v)/b(v)$ se acepta $z = v$ como muestra de $h(z)$. El método funciona porque:

$$\begin{aligned}
 P(z \leq t) &= P[\{(u, v) : v \leq t\} | \{(u, v) : u \leq h(v)/b(v)\}] \\
 &= \left[\int_{v \leq t} \int_0^{h(v)/b(v)} g(v) du dv \right] / \left[\int_{v \leq \infty} \int_0^{h(v)/b(v)} g(v) du dv \right] \\
 &= \left[\int_{v \leq t} h(v) / (\int b(s) ds) dv \right] / \left[\int_{v \leq \infty} h(v) / (\int b(s) ds) dv \right] \\
 &= \int_{v \leq t} h(v) dv
 \end{aligned}$$

Para que este método funcione bien, $h(v)/b(v)$ debe ser cercano a uno sobre las regiones donde $g(v)$ asigne una alta probabilidad de ocurrencia y además $g(v)$ sea fácil de muestrear.

Envolvente de $h(z)$

La densidad se escribe como $h(z/x) = [P(z, x)]^2 \phi(z) / c(x)$, donde $c(x) = \int [P(z, x)]^2 \phi(s) ds$ y donde $P(z, x) = \sum_{|\alpha|}^{Kz} a_\alpha(x) z^\alpha$. Dado que se cumple que:

$$\sum_{|\alpha|}^{Kz} a_\alpha(x) z^\alpha \leq \sum_{|\alpha|}^{Kz} |a_\alpha(x)| |z|^\alpha \tag{B1}$$

donde $|z|$ denota el vector z donde cada elemento es reemplazado por su valor absoluto, tal que:

$$b(z/x) = \left[\sum_{|\alpha|}^{Kz} |a_\alpha(x)/c(x)| |z|^\alpha \right]^2 \phi(z) \quad (\text{B2})$$

es una envolvente de $h(z/x)$.

Simulando una Innovación SNP

La idea general consiste en desarrollar un algoritmo para muestrear de $g(v)$ y utilizarlo para simular de $h(z/x)$ por el método de rechazo. Gallant et al (1990) plantean una representación de $b(z/x)$ en B2 como una suma ponderada de densidades CHI:

$$\chi(s, v) = \frac{2^{1-v/2}}{\Gamma(v/2)} s^{v-1} \exp(-s^2/2) \quad (\text{B3})$$

lo cual se verifica reescribiendo $b(z/x)$ del siguiente modo:

$$b(z/x) = \sum_{|\alpha|=0}^{Kz} \sum_{|\gamma|=0}^{Kz} \frac{|a_\alpha(x)a_\gamma(x)|}{c(x)} \prod_{i=1}^M \Gamma[(\alpha_i + \gamma_i + 1)/2] \prod_{i=1}^M \chi(|z|, \alpha_i + \gamma_i + 1) \quad (\text{B4})$$

Habiendo observado que $b(z/x)$ es una suma ponderada de funciones de densidad CHI, puede implementarse un algoritmo para muestrear de $g(v)$ sin necesidad de derivar una expresión cerrada para $g(v)$. Primero se normalizan los pesos en B4 tal que sumen 1 y a cada peso se le define como $\omega_{\alpha\gamma}$, donde:

$$\omega_{\alpha\gamma} = \frac{|a_\alpha(x)a_\gamma(x)| \prod_{i=1}^M \Gamma[(\alpha_i + \gamma_i + 1)/2]}{\sum_{|\alpha|=0}^{Kz} \sum_{|\gamma|=0}^{Kz} \left(|a_\alpha(x)a_\gamma(x)| \prod_{i=1}^M \Gamma[(\alpha_i + \gamma_i + 1)/2] \right)} \quad (\text{B5})$$

Los pesos definen una distribución $F(\alpha, \gamma)$ sobre $\{(\alpha, \gamma) : 0 \leq |\alpha| \leq Kz, 0 \leq |\gamma| \leq Kz\}$. El algoritmo que presentan Gallant et al.(1990) para simular de $g(v)$ como paso intermedio para simular de $h(z)$ es el siguiente:

1. Samplear (α, γ) de $F(\alpha, \gamma)$. Para ello se genera una variable u que se distribuye $U[0, 1]$ y encontrando el mayor T tal que $\sum_{\tau=1}^T \omega_\tau \leq u$ y tomando $(\alpha, \gamma) = (\alpha, \gamma)_T$ como muestra de $F(\alpha, \gamma)$.

2. Generar la secuencia s_1, \dots, s_M muestreando independientemente de $\chi(s, \alpha_1 + \gamma_1 + 1), \dots, \chi(s, \alpha_M + \gamma_M + 1)$.
3. Aleatoriamente cambiar el signo de s_i .
4. Generar u que se distribuye $U[0, 1]$. Si $u > h(v/x)/b(v/x)$ entonces rechazar el par (u, v) e intentar de nuevo. Si $u \leq h(v/x)/b(v/x)$ aceptar $z = v$ como una muestra de $h(z)$.

Se utilizó el algoritmo de Monahan (1987) para simular de la densidad CHI en B3.