



Universidad de Chile

Facultad de Ciencias Económicas y Administrativas

Escuela de Economía y Administración

MÉTODOS ECONOMETRÍCOS PARA EL PRONÓSTICOS DE DELITOS EN EL GRAN SANTIAGO

**SEMINARIO PARA OPTAR AL TÍTULO DE INGENIERO COMERCIAL
CON MENCIÓN EN ECONOMÍA**

AUTOR: Rodrigo Toledo Astudillo

PROFESOR GUÍA: José Miguel Benavente

Santiago de Chile, Agosto de 2005

DEDICATORIA

A mi hermana Yuly

AGRADECIMIENTOS

A Carabineros de Chile por su disposición y gentileza para colaborar con la información para el desarrollo del estudio.

A Fondef que gracias a su financiamiento, hizo posible que esta investigación se concretara.

Al profesor guía José Miguel Benavente, por sus valiosos comentarios que permitieron mejorar continuamente este documento.

A mis padres y amigos que estuvieron constantemente brindándome su apoyo y consejos en este periodo.

RESUMEN

El presente estudio, tiene como objetivo probar cuan efectivos pueden ser los modelos multiecuacionales del tipo vectores auto regresivos (VAR) y los modelos de sistemas de ecuaciones, para desarrollar pronósticos de corto plazo de delitos, aplicado a la comuna de Santiago en el período que comprende el primero de enero de 2001 y el 30 de junio de 2004.

Se encontró que los modelos apropiados para la formulación de pronósticos difieren, dependiendo del sector que se esté tratando. Los errores de pronósticos para las series diarias son cercanos al 27 %. Similares a los que se han obtenido en estudios para otros países como Inglaterra y Estados Unidos.

ABSTRACT

In this study, we aim to test how effective are multiecuacionales models like auto regressive vectors (VAR) and models of systems of equations, for forecast by short term of crimes, applied to the Santiago in the period between January 1st of 2001 and June 30th of 2004. We show that bests forecast models are different, depending on the sector that is under study, the errors of daily series near 27%, similar to studies for other countries like England and the United States.

Índice

	Pág.
Capítulo 1	
1. Introducción.....	10
Capítulo 2	
1. Análisis Estadístico de las series.....	13
2. 2.0 consideraciones iniciales acerca de los datos.....	13
2.0.1 Selección de delitos.....	14
2.0.2 Descripción de las series.....	14
3. 2.1 Introducción al estudio estadístico.....	14
4. Análisis Gráfico.....	16
5. 2.2.1 Series en nivel diario.....	16
2.2.2 Series en nivel semanal.....	18
2.2.3 Series en nivel mensual.....	19
6. 2.3 Estadísticas Descriptivas.....	20
2.3.1 Series en nivel diario.....	21
2.3.2 Series a nivel semanal.....	23
2.3.3 Series en nivel mensual.....	23
7. 2.4 Patrones de correlaciones simples de las series.....	24
2.4.1 Series en nivel diario.....	24
2.4.2 Series en nivel semanal.....	26
2.4.3 Series en nivel mensual.....	26
8. 2.5 Pruebas de estacionariedad.....	27
9. 2.6 Estudio de Tendencias.....	29

2.6.1	Series en nivel diario.....	30
2.6.2	Series en nivel semanal.....	31
2.6.3	Series en nivel mensual.....	31
10. 2.7	Tests de igualdad de medias.....	31
11. 2.8	Conclusiones.....	32
2.8.1	Series en nivel diario.....	32
2.8.2	Series en nivel semanal.....	34
2.8.3	Series en nivel mensual.....	34

Capítulo 3

1.	Modelos de Vectores Auto regresivos (VAR) y Sistemas de ecuaciones.....	35
2.	3.1 Justificación estadística-econométrica para la utilización de modelos VAR.....	35
3.	3.2 Metodología empleada y justificaciones teóricas.....	37
	3.2.1 Metodología.....	37
4.	3.3.1 Modelo VAR en nivel diario.....	38
	3.3.2 Modelo VAR en nivel semanal.....	42
	3.3.3 Modelo VAR en nivel mensual.....	45
5.	3.4 Funciones de impulso-respuesta diaria.....	47
6.	3.5 Sistemas de ecuaciones.....	50
	Metodología.....	50
	3.5.1 Sistema diario.....	51
	3.5.2 Sistema semanal.....	52
	3.5.3 Sistema mensual.....	53

Capítulo 4

1.	Pronósticos	
	4.1 Introducción.....	56
2.	4.2 Modelos alternativos de predicción.....	57
	4.2.1 Suavización Exponencial Simple.....	57
	4.2.2 Suavización Exponencial con Tendencia.....	58
	4.2.3 Suavización Exponencial con Estacionalidad.....	59
	4.2.4 Suavización Exponencial con Estacionalidad Aditiva.....	59
	4.2.5 Suavización con Estacionalidad Multiplicativa.....	60
3.	4.3 Resultados de los Modelos de Suavización Exponencial.....	60
5.	4.4 Pronósticos con los modelos VAR y de sistemas de ecuaciones.....	62
	4.4.1 Resultados de los modelos VAR y sistemas de ecuaciones.....	63
6.	4.5 Proyecciones dentro de muestra.....	65
7.	4.6 Pronósticos diarios fuera de muestra para el sistema de ecuaciones.....	67
8.	4.7 Pronósticos a partir de julio de 2004 con los modelos VAR y de Sistema de ecuaciones.....	70

Capítulo 5

1. Conclusiones y comentarios finales.....	72
• Referencias Bibliográficas.....	75
• Anexo.....	76
• 2.3.1 Correlaciones en nivel diario rezagadas.....	77
• Correlaciones en nivel mensual rezagadas.....	77
• Funciones de impulso-respuesta en nivel diario.....	78
• Funciones de impulso-respuesta en nivel mensual.....	82
• Sistema diario.....	87
• Sistema mensual.....	89

Índice de Gráficos y Tablas	pág.
Gráfico 2.2.1a (Series diarias para los 5 cuadrantes de la comuna de Santiago Enero-Febrero 2001).....	17
Gráfico 2.2.1b (Series diarias para los 5 cuadrantes de la comuna de Santiago Agosto-Septiembre 2003).....	18
Gráfico 2.2.2 (Series semanales para los 5 cuadrantes de la comuna de Santiago).....	19
Gráfico 2.2.3 (Series mensuales para los 5 cuadrantes de la comuna de Santiago).....	20
Cuadro 2.3.1 (Estadísticas descriptivas a nivel diario).....	21
Cuadro 2.3.1a (cantidad de delitos en el X% de los días del periodo).....	22
Cuadro 2.3.1b (Cantidad acumulada de delitos por intervalos en RM).....	23
Cuadro 2.3.2 (Estadísticas descriptivas a nivel semanal).....	23
Cuadro 2.3.3 (Estadísticas descriptivas a nivel mensual).....	24
Cuadro 2.4.1 Matriz de correlaciones a nivel diario.....	24
Cuadro 2.4.2 Matriz de correlaciones a nivel semanal.....	26
Cuadro 2.4.3 Matriz de correlaciones a nivel mensual.....	26
Cuadro 2.5.1 Resultados pruebas de raíz unitaria series diarias.....	28
Cuadro 2.5.2 Resultados pruebas de raíz unitaria series semanales.....	29
Cuadro 2.5.3 Resultados pruebas de raíz unitaria series mensuales.....	29
Cuadro 2.6.1 (Significancia de tendencias diarias y ajustes).....	30
Cuadro 2.6.2 (Significancia de tendencias mensuales).....	31
Cuadro 2.7.1(Resultados test de medias).....	32
Cuadro 2.8.1 Promedio de delitos diario por día de la semana.....	33
Cuadro 2.8.2 Porcentaje promedio de ocurrencia de delitos.....	33
Cuadro 3.3.1 Resultado estimación VAR diario.....	40
Cuadro 3.3.1a.....	42
Cuadro 3.3.2 Resultado estimación VAR semanal.....	44
Cuadro 3.3.3 Resultado estimación VAR mensual.....	46
Gráfico 3.4 (Efectos de un shock en C1, sobre todos los cuadrantes).....	49
Cuadro3.5.1 R- cuadrado ajustado del VAR a nivel diario.....	52

Cuadro 3.4.2 cuadrado ajustado del VAR a nivel semanal.....	52
Cuadro 3.4.3 Ajustes alcanzados por ecuación.....	54
Cuadro 4.3.1 Error de Pronósticos diarios.....	61
Cuadro 4.3.2 Error de Pronósticos semanales.....	62
Cuadro 4.3.3 Error de pronósticos mensuales.....	62
Cuadro 4.4.2 Error porcentual medio de predicción en nivel diario.....	64
Cuadro 4.4.3 Error absoluto medio de predicción en nivel mensual.....	64
Gráfico 4.5.1 (Pronóstico diario cuadrante 2 utilizando método de suavización exponencial).....	66
Gráfico 4.5.2 (Pronóstico diario Cuadrante 1 utilizando VAR).....	67
Cuadro 4.6.1 Error absoluto medio de pronóstico por cuadrante utilizando sistema (fuera de muestra).....	68
Gráfico 4.6.1 (Pronóstico diario fuera de muestra utilizando sistemas de ecuaciones para el cuadrante 1).....	68
Cuadro 4.6.2 Error absoluto medio de pronóstico por cuadrante utilizando VAR (fuera de muestra).....	69
Gráfico 4.6.2 (Pronóstico diario fuera de muestra utilizando modelos VAR para el cuadrante).....	70
Gráfico 4.7.1 (Pronóstico diario fuera de muestra utilizando modelos VAR para el cuadrante 1).....	71
Gráfico 4.7.2 (Pronóstico diario fuera de muestra utilizando sistemas de ecuaciones para el cuadrante 1).....	71

Capítulo 1

Introducción

En los últimos años, países como Inglaterra y Estados Unidos han estado investigando la utilidad de los modelos econométricos para la formulación de pronósticos de corto plazo de delitos, alcanzándose excelentes resultados en esta materia. Los modelos han permitido detectar con mayor claridad donde se llevarán a cabo los crímenes, los sectores que son más propensos a sufrir ataques en las distintas fechas, las horas y días de la semana en donde se concentran los asaltos según zonas geográficas, entre otras. El conocimiento de lo anterior permite realizar una asignación más eficiente y ordenada de los policías para disuadir los asaltos. En algunas regiones de Inglaterra, en donde se ha implementado la “Operación Fénix”¹ se ha reducido notablemente la tasa de criminalidad. Los policías son enviados en multitudes hacia los lugares en donde los modelos indican que se perpetrarán la mayor cantidad de delitos, y de esta forma se logra prevenir que estos se materialicen.

La idea que subyace detrás de la utilización de modelos econométricos para realizar pronósticos de delitos, es que los crímenes tienden a presentar patrones de comportamiento definidos a lo largo del tiempo, los cuales pueden ser estudiados y capturados a través de técnicas estadísticas y matemáticas. En particular, como se desconoce las causas que motivan a los delincuentes a delinquir, o al menos, no se tiene un registro de ellas en las bases de datos, sólo es posible estudiar el comportamiento de los delitos considerando como estos se distribuyen a través del tiempo. Se estudian los patrones de comportamiento, las estacionalidades, tendencias, entre otras variables que tienen relación con el tiempo.

Existen múltiples técnicas de series temporales. Estas se pueden dividir según su estructura. En modelos uniecuacionales; estos describen el comportamiento de una serie a lo largo del tiempo sin tener en consideración lo que pasa en las otras, es decir, se asume que cada una de las series se afectan independientemente y no se alteran por sucesos que provoquen alteraciones en otras series; dentro de estos modelos se puede encontrar todos aquellos pertenecientes a la familia ARIMA.

¹ Sistema de vigilancia policial

También es posible clasificarlos en modelos multiecuacionales los que consideran que existe una interrelación en las distintas series que son materia de estudio, afectándose cada una de ellas ante cambios en las otras series.

El presente estudio tiene como objetivo probar cuan efectivos pueden ser los modelos multiecuacionales como el de vectores auto regresivos (VAR) y los modelos de sistemas de ecuaciones para el desarrollo de pronósticos de corto plazo, aplicado a la predicción de crímenes en la comuna de Santiago en el período que comprende el primero de enero de 2001 y el 30 de junio de 2004.

Para lograr lo anterior el estudio se estructura de la siguiente manera. En el capítulo 2 se hace una revisión estadística de las series de delitos correspondientes a la comuna de Santiago, analizando estacionalidades, tendencias, patrones de correlación y estadísticas descriptivas principales. Se detectó que las series presentan estacionalidades a nivel diario, es decir, la cantidad de delitos tiende aumentar a medida que transcurre la semana, de lunes a viernes y declinan los días sábados y domingos. En algunos sectores se observó disminuciones significativas en la cantidad de delitos para los fines de mes en comparación con el resto de los días del mes. A su vez, se detectaron correlaciones negativas para algunos sectores, entre otros hallazgos.

En el capítulo 3 se realizan estimaciones de las series de delitos utilizando las metodologías VAR y de sistemas de ecuaciones. Fue posible detectar que las series presentan estacionalidades y tendencias a nivel diario. A niveles semanales y mensuales no presentan comportamientos claros. Los ajustes de los modelos VAR y de los Sistemas de ecuaciones difieren para las distintas zonas geográficas analizadas; en algunos casos es preferible el VAR mientras que en otros el sistema ajusta mejor.

En el capítulo 4 se evalúa la capacidad de predicción dentro y fuera de muestra de los modelos VAR y de sistemas de ecuaciones para las distintas zonas de la comuna de Santiago. Además, se introducen los métodos de suavización exponencial como forma alternativa para generar pronósticos. Se encontró que no existe un método único que permita proyectar la cantidad de delitos para las

distintas zonas de la comuna de Santiago, algunos métodos entregan mejores resultados que otros para algunas zonas, mientras que fallan para otras. Los métodos de suavización exponencial fueron de gran ayuda para realizar pronósticos de las series a nivel semanal.

Finalmente en el capítulo 5 se exponen comentarios y conclusiones finales derivadas del estudio.

Capítulo 2

Análisis Estadístico de las series

2.0 consideraciones iniciales acerca de los datos

La comuna de Santiago tiene una superficie de 5.500 Km² aproximadamente y se encuentra dividida en 5 cuadrantes. El concepto de cuadrantes surgió en 1999 cuando se implementó el Plan Cuadrante en la comuna de Santiago. Este plan apuntó básicamente a estrechar las relaciones entre Carabineros de Chile y la comunidad y consistió en que el territorio jurisdiccional de cada comisaría se subdividió en sectores claramente delimitados y de los cuales se tiene información precisa acerca del número y características de la población, actividades económicas y sociales más importantes, así como la realidad delictual de dicho territorio. A cada una de las áreas de acción descritas anteriormente se le asignó el nombre de Cuadrante.

La información respecto al momento y lugar de ocurrencia de los delitos considerados en el presente estudio se obtuvo de la base de datos AUPOL, que maneja Carabineros de Chile desde 2001. El desarrollo del Sistema de Automatización de Unidades Policiales (AUPOL) forma parte de un proyecto de modernización institucional de Carabineros de Chile en el área de Sistemas de Información Operativa, en línea con la iniciativa de Gobierno de modernizar las instituciones del sector público, iniciativa conocida con el nombre de "Gobierno Electrónico". El AUPOL consiste, básicamente, en un sistema computacional que registra de manera estandarizada e instantánea el momento y lugar de ocurrencia de un delito o una denuncia. La información que automáticamente se transfiere a un archivo histórico, permite realizar un monitoreo de la actividad criminal a lo largo de todo el país. Esta tecnología permite, entre otras cosas, el envío electrónico de partes al ministerio público y es la base de otros sistemas de información que se elaboran a partir de este sistema de información.

A partir del AUPOL se desarrolla el Sistema Estadístico Institucional de Carabineros de Chile, que permite el procesamiento de datos de manera automática asociada, por ejemplo a cantidad y evolución de un determinado delito en el país, zona, prefectura, comuna, ciudad y cuadrante; lo cual permite contar con información en tiempo real y responder de mejor manera a las eventualidades. El AUPOL también se encuentra conectado al Sistema de Información Geográfica Delictual, que posibilita el análisis respecto a la movilidad territorial y horaria en un mapa digital.

2.0.1 Selección de delitos²

La base de datos de AUPOL cuenta con información de diversos tipos de delitos. Entre ellos están las infracciones a ley del tránsito, violencia intra-familiar, daños a la propiedad pública y privada, desórdenes en la vía pública, hurtos, robos entre otros.

Para efectos del presente estudio, solo consideraremos los delitos de mayor connotación social. Los delitos de mayor connotación social, son aquellos que ocurren con mayor frecuencia y representan un alto porcentaje de participación en el total de los delitos ocurridos. Los principales son: el robo con violencia y fuerza (52%), el hurto (23%), violación, homicidios y las lesiones (22%).

2.0.2 Descripción de las series

Se construyeron seis series temporales denominadas C1, C2, C3, C4, C5 y RM, las cuales representan al cuadrante 1, 2, 3, 4, 5 de la comuna de Santiago y RM que considera a las 52 comunas de la región metropolitana. El periodo considerado en el estudio abarca desde el primero de enero de 2001 hasta el 30 de junio de 2004.

2.1 Introducción al estudio estadístico

Como se mencionó, el presente capítulo tiene como finalidad realizar un estudio preliminar de los datos, describiendo en detalle el comportamiento de las series de delitos a lo largo del tiempo a través de un conjunto de herramientas estadísticas

² Los tipos de delitos incluidos en el estudio se detallan en el anexo

que permitirán caracterizar a priori las series, para luego en fases posteriores llegar finalmente a la modelación en términos matemáticos de ellas.

Para obtener una cantidad de información razonable de los datos, se hará el estudio en tres niveles diario, semanal y mensual³.

Algunos de los aspectos que se considerarán en el estudio preliminar son:

Estacionalidades: Con esto se comprueba si las series presentan algún comportamiento cíclico en el tiempo. Por ejemplo veremos que RM presenta un aumento explosivo de la cantidad de delitos, en los fines de cada mes, en cambio otros cuadrantes de la comuna de Santiago no parecen ser afectados, teniendo un comportamiento similar en el transcurso del mes.

Tendencias: Con esto se observará si existe alguna tendencia alcista o a la baja en la cantidad de delitos en los años 2001 a 2004, es decir, si ha habido o no, un aumento sostenido a lo largo del período analizado. Se verá que existen tendencias mensuales que declinan suavemente, etc.

Relación entre las series: Se detectará que series presentan interrelación estadística contemporánea o rezagada, para luego especificar modelos de ecuaciones utilizando esta información.

Estacionariedad: La estacionariedad de las series es una condición necesaria para modelar las series de tiempo, esta condición asegura la convergencia de los parámetros de los modelos ajustados.

Estadísticas Descriptivas: Las estadísticas descriptivas permiten hacer comparaciones estadísticas entre las series, se utilizará la media, mediana, desviación estándar, coeficiente de variación⁴; entre otros.

³ Este seminario no estudia el comportamiento de la serie de delitos al interior de un día, para un análisis detallado se recomienda leer "Modelos de series de tiempo para proyectar delitos en la Región Metropolitana" López 2005

⁴ El coeficiente de variación nos da una medida para comparar la dispersión de los datos que existe en las series.

La fórmula es $CV = SD/M$

CV: coef. de variación , SD: desviación estándar , M: media

El resto de este capítulo se estructura de la siguiente manera. En primer lugar se realiza un estudio gráfico de las series y luego se hace una revisión de estadísticas descriptivas. Posteriormente, se hacen pruebas de correlación contemporánea y con rezagos de cada una de las series, para luego realizar tests de Dickey –Fuller aumentados para comprobar si las series son estacionarias y por ende susceptibles de ser modeladas con la metodología de series temporales. Enseguida, se hace un estudio para detectar posibles tendencias en las series. Finalmente se hacen tests de hipótesis para comprobar si las medias de cada una de las series permanecen iguales en los distintos años que comprende el estudio.

2.2 Análisis Gráfico

El análisis gráfico es un estudio previo, donde se identifican a priori una sucesión de elementos que permiten caracterizar a las series, dentro de ellos se pueden mencionar los siguientes: estacionalidades, interrelaciones entre las series, tendencias, estacionariedad, patrones de comportamiento, posibles outliers⁵, entre los más importantes.

Una vez detectados a priori estos elementos, se prosigue con la comprobación estadística, utilizando la información proporcionada por las matrices de correlaciones, tests de estacionalidad, tests de estacionariedad y estadísticas descriptivas.

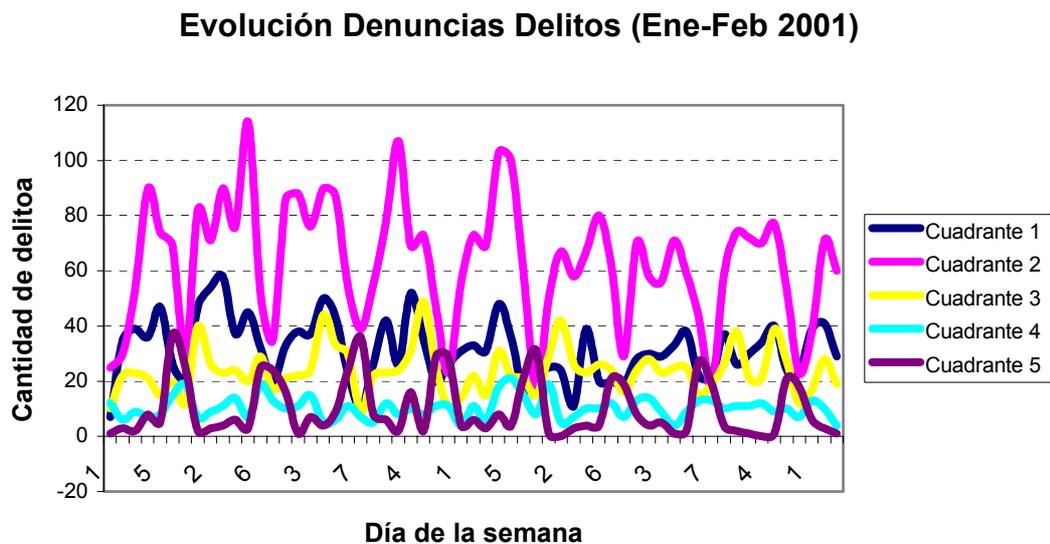
2.2.1 Series en nivel diario

Debido a la extensión de las series, cuando se consideran en nivel diario no son posibles de visualizar claramente en un gráfico, por lo tanto se escogió un extracto de ellas para presentar el comportamiento que tienen los datos en nivel diario. El examen se hizo para cada una de las series tomando intervalos de dos meses, para comprobar si los patrones se mantenían a lo largo del tiempo y si las estacionalidades detectadas no variaban. En efecto, las características de las series no eran muy disímiles de un período a otro, lo que implica que las conclusiones

⁵ Se denomina **outlier** a los datos que se escapan demasiado del comportamiento que presentan los restantes que existen en la muestra, por ejemplo si en el mes de febrero de 2002 hubieran 1.000.000 de delitos este mes sería un outlier y habría que identificar el porqué de aquella anomalía.

obtenidas de los gráficos 2.2.1a y 2.2.1b que se observan a continuación se replican en promedio para todas las series.

Gráfico 2.2.1a (Series diarias para los 5 cuadrantes de la comuna de Santiago Enero-Febrero 2001)

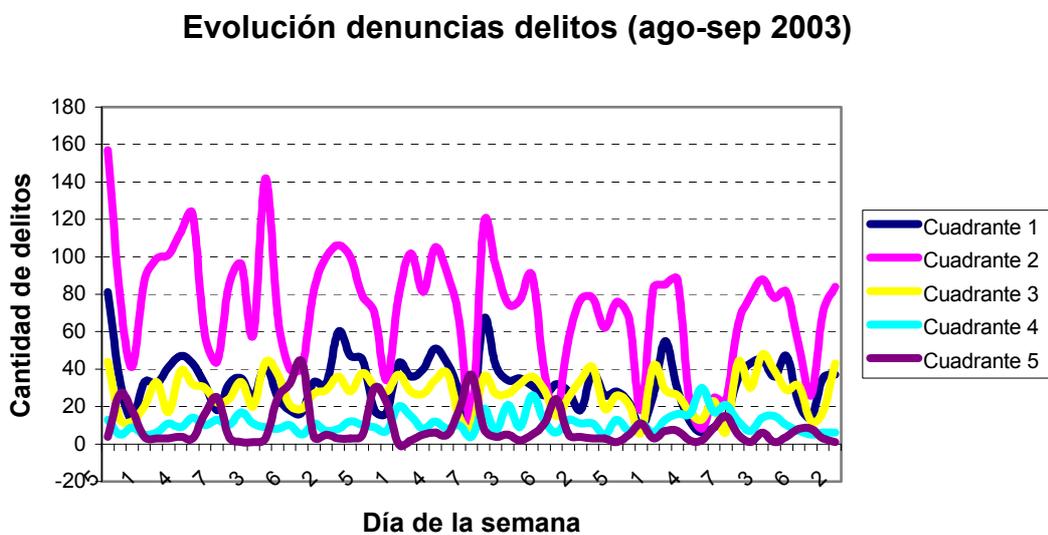


De la inspección de los gráficos 2.2.1a y 2.2.1b es posible percatarse de la existencia de un fuerte patrón cíclico para todos los cuadrantes siendo el intervalo del ciclo de un período constante. La cantidad de delitos va en ascenso a medida que aumenta el día de la semana y declina los fines de semana, alcanzando generalmente un peak los días viernes. A priori, se puede decir que hay tendencias crecientes de lunes a viernes y luego tendencias decrecientes para sábados y domingos. Para corroborar esta primera impresión se realizan tests de hipótesis contrastando la hipótesis nula de existencia de tendencia a un nivel de significancia del 5%, los resultados se resumen en la sección 2.5.1.

De los gráficos anteriores, se puede ver que todas las series tienden a comportarse en torno a una media no existiendo datos muy dispersos, lo que da indicios de que las series son estacionarias. El C1 mantiene una correlación contemporánea positiva con el C2, con C3 también presenta una correlación positiva pero que tiende a rezagarse, con el C4 y C5 no presenta una correlación clara y esta parece ser negativa si es que realmente existiese. El C2 presenta una correlación negativa

con el C5 y una correlación negativa pero no muy distinguible con el C4, con el C3 se observa una correlación positiva pero que tiende a rezagarse un período. El C3 presenta correlación negativa con el C5 pero rezagada, y nada muy claro con el C4.

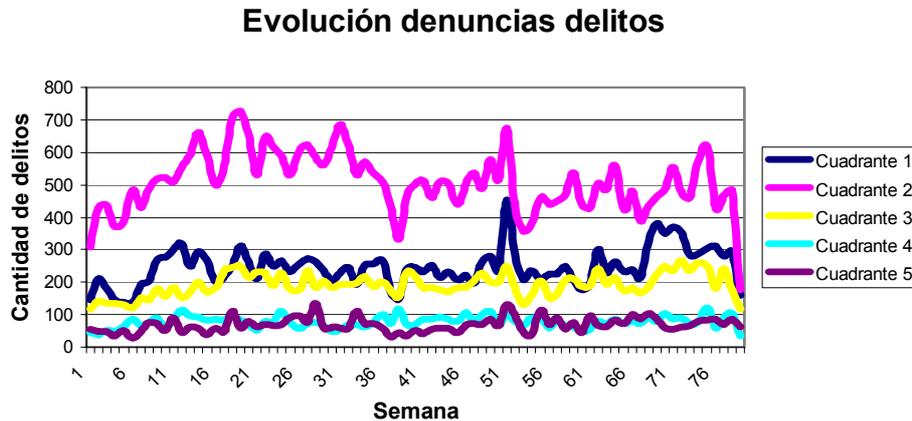
Gráfico 2.2.1b (Series diarias para los 5 cuadrantes de la comuna de Santiago Agosto-Septiembre 2003)



2.2.2 Series en nivel semanal

A nivel semanal las series no presentan tendencias a simple vista. Del gráfico 2.2.2 se observa que las series, tienden a concentrarse alrededor de un valor promedio, ello sugiere que las series son estacionarias; las correlaciones observadas en nivel diario tienden a mantenerse. Sin embargo, es más difícil darse cuenta de este fenómeno visualmente, por lo que será necesario verificarlo a través de las matrices de correlación. En general, las series parecieran comportarse de manera estable sin tener mucha volatilidad, ni sobresaltos significativos en las semanas, salvo el C2 que experimenta un fuerte cambio al alza en la primera semana de 2003 hasta principios de mayo, para luego retomar los valores que venia presentando en las semanas anteriores.

Gráfico 2.2.2 (Series semanales para los 5 cuadrantes de la comuna de Santiago)



2.2.3 Series en nivel mensual

Las series mensuales para C3, C4 y C5 no presentan mucha variación a lo largo del tiempo (ver gráfico 2.2.3) Se observa poca oscilación mensual en torno a la media, lo que indica que no existe evidencia de estacionalidad mensual, es decir, el comportamiento de los delitos en estos cuadrantes no depende de la fecha, por lo tanto la cantidad de delitos es muy similar si es febrero o diciembre o cualquier otro mes del año. El comportamiento de los cuadrantes anteriores es acotado a nivel mensual siendo poca la variación que presentan en los distintos meses del año.

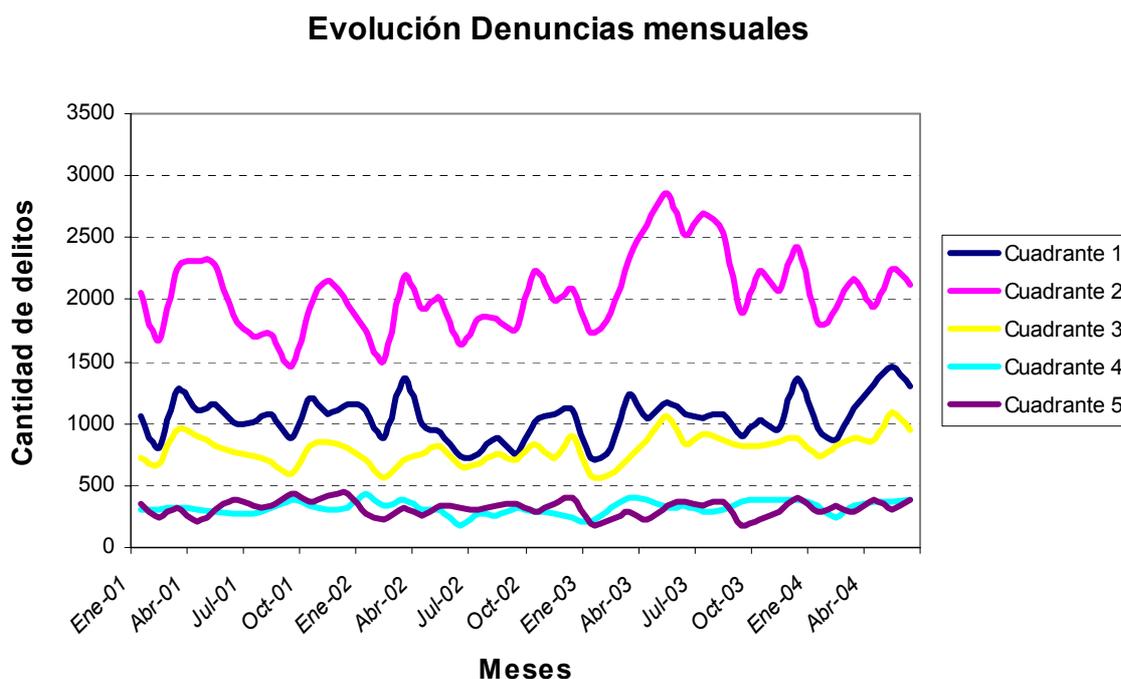
Los cuadrantes C1 y C2 son los que presentan una mayor variabilidad, de la inspección gráfica se puede decir que existe un leve patrón estacional; los meses de enero y febrero experimentan un descenso de los delitos, para luego aumentar en los meses de marzo y abril. Posteriormente, a partir de mayo, se observa un nuevo declive hasta el mes de septiembre, para luego retomar una tendencia alcista en los últimos meses del año.

Cabe destacar que a partir del mes de enero de 2003 ocurre un aumento brusco en la cantidad observada de delitos en el C2, la que comienza con un ligero aumento en enero para luego acentuarse en los meses de marzo y abril. Después, la serie

tiende a suavizarse a partir de mayo de 2003, retomando el patrón estacional descrito previamente.

Se puede observar a priori que las series en nivel mensual también son estacionarias y no presentan tendencias a simple vista. A diferencia de las series en nivel diario estas están casi todas correlacionadas positivamente entre ellas, exceptuando la serie C5 que pareciera ser acíclica.

Gráfico 2.2.3 (Series mensuales para los 5 cuadrantes de la comuna de Santiago)



2.3 Estadísticas Descriptivas

En esta sección se pasa revista a los principales indicadores estadísticos que caracterizan a las series. Para ello se utilizarán la media, mediana, desviación estándar, coeficiente de variación, máximos y mínimos.

2.3.1 Series en nivel diario

Del cuadro 2.3.1 se puede inferir que la mayor cantidad de delitos registrados en la comuna de Santiago se encuentra en el C2 con un promedio de 67.5 delitos por día, y el mínimo en el C5 con un promedio de 10.3 delitos por día. La mayor variabilidad de las series, medida por el coeficiente de variación, se encuentra en el C5 con un Coeficiente de variación de 1.05 y la menor en el C3 con un 0.39. Esto quiere decir que la dispersión de los datos en torno a la media es mayor para el C5 en comparación con el C3, sugiriendo que los pronósticos asociados al C5 tendrán una mayor probabilidad de error, dada la volatilidad de dicha serie. La mediana muestra la cantidad de delitos que se encuentra en el valor central de la distribución; para el C2 es 68, esto implica que el 50% de los días de 2001 al primer semestre de 2004 tienen una cantidad superior a 68 delitos. Los valores máximos reflejan cual es la cantidad mayor de delitos observada en el período por cuadrante. Por ejemplo, para el C1 un día se registraron 95 delitos, análogamente para el valor mínimo. (ver cuadro 2.3.1)

De los análisis de distribución se obtuvieron los siguientes resultados:

Cuadro 2.3.1 (Estadísticas descriptivas a nivel diario)

	C1	C2	C3	C4	C5	RM
Media	33.7	67.5	25.5	10.4	10.3	882.5
Mediana	32.0	68.0	25.0	10.0	5.0	857.0
Máximo	95.0	257.0	74.0	32.0	54.0	2030.0
Mínimo	3.0	7.0	4.0	0.0	0.0	285.0
Std Desv	13.4	28.0	10.0	5.1	10.9	193.6
Coef var	0.40	0.41	0.39	0.49	1.05	0.22

Fuente: Cálculo basado en los datos de AUPOL

Para el C1 se observa que el 25% de los días del período tiene una cantidad de delitos menor igual a 24 delitos. El 90% de los días una cantidad menor o igual a 52. En el C2 el 25% de los días tiene una cantidad menor a 50 delitos, mientras que el 88% de los días tiene una cantidad menor a 100 delitos.

En el C3 el 25% de los días registra una cantidad de delitos menor o igual a 19, mientras que el 90% de los días tiene una cantidad menor o igual a 40 delitos. En el C4 el 25% de los días tiene una cantidad de delitos inferior a 7, mientras que en el 90% de los días se observa una cantidad delitos menor a 18.

En el C5 se observa que el 25% de los días tiene una cantidad de delitos inferior a 3, el 2.51% de los días no registra delitos, el 90% de los días tiene una cantidad de delitos menor a 28. Esta última medida es útil para saber cuanto es en realidad la cantidad de delitos que ocurre en un día cualquiera. Por ejemplo, para el C2 se tiene un máximo en el periodo de 257, y además se sabe que el 90% de los días tiene una cantidad de delitos no superior a 100, esto refleja que dicho máximo es un outlier no siendo representativo del comportamiento habitual de la serie.

Cuadro 2.3.1a (cantidad de delitos en el X% de los días del periodo)

	25%	50%	90%
c1	24	32	52
c2	50	68	100
c3	19	25	40
c4	7	10	18
c5	3	5	28

Fuente: Cálculo basado en los datos de AUPOL

El cuadro 2.3.1b presenta la misma información que el cuadro (2.3.1a) pero para el caso de RM, la lectura del cuadro es la siguiente. Si se considera la segunda fila el primer elemento de ella [500, 1000) representa el intervalo de 500 a 1000 delitos, luego en la celda día aparece 1043, esto significa que 1043 días tienen una cantidad de delitos que oscila entre 500 y 1000. La celda siguiente indica el porcentaje de días que tienen una cantidad de delitos entre 500 y 1000, la celda días acumulados muestra la cantidad de delitos acumulados, es decir es la suma de los días en el primer intervalo más los días en el segundo intervalo, y finalmente se muestra el porcentaje de días que capturan el intervalo [0, 1000).

De lo anterior se puede decir que en la Región Metropolitana el 82% de los días se observa una cantidad de delitos que no supera los 1000 delitos, y que solo 33 días (2.58%) del periodo tiene una cantidad superior a los 1500 delitos.

Cuadro 2.3.1b (Cantidad acumulada de delitos por intervalos en RM)

Cantidad de delitos	Días	Porcentaje	Días acumulados	Porcentaje acumulado
[0, 500)	4	0.31	4	0.31
[500, 1000)	1043	81.68	1047	81.99
[1000, 1500)	197	15.43	1244	97.42
[1500, 2000)	29	2.27	1273	99.69
[2000, 2500)	4	0.31	1277	100
Total	1277	100	1277	100

Fuente: Cálculo basado en los datos de AUPOL

2.3.2 Series a nivel semanal

A nivel semanal ocurre una cantidad promedio de delitos de 239.8 ; 470.8 ; 181.5 ; 73.7 ; 72.8 en los cuadrantes C1, C2, C3, C4 y C5 respectivamente. La menor variabilidad (medida por el coeficiente de variación) se observa en RM con un 0.12, seguida de los cuadrantes C3, C2, C1, C4 y C5. (Ver cuadro 2.3.2)

Cuadro 2.3.2 (Estadísticas descriptivas a nivel semanal)

	C1	C2	C3	C4	C5	RM
Media	239.8	470.8	181.5	73.7	72.8	6,227.5
Mediana	235.0	469.0	180.0	71.0	70.0	6,247.0
Máximo	452.0	728.0	277.0	126.0	136.0	8,179.0
Mínimo	110.0	177.0	117.0	26.0	28.0	2,639.0
Desv. Estándar	51.0	87.8	33.4	17.5	20.8	735.6
Coef. var	0.21	0.19	0.18	0.24	0.29	0.12

Fuente: Cálculo basado en los datos de AUPOL

2.3.3 Series en nivel mensual

De manera similar a los casos anteriores la media mensual refleja la cantidad promedio de delitos que se da en un mes cualquiera. Se puede ver que la cantidad de delitos promedio que ocurre en la región metropolitana alcanza los 27133 y que el 50% de los meses concentran una cantidad de delitos superior a los 27236 (ver cuadro 2.3.3), existiendo una baja variabilidad en comparación con la que existe en los cuadrantes de la comuna de Santiago. Además la variabilidad de las series en, esto indica que se podrían tener buenos pronósticos de tipo mensual.

Cuadro 2.3.3 (Estadísticas descriptivas a nivel mensual)

	C1	C2	C3	C4	C5	RM
Media	1045.0	2051.4	791.0	321.0	317.3	27133.9
Mediana	1054.5	1998.5	811.5	317.0	325.5	27236.0
Máximo	1453.0	2860.0	1091.0	426.0	434.0	31770.0
Mínimo	720.0	1468.0	555.0	183.0	173.0	21020.0
Desv. Est	181.1	316.2	117.8	53.5	65.4	2859.3
Coef. Var	0.17	0.15	0.15	0.17	0.21	0.11

Fuente: Cálculo basado en los datos de AUPOL

2.4 Patrones de correlaciones simples de las series

El siguiente punto que se analizará es la correlación que existe entre las denuncias de delitos para los distintos cuadrantes de la comuna de Santiago. Esto dará un indicio de que series tienen una interrelación. Como se ha mencionado anteriormente, la relación entre ellas puede ser explicada por una gran cantidad de fenómenos; como puede ser su cercanía geográfica, la cantidad de centros comerciales que hay en cada uno de los cuadrantes, el número de cuarteles de carabineros, entre otras variables. No obstante, las series no contienen dicha información, por lo tanto el estudio sólo se reduce a observar que series se relacionan entre sí, sin importar el motivo de tal relación, solo se toma como guía lo que indican los datos. Este punto dará indicios de cómo deben ser modelados los sistemas multiecuacionales. Mientras mayor sea la correlación positiva o negativa que tenga un cuadrante con alguno de los otros, se sabrá que la relación entre ellos es más fuerte y por lo tanto la ecuación que describa a ese cuadrante debe incluir a sus vecinos.

2.4.1 Series en nivel diario

Cuadro 2.4.1 Matriz de correlaciones a nivel diario

	C1	C2	C3	C4	C5	RM
C1	1.00	0.61	0.48	0.36	-0.28	0.47
C2	0.61	1.00	0.57	0.32	-0.46	0.40
C3	0.48	0.57	1.00	0.22	-0.25	0.39
C4	0.36	0.32	0.22	1.00	-0.13	0.40
C5	-0.28	-0.46	-0.25	-0.13	1.00	0.21
RM	0.47	0.40	0.39	0.40	0.21	1.00

Fuente: Cálculo basado en los datos de AUPOL

Del cuadro 2.4.1 es posible percatarse que el cuadrante 1 presenta una alta correlación con el cuadrante 2 siendo esta igual a 0.61, y una correlación menor con el C3 y RM de 0.48 y 0.47 respectivamente. Esto sugiere que a priori, la ecuación que describe al cuadrante 1 debería incluir dentro de sus variables explicatorias a los cuadrantes 2,3 y RM en su forma contemporánea o algunos de sus rezagos.

Del mismo cuadro se puede inferir que el cuadrante 2 presenta una correlación positiva con los cuadrantes 1 y 3, también se aprecia una correlación negativa con el C5 igual a -0.46. El C3 se relaciona con los cuadrantes 1 y 2. El C4 no presenta una relación muy alta con ninguno de los cuadrantes siendo la mayor correlación observada la que tiene con RM que es igual a 0.38. El C5 se correlaciona negativamente con C1, C2, C3 y C4 y positivamente con RM, sin embargo las correlaciones son bajas destacándose sólo la que mantiene con el C2⁶.

Debido a la baja correlación existente entre los cuadrantes en nivel diario se sospecha que la modelación del VAR no entregará muy buenos ajustes. Por ello, para predecir de mejor forma el comportamiento en los cuadrantes 4 y 5 será necesario aplicar otro tipo de metodología como modelos ARIMA⁷. Otro punto que se puede observar es que existe una baja correlación de RM con el resto de los cuadrantes, no siendo apropiado incluir este en cada uno de las ecuaciones por ser un mal predictor del comportamiento de las series en nivel diario.

⁶ Se hicieron pruebas de correlación para los rezagos de las series y no se encontraron correlaciones significativas, salvo para el C2 con su rezago 7 C2(-7) y el C5 con C5(-7) con un valor igual a 0.58 y 0.65 respectivamente. Ver anexo donde se presenta un cuadro con las correlaciones más fuertes encontradas para cada cuadrante.

⁷ Ver López 2005.

2.4.2 Series en nivel semanal

Cuadro 2.4.2 Matriz de correlaciones a nivel semanal

	C1	C2	C3	C4	C5	RM
C1	1.00	0.50	0.58	0.39	0.21	0.57
C2	0.50	1.00	0.65	0.22	0.02	0.67
C3	0.58	0.65	1.00	0.23	0.10	0.62
C4	0.39	0.22	0.23	1.00	0.01	0.40
C5	0.21	0.02	0.10	0.01	1.00	0.19
RM	0.57	0.67	0.62	0.40	0.19	1.00

Fuente: Cálculo basado en los datos de AUPOL

En nivel semanal las correlaciones contemporáneas más fuertes son:

El C1, con C3 y RM con valores de 0.58 y 0.57 respectivamente. El C2, con C3 y RM tomando valores de 0.65 y 0.67 respectivamente. El C3 con C1, C2 y RM. Tanto El C4 como el C5 no presentan correlaciones significativas con ninguno de los cuadrantes. Los cuadrantes con sus variables rezagadas tampoco son muy significativos las únicas relaciones que se pueden destacar son la que presentan el C1 y el C2 con sus valores rezagados en un período estos valores son de 0.51 y 0.60 respectivamente. De lo anterior podemos concluir que en el nivel semanal los cuadrantes de Santiago tienden a comportarse de manera más parecida a como lo hace RM, salvo el C5 que sigue manteniendo una baja correlación con RM.

2.4.3 Series en nivel mensual

Cuadro 2.4.3 Matriz de correlaciones a nivel mensual

	C1	C2	C3	C4	C5	RM
C1	1.00	0.54	0.66	0.52	0.36	0.65
C2	0.54	1.00	0.74	0.20	0.06	0.69
C3	0.66	0.74	1.00	0.21	0.23	0.76
C4	0.52	0.20	0.21	1.00	-0.05	0.41
C5	0.36	0.06	0.23	-0.05	1.00	0.37
RM	0.65	0.69	0.76	0.41	0.37	1.00

Fuente: Cálculo basado en los datos de AUPOL

Cuando se examina la cantidad de delitos en nivel mensual, mejoran las correlaciones entre algunos cuadrantes. Del cuadro 2.4.3 se observa que el C1 se correlaciona con C2, C3, C4 y RM, siendo las correlaciones de 0.54; 0.66 ; 0.52 y

0.65 respectivamente. El C2 presenta una fuerte correlación positiva con el C3 y RM alcanzando un valor de 0.74 y 0.69 respectivamente y una correlación algo menor con el C1 de 0.54. El C3 también muestra patrones de correlación altos con C1, C2 y RM los valores de estas son de 0.66 ; 0.74 y 0.76 respectivamente. El C4 sólo se correlaciona con el C1 alcanzando un valor de 0.52. C5 no presentan correlaciones altas con ninguno de los cuadrantes.

Con las variables rezagadas el C1 mantiene correlaciones con C2(-12) y C3(-12) tomando valores de 0.76 y 0.68 respectivamente. El C2 por su parte se correlaciona con C2(-1), C1(-10) y C1(-11) siendo los valores de 0.54 ; -0.64 ; -0.51. El C3 se relaciona con el RM(-1), C2(-12) y RM(-12) tomando los valores de 0.50 ; 0.56 y 0.58 respectivamente. El C4 solo presenta una correlación negativa con C4(-10) igual a -0.51. El C5 no presenta correlaciones superiores a 0.5 con ninguna de las variables rezagadas⁸.

En síntesis se puede concluir que a medida que se van agregando más los niveles temporales de las series, el comportamiento de cada una de ellas tiende a parecerse aún más a la conducta que presenta la serie que describe a la región metropolitana, lo que daría indicios de que al tomar ciertas comunas de la región como muestra representativa, se podría replicar el comportamiento aproximado de la región si es que el nivel temporal de las series es mensual o anual.

Esto puede reducir el costo de los estudios, bastaría considerar sólo una muestra apropiada para establecer el comportamiento del agregado de la región metropolitana.

2.5 Pruebas de estacionariedad

Antes de poder especificar un sistema de ecuaciones como un modelo VAR, es necesario verificar la estacionariedad de las series, o en su defecto detectar si es que existen vectores de cointegración que hagan que el sistema sea estacionario. Según Box y Jenkins (1970), esta es una condición altamente necesaria que deben cumplir las series de tiempo para que sean susceptibles de ser modeladas a través de alguna ecuación, por lo tanto es una primera exigencia, para luego poder realizar pronósticos a partir de las series.

⁸ Ver anexo

La estacionariedad de las series, asegura que sus características permanecen en el tiempo, tendiéndose a comportar de forma acotada, es decir, las series no divergen a medida que transcurre el tiempo, siendo posible encontrar valores medios en los que esta fluctúa. Esto significa que tanto la varianza como la media de la serie no dependen del tiempo (estacionariedad débil). Cuando se cumple la condición de estacionariedad, se asegura que la serie no se escapa de un cierto rango, además se evita el problema de la regresión espúrea⁹.

Una vez verificada la estacionariedad se realiza un estudio acerca de sí la serie presenta patrones definidos que caracterizan sus ciclos. Si es que se da lo anterior, es posible proyectar con una cierta seguridad el comportamiento futuro de la serie.

Cuadro 2.5.1 Resultados pruebas de raíz unitaria series diarias

Cuadrante	ADF Test Statistic	1% Critical Value*
c1	-16.87	-3.44
c2	-19.13	-3.97
c3	-30.60	-3.97
c4	-13.35	-3.97
c5	-27.60	-3.44
rm	-13.30	-3.97

Fuente: Cálculo basado en los datos de AUPOL

Para verificar la estacionariedad de cada una de las series se utilizó el test de Dickey- Fuller aumentado¹⁰. Los resultados que se presentan en las siguientes tablas, muestran que todas las series que describen a los cuadrantes son estacionarias, debido a que se rechaza la hipótesis nula de raíz unitaria al 1% de significancia en todos los casos¹¹.

⁹ Este problema se origina cuando dos series que no tienen relación alguna, se relacionan por existir alguna tendencia en común en el período.

¹⁰ Existen múltiples formas alternativas de comprobar la estacionariedad de la series, sin embargo la prueba de dickey fuller aumentada es la más usada.

¹¹ La hipótesis nula de raíz unitaria se rechaza cuando el valor absoluto del ADF Test Statistic es mayor que el valor absoluto del valor crítico a un nivel de significancia dado. En el caso de nivel diario se rechaza la no estacionariedad al 1% para todas las series. Las series en nivel diario son estacionarias en tendencia. Si el lector quiere consultar como se construye el test se recomienda leer “Métodos econométricos” Johnston y Dinardo (1997).

Cuadro 2.5.2 Resultados pruebas de raíz unitaria series semanales

Cuadrante	ADF Test Statistic	1% Critical Value*
c1	-7.59	-3.47
c2	-3.56	-3.47
c3	-5.50	-3.47
c4	-3.94	-3.47
c5	-6.73	-3.47
rm	-4.65	-3.47

Fuente: Cálculo basado en los datos de AUPOL

En nivel semanal todas las series son estacionarias al 1% de nivel de significancia.

Cuadro 2.5.1 Resultados pruebas de raíz unitaria series mensuales

Cuadrante	ADF Test Statistic	5% Critical Value
c1	-3.18	-2.94
c2	-3.58	-2.93
c3	-4.32	-3.52
c4	-3.58	-2.93
c5	-5.14	-2.93
Rm	-4.52	-3.52

Fuente: Cálculo basado en los datos de AUPOL

A nivel mensual se encontró que las series no eran estacionarias a un nivel de significancia del 1%, sin embargo, se alcanza la estacionariedad para todas las series cuando se acepta un 5% de nivel de significancia (ver cuadro 2.5.3).

2.6 Estudio de Tendencias

Para analizar si las series presentan tendencias, se realizaron regresiones tomando como variable dependiente las denuncias de delitos en cada cuadrante que se quería examinar regresionándolo contra una constante y variables de tendencia, la estimación se hizo a través de mínimos cuadrados ordinarios, sin incluir más variables explicativas en el modelo. La construcción de la tendencia varía según el nivel temporal de las series.

En general la ecuación utilizada para analizar la significancia de las tendencias es:

$$C_{i,t} = \alpha + \beta_1 t + \beta_2 t^2 + \omega_t$$

Donde:

- $C_{i,t}$ = Cantidad de delitos en el cuadrante i en el nivel temporal t
- α = Constante

- B_i = Coeficiente asociado a las variables de tendencia
- ω_t = Perturbación aleatoria
- t y t_1 = Tendencias (la construcción de cada una de ellas depende del nivel temporal del cual se trate)
- t^2 = tendencia t al cuadrado

2.6.1 Series en nivel diario

Cuadro 2.6.1 (Significancia de tendencias diarias y ajustes)

	Constante	T	T1	T ²	R ² ajustado
RM	877.55	69.61	-8.26	-6.68	0.18
C1	28.11	9.28	-0.22	-1.37	0.20
C2	44.12	27.07	-0.32	-3.99	0.39
C3	19.66	6.94	-0.08	-1.01	0.16
C4	9.83	1.96	-0.09	-0.29	0.07
C5	12.31	-6.72	-0.01	1.26	0.53

Fuente: Cálculo basado en los datos de AUPOL

Para efectos de cálculos de la tendencia se consideró la serie t que es un incremento en una unidad a medida que transcurre la semana, esto quiere decir que todos los lunes de la serie van acompañados de un 1 los martes de un 2 y los domingos de un 7. También se incluyó la tendencia t_1 que captura si es que existe una tendencia de tipo mensual, esta se construyó asignando un 1 al primer día del mes, un 2 al segundo día y 30 o 31 al último, según corresponda. Al evaluar la significancia de la tendencia t y de t_1 en los cuadrantes de Santiago y RM se obtuvieron los siguientes resultados.

Para el C1 las tendencias t , t^2 y t_1 son significativas al 1% lográndose un ajuste del 20% al incluir solo una constante y estas tres variables de tendencias (este 20% significa que con sólo las variables de tendencia y una constante se puede explicar el 20% del comportamiento de la serie) ver cuadro 2.6.1. Para el C2 las tendencias t , t^2 y t_1 también son significativas alcanzándose un ajuste del 39%. En el C3 también son significativas dichas tendencias alcanzando un ajuste del 16%. Lo mismo para el C4 pero sólo se alcanza un ajuste del 7%. En el C5 t y t^2 son significativas obteniéndose un ajuste del 53%, sin embargo t_1 no es significativa para el C5.

2.6.2 Series en nivel semanal

A nivel semanal las series no presentan tendencias significativas.

2.6.3 Series en nivel mensual

Cuadro 2.6.1 (Significancia de tendencias mensuales)

	C1		C2		C3		C4		C5	
	coef	p-value	coef	p-value	coef	p-value	coef	p-value	coef	p-value
Constante	960.4	0	1657.1	0	597.53	0	308.55	0	267.7	0
T	6.87	0.85	78.54	0.18	39	0.06	-4.97	0.63	10.58	0.36
T ²	-0.2	0.94	-5.6	0.21	-2.53	0.11	0.35	0.66	-0.1	0.91
T1	2.46	0.3	8.81	0.03	3.69	0.01	1.2	0.09	-0.45	0.55

Fuente: Cálculo basado en los datos de AUPOL

La construcción de las variables temporales en el caso mensual son. Una serie t que lleva un 1 cuando se trata del mes de enero, un 2 cuando es febrero y 12 si es diciembre; la otra variable de tendencia que se estudia es la serie t1 que toma valores de 1 a 42 representando los 42 meses de los cuales se compone la serie. Las series a nivel mensual para los cuadrantes de la comuna de Santiago no presentan tendencias, rechazándose la hipótesis al 10% de significancia en todos los casos, salvo en el C3 en donde es significativa al 6%.

Se probó también si es que existía alguna tendencia para la serie completa desde enero de 2001 hasta junio de 2004, rechazándose la hipótesis nula para los cuadrantes 1 y 5. Para el resto de los cuadrantes, se manifiesta una tendencia positiva a medida que aumenta el tiempo.

2.7 Tests de igualdad de medias

En general, se obtiene que no existe igualdad de medias cuando se consideran por separado las series, esto quiere decir que, estadísticamente las medias de denuncias de 2001, 2002 y 2003 no son iguales conjuntamente para todos los cuadrantes de la comuna de Santiago y RM a un nivel de significancia del 5%. Sólo hay evidencia estadística de igualdad de medias en las series para C2, C4, C5 Y

RM; en el período 2001-2002, y en el C5 para el período 2002-2003 (ver cuadro 2.7.1). Esto quiere decir por ejemplo, que para el C2 la media de la cantidad de delitos en el período 2001 es estadísticamente igual a la que se observa en el 2002 en dicho cuadrante, en cambio para el C1 no ocurre lo mismo, no existe evidencia estadística que avale la igualdad de medias entre el período 2001 y el 2002.

Cuadro 2.7.1 (Resultados test de medias)

	2001-2002-2003	entre 2001-2002	entre 2002-2003
C1	no	no	no
C2	no	si	no
C3	no	no	no
C4	no	si	no
C5	no	si	si
RM	no	si	no

Fuente: Cálculo basado en los datos de AUPOL

2.8 Conclusiones

Las series difieren en las características según el nivel temporal que se esté considerando, a continuación se hará un resumen y se derivaran conclusiones a partir del estudio previo para cada nivel temporal.

2.8.1 Series en nivel diario

Las estacionalidades se presentan principalmente en las series de tipo diario, en el cuadro 2.8.1 se presentan la cantidad diaria promedio de delitos que ocurren en los distintos cuadrantes de la comuna de Santiago y región metropolitana, es posible observar que los delitos van en aumento a medida que transcurre la semana los días miércoles y jueves son casi iguales, luego ocurre una nueva alza los días viernes para declinar progresivamente los sábados y domingos, exceptuando el C5 que presenta un comportamiento contrario al descrito anteriormente.

Cuadro 2.8.1 Promedio de delitos diario por día de la semana

	lun	mar	mié	jue	vie	sáb	dom
C1	28.6	31.2	32.2	32.0	35.8	25.6	18.8
C2	54.9	63.9	65.2	65.3	74.8	49.0	24.3
C3	21.6	23.6	23.9	23.3	26.6	21.2	13.6
C4	9.5	9.0	10.2	10.0	10.5	9.0	6.9
C5	5.3	3.8	4.1	4.1	4.4	19.3	18.7
RM	743.7	768.8	786.2	790.5	859.8	879.4	767.2

Fuente: Cálculo basado en los datos de AUPOL¹²

El cuadro 2.8.2 muestra el porcentaje promedio de delitos ocurrido por cuadrante según el día de la semana. Este puede ser un primer indicio de como pueden ser asignados los efectivos policiales en la comuna de Santiago¹³, claramente de este cuadro se puede deducir que no es correcto asignar una cantidad igual de policías para cada uno de los cuadrantes, tampoco, la asignación debiera ser la misma en los distintos días de la semana. Por ejemplo, si la cantidad total de policías que dispone la comuna de santiago es de 1000 para cada día de la semana, entonces la asignación correcta de ellos si consideramos un día lunes:

C1=239 ; C2=460 ; C3=180 ; C4=79; C5=44

En cambio en un día domingo la asignación correcta debiera ser:

C1=228 ; C2=295 ; C3=165 ; C4=84 ; C5=227

Cuadro 2.8.2 porcentaje promedio de ocurrencia de delitos

	LUN	MAR	MIÉ	JUE	VIE	SÁB	DOM
C1	23.85	23.73	23.75	23.76	23.54	20.63	22.84
C2	45.79	48.59	48.08	48.48	49.18	39.48	29.53
C3	18.02	17.95	17.63	17.30	17.49	17.08	16.52
C4	7.92	6.84	7.52	7.42	6.90	7.25	8.38
C5	4.42	2.89	3.02	3.04	2.89	15.55	22.72
TOTAL	100	100	100	100	100	100	100

Fuente: Cálculo basado en los datos de AUPOL

Como se vio anteriormente en la sección 2.4.1 El C2 se correlaciona positivamente con el C1 y C3, esto es una señal de que la cantidad de delitos tiende a moverse en la misma dirección en estos cuadrantes. En cambio se correlaciona negativamente con el C5, esto quiere decir que cuando aumenta la cantidad de delitos en el C2 se reducen en el C5. Este fenómeno es útil, para una apropiada asignación de

¹² Cuadro construido con información de 2001 a 2003

¹³ Se está asumiendo que la tasa de denuncia es bien similar para todos los cuadrantes.

policías, esta información conjuntamente con la del cuadro 2.8.2 permite inferir que los días de semana lunes a viernes es conveniente desplazar carabineros del C5 al C2 de manera creciente, es decir, el viernes más que el martes y el martes más que el lunes; en cambio el fin de semana, sería pertinente una reasignación desde el C2 al C5. El resto de los cuadrantes no presenta correlaciones fuerte con los demás.

En cuanto a las tendencias observamos que para los cuadrantes C1, C2, C3 y C4, t_1 es negativa¹⁴, esto refleja que a medida que aumenta el día del mes los delitos van en descenso, sin embargo, los coeficientes son bien bajos lo que da cuenta de que este descenso es bien suave. El coeficiente asociado a t_1 representa la cantidad promedio de descenso de delitos cuando se pasa de un día a otro.

2.8.2 Series en nivel semanal

Las series en nivel semanal no presentan características especiales en el tiempo, es decir, no existen diferencias significativas en las distintas semanas del mes, no presentan tendencias ni patrones claros. Como esto es una limitación para la modelación de las series semanales a través de la metodología de series temporales, en el capítulo 3 se sugerirá una forma alternativa de modelar el VAR semanal.

2.8.3 Series en nivel mensual

Las series en nivel mensual presentan características distintas para los distintos cuadrantes. En los cuadrantes 1 y 2 se observa un declive de los delitos en los meses de enero y febrero, un aumento fuerte a partir de marzo que se mantiene hasta los primeros días de mayo, luego retoma nuevamente una tendencia a la baja de los delitos hasta septiembre y en octubre recobra una tendencia al alza hasta finalizar el año. Este patrón puede ser capturado a través de variables dummies en los sistemas de ecuaciones para verificar si estas diferencias y estacionalidades encontradas a priori tienen significancia estadística. Esto se analiza en el siguiente capítulo.

¹⁴ Ver anexo al capítulo 2 sección 2.5

Capítulo 3

Modelos de Vectores Auto regresivos (VAR) y Sistemas de ecuaciones

El objetivo de este capítulo, es modelar matemáticamente el comportamiento de las distribuciones de delitos para los distintos cuadrantes de la comuna de Santiago, con la finalidad ulterior de proyectar sobre la base de los mejores modelos obtenidos, la cantidad de delitos esperada para cada cuadrante en una fecha determinada, de modo de suministrar esta información para realizar asignaciones más eficientes de los recursos policiales, en los distintos cuadrantes de la comuna de Santiago.

En esta sección nos centraremos en modelos VAR y sistemas de ecuaciones, discutiendo principalmente los siguientes tópicos:

- i) Justificación estadística-econométrica para la utilización de modelos VAR
- ii) Metodología empleada y justificaciones teóricas
- iii) Especificación del VAR diario, semanal y mensual.
- iv) Resultados y ajustes.
- v) Sistemas de ecuaciones: Especificación y resultados

La especificación de los VAR y de los sistemas de ecuaciones en los tres niveles temporales diario, semanal y mensual; responde a las principales características detectadas en el capítulo anterior de este seminario.

3.1 Justificación estadística-econométrica para la utilización de modelos VAR

Del análisis gráfico y estadístico, fue posible detectar que la distribución de delitos se comportaba de manera distinta en cada uno de los cuadrantes de la comuna de Santiago presentando fuertes estacionalidades y similitudes en algunos casos mientras que en otros los patrones diferían del resto.

Los modelos VAR capturan aquellas características que son comunes a todas las series, modelándolas como un sistema de ecuaciones. Por lo tanto, cada ecuación que se incluye en el VAR es una forma representativa del conjunto en general, sin aislar o especificar detalles de cada una de las series en las ecuaciones, para incluir características particulares de las series el procedimiento apropiado es una estimación por sistema de ecuaciones. Otra forma, es estudiar el comportamiento por separado de cada una de las series, empleando metodologías como el ARIMA.

Dado que existe una aparente interrelación en los delitos ocurridos en los distintos cuadrantes de la comuna de Santiago, es lógico asumir una especificación en la cual, cada serie de tiempo, representativa de cada uno de los cuadrantes, dependa además de sus valores rezagados, de lo que ocurre en los otros cuadrantes. Esta relación puede tener diversas explicaciones, por ejemplo aumentos en la dotación de carabineros en un cuadrante pueden haber desplazado a los delincuentes hacia otros lugares, o la mayor existencia de comercio en algún día u hora en un cuadrante, puede movilizar a los delincuentes no tan sólo a ese sector, sino que, también a los cuadrantes más próximos, entre otras posibilidades. Sin embargo, como no se cuenta con ese tipo de información para usarlos como variables explicativas de los modelos, no interesa cuales son las causas de éstas posibles interrelaciones que pueden existir entre los cuadrantes, si no que, se procederá a testear si es que existen relaciones de causalidad estadísticamente significativas y se observará como se comportan a través del tiempo. Dada esta carencia de variables explicatorias adicionales resulta útil plantear un modelo del tipo vectores auto regresivos¹⁵.

Por otro lado, como se vio en el apartado anterior, las series de delitos presentan un comportamiento distinto según sea el día de la semana o la hora del día, presentando estacionalidad y tendencias, esto favorece la utilización de modelos VAR debido a que el tiempo pasa a ser un buen descriptor del comportamiento de la distribución.

¹⁵ Si el lector no está familiarizado con los modelos VAR, se recomienda leer “métodos de econometría” Johnston-Dinardo Cap. 9

3.2 Metodología empleada y justificaciones teóricas

La metodología de predicción a través de modelos VAR, es altamente utilizada porque en general, las proyecciones que se obtienen a través de este método son bastante buenas en comparación con el uso de formas alternativas, además, el proceso de estimación en la mayoría de los casos es altamente sencillo no se requiere de técnicas muy sofisticadas, muchas veces basta con el procedimiento de mínimos cuadrados ordinarios. La metodología VAR no necesita de información adicional que explique el comportamiento de las series, es decir, no se necesita de una construcción teórica que explique el porqué van ocurriendo cambios en las distintas series, con esta metodología el análisis se centra en estudiar el comportamiento mismo de las series a través del tiempo en donde la variable dependiente es explicada por sus valores pasados y también por los rezagos de otras series que están relacionadas, y algún término de error o innovación. Por lo tanto, el estudio es puramente estocástico, no dependiendo de otros factores endógenos, es decir, solo basta con la información que entrega cada una de las series.

3.2.1 Metodología

La representación matemática del VAR viene dada por:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + Bx_t + \varepsilon_t$$

Donde:

- y_t : Es un vector de $k \times 1$ variables endógenas
- A_1, \dots, A_p, B : Son matrices de coeficientes a estimar de dimensión $k \times k$
- ε_t : Es un vector de innovaciones, que presenta correlaciones contemporáneas, pero que no está correlacionado con sus valores rezagados, y tampoco presenta correlación con las variables explicativas (las que están al lado derecho de la igualdad)

Estimación del VAR:

Como los modelos construidos para los distintos niveles son estables, es decir, todas las raíces son menores que 1. La estimación se hará por mínimos cuadrados ordinarios para el sistema completo.

3.3.1 Modelo VAR en nivel diario

Para realizar la especificación del VAR, se considerarán las características que son comunes a la mayor parte de las series, por ejemplo correlaciones con el rezago t , presencia de tendencias, etc. Estos elementos serán los que determinarán la configuración final del VAR, entre otras. De la inspección gráfica, estudio de correlaciones simples, tendencias y estacionalidades fue posible detectar que las series en nivel diario presentan en general, un aumento en los primeros días de la semana, a partir de lunes hasta llegar a un máximo el día viernes para luego declinar el día domingo, este fenómeno será capturado a través de variables dummies días. Las correlaciones de las variables endógenas con los rezagos de las mismas en general son malas, por lo tanto la extensión del VAR se limita a un período rezagado, esto se determinó a través de la minimización de los criterios de información de **akaike** y **schwartz**.

Como existían algunos rezagos de ciertos cuadrantes que eran significativos para algunas ecuaciones del var y no para todas, se incluyeron estas variables como exógenas, previo a esto se verificó si es que existía causalidad a la Granger entre los rezagos y las variables dependientes. Todos los rezagos incluidos como variables exógenas presentaban causalidad a la granger al 5% de significancia con los 5 cuadrantes salvo el rezago C4(-7) que sólo presentaba causalidad con el C5.

El mejor modelo que encontramos para nivel diario tiene la siguiente especificación, los criterios utilizados para la elección de este modelo fueron principalmente el r cuadrado ajustado sumado a los criterios de **akaike** y **schwartz**.

Especificación del VAR diario

A continuación se describen las ecuaciones para cada uno de los cuadrantes. En lado izquierdo de la igualdad está la variable dependiente C_{it} , al lado derecho el primer término representa la sumatoria del primer rezago de cada uno de los cuadrantes, el segundo término es la sumatoria de las variables ficticias (Dummies) días, el tercer término captura todas los rezagos que fueron incluidos como variables exógenas, el cuarto término refleja la incorporación de una tendencia¹⁶ y finalmente el término de perturbación:

Var diario

$$C_{1,t} = a_1 + \sum_{i=1}^5 \alpha_{1i} C_{i,t-1} + \sum_i^6 \beta_{1i} D_i + \sum_i \xi_{1i} R + \delta_{1i} t + v_1 \quad (1)$$

$$C_{2,t} = a_2 + \sum_{i=1}^5 \alpha_{2i} C_{i,t-1} + \sum_i^6 \beta_{2i} D_i + \sum_i \xi_{2i} R + \delta_{2i} t + v_2 \quad (2)$$

....

$$C_{5,t} = a_5 + \sum_{i=1}^5 \alpha_{5i} C_{i,t-1} + \sum_i^6 \beta_{5i} D_i + \sum_i \xi_{5i} R + \delta_{5i} t + v_5 \quad (5)$$

Donde:

- $C_{i,t}$ = cantidad de delitos ocurrida en el cuadrante i en el día t
- $RM_{t-1}; C_{2,t-7}; C_{4,t-7}; C_{5,t-7}; C_{2,t-365}; C_{4,t-365}; RM_{t-365}$ representan rezagos de las series
- D_i dummies días
- t variable de tendencia

¹⁶ La variable t toma el valor de 1 a 31 según el día del mes que corresponda.

Determinación de estabilidad del VAR: Todas las raíces permanecen dentro del círculo unitario, es decir, tienen un módulo menor que uno, esto garantiza que el VAR es estable, y por lo tanto se pueden aceptar las conclusiones que se obtengan al aplicar la función de impulso respuesta.

El cuadro 3.3.1 muestra las estimaciones para cada una de las ecuaciones del VAR, la lectura del cuadro es la siguiente. Cada columna representa la variable dependiente que es explicada por todas las variables que están en las filas, siendo los valores de las celdas los estimadores de los parámetros.

Resultados

Cuadro 3.3.1 Resultados estimación VAR diario

	C1	C2	C3	C4	C5
C1(-1)	0,32	0,07	0,07	0,04	0,03
C2(-1)	0,07	0,32	0,07	-0,01	0,02
C3(-1)	0,09	0,03	0,06	0,00	0,03
C4(-1)	0,06	-0,15	0,07	0,12	-0,01
C5(-1)	0,17	0,02	0,07	-0,01	0,18
C	-5,23	3,37	-2,12	3,77	11,02
RM(-1)	-0,02	-0,01	0,00	0,00	0,00
C2(-7)	0,03	0,21	0,05	0,01	0,01
C4(-7)	0,16	-0,01	-0,07	0,08	-0,02
C5(-7)	0,08	0,03	0,04	-0,03	0,06
DD1	20,63	44,03	13,24	3,47	-17,67
DD2	17,39	35,11	11,20	2,05	-13,53
DD3	14,71	29,71	11,13	3,40	-13,42
DD4	13,84	28,49	9,71	3,07	-13,27
DD5	16,56	36,66	12,59	2,87	-12,47
DD6	0,35	6,28	4,78	1,58	4,46
T1	-0,13	-0,21	-0,01	-0,06	0,03
C2(-365)	0,09	0,09	0,03	0,01	-0,03
C4(-365)	-0,09	-0,16	-0,18	0,06	-0,02
C5(-365)	-0,01	-0,19	-0,04	-0,01	0,10
RM(-365)	0,02	0,02	0,01	0,01	0,01
R-squared	0,45	0,58	0,31	0,19	0,68
Adj. R-squared	0,44	0,57	0,30	0,17	0,67
Sum sq. resids	101125,10	294968,40	69014,16	22968,32	35582,45
S.E. equation	10,65	18,19	8,80	5,08	6,32
F-statistic	36,37	61,45	20,30	10,31	93,44
Log likelihood	-3441,14	-3929,29	-3266,92	-2765,23	-2964,84
Akaike AIC	7,59	8,66	7,21	6,11	6,55
Schwarz SC	7,70	8,77	7,32	6,22	6,66
Mean dependent	34,05	68,86	26,21	10,69	10,08
S.D. dependent	14,20	27,76	10,50	5,57	11,00

El ajuste de la ecuación por cuadrante viene dada por el R- cuadrado ajustado vemos que los cuadrantes que presentan un mejor ajuste son el C5, C2 seguido del C1 con un porcentaje de ajuste del 67%, 57% y 44% respectivamente, esto quiere decir que el sistema de ecuaciones especificado describe un 67% del comportamiento del C5.

La interpretación de los parámetros estimados es la siguiente. Al existir un cambio de una unidad en la variable explicativa ocurre un cambio de X unidades en la variable dependiente siendo X el valor del parámetro, por ejemplo, el coeficiente del $C_{1,t-1}$, para el C1 es de 0.32, lo que implica que un aumento de un delito en el $C_{1,t-1}$ produce aumentos de 0.32 delitos en C1.

Por otra parte se vio que existe una clara correlación entre variables contemporáneas, pero como el objetivo de este tipo de modelación es realizar pronósticos, no se pueden incluir variables contemporáneas en la explicación del modelo. Sin embargo, si es que se tiene un buen modelo que ajuste al predictor contemporáneo se puede realizar una buena proyección al incorporar la variable contemporánea como regresor del modelo.

Sólo como referencia, se mencionaran los R –cuadrados ajustados obtenidos al incorporar RM como variable exógena dentro del modelo var. El cuadro 3 compara los ajustes obtenidos del modelo que no contiene ninguna variable contemporánea y el que incluye a RM en su forma contemporánea, se ve que al incluir esta última mejora los ajustes para todos los cuadrantes, explicando ahora el sistema alrededor de 12 puntos porcentuales más.

Cuadro 3.3.1a

	c1	c2	c3	c4	C5
Mod con rezagos					
Adj. R-squared	0.44	0.57	0.30	0.17	0.67
Incluye contemporanea					
var					
Adj. R-squared	0.59	0.73	0.46	0.27	0.66

Fuente: Cálculo basado en los datos de AUPOL

3.3.2 Modelo VAR en nivel semanal

Para el VAR semanal fue difícil encontrar una especificación adecuada que describiera bien el comportamiento de los cuadrantes, como se mencionó en el capítulo anterior las series en nivel semanal no presentan patrones distinguibles dentro del periodo, las hipótesis de tendencia fueron rechazadas y no hay evidencia de comportamientos distintos a lo largo de las semanas que justifique la inclusión de variables dummies. Los ajustes de los modelos semanales fueron bastante bajos alcanzándose como máximo un 51% para el C2 que fue el que presentó un mejor ajuste versus un 17% para el C5 que mostró el peor ajuste en el sistema (el cálculo se hizo considerando un VAR con un rezago de 10 períodos)

Para obtener buenos pronósticos en nivel semanal será necesario ajustar un modelo que considere alguna variable contemporánea. Del estudio de correlaciones simples realizado en el capítulo anterior vimos que RM presentaba una alta correlación con casi todos los cuadrantes. Por lo tanto, el camino a seguir será ajustar un VAR que considere dentro de las variables exógenas a RM. Y Posteriormente hacer un pronóstico de RM en nivel semanal con alguna otra metodología distinta al var.

Especificando el VAR semanal

$C_{i,t}$ = cantidad de delitos ocurrida en el cuadrante i en la semana t

Con el objeto de alcanzar mejores ajustes en el sistema de ecuaciones VAR en nivel semanal, se hizo una transformación logarítmica de las series, entonces se llamará $c_{i,t} = \log(C_{i,t})$. La transformación de las series a logaritmo no invalida las conclusiones obtenidas en un principio, es decir, no se pierde la estacionariedad ni

los patrones que tiene cada una de las series, porque la transformación logarítmica es una transformación monótona de las series.

A continuación se describen las ecuaciones para cada uno de los cuadrantes. En el lado izquierdo de la igualdad está la variable dependiente $\log(C_{it})$, al lado derecho el primer término representa la sumatoria del primer rezago de cada uno de los cuadrantes, el segundo y tercer término es la sumatoria de las variables ficticias (Dummies) semanas y mes respectivamente, el cuarto término captura todas los rezagos que fueron incluidos como variables exógenas, el quinto término representa la variable de tendencia

VAR semanal

Entonces el sistema queda especificado como:

$$c_{1,t} = a_1 + \sum_{i=1}^5 \alpha_{1i} c_{i,t-1} + \sum_i^4 \beta_{1i} Ds_i + \sum_i^{11} \lambda_i Dm_i + \sum_i \xi_{1i} R + \delta_{1i} t + \nu_1 \quad (1)$$

$$c_{2,t} = a_2 + \sum_{i=1}^5 \alpha_{2i} c_{i,t-1} + \sum_i^4 \beta_{2i} Ds_i + \sum_i^{11} \lambda_{2i} Dm_i + \sum_i \xi_{2i} R + \delta_{2i} t + \nu_2 \quad (2)$$

....

....

$$c_{5,t} = a_5 + \sum_{i=1}^5 \alpha_{5i} c_{i,t-1} + \sum_i^4 \beta_{5i} Ds_i + \sum_i^{11} \lambda_{5i} Dm_i + \sum_i \xi_{5i} R + \delta_{5i} t + \nu_5 \quad (5)$$

Donde:

- C_{it} : cantidad de delitos en el cuadrante i, en la semana t
- c_{it} : $\log(C_{it})$
- Ds_i : Dummies semana i^{17} i= 1,2,...,5
- Dm_i : Dummies mes i i= 1, 2,...,12

¹⁷ En la estimación se excluye una de las dummies para que el sistema sea l.i y por lo tanto invertible.

- Variables exógenas incluidas: Ds_i , Dm_i , $\log(rm)$, $\log(rm_{t-11})$, $\log(rm_{t-15})$, $\log(rm_{t-16})$, $\log(rm_{t-24})$
- v_i : término de error

Finalmente se estimó un VAR con tres rezagos incorporando a RM en su forma contemporánea, alcanzándose ajustes de 55%, 74%, 58%, 44% y 10% para el C1,C2,C3,C4 Y C5 respectivamente Ver cuadro (3.3.2).

La interpretación de los parámetros es la siguiente. Al existir un cambio porcentual de un punto en la variable explicativa ocurre un cambio de X% en la variable dependiente siendo X el valor del parámetro, por ejemplo, el coeficiente del $\text{Log}(C_{1,t-2})$ para el $\text{Log}(C_2)$ es de -0.10, lo que implica que un aumento en un punto porcentual en $C_{1,t-2}$ produce una disminución de 0.10 puntos porcentuales en C_2 .

Resultados

Cuadro 3.3.2 Resultados estimación VAR semanal

	LOGC1	LOGC2	LOGC3	LOGC4	LOGC5
LOGC1(-1)	0.272*	-0.02	-0.07	0.106	0.239
LOGC1(-2)	-0.03	-0.10	-0.05	0.184	-0.17
LOGC1(-3)	0.123	-0.02	0.083	0.118	-0.03
LOGC2(-1)	-0.00	0.197*	0.048	0.059	0.005
LOGC2(-2)	-0.08	0.109	-0.13	-0.03	0.013
LOGC2(-3)	-0.13	0.225*	0.087	-0.09	-0.11
LOGC3(-1)	0.018	0.067	0.331*	-0.17	0.107
LOGC3(-2)	-0.03	-0.13	0.107	-0.27	0.125
LOGC3(-3)	-0.00	-0.07	-0.07	-0.16	-0.04
LOGC4(-1)	0.042	0.008	-0.05	0.098	-0.08
LOGC4(-2)	0.169*	0.031	0.056	0.043	-0.02
LOGC4(-3)	0.000	-0.02	0.041	0.168*	-0.03
LOGC5(-1)	0.041	-0.06*	0.011	-0.10	0.167
LOGC5(-2)	-0.00	-0.04	-0.05	0.040	0.169
LOGC5(-3)	0.021	-0.04	0.034	-0.11	0.035
C	-3.41	-0.17	0.321	-10.3*	-2.39
LOGRM(-16)	-0.36*	-0.29*	-0.58*	-0.00	-0.31
T	0.000	0.000	0.000*	-0.00	-0.00
LOGRM(-11)	-0.13	-0.18	-0.26*	0.325	0.100
LOGRM(-24)	0.009	-0.15	0.004	0.505*	0.109
DM8	-0.05	-0.12*	-0.01	-0.04	0.016
DM9	-0.16*	-0.10*	-0.03	0.091	-0.06
LOGRM(-15)	0.438*	0.259*	0.552*	-0.08	0.250
LOGRM	0.862*	0.959*	0.645*	1.025*	0.465*
DS1	0.063	0.083*	0.018	0.040	-0.04

DS3	0.050	0.072*	0.053	0.001	0.038
DS4	0.078	0.052	0.041	0.016	0.063
R-squared	0.62	0.78	0.65	0.53	0.25
Adj. R-squared	0.55	0.74	0.58	0.44	0.10
Sum sq. resid	2.94	1.39	1.91	4.96	10.57
S.E. equation	0.15	0.10	0.12	0.19	0.28
F-statistic	8.37	18.28	9.54	5.83	1.68
Log likelihood	91.74	151.19	125.94	49.99	-10.13
Akaike AIC	-0.81	-1.56	-1.24	-0.29	0.47
Schwarz SC	-0.29	-1.04	-0.72	0.23	0.99
Mean dependent	5.45	6.13	5.18	4.27	4.25
S.D. dependent	0.22	0.20	0.19	0.26	0.30

*parámetros significativos al 5%

El VAR semanal no cumple con las condiciones de estabilidad porque al menos una de sus raíces está fuera del círculo unitario lo que invalida el uso de los parámetros estimados previamente.

3.3.3 Modelo VAR en nivel mensual

El principal problema que se tiene al intentar modelar un sistema de ecuaciones VAR en nivel mensual es que el tamaño de la muestra con la que se cuenta es muy reducido como para especificar un sistema que dependa de muchas variables, debido a que al aumentar la cantidad de regresores se pierden grados de libertad. Por otro lado, no se pueden considerar como variables explicativas rezagos de las series que estén muy distantes en el tiempo por que se pierden datos para el ajuste y como la serie es corta, sólo se cuenta con 42 observaciones (Ene de 2001 a Jun de 2004) no es posible especificar el VAR ni con muchas variables ni con rezagos que disten demasiado del nivel contemporáneo.

Dadas estas limitaciones se tuvo que ajustar un VAR que considera dentro de sus variables explicativas a RM en su forma contemporánea, similar a lo que se hizo en nivel semanal pero por causas distintas.

Especificación del VAR mensual

$$C_{1,t} = a_1 + \sum_{i=1}^5 \alpha_{1i} C_{i,t-1} + \beta_1 RM_t + \varphi_1 C_{2,t-3} + v_1 \quad (1)$$

$$C_{2,t} = a_2 + \sum_{i=1}^5 \alpha_{2i} C_{i,t-1} + \beta_2 RM_t + \varphi_2 C_{2,t-3} + v_2 \quad (2)$$

....

....

$$C_{5,t} = a_5 + \sum_{i=1}^5 \alpha_{5i} C_{i,t-1} + \beta_5 RM_t + \varphi_5 C_{2,t-3} + v_5 \quad (5)$$

Donde:

- $C_{i,t}$: Cantidad de delitos ocurridos en el mes t en el cuadrante i
- RM_t : Cantidad de delitos ocurridos en el mes t en la Región metropolitana
- v_i : Perturbación aleatoria

Resultados

Cuadro 3.3.3 Resultados estimación VAR mensual

	C1	C2	C3	C4	C5
C1(-1)	0.24	-0.39	-0.01	0.08	-0.03
C2(-1)	-0.16	0.62*	0.04	-0.05	-0.03
C3(-1)	0.03	-0.57	0.15	-0.08	0.08
C4(-1)	0.52	0.64	0.24	0.27	-0.14
C5(-1)	-0.24	-0.83	0.01	-0.16	0.18
C	-90.66	118.75	-158.28	79.49	171.35
RM	0.05*	0.07*	0.03*	0.01*	0.01*
C2(-3)	-0.08	-0.12	-0.06	0.02	-0.07*
R-squared	0.61	0.71	0.68	0.39	0.33
Adj. R-squared	0.52	0.65	0.61	0.26	0.18
Sum sq. resids	486293	1115252	168125	70957	112460
S.E. equation	125.25	189.67	73.64	47.84	60.23
F-statistic	6.84	11.08	9.52	2.86	2.18
Log likelihood	-239.24	-255.43	-218.53	-201.71	-210.69
Akaike AIC	12.68	13.51	11.62	10.75	11.21
Schwarz SC	13.02	13.85	11.96	11.10	11.56
Mean dependent	1045.05	2055.90	791.69	322.00	318.44
S.D. dependent	180.46	320.64	118.04	55.45	66.46

*parámetros significativos al 5%

Estabilidad del VAR mensual

El VAR estimado previamente cumple con la condición de estabilidad, debido a que todas las raíces de la matriz son menores que uno.

La interpretación de los parámetros estimados es la siguiente. Se aprecia del (cuadro 3.3.3) que el cuadrante 1 está explicado por el primer rezago de cada uno de los cuadrantes más tres variables exógenas C , RM y $C_{2,t-3}$. Si consideramos el coeficiente asociado a RM (0.05), esto quiere decir que si en RM ocurre un delito hoy entonces en el $C1$ ocurrirán $=0.05$ delitos, de esta manera si proyecta que para RM habrán 30000 delitos en el periodo $t+1$, entonces según el coeficiente de RM asociado al $C1$, habrán $30000 \cdot 0.05 = 1500$ delitos en el $C1$, asumiendo que ninguna otra variable está cambiando.

3.4 Funciones de impulso-respuesta diaria

Cuando ocurre un shock en una variable endógena i , no sólo afecta a la ecuación que la describe, el efecto se propaga a las restantes variables endógenas del modelo. Las funciones de impulso respuestas capturan este fenómeno, por ejemplo, si introducimos un shock en una de las variables endógenas queremos ver como se ven afectadas las restantes variables del modelo, en su formas contemporáneas como en sus valores futuros. A continuación presentamos el caso en que se introduce una innovación de una desviación estándar sobre el $C1$, y veremos como se afectan las otras variables endógenas en el modelo, cuanto tiempo demoran en desaparecer los efectos del shock, en que ecuaciones es más severo, entre otras.

Los resultados que se presentan en el gráfico 3.4 muestran que el impacto del shock en el $C1$ va desapareciendo lentamente hasta ser nulo a partir del adelanto¹⁸ siete para todos los cuadrantes, como era de esperar la variable que se ve más afectada es el $C1$, luego el $C2$ y $C4$. Estas funciones nos dan indicios de como se verán afectadas las series cuando se introduce un cierta anomalía en alguna de

¹⁸ Se llama adelanto a una variable que está en tiempos futuros, por ejemplo si hoy estamos en t , a un valor en $t+3$ se le llamará adelanto tres.

las ecuaciones del var. Por ejemplo, si es que ocurriese algún fenómeno extraño que produzca que el C1 aumente demasiado en un cierto momento, esto repercutirá más en el cuadrante 2 que en el resto de los cuadrantes, por lo tanto, si se sospecha que hay un shock o una anomalía en una de las observaciones para una de las variables endógenas, deberán tenerse en cuenta los efectos que este produce en las restantes variables del sistema VAR, cuando se realicen las proyecciones con los modelos, para no sobrestimar lo que sucede en el resto de los cuadrantes.

Cuando se introduce un cambio en el C2¹⁹, se puede observar que los cuadrantes más afectados son el C1 y el C3, el C1 por ejemplo, muestra un impacto mayor en t+2 en comparación con t+1, luego declina el efecto del shock para los valores del C1 que están más adelante, haciéndose cero cerca del noveno adelanto. En los cuadrantes 4 y 5 los efectos se hacen sentir en los primeros días desapareciendo a partir del tercer adelanto para el C4 y del adelanto seis para el C5.

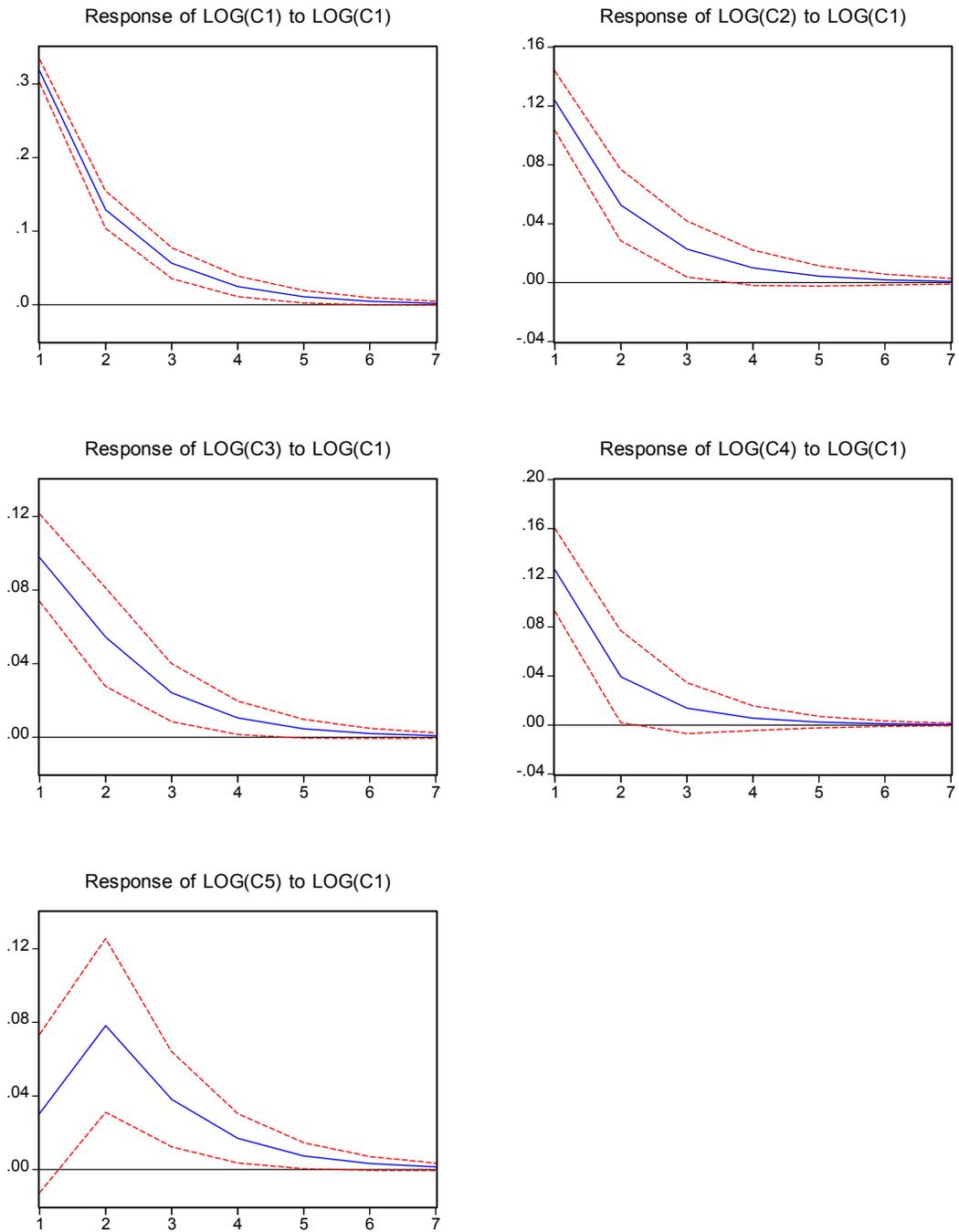
Los shocks en el C3 tienen un menor impacto en el resto de los cuadrantes, además suelen desaparecer más rápido que los shocks en ocurridos en el C1 y C2.

Cuando se introduce un shock en el C4, provoca efectos negativos en el C1 y C2 los efectos tienden a cero a partir del octavo adelanto, esto indica que se estarían subestimando las proyecciones de aquellos cuadrantes cuando alguna anomalía se presenta en el C4, el C3 se afecta positivamente en los primeros adelantos y luego negativamente a partir del adelanto 3, hasta que los efectos desaparecen alrededor del séptimo adelanto. Cuando el shock ocurre en el C5 la propagación de los efectos a los otros cuadrantes no es tan grande, los cuadrantes 1 y 3 responden positivamente atenuándose el impacto a partir del quinto adelanto, el C2 presenta cambios negativos que tienden a cero a partir del quinto adelanto, el C4 se afecta negativamente para sus primeros tres adelantos y luego positivamente para los adelantos 4 y 5 y luego los efectos tienden a desaparecer.

¹⁹ Para visualizar los impactos de innovaciones ver los gráficos proporcionados en el anexo

Gráfico 3.4 (Efectos de un shock en C1, sobre todos los cuadrantes) ²⁰

Response to Cholesky One S.D. Innovations \pm 2 S.E.



Las funciones de impulso-respuesta en nivel mensual se presentan en el anexo.

²⁰ Las funciones de impulso respuesta fueron construidas en su forma logarítmica, como se explicó esto no invalida las conclusiones obtenidas.

3.5 Sistemas de ecuaciones

Los sistemas de ecuaciones son un modo alternativo de modelar series que presentan interacción, la ventaja que tienen por sobre los VAR es que cada una de las ecuaciones del sistema puede ser especificada según las características que cada serie presente. En cambio en el modelo VAR, todas las ecuaciones que se incluyen tienen la misma estructura.

Estimar un modelo por medio de sistema de ecuaciones tiene ventajas tales como determinar restricciones a los parámetros afectando al conjunto, y también se pueden calcular la correlación entre los residuos de las ecuaciones. Sin embargo, la gran desventaja que tienen, es que si el investigador se equivoca en la especificación de alguna de las ecuaciones del sistema, esto se transmitirá a las demás ecuaciones ensuciando las estimaciones.

Metodología

Forma genérica

$$f (y_t, x_t, \theta) = \varepsilon_t$$

Donde:

- y_t : vector de variables endógenas
- x_t : vector de variables exógenas
- θ : vector de parámetros

Existen diversas formas de estimar un sistema entre ellas destacan mínimos cuadrados ordinarios, mínimos cuadrados en dos y tres etapas, S.U.R (regresiones aparentemente no relacionadas), máxima verosimilitud con información completa, GMM etc.

3.5.1 Sistema diario²¹

El sistema diario incluye las variables ficticias que captan la estacionalidad por día de la semana, estas resultaron ser estadísticamente significativas, el resto de las variables exógenas del modelo se escogió según los patrones de correlaciones entre las distintas series

Especificación sistema diario

$$\begin{aligned} \mathbf{C1} &= C(1) + C(3)*C2(-35) + C(5)*C1(-21) + C(44)*DD1 + C(45)*DD2 + C(46)*DD3 \\ &+ C(47)*DD4 + C(48)*DD5 + C(49)*DD6 + C(68)*RM(-365) + C(90)*DFM + \xi_1 \\ \mathbf{C2} &= C(2) + C(10)*C2(-28) + C(11)*C2(-21) + C(15)*C2(-105) + C(51)*DD2 + C(53)*DD4 \\ &+ C(54)*DD5 + C(69)*RM(-365) + C(91)*DFM + \xi_2 \\ \mathbf{C3} &= C(25) + C(28)*C2(-14) + C(56)*DD1 + C(57)*DD2 + C(58)*DD3 + C(59)*DD4 \\ &+ C(60)*DD5 + C(61)*DD6 + C(70)*RM(-365) + \xi_3 \\ \mathbf{C5} &= C(33) + C(34)*C5(-14) + C(62)*DD1 + C(63)*DD2 + C(64)*DD3 + C(65)*DD4 \\ &+ C(66)*DD5 + C(71)*RM(-365) + \xi_5 \\ \mathbf{C4} &= C(72) + C(73)*C2(-161) + C(74)*C2(-308) + C(75)*C2(-357) + C(76)*C2(-364) \\ &+ C(94)*DFM + \xi_4 \end{aligned}$$

Donde:

- **C1,..C5:** es la cantidad de delitos diaria ocurrida en cada uno de los cinco cuadrantes de la comuna de Santiago
- C(i): son los coeficientes de las variables explicativas
- Cj(-k): es la cantidad de delitos ocurrida en el cuadrante j hace k días j=1,...5
- RM(-k): rezago k de la variable RM
- DD1: variable ficticia que toma el valor uno cuando se trata del día lunes y cero en cualquier otro caso
- DD2: variable ficticia que toma el valor uno cuando se trata del día martes y cero en cualquier otro caso.
- DFM: variable ficticia que toma el valor de 1 cuando el día del mes es 30 o 31 y cero en otro caso.
- ξ_i : Perturbación aleatoria

²¹ Los resultados se presentan en el anexo al capítulo 3.

Ajustes en nivel diario

Cuadro3.5.1

	Adjusted R-squared
C1	0.34
C2	0.48
C3	0.25
C4	0.13
C5	0.65

Fuente: Cálculo basado en los datos de AUPOL

En nivel diario el VAR presenta mejores ajustes que el sistema de ecuaciones.

En el sistema diario se incorporó la variable dummy fin de mes, esta asigna un valor de 1 cuando el día del mes es 30 o 31, está resultó ser significativa para los cuadrantes 1,2 y 4. Esto quiere decir que estadísticamente existen diferencias los fines de mes para los cuadrantes mencionados, mientras que para el C3 y C5 no existe evidencia estadística para avalar un comportamiento distinto en los fines de mes.

3.5.2 Sistema semanal

Dada la inexistencia de variables que describan el comportamiento de las series en nivel semanal, no se especificará un sistema de ecuaciones para el nivel semanal. Sin embargo, se presentan los resultados obtenidos para la mejor especificación encontrada en nivel semanal, como estos resultados son poco satisfactorios, se omitirá en esta sección el análisis de las series en dicho nivel. Los ajustes en términos de R-cuadrado ajustado para las distintas series fueron los siguientes:

Cuadro 3.5.2

	Adjusted R-squared
C1	0.25
C2	0.41
C3	0.30
C4	0.27
C5	0.09

Fuente: Cálculo basado en los datos de AUPOL

3.5.3 Sistema mensual²²

Se probó la estacionalidad aparente de los cuadrantes 1 y 2 en nivel mensual, como se apreció en el análisis gráfico, parecía que existía un comportamiento cíclico dependiendo de los meses del año, detectándose una disminución de los delitos de enero a febrero, una tendencia alcista entre marzo y abril, una caída en los meses de mayo hasta septiembre y nuevo aumento a partir de octubre hasta fin de año. Sin embargo, los parámetros asociados a las variables ficticias fueron no significativas, lo que da cuenta que no existe diferencia estadísticamente significativa en nivel mensual para los cuadrantes 1 y 2 según los meses del año. Para el C4 se encontró que en los meses de octubre a diciembre si existen diferencias significativas con respecto al resto de los meses, para el C5 las diferencias significativas se dan para los dos primeros meses del año. Las correlaciones rezagadas más fuertes encontradas son con rezagos de 12 meses, lo que refleja un componente de estacionalidad anual para el c1, c3 y c5 sin embargo ninguna de estas correlaciones supera el valor de 0.5.

Para efectos de pronósticos se incorporó la variable desocupados²³, está variable en sus rezagos dos y tres es altamente significativa en los cuadrantes 2 y 3 lo que mejora considerablemente los ajustes de los modelos mensuales. La lógica de la incorporación de esta variable es evidente, se debiera esperar que la cantidad de desocupados de la economía influya en la cantidad de delitos ocurrida.

La especificación del modelo de sistema de ecuaciones en nivel mensual es:

Especificación Sistema mensual

$$C1 = C(1) + C(7) * C2(-12) + \xi_1$$

$$C2 = C(2) + C(21) * C1(-10) + C(22) * C1(-11) + C(23) * DESOCUPADOS(-2) + C(24) * DESOCUPADOS(-3) + \xi_2$$

²² Los resultados del sistema se presentan en el anexo al capítulo 3.

²³ Esta serie fue extraída de las estadísticas económicas del Banco Central de Chile.

$$C3=C(3)+C(28)*C1(-12)+C(29)*C2(-12)+C(31)*RM(-12)+C(32)*RM(-1)+C(32)*DESOCUPADOS(-2)+C(33)*DESOCUPADOS(-3) + \xi_3$$

$$C4= C(4)+C(47)*C5(-8)+C(48)*C4(-10)+C(60)*DE2 + \xi_4$$

$$C5= C(5)+C(64)*C3(-7)+C(65)*C1(-8)+C(71)*DE1+\xi_5$$

Donde:

- **C1,..C5:** es la cantidad de delitos mensual ocurrida en cada uno de los cinco cuadrantes de la comuna de Santiago
- C(i): son los coeficientes de las variables explicativas
- Cj(-k): es la cantidad de delitos ocurrida en el cuadrante j hace k meses j=1,...5
- RM(-k): rezago k de la variable RM
- DESOCUPADOS(-K): cantidad de desocupados hace k meses
- DE1: variable ficticia que toma el valor uno cuando se trata del mes de enero o febrero y cero en cualquier otro caso
- DE2: variable ficticia que toma el valor uno cuando se trata del mes de marzo o abril y cero en cualquier otro caso.
- ξ_j : Perturbación aleatoria

Ajustes alcanzados por ecuación

Cuadro 3.5.3

	Adjusted R-squared
C1	0.56
C2	0.71
C3	0.58
C4	0.37
C5	0.44

Fuente: Cálculo basado en los datos de AUPOL

Si se comparan los resultados obtenidos con el VAR mensual con los que provienen del sistema mensual, se ve que este último es bastante mejor, lo que debiera reflejarse en mejores pronósticos.

Capítulo 4

Pronósticos

4.1 Introducción

En este capítulo se obtienen pronósticos con los diferentes modelos estudiados previamente. En primer lugar se introducen modelos alternativos de predicción, estos pertenecen a la familia de los métodos de suavización exponencial. Se incorporan porque no es necesario saber con precisión cual es el modelo que subyace detrás de los datos, los pronósticos obtenidos suelen ser buenos cuando se trata de series en que el tiempo es casi el único determinante de su comportamiento, se van ajustando paulatinamente a los cambios ocurridos en las últimas observaciones, es decir, capturan bien la volatilidad que se va generando por la introducción de algún shock y por último son fáciles de calcular.

En segundo lugar se compara la capacidad de predicción de los distintos modelos, para ello se hace un ranking de los mejores modelos, utilizando como criterio el error absoluto medio de predicción²⁴.

$$MAPE = \frac{\sum_{t=T+1}^{T+k} \left| \frac{\hat{y}_t - y_t}{y_t} \right|}{k + 1}$$

donde:

\hat{y}_t : Valor proyectado de acuerdo al modelo.

y_t : Valor de la variable que se desea proyectar.

k: número de proyecciones realizadas.

En tercer lugar se realizan pronósticos para algunos cuadrantes con los modelos que presentan un menor error promedio para cada caso. Los pronósticos se efectúan dentro de muestra y fuera de muestra. Finalmente se hace una

²⁴ Este consiste en un promedio del error porcentual que se obtiene, de comparar el pronóstico con el valor verdadero de la observación. (este cálculo se realiza dentro de muestra)

proyección a partir de la última observación registrada en este estudio para el cuadrante 1 utilizando el VAR y el sistema de ecuaciones.

4.2 Modelos alternativos de predicción

En esta sección se presentan técnicas adicionales a las estudiadas en el capítulo anterior, las cuales serán de utilidad para realizar pronósticos. Estas son útiles en casos como el estudio a nivel semanal, en donde, no fue posible encontrar una especificación adecuada a través de los modelos multiecuacionales, porque en este nivel se carecía de elementos como estacionalidades, tendencias, patrones; que permitiesen caracterizar las series y modelarlas.

Los métodos de suavización exponencial, son útiles para la predicción de corto plazo, no es necesario especificar un modelo de ecuaciones que describa a las series para poder utilizarlos. Los pronósticos se calculan considerando sólo variables como el error de pronóstico anterior y el valor pasado de la variable, tendencias y estacionalidades de las series. No importando la naturaleza de la distribución que determina a las series.

4.2.1 Suavización Exponencial Simple (1 parámetro)

La suavización exponencial simple es utilizada cuando las series presentan un comportamiento aleatorio, hacia arriba y abajo con una media definida, sin presentar patrones estacionales ni tendencias claras.

Para esto se introduce un valor α que corresponde al grado de importancia que se le da a la historia de la cantidad de delitos ocurrida en el cuadrante i versus el pronóstico que se obtuvo para ella, por lo que se puede tener distintos pronósticos dependiendo de su valor. De esta forma, el modelo resultante es:

$$P_{t+1} = \alpha C_{i,t} + (1 - \alpha)P_t$$

donde:

$C_{i,t}$ = cantidad de delitos ocurridas en el cuadrante i en el periodo t
 P_{t+1} = cantidad de delitos pronosticada para el cuadrante i en el periodo t+1
 α = Ponderador que se le asigna a la cantidad observada de delitos en t versus la que se le da al pronóstico

Para escoger el valor de α se puede realizar una prueba considerando los resultados obtenidos para distintas opciones y midiendo los grados de error producidos en cada caso. El software utilizado para los cálculos²⁵, estima el valor de α minimizando la suma de los errores de predicción al cuadrado.

El método anterior no admite la posibilidad de que las series puedan presentar patrones estacionales o tendencias, es por eso que se introducen otros nuevos modelos de suavizamiento que si recogen estas características

4.2.2 Suavización Exponencial con Tendencia

Una tendencia es un comportamiento sostenido que presenta una serie a lo largo del tiempo, esta puede ser ascendente o descendente. Las ecuaciones que describen el método de suavización exponencial con tendencia se presentan a continuación:

$$A_t = \alpha C_{i,t} + (1 - \alpha)(A_{t-1} + t_{t-1})$$

Donde,

$$t_t = \beta(A_t - A_{t-1}) + (1 - \beta)t_{t-1}$$

Donde:

$(A_t - A_{t-1})$ = tendencia observada

²⁵ Eviews 5.0

(t_{t-1}) = tendencia obtenida
 β = ponderador

A partir de esto, el pronóstico para el futuro (para k periodos en el futuro) queda determinado por:

$$P_{t+k} = A_t + K * t_t \quad t = 1, 2, 3, \dots, 7$$

4.2.3 Suavización Exponencial con Estacionalidad

Existe estacionalidad cuando el comportamiento de una serie presenta patrones que se replican cada cierto período de tiempo.

4.2.4 Suavización Exponencial con Estacionalidad Aditiva: Si la estacionalidad no cambia con la tendencia. El modelo es el siguiente:

$$A_t = (\alpha (C_{i,t} - E_{t-1}) + (1 - \alpha)(A_{t-1} + t_{t-1}))$$

Donde:

$$t_t = \beta(A_t - A_{t-1}) + (1 - \beta)t_{t-1}$$

El índice estacional para el periodo t será:

$$E_t = \gamma(C_{i,t} - A_t) + (1 - \gamma)E_{t-L}$$

En este caso, se supone que el ciclo de estacionalidad contiene L periodos. Existen L índices de estacionalidad, uno para cada periodo. Por ejemplo si la cantidad observada de delitos es en nivel diario y el ciclo de estacionalidad se observa todas las semanas, entonces L=7. Cada día, uno de los índices se actualizará obteniéndose un valor junto con la tendencia y el promedio.

El pronóstico para k periodos en adelante será:

$$P_{t+k} = A_t + Kt_t + R_{t-L+k}$$

4.2.5 Suavización con Estacionalidad Multiplicativa: Si la estacionalidad varía con la tendencia entonces es apropiado ajustar el siguiente modelo:

$$A_t = \alpha \frac{C_{i,t}}{E_{t-L}} + (1 - \alpha)(A_{t-1} + t_{t-1})$$

Donde:

$$t_t = \beta(A_t - A_{t-1}) + (1 - \beta)t_{t-1}$$

$$E_t = \gamma \left(\frac{C_{i,t}}{P_{t+1}} \right) + (1 - \gamma)E_{t-L}$$

$$P_{t+k} = A_t + Kt_t + E_{t-L+k}$$

La estacionalidad de la variable se corrige por un factor γ que indica el grado de importancia que se le asigna a la estacionalidad, y se calcula ponderando por este factor a la razón entre la cantidad observada de delitos efectiva obtenida para un período y la estimación realizada para el período siguiente y, por el complemento del factor, a la estacionalidad calculada para el período anterior.

4.3 Resultados de los Modelos de Suavización Exponencial

En este ítem se evalúa la capacidad de predicción de los distintos modelos de suavización exponencial expuestos anteriormente, el objetivo es escoger un método, para luego hacer comparaciones con los modelos de sistemas, desarrollados en el capítulo anterior. Similarmente a como se hará más adelante en

este capítulo, se hizo un ranking para escoger el mejor método de suavización exponencial, utilizando como medida el error porcentual medio de predicción²⁶.

La información contenida en los cuadros es la siguiente:

- SM_AD: Suavización exponencial con estacionalidad aditiva
- SM_D : Suavización exponencial con tendencia
- SM_MU: Suavización exponencial con estacionalidad multiplicativa
- SM_NS: Suavización exponencial sin estacionalidad
- SM_S : Suavización exponencial simple

En los cuadros 4.3.1 4.3.2 y 4.3.3 se presenta el error porcentual medio que se incurre por cuadrante cuando se utilizan las diferentes técnicas de suavizamiento exponencial, para los tres nivel temporales especificados. Se puede observar que a medida que los datos se van agregando, es decir, cuando examinamos el paso de cantidad de delitos diarios a semanales y de semanales a mensuales se obtienen mejores pronósticos con todos los métodos. Para todos los niveles temporales se observa que el mejor método es el de suavización exponencial con estacionalidad aditiva.

En nivel diario el C2 es el que presenta un menor error de pronóstico alcanzando en promedio un 27% de error (ver cuadro 4.3.1), por otra parte en el C5 se está cometiendo el error más alto, siendo este de un 73%, el error asociado al cálculo de pronósticos de la región metropolitana es del 10%.

Cuadro 4.3.1 Error de Pronósticos diarios

	SM_AD	SM_D	SM_MU	SM_NS	SM_S
c1	0,30	0,38	0,30	0,38	0,38
c2	0,27	0,49	0,27	0,50	0,49
c3	0,31	0,39	0,31	0,40	0,39
c4	0,47	0,49	0,50	0,49	0,00
c5	0,73	1,94	1,92	1,95	0,00
rm	0,10	0,12	0,10	0,12	0,12

Fuente: Cálculo basado en los datos de AUPOL

²⁶ Este indicador también es conocido como MAPE.

En el nivel semanal los mejores pronósticos son para el C2 y C3, alcanzando ambos un error porcentual medio del 13% con el método de suavización exponencial aditiva (ver cuadro 4.3.2), para el caso de RM se comete un error de predicción del 8%, sin embargo, con el método de suavización exponencial simple para RM, se logra un 7% de error.

Cuadro 4.3.2 Error de Pronósticos semanales

	SM_AD	SM_D	SM_MU	SM_NS	SM_S
c1	0,15	0,16	0,15	0,15	0,15
c2	0,13	0,14	0,13	0,13	0,13
c3	0,13	0,14	0,13	0,13	0,13
c4	0,17	0,18	0,17	0,18	0,18
c5	0,24	0,26	0,24	0,24	0,24
rm	0,08	0,08	0,08	0,08	0,07

Fuente: Cálculo basado en los datos de AUPOL

En nivel mensual los mejores pronósticos son para el C3, seguido indistintamente del C2 y C1 (ver cuadro 4.3.3), obteniéndose errores del 7, 8, 8 por ciento respectivamente. El error que se alcanza en el caso de RM es bastante bajo, siendo éste de un 4%.

Cuadro 4.3.3 Error de pronósticos mensuales

	SM_AD	SM_D	SM_MU	SM_NS	SM_S
c1	0,08	0,15	0,08	0,14	0,15
c2	0,08	0,12	0,08	0,11	0,11
c3	0,07	0,11	0,08	0,11	0,11
c4	0,11	0,13	0,12	0,13	0,13
c5	0,13	0,19	0,12	0,19	0,19
rm	0,04	0,08	0,04	0,08	0,09

Fuente: Cálculo basado en los datos de AUPOL

4.4 Pronósticos con los modelos VAR y de sistemas de ecuaciones

El procedimiento efectuado para obtener los pronósticos dentro de muestra, a través de las metodologías VAR y Sistemas de ecuaciones fue el siguiente:

- i) Se estimaron los parámetros de los modelos para las series desde enero de 2001 hasta el 30 de junio de 2004 (en el capítulo 3).
- ii) Luego se hizo el cálculo de pronósticos para los distintos cuadrantes utilizando como parámetros las ecuaciones las encontradas en i).
- iii) Finalmente se comparó la diferencia del pronóstico con el valor verdadero y se calculó el error porcentual medio de predicción.

En el caso del VAR mensual como se había incorporado RM en forma contemporánea se hizo primero un pronóstico de RM.

4.4.1 Resultados de los modelos VAR y sistemas de ecuaciones

Como se hizo mención en los capítulos anteriores, no fue posible encontrar una buena especificación en nivel semanal, por lo tanto el análisis en esta sección se reduce solo al comportamiento de las series en nivel diario y mensual.

Del cuadro 4.4.2 podemos observar que en general los mejores resultados se obtienen con el sistema de ecuaciones, sin embargo el VAR es más apropiado para el caso del cuadrante 1 alcanzándose un error del 27%.

Si se comparan los tres modelos se puede decir que para el caso del C1, es conveniente la utilización del VAR, para el C2, C3 y C4 el método de suavización exponencial con estacionalidad aditiva es el que logra mejores resultados, aunque marginalmente por sobre el sistema de ecuaciones, para el cuadrante 5 el sistema es el que logra mejores resultados.

Cuadro 4.4.2 Error porcentual medio de predicción en nivel diario

	C1	C2	C3	C4	C5
Sistema	0.30	0.31	0.32	0.49	0.67
Var	0.27	0.56	2.02	0.63	2.16

Fuente: Cálculo basado en los datos de AUPOL

En nivel mensual es menos clara la elección entre el VAR y el sistema de ecuaciones, este domina ligeramente por sobre el VAR en todos los cuadrantes exceptuando el C3 en que el VAR es ampliamente preferido ver cuadro (4.4.3).

Si comparamos nuevamente las tres formas alternativas, se puede llegar a la conclusión siguiente. El método que mejor ajusta a C1 y C5 es el de suavizamiento exponencial con tendencia aditiva, para el C2 todos los métodos tienen en el mismo error absoluto medio igual al 8%, para el C3 el mejor modelo es el VAR con el cual se alcanza un error de sólo un 6%, para el C4 se puede usar indistintamente el sistema de ecuaciones o el suavizamiento exponencial con estacionalidad aditiva.

Cuadro 4.4.3 Error absoluto medio de predicción en nivel mensual

	C1	C2	C3	C4	C5
Sistema	0.10	0.08	0.65	0.11	0.16
Var	0.12	0.08	0.06	0.14	0.17

Fuente: Cálculo basado en los datos de AUPOL

Al comparar las proyecciones de los distintos modelos estudiados en esta sección es posible percatarse que no existe sólo un modelo que sea capaz de entregar los mejores pronósticos para los distintos niveles temporales estudiados, por lo tanto la mejor estrategia a seguir para realizar pronósticos es utilizar todas las técnicas y preferir aquella que mejor predice según el cuadrante que se quiera estudiar.

El resultado anterior sugiere que si bien el comportamiento sistémico e interrelacionado que debieran presentar los distintos cuadrantes de la comuna de Santiago parece ser real, en muchos casos no lo es y sólo se da para algunos

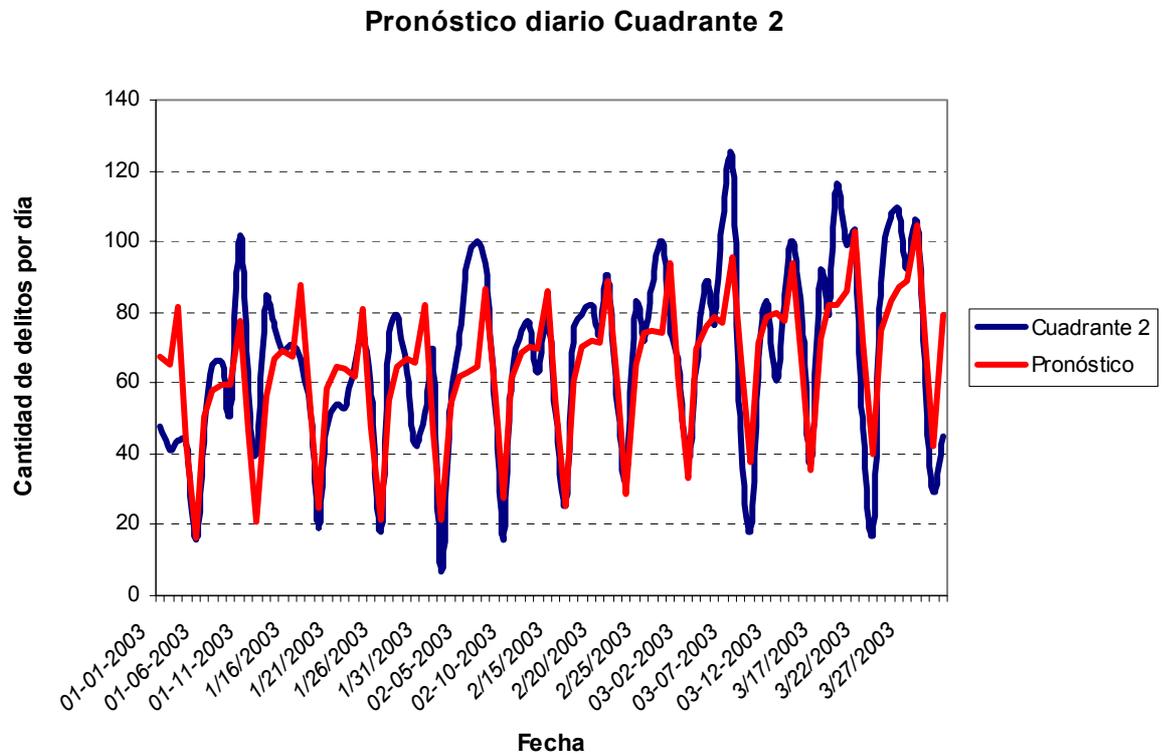
cuadrantes, en términos prácticos a veces es conveniente ignorar el comportamiento conjunto de las series utilizando técnicas como el suavizamiento exponencial que sólo considera a la serie que se está analizando sin tomar en consideración las otras.

4.5 Proyecciones dentro de muestra

En esta sección se presentan gráficos para visualizar como es el comportamiento de la proyección utilizando el mejor modelo encontrado para un cierto cuadrante, contrastándolo con el curso real de la serie para dicho cuadrante.

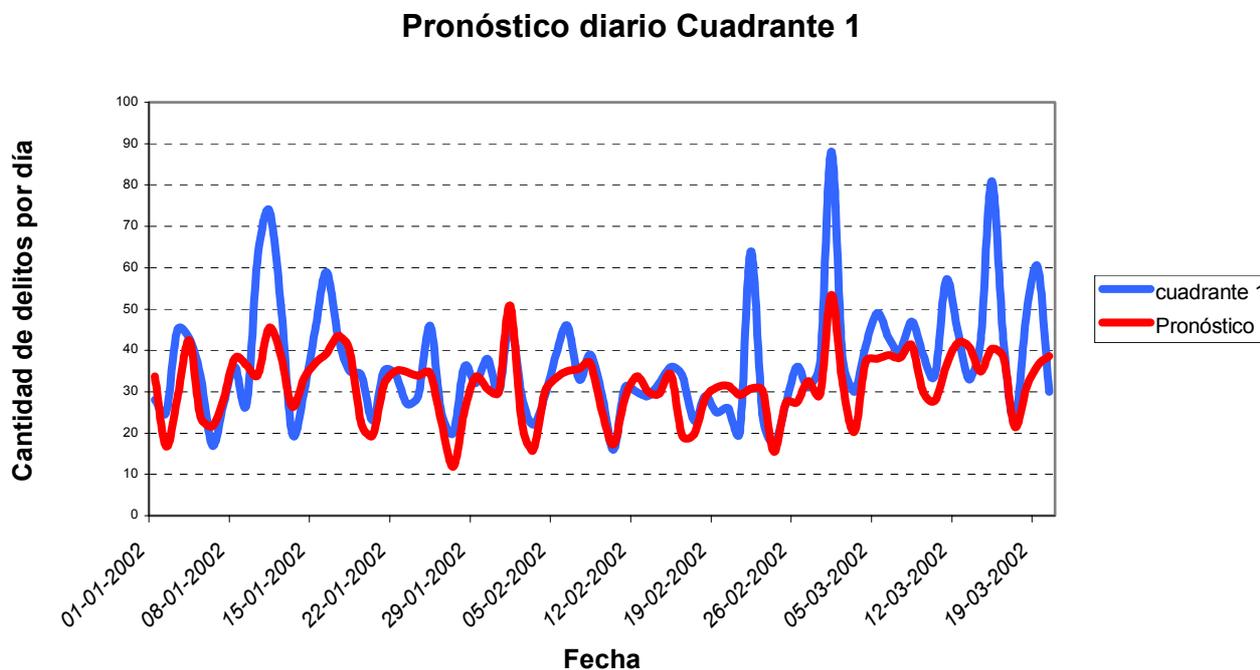
En nivel diario se presenta el pronóstico obtenido para el cuadrante 2 utilizando el método de suavización exponencial con estacionalidad aditiva para el primer semestre de 2003 (gráfico 4.5.1). Vemos que el pronóstico se comporta de manera acotada tendiéndose a comportar de manera rezagada con respecto a la serie real (esto viene dado por la estructura de modelación del método).

Gráfico 4.5.1 (Pronóstico diario cuadrante 2 utilizando método de suavización exponencial)



En el gráfico 4.5.2, se presenta el pronóstico obtenido a través del modelo VAR contrastado con la serie real observada para dicho cuadrante, entre el primero de enero de 2002 y el 20 de marzo del mismo año.

Gráfico 4.5.2 (Pronóstico diario Cuadrante 1 utilizando VAR)



4.6 Pronósticos diarios fuera de muestra para el sistema de ecuaciones

En este punto se analizará la capacidad de predicción fuera de muestra de los modelos de sistemas de ecuaciones y de los modelos VAR en nivel diario, no se verá el caso de estos modelos en nivel mensual, porque las series de datos son demasiado cortas para llegar a conclusiones razonables.

Para realizar los pronósticos fuera de muestra se consideró como base, los parámetros estimados con la información para el período comprendido entre el primero de enero de 2001 y el 31 de diciembre 2003. Luego se procedió a formular pronósticos para el primer semestre del año 2004 a partir de los parámetros estimados para el periodo señalado anteriormente.

El cuadro 4.6.1 muestra los errores absolutos medios de predicción por cuadrantes cuando se utiliza el sistema de ecuaciones para hacer las proyecciones. Vemos que el menor error se obtiene para el cuadrante 3, siendo éste de un 28%. Si comparamos los resultados en términos de error y los comparamos con los del

sistema en nivel diario que considera a toda la muestra para el cálculo de los parámetros (cuadro 4.4.2) estos no difieren en demasía de los calculados en esta sección, este resultado denota una cierta estabilidad en los parámetros estimados, a lo largo del tiempo.

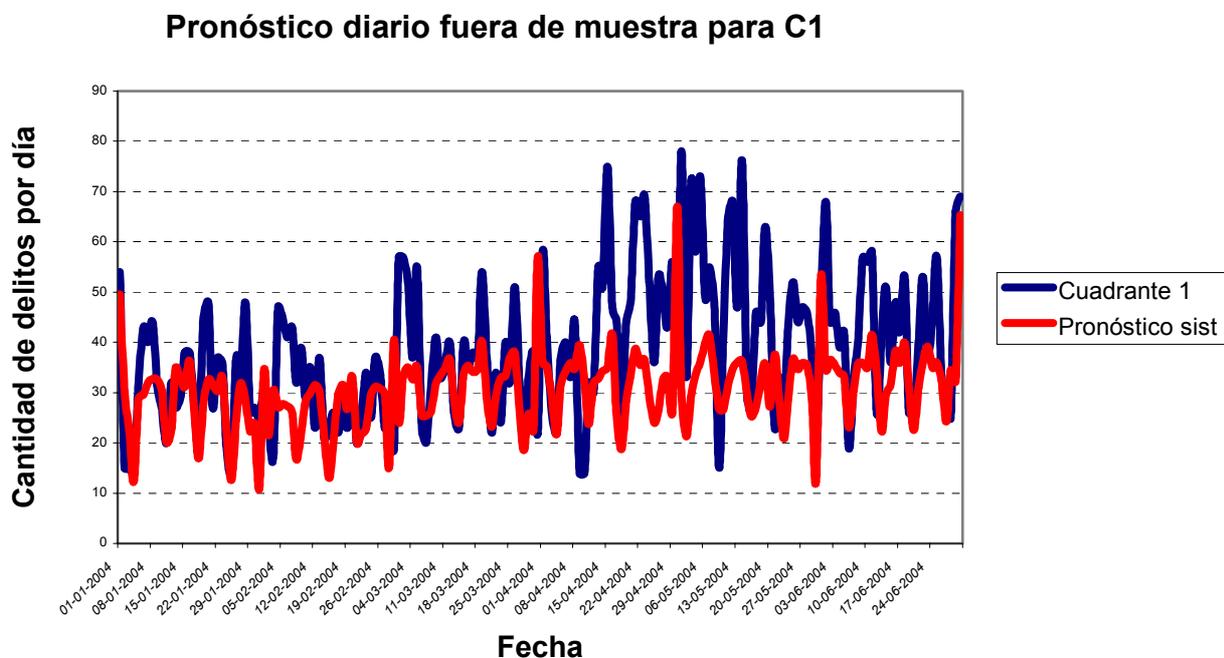
Cuadro 4.6.1 Error absoluto medio de pronóstico por cuadrante utilizando sistema (fuera de muestra)

	C1	C2	C3	C4	C5
error abs medio	0.29	0.31	0.28	0.47	0.71

Fuente: Cálculo basado en los datos de AUPOL

El gráfico 4.6.1 muestra la serie real para el cuadrante 1, y el pronóstico usando sistemas de ecuaciones con información hasta 2003.

Gráfico 4.6.1 (Pronóstico diario fuera de muestra utilizando sistemas de ecuaciones para el cuadrante 1)



Quando consideramos el VAR para los pronósticos fuera de muestra se obtienen los siguientes resultados:

Cuadro 4.6.2 Error absoluto medio de pronóstico por cuadrante utilizando VAR (fuera de muestra)

	C1	C2	C3	C4	C5
error abs medio	0.27	0.27	1.75	0.51	1.67

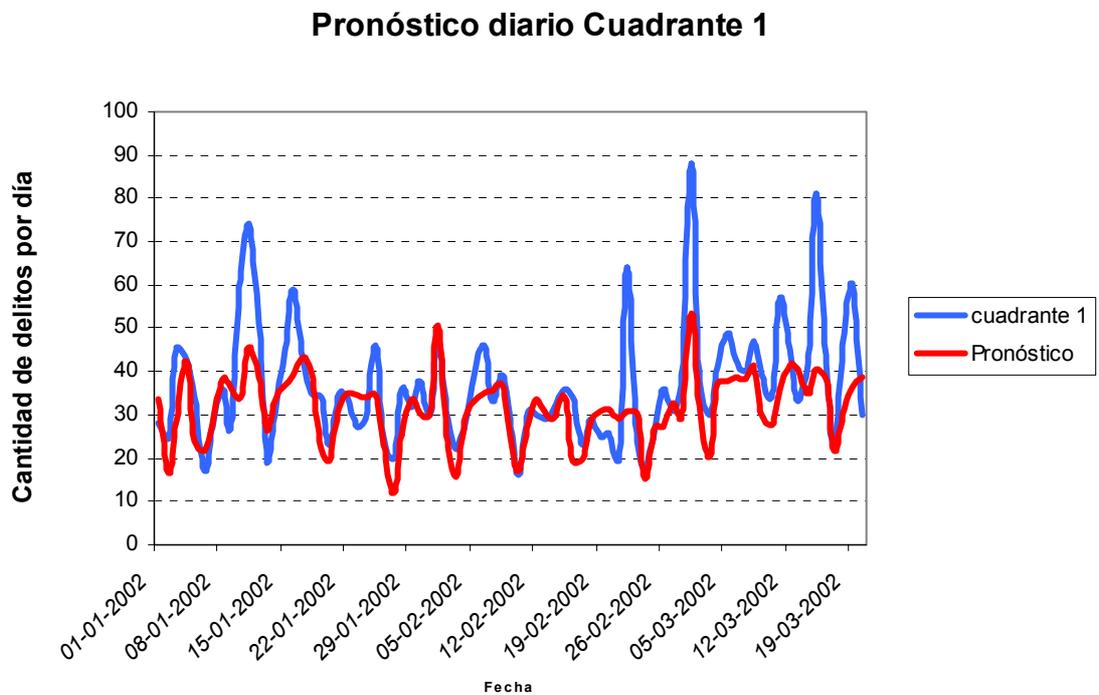
Fuente: Cálculo basado en los datos de AUPOL

Se puede apreciar que se comete un menor error para los cuadrantes 1 y 2 siendo este de un 27% para ambos. En cambio para el resto de los cuadrantes el error absoluto medio de predicción es bastante mayor en comparación con los obtenidos del sistema de ecuaciones.

Como se hizo notar en el punto de pronóstico dentro de muestra, no es posible encontrar un único modelo que prediga mejor para todas los cuadrantes, siendo apropiado utilizar distintos modelos según el cuadrante de cual se trate, en el caso de pronóstico fuera de muestra resulta apropiado utilizar modelos VAR para los cuadrantes 1 y 2 y sistemas de ecuaciones para los cuadrantes 3, 4 y 5.

El gráfico 4.6.2 muestra la cantidad observada de delitos para el cuadrante 2 de la comuna de Santiago versus su pronóstico estimado a partir de los modelos VAR con los datos desde enero de 2001 hasta diciembre de 2003.

Gráfico 4.6.2 (Pronóstico diario fuera de muestra utilizando modelos VAR para el cuadrante 1)



4.7 Pronósticos a partir de julio de 2004 con los modelos VAR y de Sistema de ecuaciones

A continuación se presentan dos pronósticos realizados para el cuadrante 1 de la comuna de Santiago utilizando los modelos VAR y los Sistemas de ecuaciones, este ejercicio se puede replicar para todos los cuadrantes de la comuna de Santiago, sin embargo los pronósticos más fiables son para los cuadrantes 1, 2 y 3.

Gráfico 4.7.1 (Pronóstico diario fuera de muestra utilizando modelos VAR para el cuadrante 1)

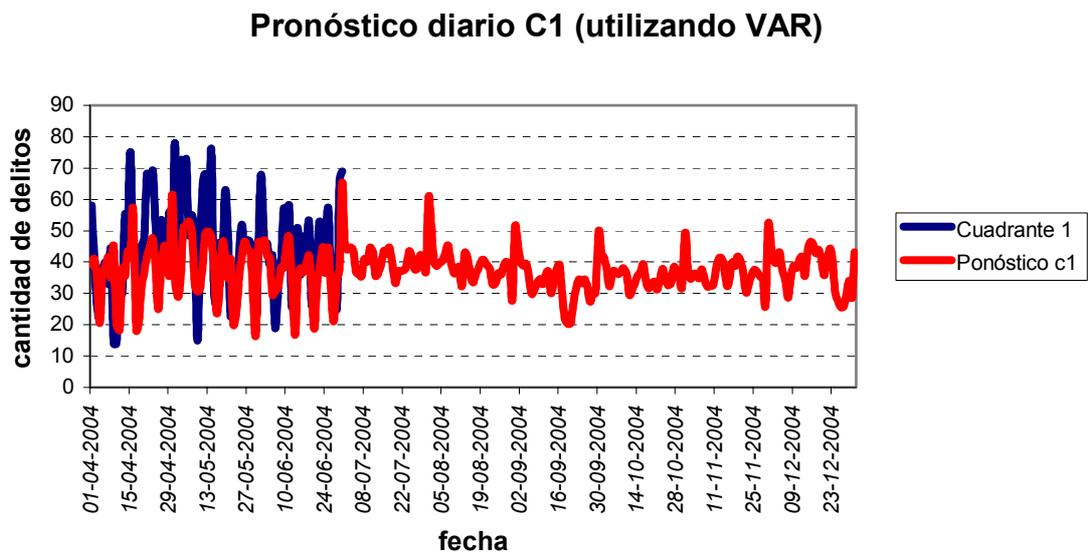
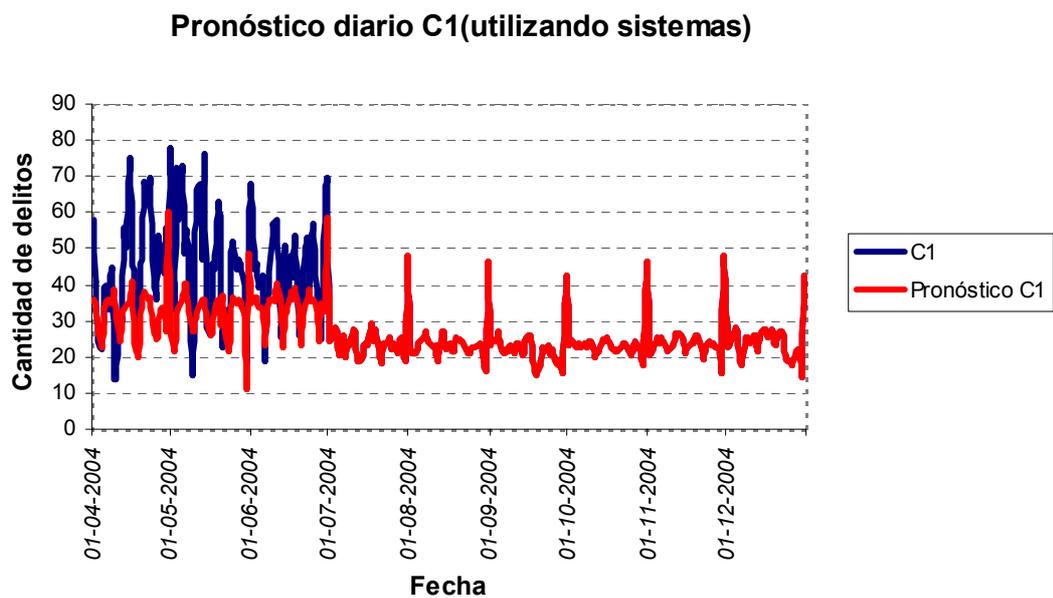


Gráfico 4.7.2 (Pronóstico diario fuera de muestra utilizando sistemas de ecuaciones para el cuadrante 1)



Capítulo 5

Conclusiones y comentarios finales

La distribución de delitos en los distintos cuadrantes de la comuna de Santiago tienen características distintas, es decir, no presentan los mismos patrones de comportamiento a lo largo del tiempo. Por este motivo, es conveniente realizar proyecciones para cada uno de los cuadrantes de la comuna de Santiago en forma independiente.

Cuando se estudian niveles temporales más agregados, por ejemplo, la cantidad de delitos ocurridos a nivel mensual, el comportamiento de las series es más similar al de la región metropolitana. Entonces, bastaría con considerar sólo una muestra de algunas comunas de Santiago para pronosticar el comportamiento de la región metropolitana a nivel mensual.

Las tendencias y estacionalidades se presentan de manera distinta dependiendo del nivel temporal que se esté analizando; se encontró que a nivel diario existe un fuerte patrón cíclico, este es un aumento de los delitos a medida que transcurre la semana, es decir, ocurre un aumento en la cantidad de delitos de lunes a viernes para luego declinar los fines de semana.

Para los niveles temporales más agregados como el de semanas no fue posible detectar patrones de comportamiento a lo largo del tiempo.

A nivel mensual se aprecia que existen tendencias y estacionalidades pero sólo para algunos cuadrantes. Por ejemplo, se detectó que en los cuadrantes 1,2 y 4 existen diferencias estadísticamente significativas en los fines de mes, se encontró que los días 30 o 31, presentan en promedio una cantidad menor de delitos en comparación con el resto de días del mes en dichos cuadrantes. Mientras que para el C3 y C5 no existe evidencia estadística para avalar un comportamiento distinto en los fines de mes. Para el C4 se observa que en los meses de octubre a diciembre ocurre un aumento estadísticamente significativo en la cantidad de delitos en comparación con el resto de los meses del año, para el C5 se obtuvo que en los

meses de enero y febrero existe en promedio una cantidad menor de delitos, en comparación con los otros meses del año.

Los modelos multiecuacionales sin son útiles en términos de pronósticos para la comuna de Santiago, en algunos cuadrantes obtenemos errores de predicción del orden del 27% en los modelos en nivel diario, lo que es un buen resultado si es que lo comparamos con los errores medios que se logran con la utilización de otras metodologías en otros países, estos son de el orden del 30% en promedio²⁷.

Una de las grandes ventajas de la utilización de los modelos VAR, es que se cuenta con las funciones de impulso respuesta, estas permiten observar que cuadrantes se ven más afectados, cuando existe alguna alteración en otro cuadrante del modelo. Esto es de utilidad para hacer las correcciones pertinentes en términos de proyección, cuando se identifica que un shock está causando efectos sobre algún cuadrante. Por ejemplo, se observa que a nivel diario un shock en el cuadrante 1 tiene un mayor impacto en los valores futuros en el cuadrante 2 que en el resto de los cuadrantes.

Los ajustes de los modelos medidos en términos de r cuadrado ajustado aunque suelen ser buenos indicadores para discriminar entre modelos alternativos, esta no es una buena medida para hacer comparaciones entre los VAR y los sistemas, porque estos obedecen a especificaciones distintas. Como se puede apreciar el VAR presenta mejores ajustes que el sistema de ecuaciones para todos los cuadrantes en nivel diario, sin embargo, los errores de predicción que se cometen con el VAR suelen ser más altos que lo obtenidos de utilizar el sistema.

Si bien algunos de los cuadrantes presentan una clara interrelación con los otros como por ejemplo el C1 con C2 y C3. El C4 no presenta una relación clara con los otros cuadrantes. Esto sugiere que para obtener buenos pronósticos es necesario especificar modelos uniecuacionales para aquellos cuadrantes en los que los ajustes a través de modelos sistémicos son bajos.

²⁷ Referencias acerca de errores de pronósticos a nivel internacional se pueden encontrar en Short-term forecasting of crime (2003)

En términos de pronósticos la utilización de uno u otro modelo de predicción, dependerá del cuadrante que se está analizando no existiendo un único modelo que permita generar las mejores proyecciones para todos los casos.

Referencias Bibliográficas

1. Benavente, J.; Contreras, D.; Montero, R. y Melo, E. (2002). Programas Antidelincuencia: Evaluando Comuna Segura. Mimeo. Departamento de Economía. Universidad de Chile.
2. Box, G. E. P., & Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control. San Francisco, Holden-Day.
3. Cohen, L. E., & Felson, M. (1979). "Social change and crime and trends: A routine activity approach". *American Sociological Review*. 44, 588-607.
4. Deadman, D. (2003). Forecasting residential burglary. *Internacional Journal of Forecasting*.
5. Deadman, D., & Pyle, D. J. (1997). "Forecasting recorded property crime using a time series econometric model". *British Journal of Criminology*.
6. Falk, J. J. (1952). "The influence of the seasons on the crime rate". *Journal of Crime and Law Criminology*, 43, 199-213.
7. Felson, M., & Poulsen, E. (2003). Simple indicators of crime by time of day. *International Journal of Forecasting*.
8. Gorr, W. L., Olligschlaeger, A. M., & Thompson, Y. (2003). Short-term time series forecasting of crime. *International Journal of Forecasting*.
9. Harries, R. (2003). Modelling and predicting recorded property crime trends in England and Wales—A retrospective. *International Journal of Forecasting*,
10. Johnston, Dinardo. Métodos de econometría (1997)
11. Quilis Enrique M. Modelos Bvar: Especificación Estimación e Inferencia (Marzo 2002). *Instituto de estudios fiscales Madrid*.

ANEXO

SELECCIÓN DE DELITOS	
ABUSOS CONTRA PARTICULARES	PRESUNTA DESGRACIA
ABUSOS DESHONESTOS	RAPTO
AMENAZAS CON ARMA BLANCA	ROBO CALIFICADO
APROPIACION INDEBIDA (INCLUYE	ROBO CON FUERZA EN LAS COSAS
ATENTADOS Y DASACATOS CONTRA L	ROBO CON HOMICIDIO
COMERCIO AMBULANTE SIN PERMISO	ROBO CON INTIMIDACION
CONSUMIDOR DE COCAINA	ROBO CON VIOLENCIA
CONSUMIDOR DE FARMACOS	ROBO DE VEHICULO
CONSUMIDOR DE MARIHUANA	ROBO EN BIENES NACIONALES DE U
CONSUMIDOR DE OTROS	ROBO EN LUGAR HABITADO
CONSUMO DE BEBIDAS ALCOHOLICAS	ROBO EN LUGAR NO HABITADO
CONSUMO PASTA BASE COCAINA	ROBO POR SORPRESA
CONSUMO Y OTRAS FALTAS LEY DE	SECUESTRO
CRIMENES Y SIMPLES DELITOS SEG	SUSTRACCION DE MENORES
CRIMENES/SIMPLES DELITOS PROVE	TENENCIA O PORTE DE ARMAS, MUN
DAÑOS	TRAFICO DE COCAINA
DELITOS CONTEMPLADOS EN OTROS	TRAFICO DE FARMACOS
DELITOS CONTRA LEY DE PROPIEDA	TRAFICO DE MARIHUANA
DELITOS INFORMATICOS LEY N° 19	TRAFICO DE OTROS
DESORDENES	TRAFICO ILICITO DE DROGAS
DESORDENES PUBLICOS	TRAFICO PASTA BASE COCAINA
EBRIEDAD	USURP.PROP.,DESCUBR.,PRODUCCIO
ELABORACION DE MARIHUANA	USURPACION
ELABORACION DE OTROS	VIOLACION
ELABORACION DE PASTA BASE COCA	VIOLACION DE MORADA
ELABORACION/PRODUCCION SUSTANC	VIOLACION SODOMITICA
ESTUPRO (ACCION PRIVADA)	ABUSO SEXUAL
HOMICIDIO	ABUSO SEXUAL IMPROPIO
HOMICIDIO CALIFICADO	ABUSO SEXUAL IMPROPIO INFANTIL
HOMICIDIO EN RIÑA O PELEA	ABUSO SEXUAL INFANTIL
HOMICIDIO SIMPLE	APREHENDIDOS POR ROBO FUSTRADO
HURTO DE HALLAZGO	DAÑOS CALIFICADOS
HURTO SIMPLE	DAÑOS SIMPLES
LESIONES	DELITOS SIGNIFICACION SEXUAL
LESIONES CON ARMA BLANCA	HURTO AGRAVADO (ART: 447 CODIG
LESIONES CON ARMA DE FUEGO	TRAFICO ILICITO DE DROGAS (ART
LESIONES CON FUEGOS ARTIFICIAL	PROSTITUCION DE MENORES
LESIONES CORPORALES	ROBO CON RETENCION DE VICTIMAS
LESIONES EN ACCIDENTE DE TRANS	ROBO CON VIOLACION
LESIONES EN ACCIDENTE DEL TRAB	ROBO DE ACCESORIOS DE VEHICULO
LESIONES GRAVES	ROBO DE VEHICULO MOTORIZADO
LESIONES GRAVES O GRAVISIMAS	ROBO EN LUGAR HABITADO O DESTI
LESIONES LEVES	PORTE DE DROGAS (ART. 41)
LESIONES MENOS GRAVES	CONSUMO DE DROGAS (ART. 41)

<i>MALTRATO DE OBRA A CARABINERO</i>	<i>CONTROL DE IDENTIDAD</i>
<i>MALTRATO DE OBRA A MIEMBRO FF.</i>	<i>VIOLACION DE MENOR</i>
<i>MALTRATO DE OBRA A PERSONAL DE</i>	<i>PORTAR OBJETOS ROBADOS</i>
<i>OFENSAS AL PUDOR Y BUENAS COST</i>	<i>USURPACION DE PROPIEDAD, DESCU</i>
<i>PORTE ARMA BLANCA</i>	<i>USURPACION NO VIOLENTA (ART. 4</i>
<i>PORTE PASTA BASE COCAINA</i>	<i>ROBO CON CASTRACION, MUTILACIO</i>

2.3.1 Correlaciones en nivel diario rezagadas

C1		C2		C3		C5	
C2(-357)	0.46	C2(-7)	0.61	C2(-7)	0.40	C5(-14)	0.68
C2(-364)	0.44	C2(-28)	0.58	C2(-364)	0.38	C5(-7)	0.67
C2(-7)	0.41	C2(-21)	0.56	C2(-14)	0.38	C5(-231)	0.66
C2(-350)	0.40	C2(-14)	0.54	C2(-175)	0.37	C5(-364)	0.66
C1(-7)	0.40	C2(-35)	0.54	C2(-28)	0.37	C5(-245)	0.66
C2(-35)	0.39	C2(-105)	0.54	C2(-105)	0.37	C5(-308)	0.66
C2(-343)	0.39	C2(-42)	0.53	C2(-161)	0.37	C5(-287)	0.65
C2(-308)	0.39	C2(-56)	0.52	C2(-357)	0.37	C5(-35)	0.65
C2(-259)	0.39	C2(-49)	0.52	C2(-21)	0.37	C5(-21)	0.65
C2(-329)	0.39	C2(-84)	0.52	C2(-182)	0.37	C5(-56)	0.65
C2(-147)	0.38	C2(-147)	0.52	C2(-154)	0.36	C5(-203)	0.65
C2(-140)	0.38	C2(-63)	0.51	C2(-42)	0.36	C5(-28)	0.65
C2(-28)	0.38	C2(-91)	0.51	C2(-147)	0.36	C5(-350)	0.65
C2(-252)	0.37	C2(-77)	0.51	C2(-56)	0.36	C5(-357)	0.65
C2(-266)	0.37	C2(-98)	0.51	C2(-343)	0.36	C5(-273)	0.65
C1(-21)	0.37	C2(-70)	0.50	C2(-196)	0.36	C5(-259)	0.65

El C4 no presentó correlaciones en nivel diario superiores a 0.35

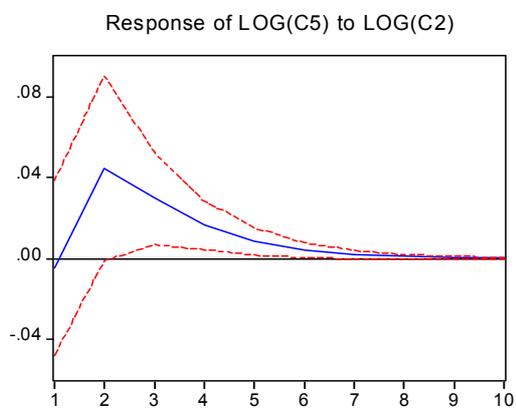
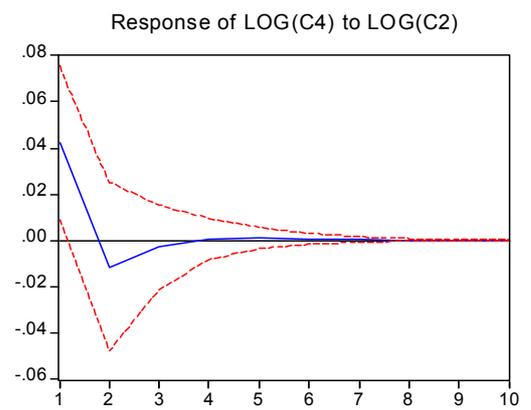
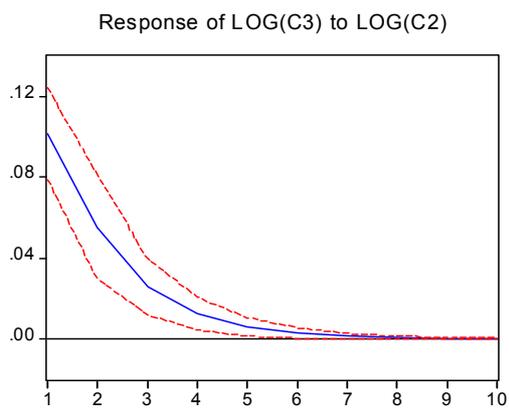
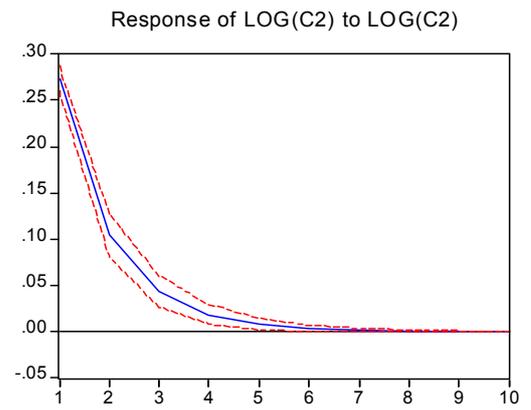
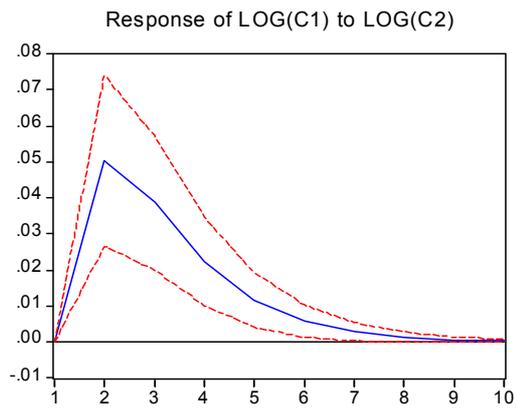
Las correlaciones en nivel semanal son demasiado bajas, por eso no se presentan en este anexo.

2.3.3 Correlaciones en nivel mensual rezagadas

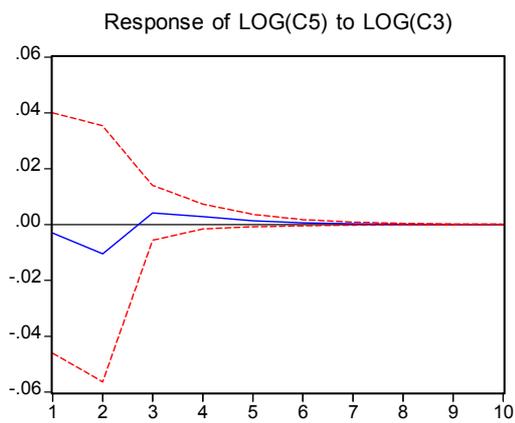
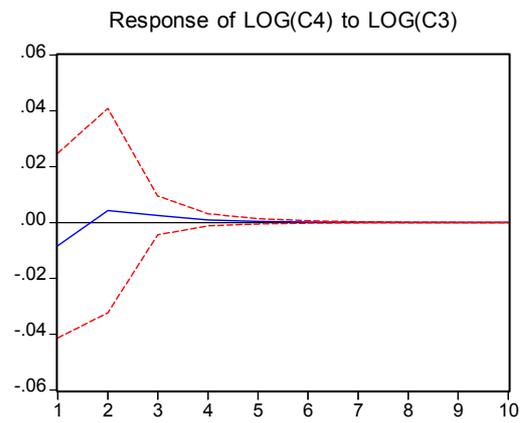
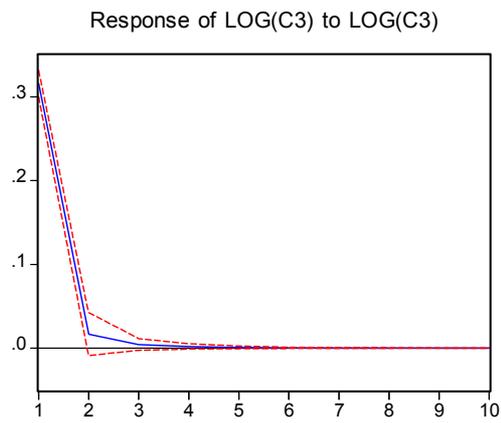
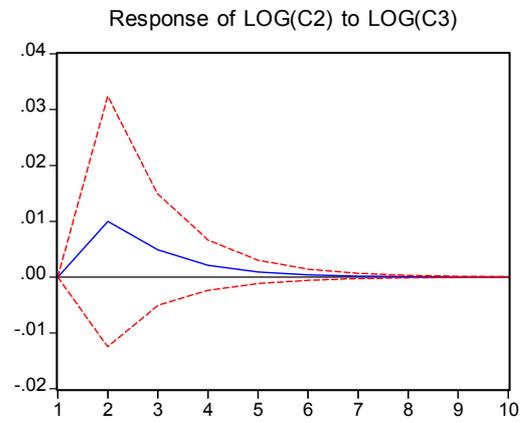
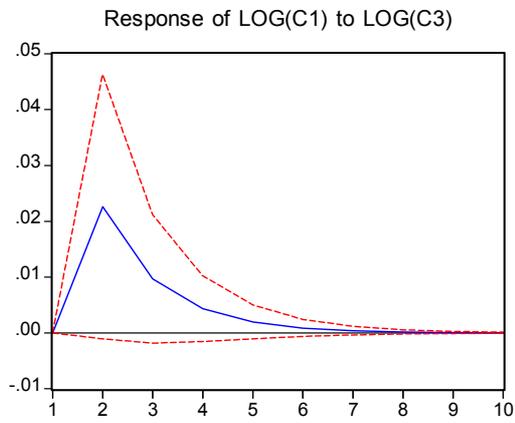
C1		C2		C3		C4		C5	
C4(-1)	0.47	C2(-1)	0.52	RM(-12)	0.58	C4(-1)	0.47	C3(-7)	0.44
C2(-11)	0.41	C2(-2)	0.41	C2(-12)	0.56	RM(-5)	0.45	C1(-8)	0.44
C1(-12)	0.42	C1(-10)	-0.64	RM(-1)	0.52	C5(-4)	0.44	C4(-8)	0.47
C2(-12)	0.76	C4(-10)	-0.43	C3(-12)	0.48	C2(-5)	0.41	C4(-9)	0.42
C3(-12)	0.68	C1(-11)	-0.51	C3(-1)	0.45	C4(-10)	-0.51	C5(-12)	0.45
RM(-12)	0.49	C4(-11)	-0.45	C1(-10)	-0.46	C5(-8)	-0.43		

FUNCIONES DE IMPULSO-RESPUESTA EN NIVEL DIARIO

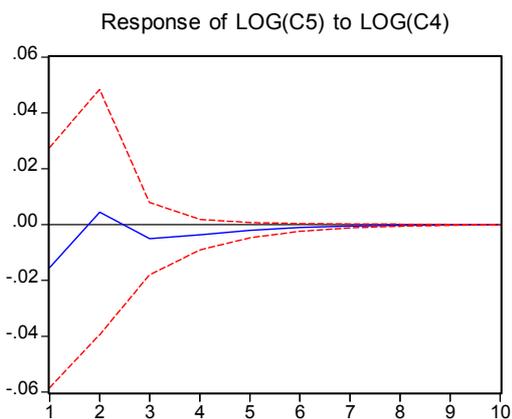
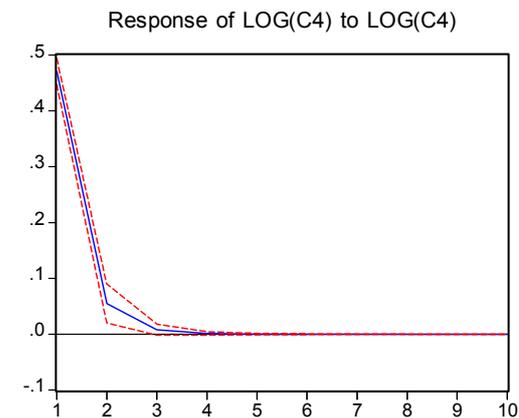
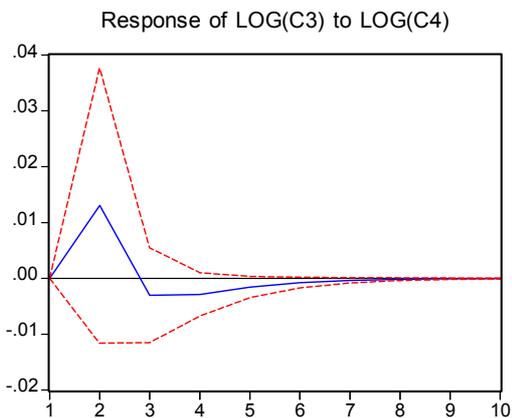
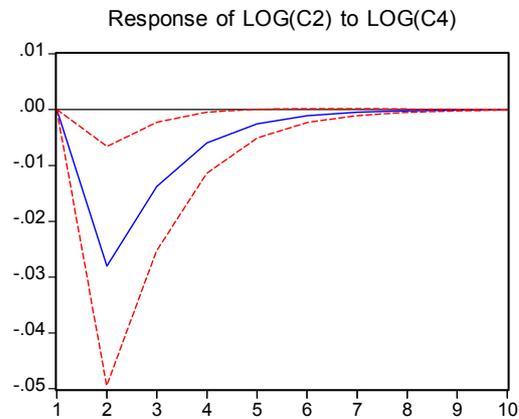
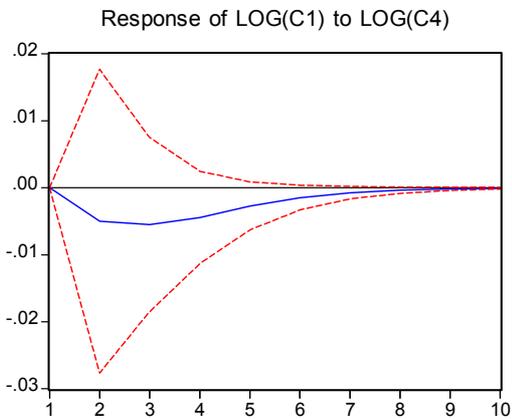
Response to Cholesky One S.D. Innovations ± 2 S.E.



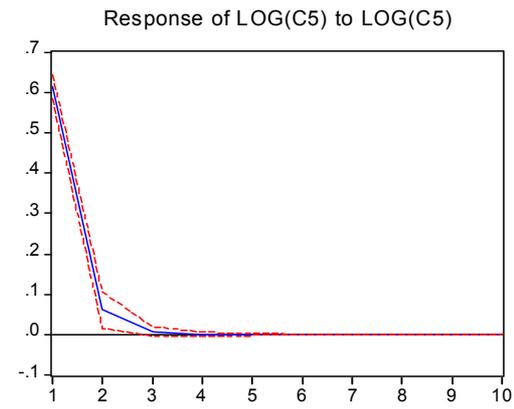
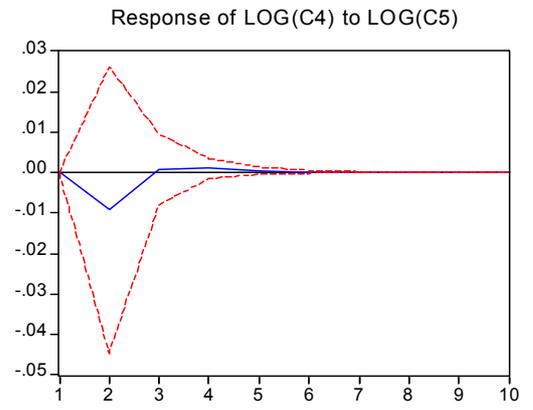
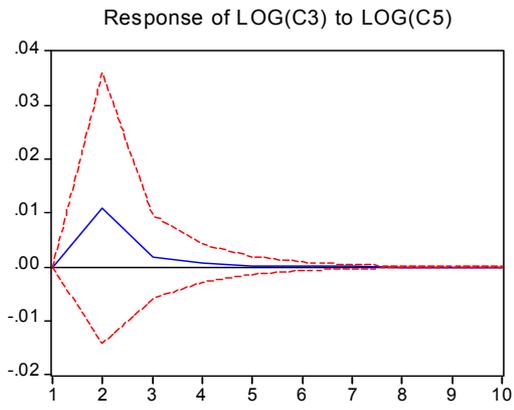
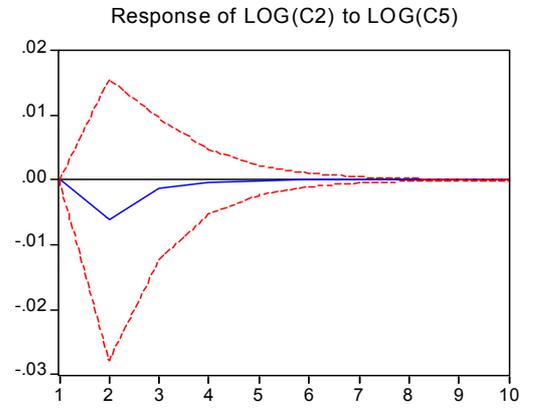
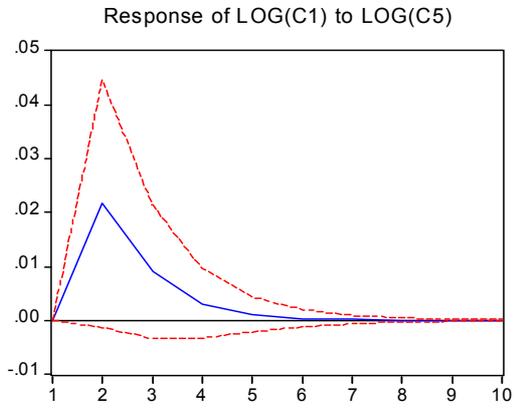
Response to Cholesky One S.D. Innovations ± 2 S.E.



Response to Cholesky One S.D. Innovations ± 2 S.E.

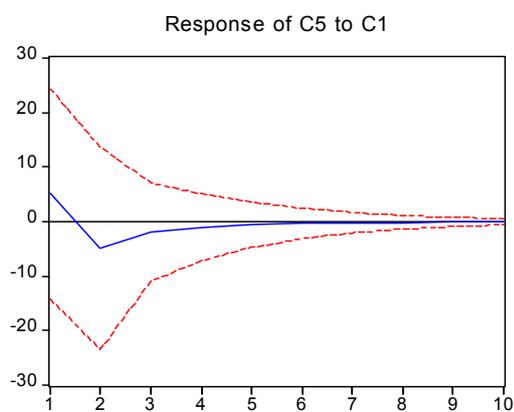
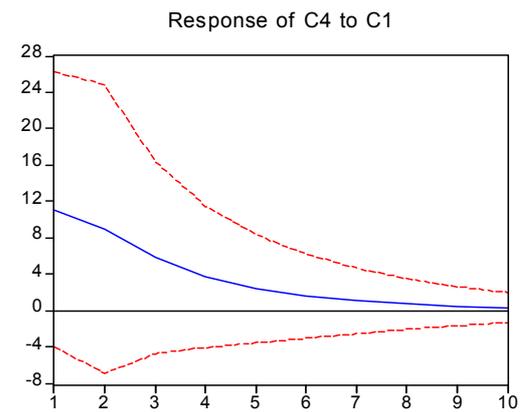
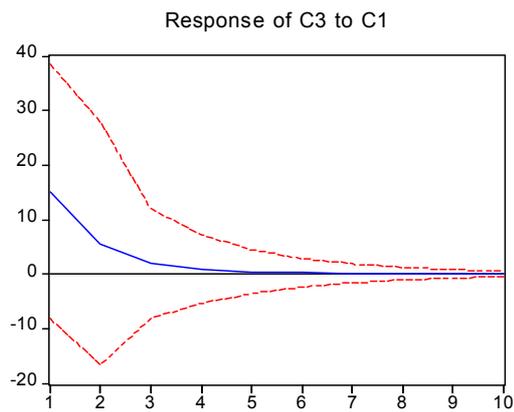
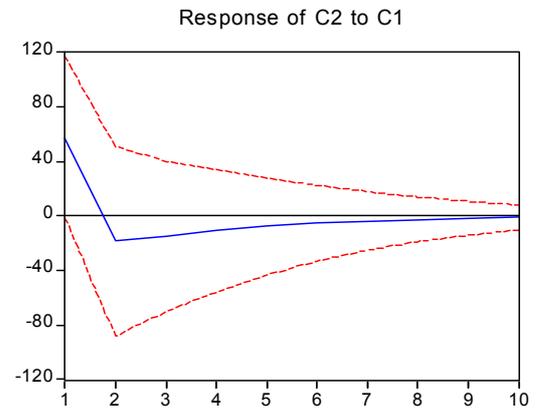
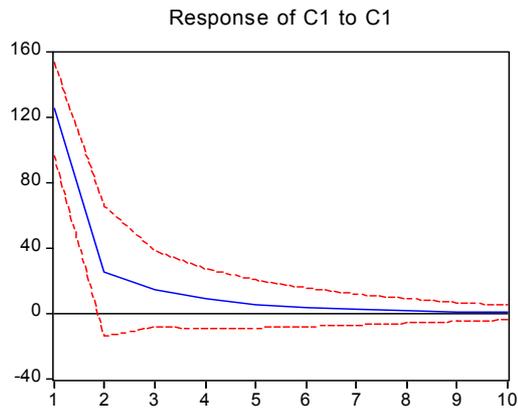


Response to Cholesky One S.D. Innovations ± 2 S.E.

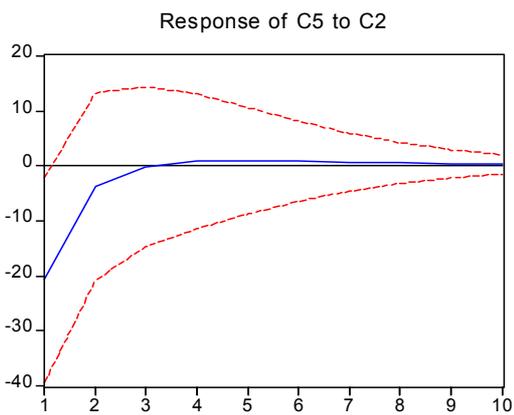
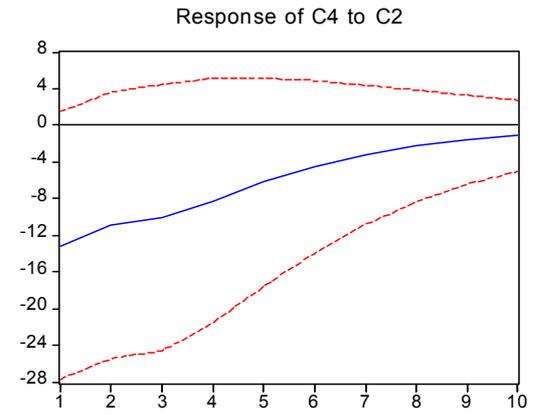
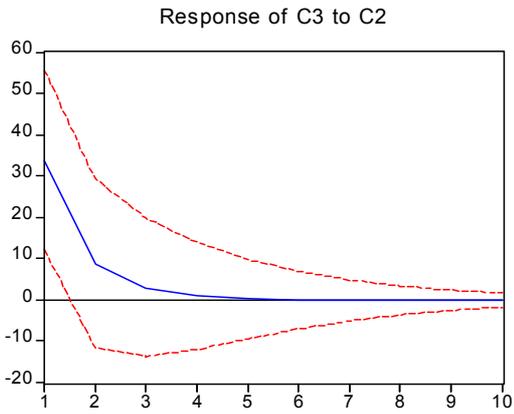
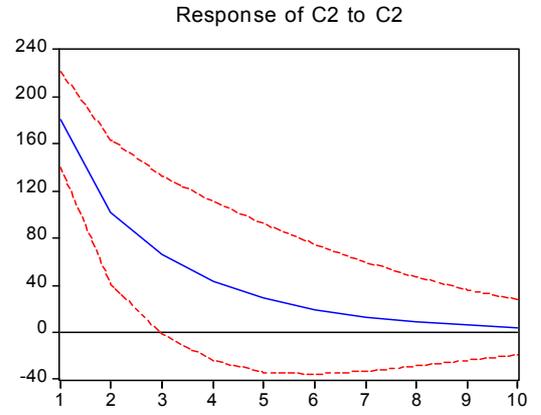
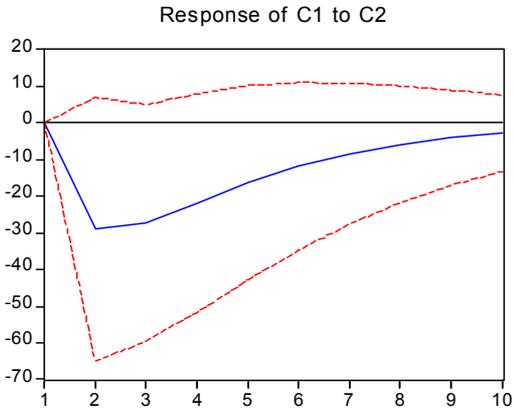


FUNCIONES DE IMPULSO-RESPUESTA EN NIVEL MENSUAL

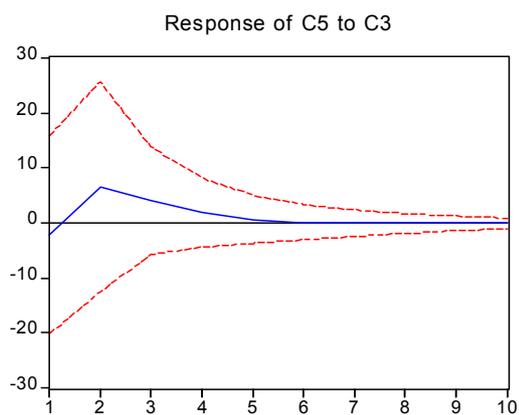
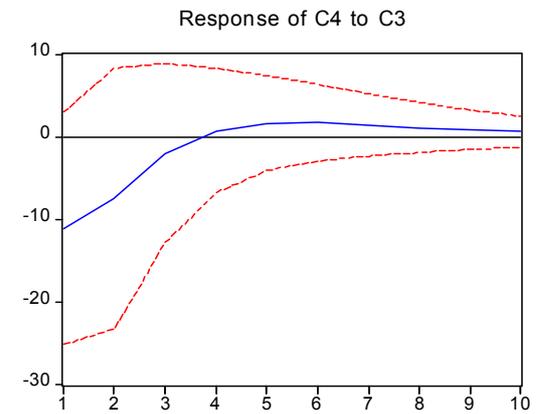
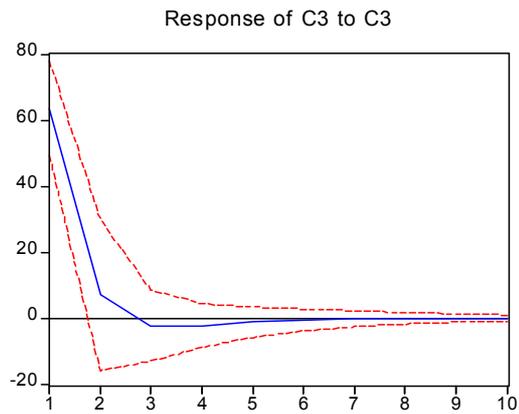
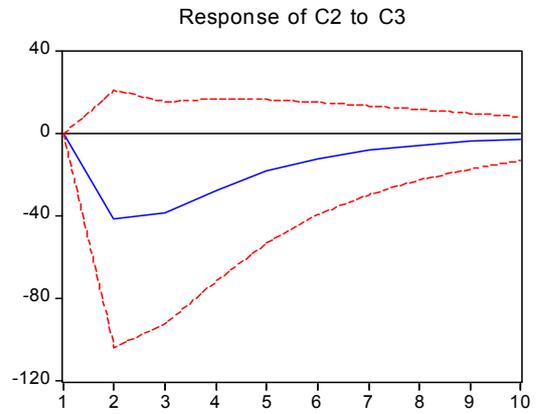
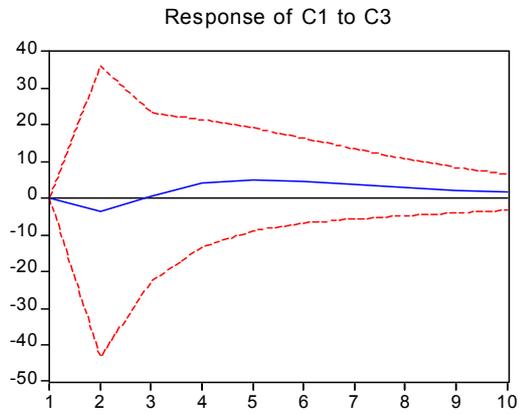
Response to Cholesky One S.D. Innovations ± 2 S.E.



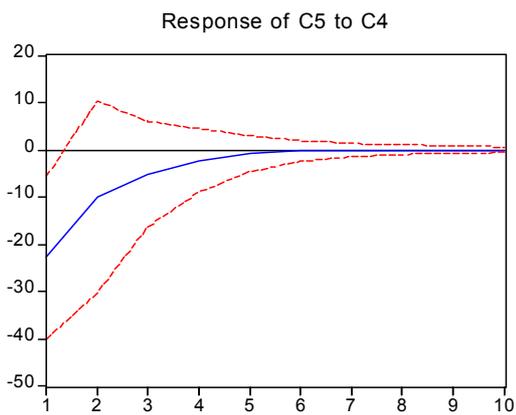
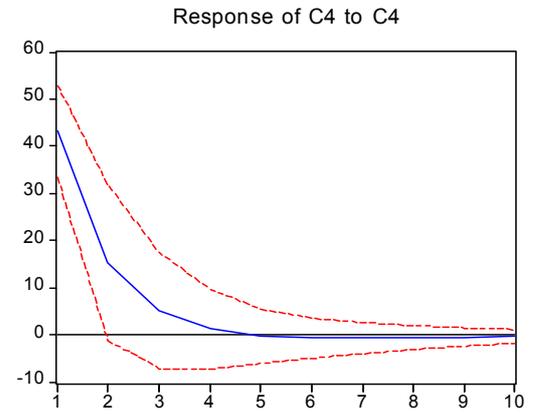
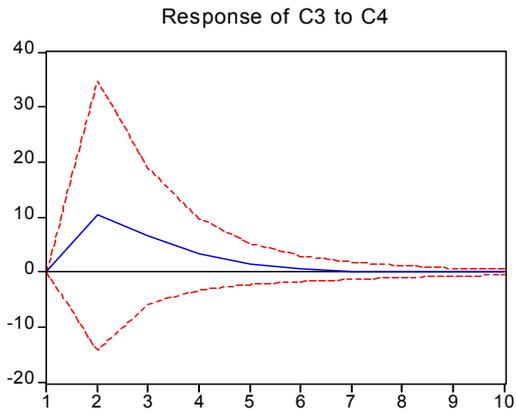
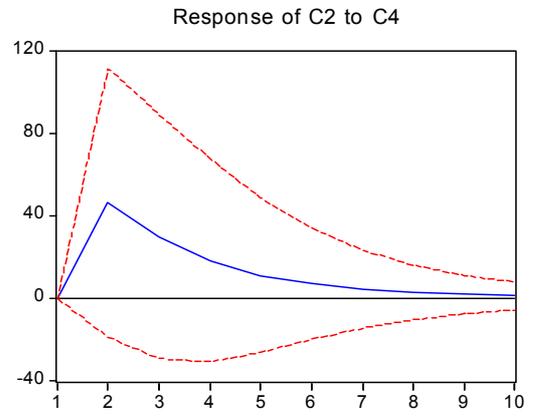
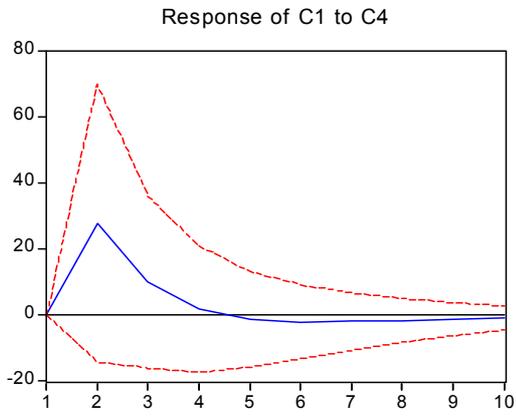
Response to Cholesky One S.D. Innovations ± 2 S.E.



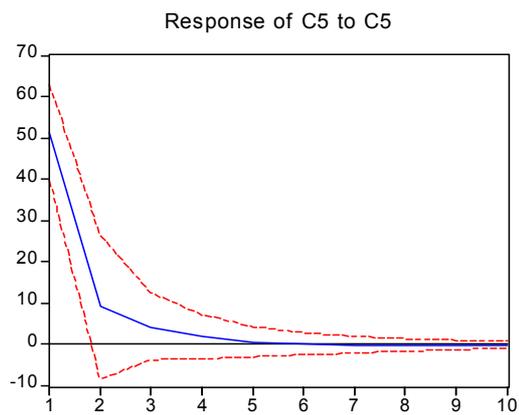
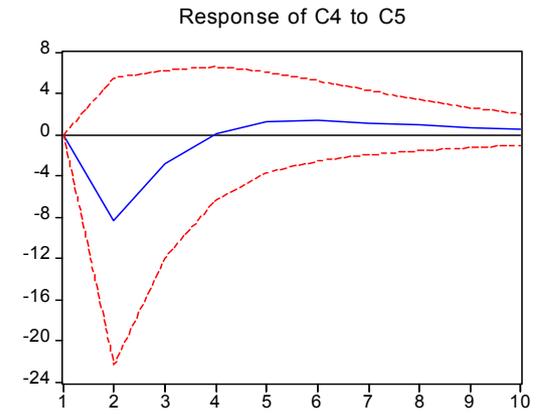
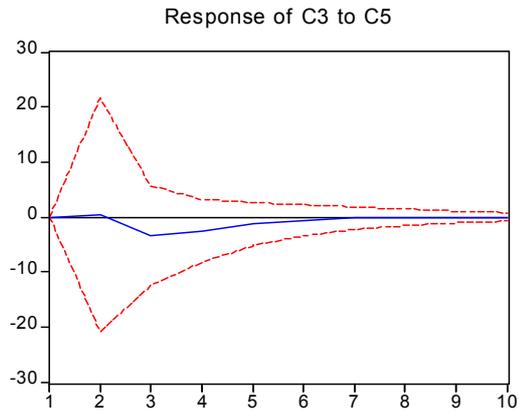
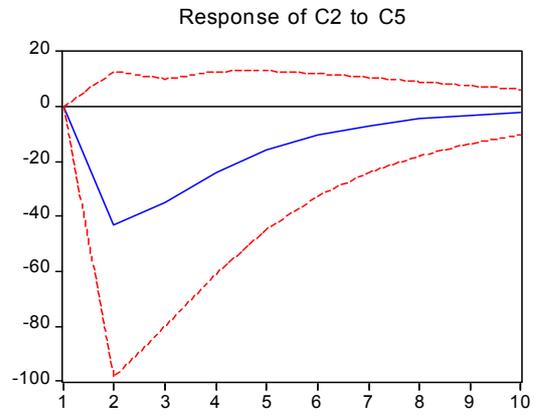
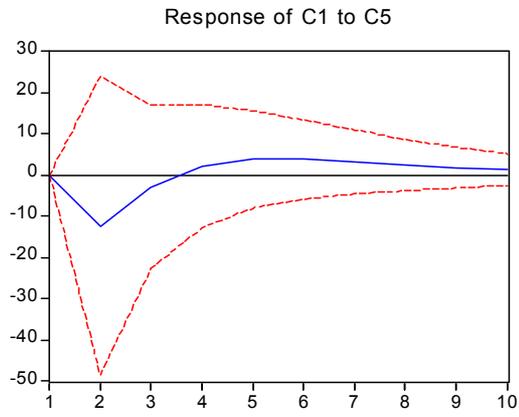
Response to Cholesky One S.D. Innovations ± 2 S.E.



Response to Cholesky One S.D. Innovations ± 2 S.E.



Response to Cholesky One S.D. Innovations ± 2 S.E.



Sistema diario

System: *SISTEMA2*

Estimation Method: Least Squares

Date: 08/09/05 Time: 01:34

Sample: 12/31/2001 6/30/2004

Included observations: 913

Total system (unbalanced) observations 4561

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-7.33	2.43	-3.02	0.00
C(3)	0.06	0.02	3.34	0.00
C(5)	0.17	0.03	5.45	0.00
C(44)	11.40	1.67	6.83	0.00
C(45)	15.50	1.81	8.58	0.00
C(46)	15.04	1.80	8.34	0.00
C(47)	14.48	1.78	8.12	0.00
C(48)	16.09	1.91	8.44	0.00
C(49)	5.69	1.54	3.69	0.00
C(68)	0.02	0.00	11.03	0.00
C(90)	-4.09	1.79	-2.29	0.02
C(2)	-1.14	4.23	-0.27	0.79
C(10)	0.27	0.03	8.47	0.00
C(11)	0.24	0.03	7.36	0.00
C(15)	0.26	0.03	8.53	0.00
C(51)	7.44	2.09	3.55	0.00
C(53)	5.93	2.07	2.86	0.00
C(54)	7.36	2.22	3.32	0.00
C(69)	0.02	0.00	4.72	0.00
C(91)	-9.87	3.12	-3.16	0.00
C(25)	2.22	1.86	1.19	0.23
C(28)	0.06	0.01	3.89	0.00
C(56)	8.76	1.29	6.77	0.00
C(57)	11.06	1.40	7.92	0.00
C(58)	12.25	1.39	8.84	0.00
C(59)	10.91	1.37	7.98	0.00
C(60)	13.52	1.45	9.33	0.00
C(61)	8.14	1.21	6.72	0.00
C(70)	0.01	0.00	7.41	0.00
C(33)	15.80	1.45	10.88	0.00
C(34)	0.10	0.03	2.89	0.00
C(62)	-16.20	0.95	-17.13	0.00
C(63)	-16.98	0.98	-17.32	0.00
C(64)	-17.32	0.97	-17.78	0.00
C(65)	-17.24	0.97	-17.79	0.00
C(66)	-16.46	0.95	-17.35	0.00
C(71)	0.01	0.00	5.16	0.00
C(72)	4.69	0.56	8.45	0.00
C(73)	0.03	0.01	3.83	0.00

C(74)	0.02	0.01	2.29	0.02
C(75)	0.02	0.01	2.87	0.00
C(76)	0.02	0.01	2.59	0.01
C(94)	-2.41	0.80	-3.00	0.00

Equation: $C1 = C(1) + C(3) * C2(-35) + C(5) * C1(-21) + C(44) * DD1 + C(45) * DD2 + C(46) * DD3 + C(47) * DD4 + C(48) * DD5 + C(49) * DD6 + C(68) * RM(-365) + C(90) * DFM$

Observations: 912			
<i>R-squared</i>	0.35	<i>Mean dependent var</i>	34.05
<i>Adjusted R-squared</i>	0.34	<i>S.D. dependent var</i>	14.20
<i>S.E. of regression</i>	11.50	<i>Sum squared resid</i>	119156.70
<i>Durbin-Watson stat</i>	1.34		

Equation: $C2 = C(2) + C(10) * C2(-28) + C(11) * C2(-21) + C(15) * C2(-105) + C(51) * DD2 + C(53) * DD4 + C(54) * DD5 + C(69) * RM(-365) + C(91) * DFM$

Observations: 912			
<i>R-squared</i>	0.48	<i>Mean dependent var</i>	68.86
<i>Adjusted R-squared</i>	0.48	<i>S.D. dependent var</i>	27.76
<i>S.E. of regression</i>	20.06	<i>Sum squared resid</i>	363395.70
<i>Durbin-Watson stat</i>	1.59		

Equation: $C3 = C(25) + C(28) * C2(-14) + C(56) * DD1 + C(57) * DD2 + C(58) * DD3 + C(59) * DD4 + C(60) * DD5 + C(61) * DD6 + C(70) * RM(-365)$

Observations: 912			
<i>R-squared</i>	0.25	<i>Mean dependent var</i>	26.21
<i>Adjusted R-squared</i>	0.25	<i>S.D. dependent var</i>	10.50
<i>S.E. of regression</i>	9.12	<i>Sum squared resid</i>	75114.27
<i>Durbin-Watson stat</i>	1.76		

Equation: $C5 = C(33) + C(34) * C5(-14) + C(62) * DD1 + C(63) * DD2 + C(64) * DD3 + C(65) * DD4 + C(66) * DD5 + C(71) * RM(-365)$

Observations: 912			
<i>R-squared</i>	0.65	<i>Mean dependent var</i>	10.08
<i>Adjusted R-squared</i>	0.65	<i>S.D. dependent var</i>	11.00
<i>S.E. of regression</i>	6.49	<i>Sum squared resid</i>	38095.78
<i>Durbin-Watson stat</i>	1.63		

**Equation: C4=C(72)+C(73)*C2(-161)+C(74)*C2(-308)+C(75)*C2(-357)
+C(76)*C2(-364)+C(94)*DFM**

<i>Observations:</i> 913			
<i>R-squared</i>	0.13	<i>Mean dependent var</i>	10.68
<i>Adjusted R-squared</i>	0.13	<i>S.D. dependent var</i>	5.57
<i>S.E. of regression</i>	5.20	<i>Sum squared resid</i>	24483.87
<i>Durbin-Watson stat</i>	1.81		

Sistema mensual

<i>System:</i> SIS_MEN_3				
<i>Estimation Method:</i> Least Squares				
<i>Date:</i> 08/07/05 <i>Time:</i> 14:42				
<i>Sample:</i> 2001M09 2004M06				
<i>Included observations:</i> 34				
<i>Total system (unbalanced) observations:</i> 157				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	102.74	153.99	0.67	0.51
C(7)	0.47	0.08	6.13	0.00
C(2)	5256.48	580.50	9.06	0.00
C(21)	-0.88	0.20	-4.43	0.00
C(22)	-0.90	0.20	-4.41	0.00
C(23)	4.51	2.64	1.71	0.09
C(24)	-10.81	2.57	-4.20	0.00
C(3)	157.95	316.33	0.50	0.62
C(28)	-0.21	0.18	-1.20	0.23
C(29)	0.14	0.07	1.95	0.05
C(31)	0.02	0.01	2.08	0.04
C(32)	0.02	0.01	2.06	0.04
C(33)	-1.81	1.00	-1.81	0.07
C(4)	545.57	58.76	9.29	0.00
C(47)	-0.32	0.13	-2.54	0.01
C(48)	-0.42	0.17	-2.40	0.02
C(60)	43.58	22.12	1.97	0.05
C(5)	13.05	77.36	0.17	0.87
C(64)	0.21	0.08	2.54	0.01
C(65)	0.16	0.06	2.81	0.01
C(71)	-74.51	22.98	-3.24	0.00

Equation: C1=C(1)+C(7)*C2(-12)

Observations: 30			
R-squared	0.57	Mean dependent var	1035.20
Adjusted R-squared	0.56	S.D. dependent var	198.91
S.E. of regression	132.26	Sum squared resid	489788.30
Durbin-Watson stat	1.36		

Equation: C2=C(2)+C(21)*C1(-10)+C(22)*C1(-11)+C(23)*DESOCUPADOS(-2)+C(24)*DESOCUPADOS(-3)

Observations: 31			
R-squared	0.75	Mean dependent var	2088.87
Adjusted R-squared	0.71	S.D. dependent var	322.20
S.E. of regression	173.11	Sum squared resid	779119.90
Durbin-Watson stat	1.63		

Equation: C3=C(3)+C(28)*C1(-12)+C(29)*C2(-12)+C(31)*RM(-12)+C(32)*RM(-1)+C(32)*DESOCUPADOS(-2)+C(33)*DESOCUPADOS(-3)

Observations: 30			
R-squared	0.65	Mean dependent var	796.93
Adjusted R-squared	0.58	S.D. dependent var	126.04
S.E. of regression	81.94	Sum squared resid	161120.50
Durbin-Watson stat	2.54		

Equation: C4=C(4)+C(47)*C5(-8)+C(48)*C4(-10)+C(60)*DE2

Observations: 32			
R-squared	0.43	Mean dependent var	324.22
Adjusted R-squared	0.37	S.D. dependent var	58.59
S.E. of regression	46.47	Sum squared resid	60462.96
Durbin-Watson stat	1.89		

Equation: C5=C(5)+C(64)*C3(-7)+C(65)*C1(-8)+C(71)*DE1

Observations: 34			
R-squared	0.49	Mean dependent var	318.76
Adjusted R-squared	0.44	S.D. dependent var	68.04
S.E. of regression	51.03	Sum squared resid	78108.22
Durbin-Watson stat	1.91		