

II Jornadas Internacionales de Estadística Aplicada 5 y 6 de diciembre de 2019

Indicador de Demanda Laboral para Salta.

Una aproximación al estudio de la Volatilidad en series de alta frecuencia.

Gastón J. Carrazán Mena y Fernando H. Romero

Instituto de Investigaciones Económicas – Facultad de Cs. Económicas Universidad Nacional de Salta

Datos de contacto: gcarrazan@gmail.com / 3875161478

RESUMEN

El Indicador de Demanda Laboral (IDL) se desarrolla en el ámbito del Instituto de Investigaciones Económicas. Utiliza un procedimiento de conteo de la cantidad de avisos de solicitudes laborales que se realiza para la ciudad de Salta. La finalidad de este artículo es explorar las características de la serie temporal que compone el IDL y, en particular, aplicar un ejercicio de modelado de la volatilidad intrínseca que se le atribuye por ser una serie de alta frecuencia.

Palabras clave: Demanda Laboral. Series de Tiempo. Volatilidad. Pronósticos.

INTRODUCCIÓN

El Indicador de Demanda Laboral (IDL) busca aproximar el comportamiento de la actividad del mercado de trabajo, como también aproximar las condiciones económicas en general y lo relacionado a la demanda de trabajo privado en particular, Bauzá et al (2008). Este procedimiento se alcanza mediante mediciones sobre la cantidad de avisos que solicitan empleados (DGECE, s.f.) en un ámbito geográfico y durante un periodo de tiempo determinados.

Así el índice se puede decir que es un reflejo de la demanda laboral, permitiendo identificar cuáles son las características más solicitadas por los empleadores. De esta manera se convierte en una importante herramienta de información para quienes buscan trabajo activamente como así también para la implementación de políticas económicas relacionadas al mercado laboral.

El IDL, entonces, tiene como objetivo principal conocer y caracterizar al conjunto de empleos solicitados en la ciudad de Salta, y la manera de computarlo es:

$$IDL_t^{Salta} = \left(\frac{q_{i,t,h}}{\bar{q}_0} \right) * 100$$

donde: q representa la cantidad de pedidos de trabajo solicitados durante un domingo i del mes t del año h ; \bar{q}_0 es el promedio semanal de la cantidad de pedidos de trabajo solicitados los días domingo del año 2007.

La selección realizada sobre los días domingos como los días mensuales de carga, obedece a que es el día de mayor publicación de avisos en el principal diario local de Salta (Diario El Tribuno).

METODOLOGÍA

Descripción de la Serie del IDL

La serie se compone de 521 observaciones, tomadas en frecuencia semanal, los días Domingo de cada semana, en el período 05/01/2003 al 30/12/2012, tal como se enseña en el Gráfico 1 que acompaña.

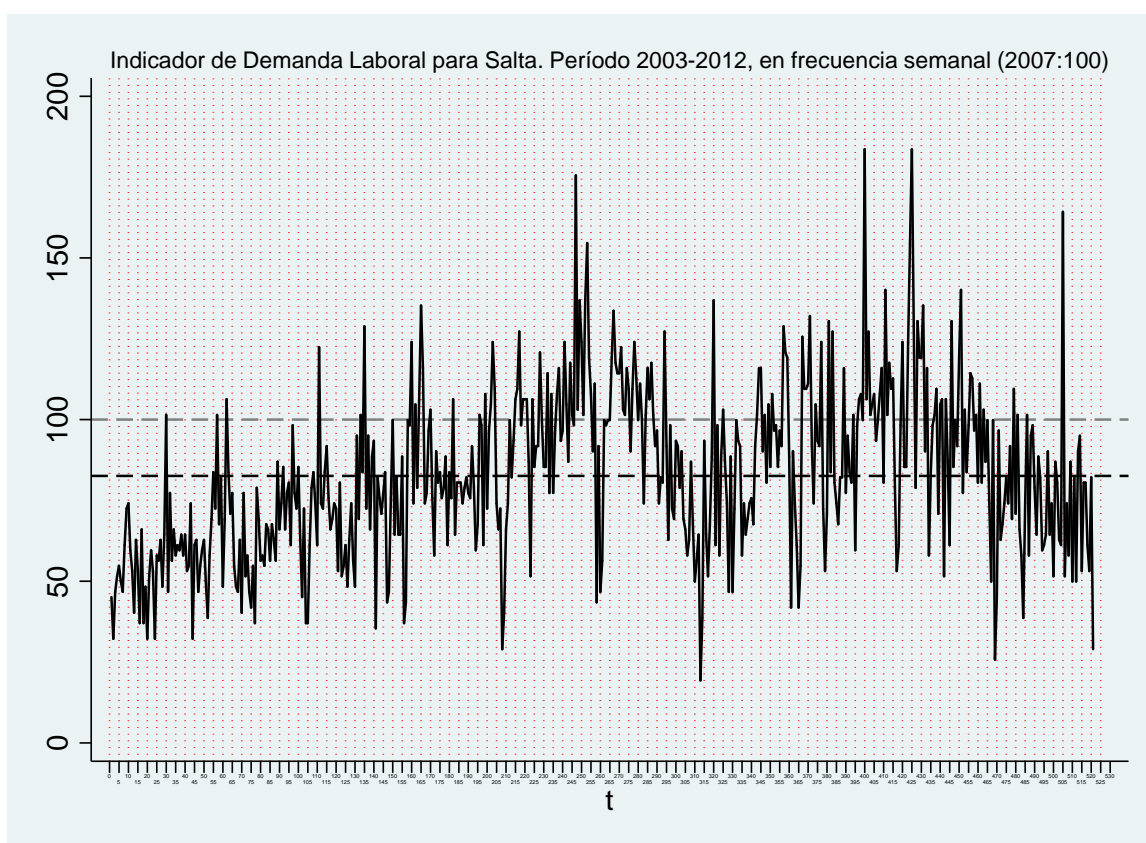


Gráfico 1: Indicador de Demanda Laboral para Salta.
Período 2003-2012, en frecuencia semanal (2007:100).
Fuente: elaboración propia en base a datos del proyecto 2610 del CIUNSa y Cuadro 1.

Como se, el período base de comparación es el promedio de los Domingos del año 2007, que se corresponde a aproximadamente 62.096 avisos. Respecto de ellos, se relativizan los avisos contados en cada Domingo del rango, construyéndose el IDL.

Variable	Obs	Mean	Std. Dev.	Min	Max
Avisos	521	51.27639	16.11295	12	114

t	521	261	150.544	1	521
base	521	62.09615	0	62.09615	62.09615
IDL	521	82.57579	25.94839	19.32487	183.5863

Cuadro1: Estadísticos Descriptivos del IDL
Fuente: elaboración propia en base a datos del Proyecto /19 CIUNSa.

Subperíodo	Fechas	Media	Varianza	n
1	5/1/2003 al 3/7/2005	62.683	263.275	130
2	10/7/2005 al 30/12/2007	88.696	590.276	130
3	6/1/2008 al 27/6/2010	89.340	545.388	130
4	4/7/2010 al 30/12/2012	89.737	775.771	131
Total	5/1/2003 al 30/12/2012	82.576	673.319	521

Cuadro 2: Medias y varianzas en subperíodos del IDL
Fuente: elaboración propia en base a datos del proyecto 2610 del CIUNSa

Aquí puede observarse que, aunque la media sea relativamente similar en 4 subperíodos delimitados de similar tamaño, la varianza no se corresponde de la misma manera. Es decir, podemos apreciar que la serie se estaciona en la media, mas no es de varianza constante.

El Correlograma y el Correlograma parcial de la serie es el que se enseña a continuación.

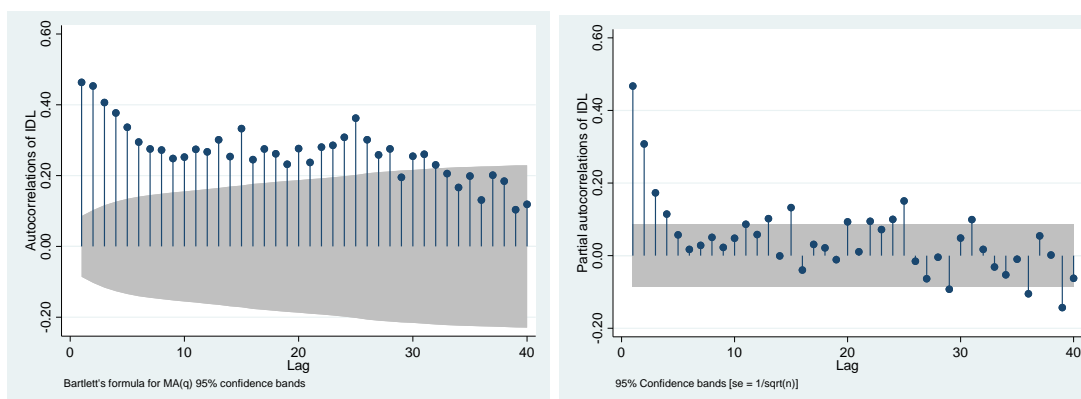


Gráfico 2: AC y PAC del IDL
Fuente: elaboración propia en base a datos del proyecto 2610 del CIUNSa u Cuadros 1 y 2.

En los mismos puede observarse que no hay un decaimiento ni geométrico u oscilante, con lo cual no podría identificarse un proceso ARMA del tipo Box Jenkins.

Al realizar un test de raíz unitaria del tipo Augmented Dickey Fuller (ADF)

Modelo	Valor tau	Valor crítico al 1% de significancia	Valor crítico al 5% de significancia	Valor crítico al 10% de significancia
--------	-----------	--------------------------------------	--------------------------------------	---------------------------------------

Sin constante, sin tendencia	-3.569	-2.580	-1.950	-1.620
Con constante, con tendencia	13.67 9	-2.334	-1.648	-1.283
Con constante, con tendencia	14.77 8	-3.960	-3.410	-3.120

Cuadro 3: ADF para el IDL

Fuente: elaboración propia en base a datos del proyecto 2610 del CIUNSa.

En todos los casos anteriores, la decisión será rechazar la hipótesis nula y, por ende, aceptar que la serie en cuestión no tiene raíz unitaria y por ende es estacionaria.

Esta conclusión va en la misma línea que lo que se puede observar por simple inspección en el Gráfico 1, donde pareciera existir una sola media para toda la secuencia.

Sin embargo, como se vio en el cuadro 2, la varianza no pareciera ser estable en el tiempo. Es por ello que se utilizará un modelo ARCH para estudiar la evolución de la Serie.

Por otra parte, la distribución de la serie denota colas pesadas, tal como puede observarse en el Gráfico 3.

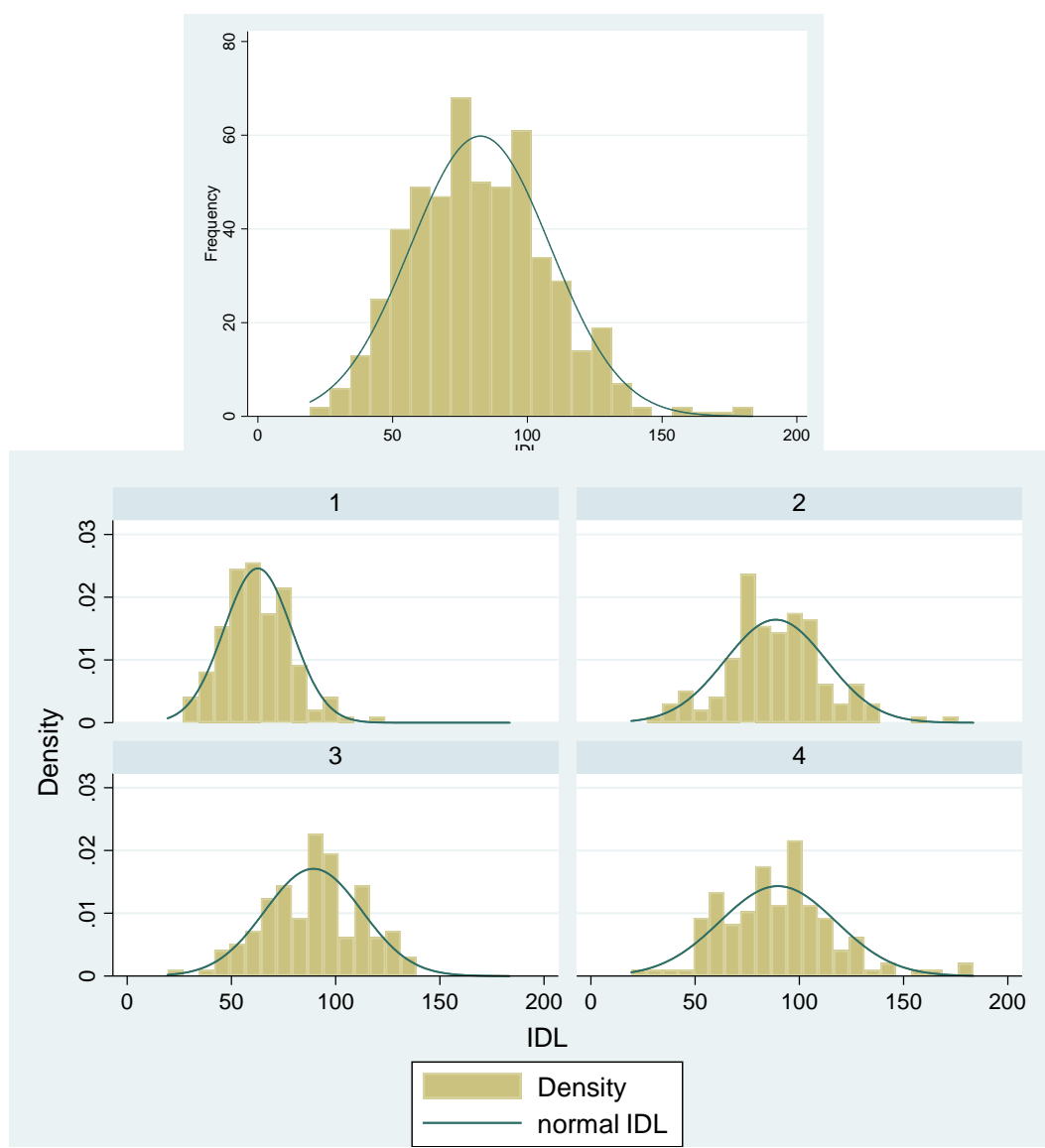


Gráfico 3: Histograma total y por subperíodos.

Donde la segunda parte del gráfico de histogramas, está dividido por los subperíodos que se enseñan en el cuadro 2.

En esta parte del gráfico a su vez puede observarse la gran dispersión de la distribución, en particular en los subperíodos 2 y 4.

DESARROLLO

Modelo ARCH

Entonces suponemos que, en lugar de una media condicional con el tiempo, lo que existe es una varianza condicionada con el tiempo. De esta manera, podemos plantear una primera especificación ARCH (1) tal que:

$$y_t = \phi + e_t$$

$$e_t | I_{t-1} \sim N(0, h_t)$$

$$h_t = \alpha_0 + \alpha_1 e_{t-1}^2, \quad \alpha_0 > 0, \quad 0 \leq \alpha_1 < 1$$

Donde denotamos que los errores dependen de un conjunto de información pasado, y la varianza tiene una estructura que depende a su vez de efectos rezagados.

Así, se puede expresar la varianza condicional normal del tipo:

$$\left(\frac{e_t}{\sqrt{h_t}} | I_{t-1} \right) = z \sim N(0, 1)$$

Donde a su vez podemos expresar

$$E(e_t) = E(z_t) E(\sqrt{\alpha_0 + \alpha_1 e_{t-1}^2})$$

$$E(e_t^2) = E(z_t^2) E(\alpha_0 + \alpha_1 e_{t-1}^2) = \alpha_0 + \alpha_1 E(e_{t-1}^2)$$

Test LM para probar la presencia de efectos ARCH

Consideremos una especificación del tipo AR(1) en los residuos.

$$y_t = \phi + e_t$$

$$e_t = \rho e_{t-1} + v_t, \quad |\rho| < 1$$

$$v_t \sim N(0, \sigma_v^2)$$

Si realizamos la estimación, rescatando los residuos, puede plantearse el test del multiplicador de Lagrange para detectar efectos ARCH, a partir de la significancia estadística individual. A saber,

$$\hat{e}_t^2 = \gamma_0 + \gamma_1 \hat{e}_{t-1}^2 + v_t$$

$$H_0 : \gamma_1 = 0 \quad H_1 : \gamma_1 \neq 0$$

En el Anexo se enseñan las regresiones auxiliares que permiten hacer la estimación y la prueba LM

$$\hat{e}_t^2 = 482 + 0.082 \hat{e}_{t-1}^2$$

$t \quad (1.86)$

$$R^2 = 0.0067$$

El estadístico-t sugiere un coeficiente de primer orden significativo

El tamaño muestral es de 521, dando una valor del LM-test de $(T - q)R^2 = 3.484$

Comparando el valor computado por el test con el valor crítico (al 5%) de una distribución $\chi^2(1)$ (3.41) conduce al rechazo de la nula.

Luego, los residuales muestran presencia de efectos ARCH(1).

Estimación ARCH(1)

```
(setting optimization to BHHH)
Iteration 0: log likelihood = -2429.7469
Iteration 1: log likelihood = -2429.6387
Iteration 2: log likelihood = -2429.6328
Iteration 3: log likelihood = -2429.6325
Iteration 4: log likelihood = -2429.6325
(switching optimization to BFGS)
Iteration 5: log likelihood = -2429.6325
```

ARCH family regression

Sample: 1 - 521	Number of obs	=	521
Distribution: Gaussian	Wald chi2(.)	=	.
Log likelihood = -2429.632	Prob > chi2	=	.

IDL	OPG		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
IDL						
_cons	81.9814	1.198773	68.39	0.000	79.63185	84.33095
ARCH						
arch						
L1.	.1630235	.0643696	2.53	0.011	.0368615	.2891855
_cons	562.9102	44.71932	12.59	0.000	475.262	650.5585

El coeficiente del ARCH es estadísticamente significativo al 5% de significancia.

Es decir, la forma de la varianza estimada será del tipo:

$$\hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 * \hat{e}_{t-1}^2 = 562.91 + 0.1630 * \hat{e}_{t-1}^2$$

$t \quad (2.53)$

CONCLUSIONES

EL artículo corrobora la existencia de heterocedasticidad en la varianza, que se le atribuye a la volatilidad intrínseca de una serie de alta frecuencia, como es la del idl semanal. Razón por la cual, intentos de modelarla utilizando el procedimiento tradicional de Box Jenkins no sería posible. En su lugar, se encuentra un primer modelo que incluye una forma autorregresiva en la varianza. Los resultados de la estimación del procedimiento conocido como ARCH(1) son satisfactorios y permitirán explorar especificaciones más complejas.

BIBLIOGRAFÍA

“La Demanda de Empleo y el Ciclo Económico en Salta, a partir del Índice de Demanda Laboral (IDL) en el período 1998-2019”. Proyecto tipo B 2610/0 del Consejo de Investigación de la Universidad Nacional de Salta (CIUNSa). Enero 2019 – Actualidad.

Notas del curso Volatilidad en series de Tiempo. Instituto de Investigaciones Estadísticas de la Universidad Nacional de Tucumán. Abril J.C. y M.M. Abril (2019).

Statistics with Stata, Hamilton (2003).

Time Series and Panel data econometrics. Pesaran M. H. (2015).

Una aproximación al estudio del Mercado Laboral y el Ciclo Económico Real, mediante un Indicador de Demanda Laboral para Salta. X Congreso Latinoamericano de Sociedades de Estadística (CLATSE). Facultad de Ciencias Económicas – Universidad Nacional de Córdoba. Carrazán G. y F. Romero (2012). Ciudad de Córdoba, Argentina. Octubre de 2012.

Anexo - Test LM

Paso 1) Estimación AR(1)

```
. arima IDL, arima(1,0,0)
```

```
(setting optimization to BHHH)
```

```
Iteration 0: log likelihood = -2371.4833
```

```
Iteration 1: log likelihood = -2371.4794
```

```
Iteration 2: log likelihood = -2371.4791
```

```
Iteration 3: log likelihood = -2371.4791
```

```
Iteration 4: log likelihood = -2371.4791
```

ARIMA regression

Sample: 1 - 521

Number of obs = 521

Wald chi2(1) = 155.12

Log likelihood = -2371.479

Prob > chi2 = 0.0000

IDL	OPG		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
IDL						
_cons	82.42339	1.96656	41.91	0.000	78.56901	86.27778
ARMA						
ar						
L1.	.4681834	.0375904	12.45	0.000	.3945075	.5418593
/sigma	22.9336	.6148017	37.30	0.000	21.72861	24.13859

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

Paso 2: rescatar los residuos (r) y regresarlos contra los residuos rezagados un período (L.r)

```
. reg r L.r
```

Source	SS	df	MS	Number of obs	=	520
Model	5621.33036	1	5621.33036	F(1, 518)	=	10.89
Residual	267307.032	518	516.036741	Prob > F	=	0.0010
				R-squared	=	0.0206
				Adj R-squared	=	0.0187
Total	272928.362	519	525.87353	Root MSE	=	22.716

r	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
r						
L1.	-.1438973	.0435987	-3.30	0.001	-.2295493	-.0582453
_cons	.1243967	.9961994	0.12	0.901	-1.832691	2.081484

Siendo rsq los residuos cuadráticos

```
. reg rsq L.rsq
```

Source	SS	df	MS	Number of obs	=	520
Model	2720968.87	1	2720968.87	F(1, 518)	=	3.47
Residual	405805669	518	783408.628	Prob > F	=	0.0629
				R-squared	=	0.0067
				Adj R-squared	=	0.0047
Total	408526638	519	787141.884	Root MSE	=	885.1

rsq	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
rsq						
L1.	.0820762	.0440402	1.86	0.063	-.0044432	.1685957
_cons	482.022	45.11355	10.68	0.000	393.394	570.6501