Ingeniería Industrial
PONTIFICIA UNIVERSIDAD CATOLICA DE VALPARAISO

ESCUELA DE COMERCIO
PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

Seminario Doctorado en Ingeniería Industrial
"Econometric Modeling and Solving Social Problems"

Hanns DE LA FUENTE-MELLA, Ing., Dr.

ESCUELA DE COMERCIO

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

NUCLEO DE INVESTIGACION DATA ANALYTICS

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

CASE 1
Analysis of the Factors of Chilean City Hall Using Econometric Modeling and Stochastic Frontier

Hanns DE LA FUENTE-MELLA, Ing. Dr.
Mauricio ALVARADO-MARTINEZ, Std.

Introduction

- The City Hall in Chile are fundamental by the **decentralization of the country**
- Represent a way **to be more connected to and with people**, their problems, needs and desires
- In Chile there are **346 City Hall** belonging to the different communes (different characteristics)
- The factors that influence the efficiency of the City Hall in Chile will be determined, based on the **quality of life index** of the cities
- For the above, an **econometric model** was developed that explains the determinants of the efficiency of the City Hall in Chile
- Therefore, the elements that make it efficient and **improve the quality of life of citizens** are fundamentals for to be identified

Mission of the Town Hall in Chile
meet the needs of the local community *and ensure their participation in the economic, social and cultural progress of the city or group of cities,* ***with a sense of efficiency and equity, improving the quality of life of its inhabitants***

Theoretical Framework

- *Quality of Life Index (QoL)*

  - "**Multidimensional concept**, includes welfare aspects (well-being) and social policies: **material and non-material**, objective and subjective, individual and collective" (*Palomba, 2002*)
  - "Quality of life is the subjective evaluation of the **satisfactory character of life as a whole**" (*Szalai, 1980*)
  - "Quality of life is the patient's appreciation of his life and **satisfaction with his current level of functioning compared to what perceives as possible or ideal**" (*Celia and Tulsky, 1990*)
  - "**It is a subjective sensation** of physical, psychological and **social well-being**. It includes, as subjective aspects: intimacy, emotional expression, perceived safety, personal productivity and objective health. As objective aspects: material well-being, harmonious relations with the physical and social environment and with the community, and health objectively perceived" (*Ardilla, 2003*)
  - "It is a composite measure of physical, mental and social well-being, **as perceived by each individual and each group; and of happiness, satisfaction and reward**" (*Levy and Anderson, 1980*)

Theoretical Framework

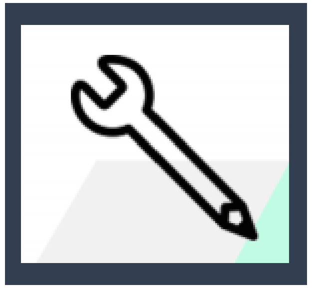
- *Quality of Life Index (QoL) (Architecture-PUC)*

  - ***Its a survey of the inhabitants of the communes (Chilean City Hall) where they are asked by:***

housing and environment

health

sociocultural conditions



business environment

work condition

connectivity

Hypothesis of this Research

1

there is not a positive relation between City Hall Budget and Own City Hall Incomes with the Efficiency

Efficiency

2

overcrowding has a negative relation with the QoL in the City Hall in Chile

2A

rate of domestic violence has a negative relation with the QoL in the City Hall in Chile

3

population density has a negative relationship with the QoL in the City Hall (*Kerstens, 1996; Sampaio de Sousa and Stosic, 2003*)

Efficiency

4

the best efficiency indexs are in the central/south area of the country

5

higher per capita income and population wealth have a negative effect on efficiency

Objective

general objetive

determine the factors that influence the quality of life of the City Hall in Chile

especific objective

estimate the efficiency of the town hall of Chile based on the index of quality of urban life

Methodology

- ➢ Descriptive Analysis
- ➢ Regressions Models (Stepwise)
- ➢ Final Regression Model
- ➢ Cluster Analysis
- ➢ Technical Efficiency (Stochastic Frontier Analysis)

- ❑ 91 city hall in Chile, 2018
- ❑ city hall over 50,000 population
- ❑ 10 metropolitan areas
- ❑ 25 intermediates cities
- ❑ all regional capitals

http://datos.gob.cl/

https://reportescomunales.bcn.cl/

Max: Las Condes
Min: Lo Espejo

Max: Lo Espejo
Min: Vitacura

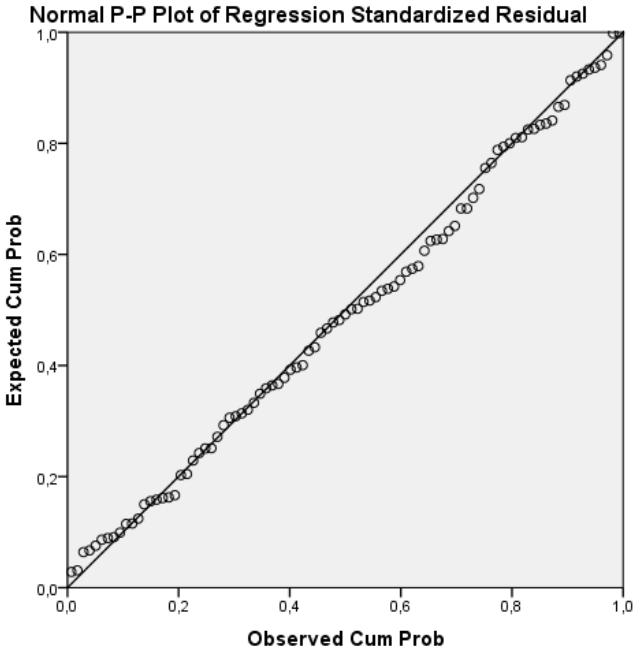
Max: Puente Alto
Min: Puerto Varas

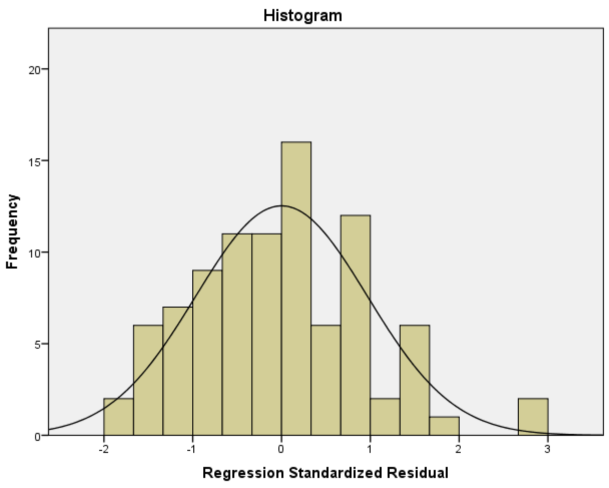
Max: Providencia
Min: Lota

Descriptive Statistics

	Mean	*Std. Deviation*	*N*
Quality of Life Index	42,736	10,0215	91
PSU	,46804	,160510	91
Population Density	3.575,76198	5.092,858885	91
Overcrowding	,19669	,044588	91
Rate of Domestic Violence	549,28571	158,273526	91
Municipal Budget (1,000 CLP$)	7.774.292,11	7.621.969,636	91
Own Incomes (1,000 CLP$)	15.619.075,68	20.172.956,959	91

Methodology (Regression Models-Stepwise)



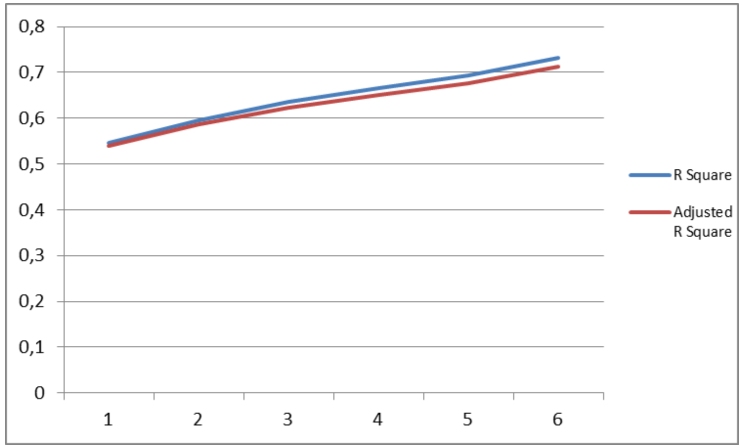


		Unstandardized Coefficients		*Standardized Coefficients*			*Collinearity Statistics*	
Model		*B*	*Std. Error*	*Beta*	*t*	*Sig.*	*Tolerance*	*VIF*
1	(Constant)	75,389	3,238		23,281	,000		
	overcrowding	-166,014	16,061	-,739	-10,337	,000	1,000	1,000
2	(Constant)	60,561	5,420		11,173	,000		
	overcrowding	-131,195	18,490	-,584	-7,095	,000	,678	1,475
	PSU	17,049	5,136	,273	3,319	,001	,678	1,475
3	(Constant)	62,952	5,243		12,008	,000		
	overcrowding	-109,437	19,076	-,487	-5,737	,000	,583	1,716
	PSU	18,718	4,942	,300	3,787	,000	,670	1,493
	rate of domestic violence	-,014	,004	-,214	-3,039	,003	,844	1,184
4	(Constant)	63,439	5,044		12,576	,000		
	overcrowding	-98,066	18,775	-,436	-5,223	,000	,556	1,798
	PSU	20,169	4,780	,323	4,220	,000	,662	1,510
	rate of domestic violence	-,016	,004	-,258	-3,713	,000	,803	1,246
	municipal budget	-2,412E-7	,000	-,183	-2,843	,006	,933	1,072
5	(Constant)	62,932	4,858		12,953	,000		
	overcrowding	-92,377	18,184	-,411	-5,080	,000	,549	1,821
	PSU	19,927	4,601	,319	4,331	,000	,662	1,511
	rate of domestic violence	-,015	,004	-,239	-3,554	,001	,794	1,259
	municipal budget	-2,347E-7	,000	-,179	-2,874	,005	,932	1,073
	population density	,000	,000	-,172	-2,800	,006	,948	1,055
6	(Constant)	59,438	4,688		12,678	,000		
	overcrowding	-79,504	17,535	-,354	-4,534	,000	,524	1,908
	PSU	16,368	4,457	,262	3,673	,000	,626	1,597
	rate of domestic violence	-,013	,004	-,209	-3,264	,002	,779	1,283
	municipal budget	-2,152E-7	,000	-,164	-2,790	,007	,927	1,079
	population density	,005	,000	-,218	-3,668	,000	,900	1,111
	own incomes	1,119E-7	,000	,225	3,434	,001	,742	1,348

Methodology (Regression Models-Stepwise)

ANOVA						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4931,256	1	4931,256	106,848	,000[b]
	Residual	4107,515	89	46,152		
	Total	9038,770	90			
2	Regression	5388,296	2	2694,148	64,946	,000[c]
	Residual	3650,474	88	41,483		
	Total	9038,770	90			
3	Regression	5738,581	3	1912,860	50,427	,000[d]
	Residual	3300,189	87	37,933		
	Total	9038,770	90			
4	Regression	6022,134	4	1505,534	42,921	,000[e]
	Residual	3016,636	86	35,077		
	Total	9038,770	90			
5	Regression	6276,820	5	1255,364	38,634	,000[f]
	Residual	2761,950	85	32,494		
	Total	9038,770	90			
6	Regression	6616,807	6	1102,801	38,248	,000[g]
	Residual	2421,963	84	28,833		
	Total	9038,770	90			

Model Summary										
					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	,739[a]	,546	,540	6,7935	,546	106,848	1	89	,000	
2	,772[b]	,596	,587	6,4407	,051	11,018	1	88	,001	
3	,797[c]	,635	,622	6,1590	,039	9,234	1	87	,003	
4	,816[d]	,666	,651	5,9226	,031	8,084	1	86	,006	
5	,833[e]	,694	,676	5,7003	,028	7,838	1	85	,006	
6	,856[f]	,732	,713	5,3696	,038	11,792	1	84	,001	1,986



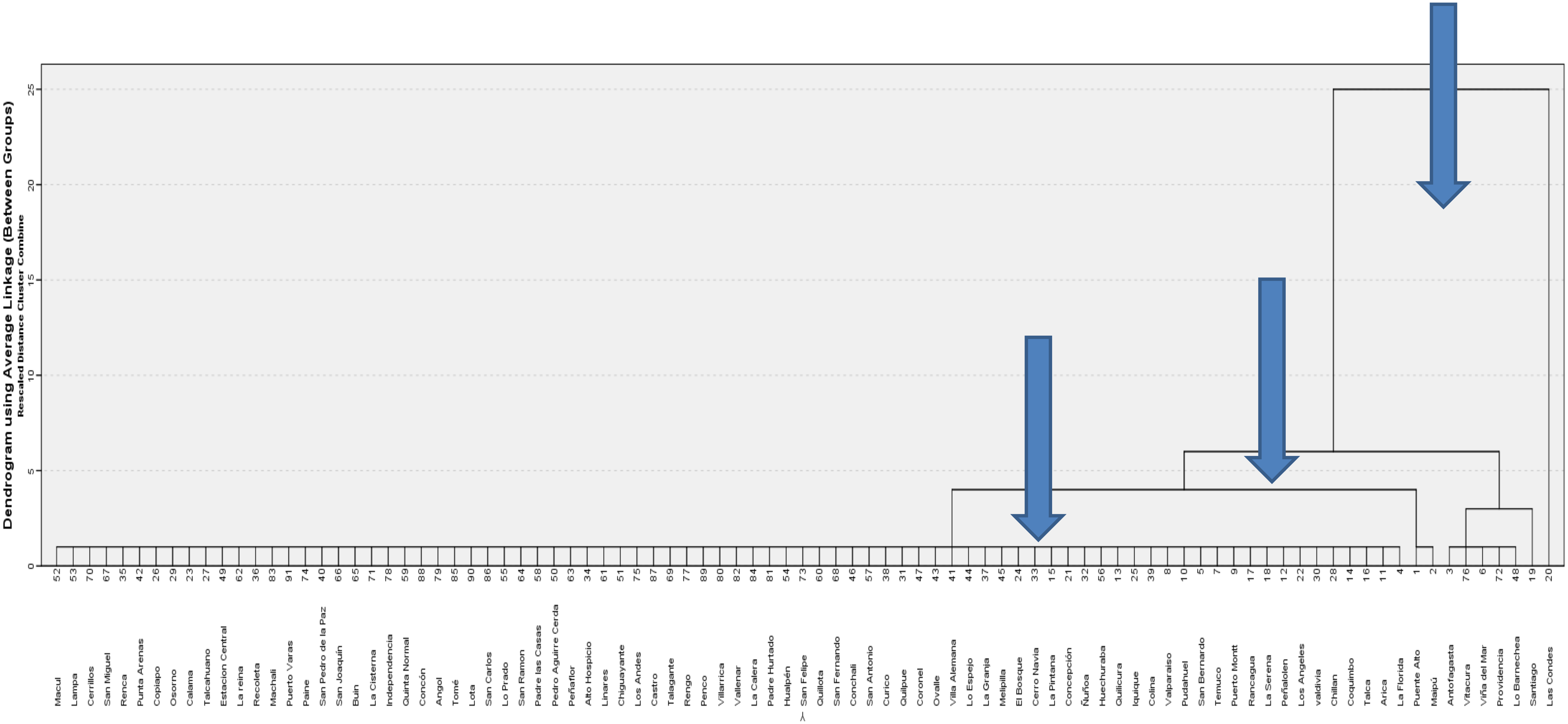
Methodology (Regression Final Iteration)

Model Summary										
					Change Statistics					
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**R Square Change**	**F Change**	**df1**	**df2**	**Sig. F Change**	**Durbin-Watson**
Quality of Life Index	,856[a]	,732	,713	5,3696	,732	38,248	6	84	,000	1,986

ANOVA						
Model		**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Quality of Life Index	Regression	6616,807	6	1102,801	38,248	,000[b]
	Residual	2421,963	84	28,833		
	Total	9038,770	90			

		Unstandardized Coefficients				95,0% Confidence Interval for B		Collinearity Statistics	
Model		**B**	**Std. Error**	**t**	**Sig.**	**Lower Bound**	**Upper Bound**	**Tolerance**	**VIF**
Quality of Life Index	(Constant)	59,438	4,688	12,678	,000	50,115	68,761		
	PSU	16,368	4,457	3,673	,000	7,505	25,230	,626	1,597
	population density	,005	,000	-3,668	,000	-,001	,000	,900	1,111
	overcrowding	-79,504	17,535	-4,534	,000	-114,374	-44,635	,524	1,908
	rate of domestic violence	-,013	,004	-3,264	,002	-,021	-,005	,779	1,283
	municipal budget	-2,152E-7	,000	-2,790	,007	,000	,000	,927	1,079
	own incomes	1,119E-7	,000	3,434	,001	,000	,000	,742	1,348

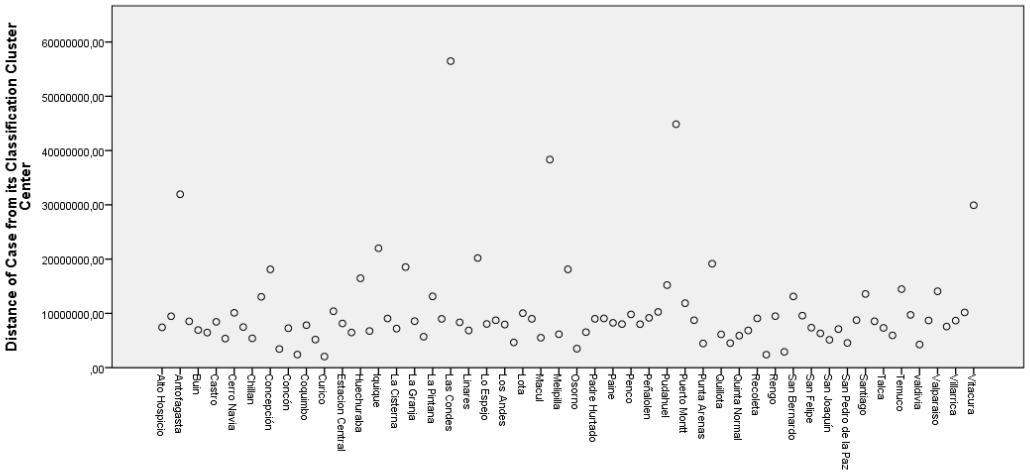
Methodology (Hierarchical Cluster-Dendogram)



Methodology (Cluster k-means)

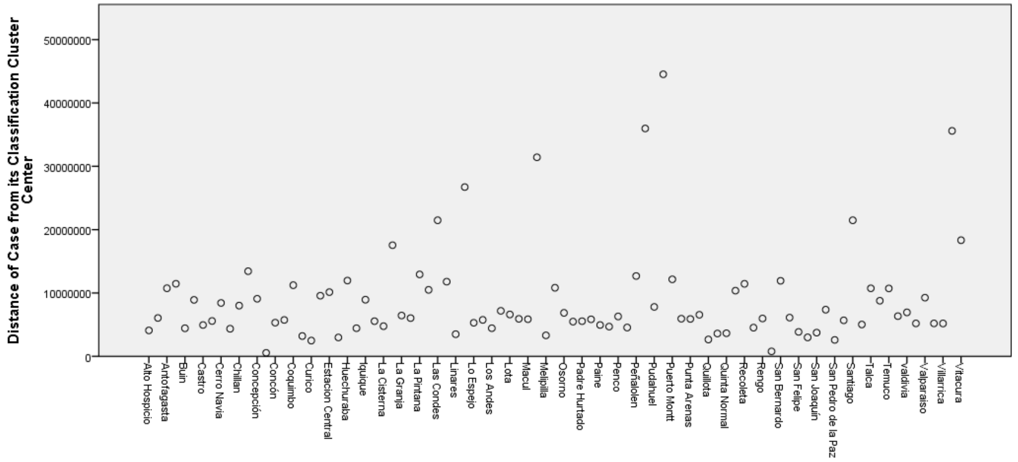
Viña del Mar
Santiago
Las Condes
Lo Barnechea
Providencia
Vitacura

Number of Cases in each Cluster		
Cluster	1	85,000
	2	6,000
Valid		91,000



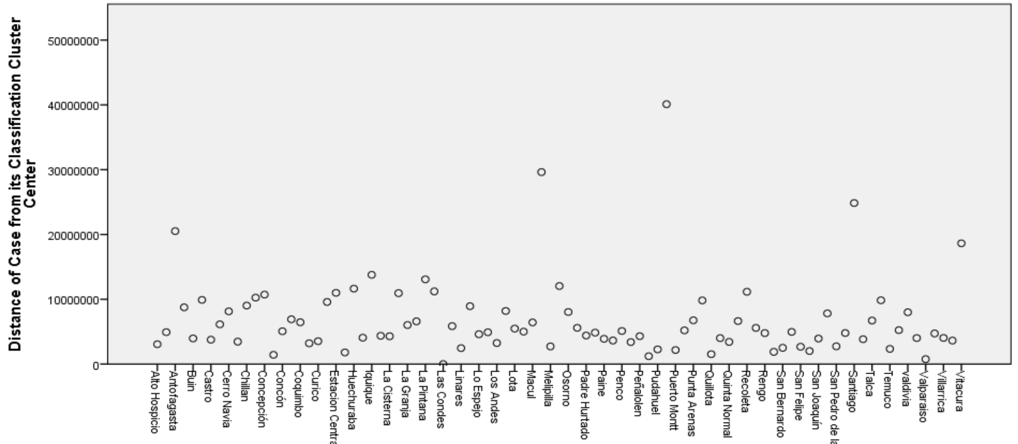
Santiago
Las Condes

Number of Cases in each Cluster		
Cluster	1	70,000
	2	19,000
	3	2,000
Valid		91,000



Las Condes

Number of Cases in each Cluster		
	1	1,000
	2	5,000
	3	22,000
	4	63,000
Valid		91,000



Methodology (Technical Efficiency (Stochastic Frontier Analysis)

Cobb-Douglas

$$\ln q_{it} = X_{it}\beta + v_{it} - u_{it}$$

Translogarithmic

$$\ln q_{it} = \beta_0 + \sum_{i=1}^{n} \beta_1 \ln X_i + 0{,}5 \sum_{i=1}^{n} \sum_{i=1}^{n} \beta_i \ln X_i \ln X_j$$

$$LR = -2(\ln(LRtest(Cobb - Douglas)) - \ln(LRtest(Trans\log))) = -2(\ln(303) - \ln(405)) = 0{,}577$$

Cobb-Douglas

$$QLI_i = e_i^{\alpha} * PSU_i^{\beta_1} * PopulationDensity_i^{\beta_2} * Overcrowding_i^{\beta_3} *$$
$$\mathrm{RateDomesticViolence}_i^{\beta_4} * MunicipalBudget_i^{\beta_5} * OwnIncomes_i^{\beta_6} * (v_{it} - u_{it})$$

Testing the absence of the technical inefficiency effects

Likelihood-ratio statistic (Kodde & Palm, 1986)

LR	Critical Value	Decision	Conclusion
305,88	7.045	Reject Ho	**Technical inefficiency effects**

Methodology (Technical Efficiency (Stochastic Frontier Analysis)

Function: *Cobb-Douglas*

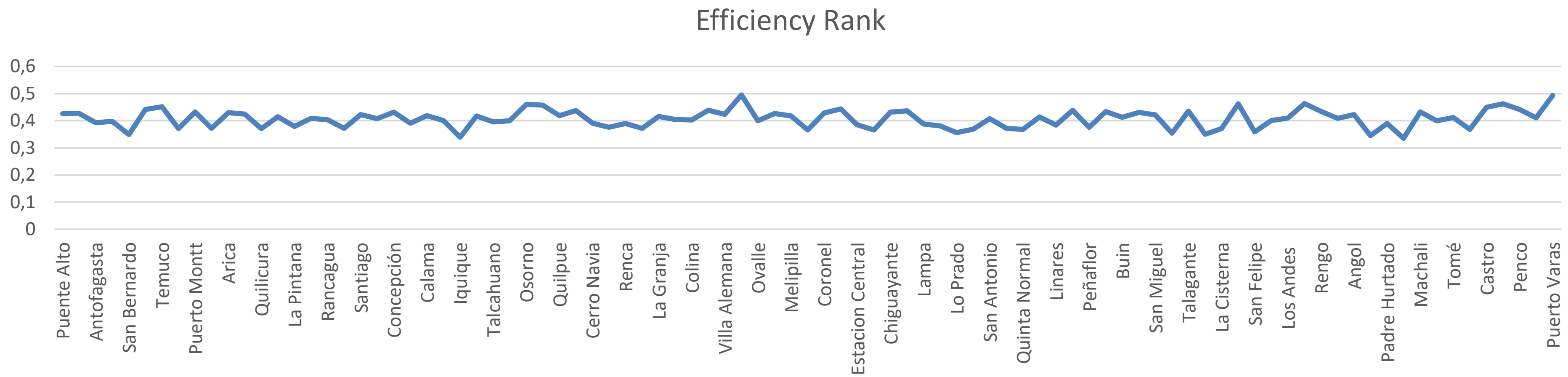
Endogenous Variable: *Quality of Life Index (QLI)*

Exogenous Variable: *PSU, Population Density Overcrowding, Rate of Domestic Violence, Municipal Budget and Own Incomes*

$$QLI_i = e_i^{\alpha} * PSU_i^{\beta_1} * PopulationDensity_i^{\beta_2} * Overcrowding_i^{\beta_3} *$$

$$\mathrm{RateDomesticViolence}_i^{\beta_4} * MunicipalBudget_i^{\beta_5} * OwnIncomes_i^{\beta_6} * (v_{it} - u_{it})$$

Methodology (Technical Efficiency (Stochastic Frontier Analysis)



id	*City Hall*	*Efficiency*	*id*	*City Hall*	*Efficiency*	*id*	*City Hall*	*Efficiency*	*id*	*City Hall*	*Efficiency*	*id*	*City Hall*	*Efficiency*
1	Punta Arenas	0,49560461	21	Machali	0,43223089	41	Quillota	0,41396319	61	Los Angeles	0,39076185	81	San Carlos	0,36811756
2	Puerto Varas	0,49313608	22	Concepción	0,43203702	42	Buin	0,41258962	62	Cerro Navia	0,39069331	82	Pedro Aguirre Cerda	0,36610374
3	Vitacura	0,46318528	23	Chiguayante	0,43157714	43	Tomé	0,41172102	63	Renca	0,39035403	83	Conchali	0,36604078
4	***Providencia***	***0,46214689***	24	San Joaquín	0,43100179	***44***	***Lota***	***0,41115038***	64	Padre Hurtado	0,38993407	84	San Felipe	0,3586453
5	Concón	0,46211268	25	Arica	0,43003703	45	Los Andes	0,41005065	65	Lampa	0,38758456	85	Lo Prado	0,35563907
6	Osorno	0,46063472	26	Coronel	0,42857925	46	Talca	0,40875832	66	Estacion Central	0,38452992	86	San Fernando	0,35407469
7	Valdivia	0,4577662	**27**	**Lo Espejo**	**0,42665622**	47	Independencia	0,40871929	67	Linares	0,38385011	87	Cerrillos	0,35051847
8	Temuco	0,45140502	28	Maipú	0,42645832	**48**	**Las Condes**	**0,40815958**	68	Hualpén	0,38081638	88	San Bernardo	0,34961307
9	Castro	0,44944759	29	Puente Alto	0,42597067	49	San Antonio	0,40799852	69	La Pintana	0,37888056	89	Villarrica	0,34476879
10	Lo Barnechea	0,44372421	30	Peñalolen	0,42430001	50	Curico	0,40501718	70	Peñaflor	0,37607275	90	Iquique	0,33964908
11	Viña del Mar	0,44152194	31	Villa Alemana	0,42379626	51	Rancagua	0,40401882	71	Alto Hospicio	0,3757423	91	Vallenar	0,33549004
12	Penco	0,44107148	32	Angol	0,42237759	52	Colina	0,40319542	72	La Serena	0,37239006			
13	La reina	0,43849692	33	Santiago	0,4223694	53	Paine	0,40103116	73	Recoleta	0,37211265			
14	San Pedro de la Paz	0,43840259	34	San Miguel	0,42121533	54	El Bosque	0,40066329	74	Padre las Casas	0,3720951			
15	Ñuñoa	0,43730231	35	Calama	0,41882841	55	La Calera	0,40022035	75	Pudahuel	0,37203603			
16	Macul	0,43686712	36	Quilpue	0,41839069	56	Ovalle	0,39988861	76	Valparaiso	0,37150118			
17	Talagante	0,43521227	37	Melipilla	0,41796852	57	Chillan	0,39980133	77	La Cisterna	0,37136703			
18	Rengo	0,4344577	38	Copiapo	0,41786668	58	La Florida	0,39752859	78	Quilicura	0,37082215			
19	San Ramon	0,43321037	39	La Granja	0,41597118	59	Talcahuano	0,39622175	79	Huechuraba	0,36857527			
20	Puerto Montt	0,43246087	40	Coquimbo	0,41449048	60	Antofagasta	0,39317003	80	Quinta Normal	0,36848419			

Efficiency Average: 0,407048582

Conclusion and Results

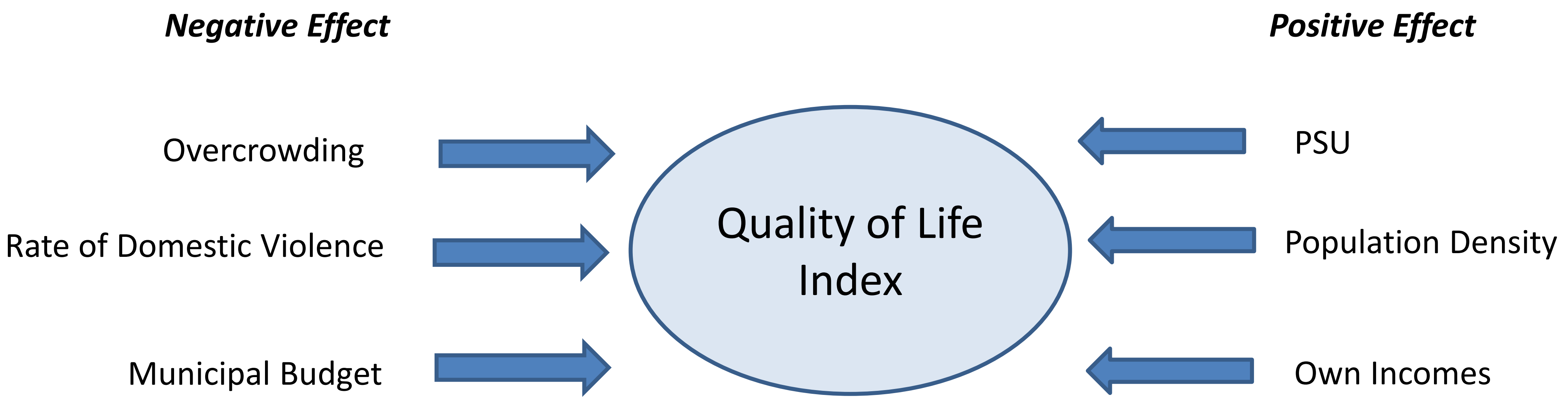
- Currently there **is a growing interest in the Town Hall** due to the search for **decentralization** in Chile
- Since it is known **the excessive centralism in the economic**, political and administrative spheres in **Chile** in comparison with other Latin American countries
- **OECD recommends strengthening the regions** so as not to continue to slow the development of the country

Efficiency of City Hall in Chile

- Therefore, to knows the factors that determine the Town Hall efficiency is of high interest since it **shows us the aspects that require special attention in new public policies regarding the management of the City Hall** as well as, and more importantly, the quality of life of the people

Conclusion and Results

- The research allows us to contrasts between different City Hall (**North, Center and South**)
- Allows us **to verify centralization of the country**
- Evidencing the variables that affect in greater proportion according to our econometric/stochastic model results



Conclusion and Results

higher per capita income
and population wealth
have a negative effect on efficiency

More Efficient

Top Ten

Concón
Vitacura
Providencia
Lo Barnechea

Valdivia
Temuco
Osorno
Castro
Puerto Varas

Punta Arenas

Cluster Result

Viña del Mar
Santiago
Las Condes
Lo Barnechea
Providencia
Vitacura

CASE 2

Forecasting Performance Measures for Traffic Safety using Deterministic and Stochastic Models

By

Alexander Paz, Ph.D., P.E.; Hanns de la Fuente-Mella, Ing., Dr. & Naveen Veeramisti, Std.
Transportation Research Center

University of Nevada, Las Vegas

United States of America

Introduction

Moving Ahead Progress in 21st Century (Map-21) https://www.fhwa.dot.gov/map21/

- Performance based Highway Safety is one of the 10 performance provisions

- Federal Highway Administration is seeking criteria to assess traffic safety regarding
  - the number of fatalities
  - the number of serious injuries
  - Fatalities per Vehicle Miles Traveled (VMT)
  - Serious injuries per VMT

- State Department of Transportation (DOTs) and Metropolitan Organizations are required to use these four measures to conduct federal-aid highway programs and assess performance

Background

- **Quality data**
  - Model Minimum uniform Crash Criteria Guidelines
  - DOTs have their own crash database. For ex: NDOT has “Nevada Citation and Accident Tracking System” database

- For many DOTs, the development of **forecasting methods** is in early stage of research

- Many traffic safety analysts at DOTs use Aspirational approaches such as zero-fatality goals or 5-Year Rolling Average models

- Many DOTs set targets using the aspirational approach rather than model-based approach

Problem Statement

The Problem:

What methodology to use, by state DOTs, for the forecasting of safety performance measures so as to set targets for the reduction of fatalities and serious injuries and reach MAP-21 policy goals?

Methodology

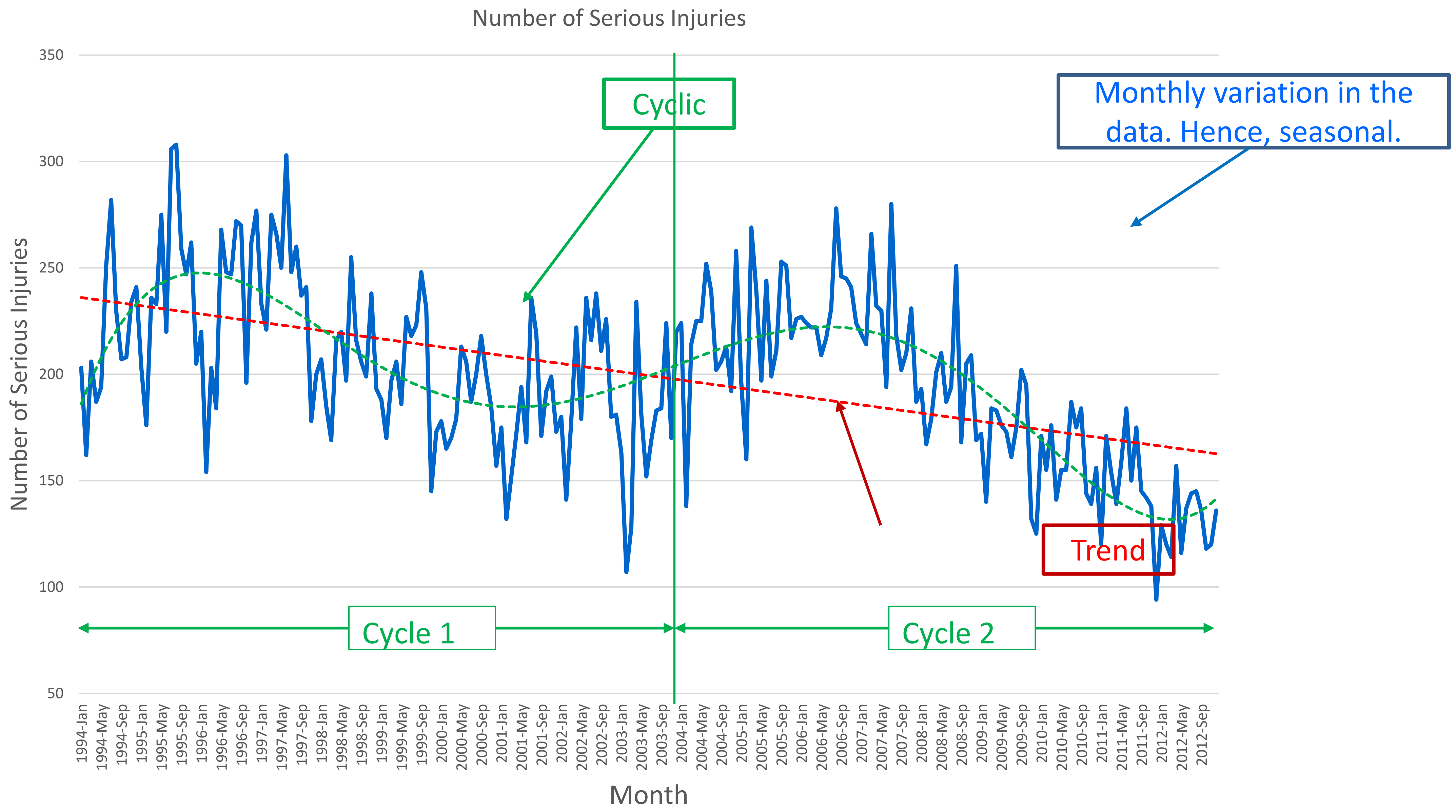
- Four independent and univariate time-series of crash datasets
  - Number of fatalities aggregated monthly (1994 – 2012)
  - Number of serious injuries aggregated monthly (1994 – 2012)
  - Fatalities/100 Million VMT aggregated annually (1995-2013)
  - Serious injuries/100 Million VMT aggregated annually (1995-2013)

- Monthly data for statewide VMT was unavailable. Hence seasonal data were not used for the last two performance measures

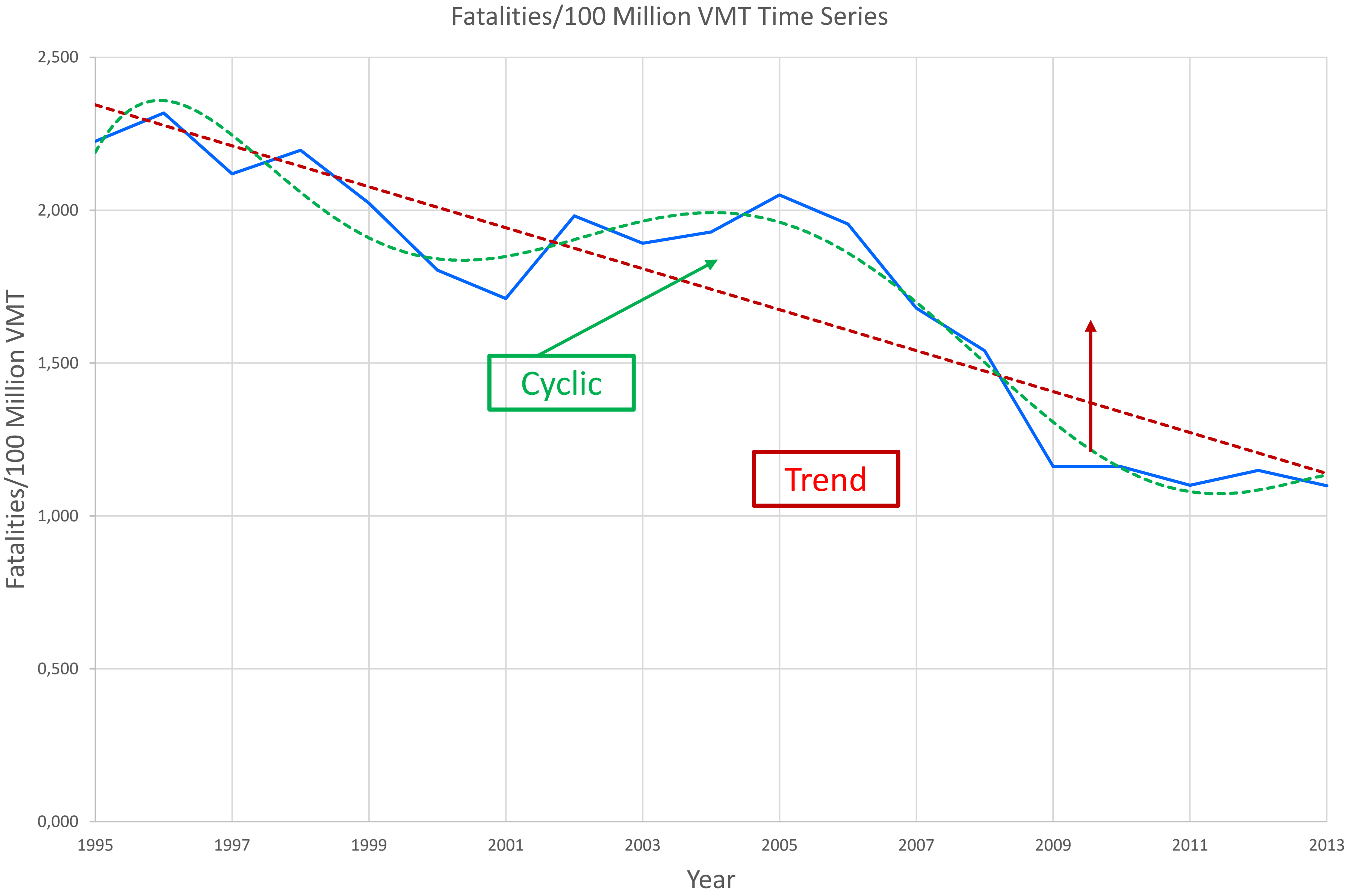
- Deterministic and Stochastic Models were tested to determine the best estimates

- Results were compared using root-mean-square error (RMSE) and mean absolute percentage error (MAPE)

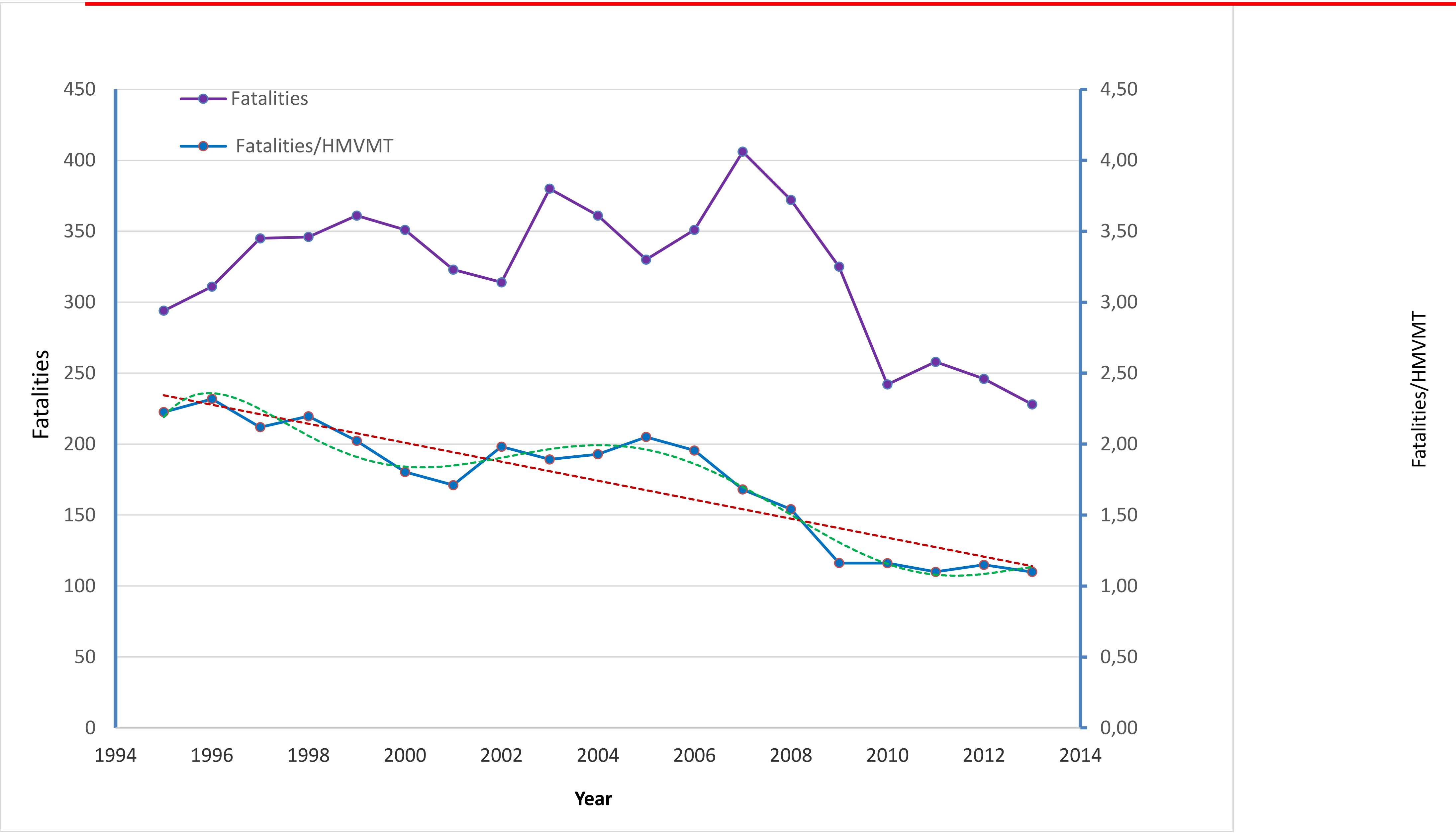
Time Series – Number of Serious Injuries



Time Series – Fatalities/100 Million VMT



Time Series – Fatalities/100 Million VMT



Deterministic Forecasting methods

Simple Exponential Smoothing: $$Y_{t+1} = \alpha X_t + (1-\alpha)Y_t \quad (1)$$

Holt:

$$Y_{t+1} = Y_t + T_t$$
$$Y_t = \alpha X_t + (1-\alpha)(Y_{t-1} + T_{t-1})$$
$$T_t = \beta(Y_t - Y_{t-1}) + (1-\beta)T_{t-1} \quad (2)$$

Brown:

$$Y_{t+1} = Y_t + T_t$$
$$Y_t = \alpha X_t + (1-\alpha)(Y_{t-1} * T_{t-1})$$
$$T_t = \beta(Y_t / Y_{t-1}) + (1-\beta)T_{t-1} \quad (3)$$

Damped-trend:

$$Y_{t+1} = Y_t + T_t$$
$$Y_t = \alpha X_t + (1-\alpha)(Y_{t-1} + \theta T_{t-1})$$
$$T_t = \beta(Y_t - Y_{t-1}) + (1-\beta)\theta T_{t-1} \quad (4)$$

where:

Y_{t+1} = forecasted value for time period t+1

Y_t = forecasted level value which represents the smoothed value up to time period t

T_t = trend estimate at time period t (slope of the trend line that we are fitting at time period t)

X_t = observed value at time period t

α, β = smoothing parameters (should be between 0 and 1)

θ = damping parameter (should be between zero and 1)

Deterministic Forecasting

Winter Additive and Multiplicative Models:

$$Y_{t+1} = (Y_t + T_t) * S_t$$

$$Y_t = \alpha \frac{X_t}{S_{t-c}} + (1-\alpha)(Y_{t-1} + T_{t-1})$$

$$T_t = \beta(Y_t - Y_{t-1}) + (1-\beta)T_{t-1} \qquad (5)$$

$$S_t = \gamma \frac{X_t}{Y_t} + (1-\gamma)S_{t-c}$$

where:

Y_{t+1} = forecasted value for the time period t+1

Y_t = forecasted Level value which represents the smoothed value up to time period t

T_t = trend estimate at time period t (slope of the trend line that we are fitting at time period t)

X_t = observed value at time period t

α, β, = smoothing parameter

S_t = seasonal parameter estimate

Results of Deterministic Forecasting

- Number of Fatalities and Serious Injuries:

Deterministic Model/ Performance Measures	Number of Fatalities		Number of Serious Injuries	
	MAPE	RMSE	MAPE	RMSE
Simple	25.698	7.361	12.309	29.891
Holt	25.469	7.374	12.185	29.942
Brown	25.691	7.419	12.681	30.520
Damped Trend	25.703	7.393	12.304	30.025
Simple Seasonal	22.770	6.729	10.724	25.638
Winter Additive	22.503	6.735	10.476	25.593
Winter Multiplicative	23.126	6.890	10.644	26.138

- Fatalities/100 Million VMT and Serious Injuries/100 Million VMT

Deterministic Model/ Performance Measures	Fatalities/ 100 Million VMT		Serious Injuries/ 100 Million VMT	
	MAPE	RMSE	MAPE	RMSE
Simple	7.714	0.166	8.987	1.262
Holt	6.888	0.158	5.916	0.916
Brown	8.659	0.176	6.300	1.015
Damped Trend	6.867	0.163	5.291	0.911

Stochastic Forecasting Model

(ARIMA (p,d,q)) Model:

$$\hat{y}_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \qquad (6)$$

$$y_t = Y_t \qquad \text{for d} = 0$$

$$y_t = Y_t - Y_{t-1} \qquad \text{for d} = 1$$

$$y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \quad \text{for d} = 2$$

The differencing, if any, must be reversed to obtain a forecast for the original series

$$\text{if d} = 0 \quad \hat{Y}_t = \hat{y}_t$$

$$\text{if d} = 1 \quad \hat{Y}_t = \hat{y}_t + Y_{t-1}$$

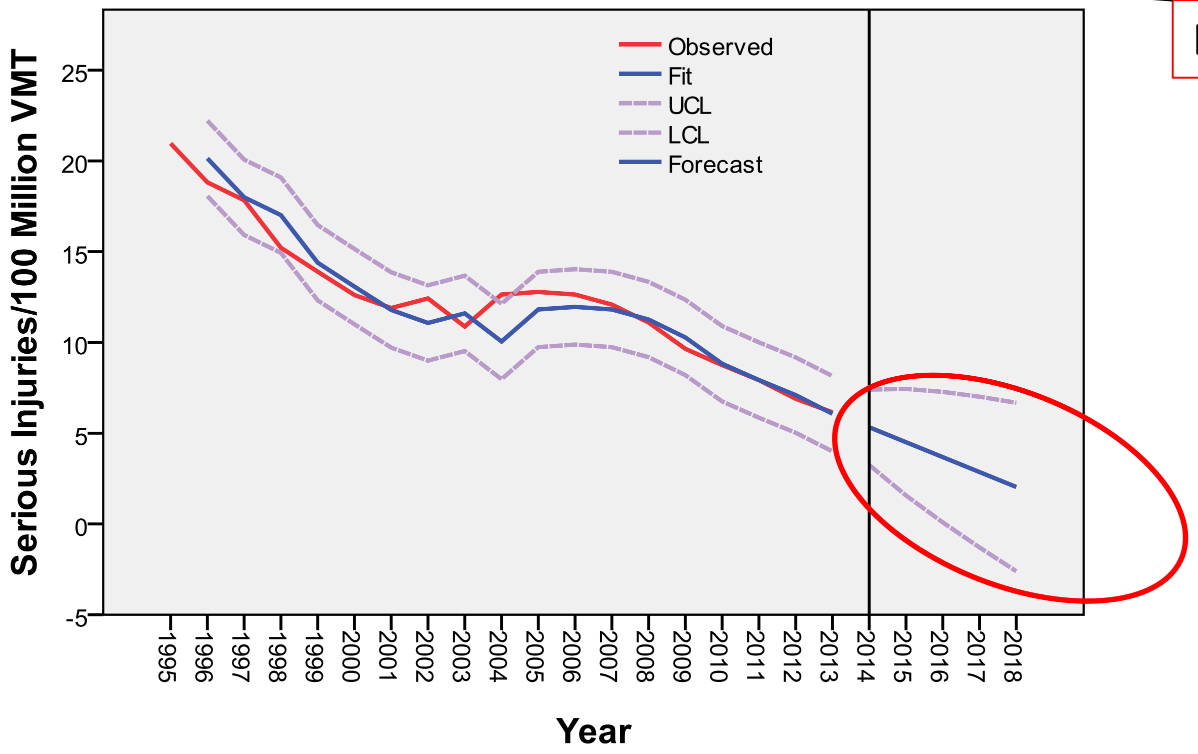
$$\text{if d} = 2 \quad \hat{Y}_t = (\hat{y}_t + Y_{t-1}) + (Y_{t-1} - Y_{t-2})$$

where:
Y_t = observed values
y_t = differenced (stationarized series)
$\hat{y}_t$ = Forecast of the stationarized series
$\hat{Y}_t$ = Forecast of the original series
Φ = autoregressive parameters
θ = moving average parameters
Φ_0 = model constant is assumed different from zero
ε_t = error
p = the number of autoregressive terms
d = the number of nonseasonal differences
q = the number of moving-average terms

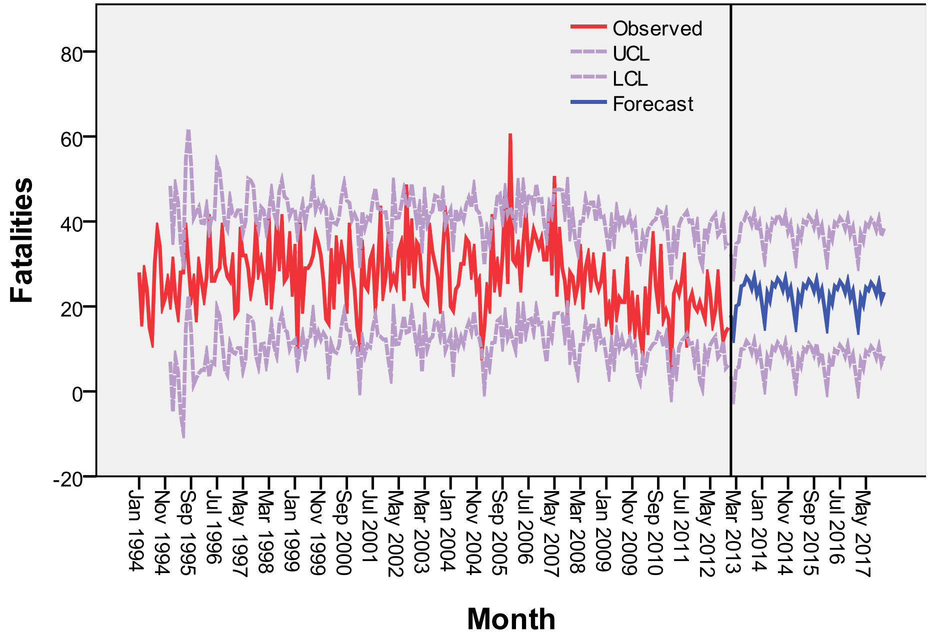
Results of ARIMA model

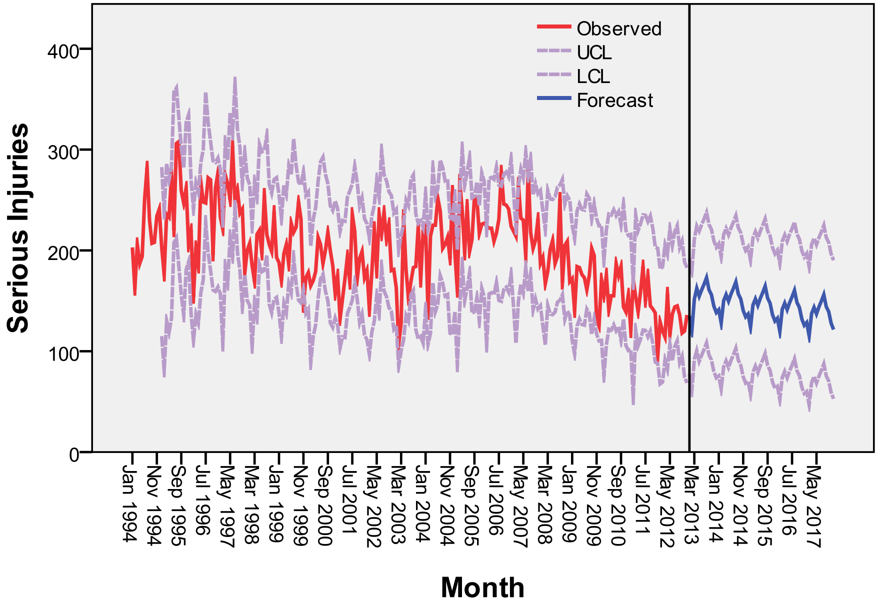
Stochastic Model/ Performance Measures	Fatalities/ 100 Million VMT		Serious Injuries/ 100 Million VMT	
	MAPE	RMSE	MAPE	RMSE
ARIMA(0,1,0)	7.254	0.158	5.232	0.984
ARIMA(0,1,1)	7.105	0.162	5.054	1.009
ARIMA(0,1,2)	7.209	0.168	4.679	0.907
ARIMA(0,1,3)	6.587	0.167	4.763	0.928

Best Fit Model



Forecast of Fatalities and Serious Injuries





Stochastic Model/ Performance Measures	Number of Fatalities		Number of Serious Injuries	
	MAPE	RMSE	MAPE	RMSE
$SARIMA(0,0,5)(0,1,1)_{12}$	25.120	7.756	12.214	30.111
$SARIMA(0,0,4)(0,1,1)_{12}$	25.910	7.519	12.303	30.113
$SARIMA(0,0,3)(0,1,1)_{12}$	25.665	7.870	12.855	31.489
$SARIMA(0,0,2)(0,1,1)_{12}$	25.662	7.892	13.406	32.663
$SARIMA(0,0,1)(0,1,1)_{12}$	27.017	7.993	13.676	33.132

Best Fit Model

Conclusions

- Historically, aspirational models or simple moving average were used. In many cases, this would led to grossly overestimating or underestimating traffic safety
- This research aimed to forecast traffic-safety performance measures using observed time series crash data
- Of all models, SARIMA(0,0,5)(0,1,1) model was determined to be best for use with Nevada data
- These forecasts could be used to determine targets for future safety-improvement programs and policies
- Policy makers need to have access to robust crash-forecast models that enable them to prioritize and implement realistic and economically viable safety policies and programs

Business Intelligence - Proof of Concept

Oracle BI Interactive Dashbo | Oracle BI Answers

utilitiesserver:9704/analytics/saw.dll?Dashboard&PortalPath=%2Fusers%2Fweblogic%2FSafety%20test%2FSafety%20Performance%20Measure&page=Safety%

ORACLE Business Intelligence | Search All

Safety Performance Measure | Home | Catalog | Favorite

Parameters

* Forecast Model SARIMA(0,0,1)(0,1,1) * Severity Serious Injuries

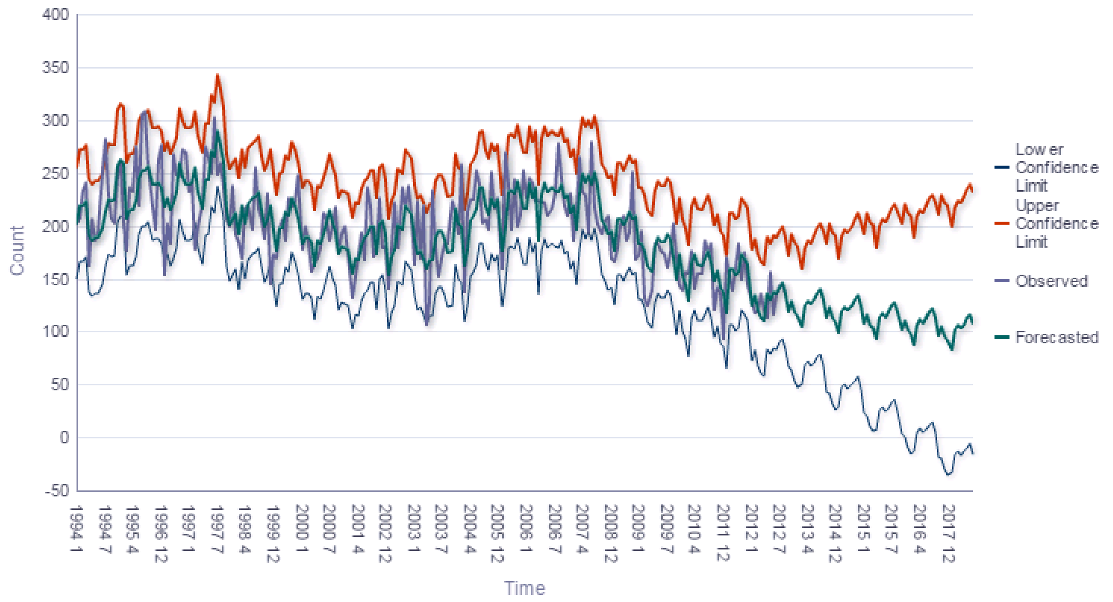
Apply | Reset

Safety Performance Measures - Forecast

Model: SARIMA(0,0,1)(0,1,1)
Performance Measure: Number of Serious Injuries

Year	Month	Observed	Forecasted	Lower Confidence Limit	Upper Confidence Limit
1994	1	203	203	150	256
	10	208	220	167	273
	11	234	220	167	273
	12	241	224	171	277
	2	162	191	138	244
	3	206	186	133	239
	4	187	190	137	243
	5	194	190	137	243
	6	251	198	145	250
	7	282	214	162	267
	8	230	225	173	278
	9	207	224	171	277
1995	1	203	225	172	277
	10	247	257	204	310
	11	262	262	209	315
	12	205	260	207	313
	2	176	207	154	260
	3	236	215	162	268
	4	233	215	162	268
	5	275	224	171	277
	6	220	246	193	299
	7	306	253	200	306
	8	308	253	200	306
	9	259	257	204	310
1996	1	220	240	187	293

Rows 1 - 25



Edit - Refresh - Export

ESCUELA DE COMERCIO

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

NUCLEO DE INVESTIGACION DATA ANALYTICS

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

CASE 3
Gobiernos Corporativos y Asimetrías de Información. Modelamiento Econométrico del Spread (Bid-Ask) para una muestra de Empresas Chilenas

Berta SILVA PALAVECINOS, Dra.

Ricardo CAMPOS ESPINOZA, PhD.

David CADEMARTORI ROSSO, Mg.

Hanns DE LA FUENTE-MELLA, Ing. Dr.

Introducción

Crisis económicas-financieras, y los efectos que éstas producen en el valor de las inversiones

- constante preocupación por **generar transparencia** sobre la gestión administrativa-contable-financiera en las entidades que emiten valores

Se ha puesto en duda la calidad de la información que reciben los inversionistas

- esto es aún más grave cuando éstos son Fondos de Pensiones, dado los **efectos sociales** que producen Estados Unidos de Norteamérica (ENRON), Europa (PARMALAT), Chile (**La POLAR**)

A mayor revelación de información se disminuyen las asimetrías de información en el mercado de capitales

- tal hecho **aumenta la liquidez** de los mercados y el **valor de la empresa**

La correcta asignación de recursos por el mercado exige que los inversionistas reciban información oportuna y de calidad

– **calidad** de los Gobiernos Corporativos

Introducción

- ¿El tamaño del corredor, medido como su patrimonio, afecta al spread?
- ¿El tamaño del corredor, medido como la cartera propia disponible, afecta al spread?
- ¿El nivel de actividad del corredor, medido como la cantidad ofertada de compra y de venta, afecta al spread?

Marco Teórico

Separación entre la propiedad y el control (no existe alineamiento de los correspondientes objetivos perseguidos)

- dilema de agenciamiento, que recoge la **Teoría de la Firma** (Jensen and Meckling, 1976)

La asimetría de información surge porque algunos participantes del mercado tienen ventajas de información sobre otros

- respecto de las compañías emisoras de valores o de otras circunstancias que puedan afectar el valor de los mismos (Ramírez and Yáñez, 2010)

Importancia que exista transparencia y revelación de información

- las empresas que adoptan **pobres prácticas** en este aspecto experimentarán graves **asimetrías de información** (Chen et al., 2007)

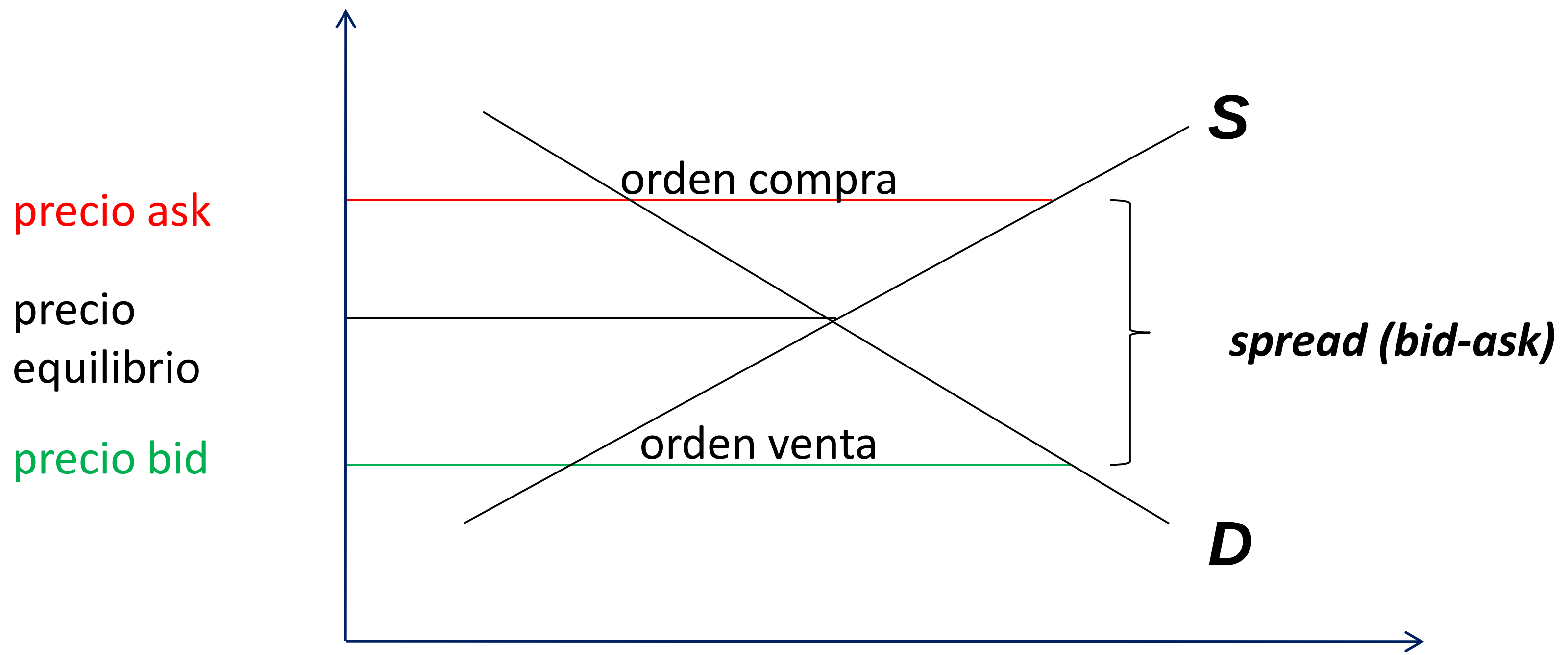
Marco Teórico

Medición de las asimetría de información a través del spread del bid-ask de los precios accionarios

- un **spread mayor** se asocia a una mayor asimetría de información (Amihud and Mendelson, 1989; Coller and Yohn, 1997; Kim and Verrecchia, 1994; Bollen, Smith and Whaley, 2004)
- hay evidencias de que el **spread es menor** cuando hay más información disponible sobre el activo (Copeland and Galai, 1983; Glosten and Milgrom, 1985)
- la presencia de **corredores con mejor nivel de información** conduce a un spread mayor (Glosten and Milgrom, 1985)
- el mercado accionario es de **carácter monopólico** (Benston and Hagerman, 1974), debido a que tiene pocos actores y que las barreras de entrada para nuevos actores no son bajas

Marco Teórico

En un mercado plenamente activo se formaría naturalmente un precio de equilibrio entre la oferta y la demanda, cuando ello no ocurre surge el precio bid y el precio ask



Marco Teórico

Principales medidas de spread que se observan en la literatura financiera

Absoluto ($)	Relativo (%)
$(PV) - (PC)$	$\frac{(PV-PC)}{\frac{(PV+PC)}{2}}$, $\ln(PV) - \ln(PC)$
En la escala original	Escala modificada
$(PV) - (PC)$, $\frac{(PV-PC)}{\frac{(PV+PC)}{2}}$	$\ln(PV) - \ln(PC)$, $\frac{(\ln PV-\ln PC)}{\frac{(\ln PV+\ln PC)}{2}}$
Relativas al mismo momento	Relativas a un momento pasado
$(PV_t) - (PC_t)$, $\frac{(PV_t-PC_t)}{\frac{(PV_t+PC_t)}{2}}$	$\frac{1}{2}\left[\left(\frac{(PV_t-PC_t)}{(PV_t+PC_t)/2}+\frac{(PV_{t-1}-PC_{t-1})}{((PV_{t-1}+PC_{t-1})/2)}\right)\right]$
Basadas sólo en puntas	Basadas en precio de transacción
$(PV) - (PC)$, $\frac{(PV-PC)}{\frac{(PV+PC)}{2}}$	$2\left[PT - \frac{(PV-PC)}{2}\right]$, $2\left[PT - \left(\frac{PV-PC}{2}\right) \Big/ \left(\frac{PV-PC}{2}\right)\right]*100$

Metodología
Modelos econométricos univariantes (Aguas Andinas S.A.)

- 21.420 observaciones
- GARCH (1,1) (desviación típica como término de la ecuación del modelo)

$$Bid\_Ask\_Aguas\_A_t = Q + 0.350061^{*}\,(\mu_{t-1}^{2} - Q_{t-1}) + 0.410890^{*}\,(Bid\_Ask\_Aguas\_A_{t-1} - Q_{t-1})$$

$$Q = 0.001205 + 0.997144^{*}\,(Q_{t-1} - 0.001205) + 0.044185^{*}\,(\mu_{t-1}^{2} - Bid\_Ask\_Aguas\_A_{t-1})$$

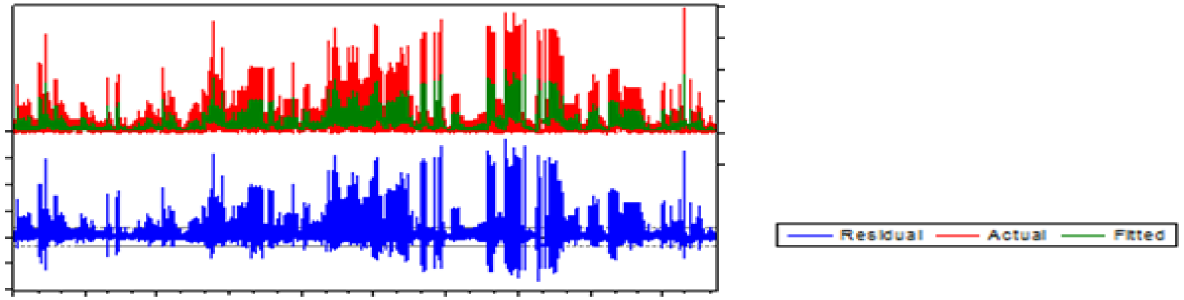
$Bid\_Ask\_Aguas\_A_t = (A-B)/((A+B)/2)$

(diferencia relativa sobre su media)

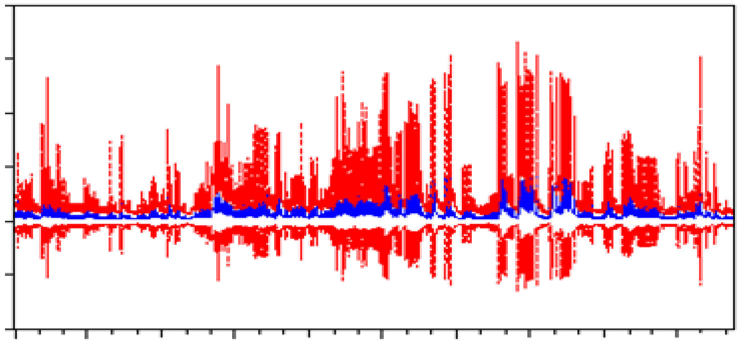
Test hipótesis distribución Extreme-Max residuos Aguas Andinas S.A.

Method	Value	Adj. Value	Probability
Cramer-von Mises (W2)	274.6647	275.0400	< 0.01
Watson (U2)	269.3139	269.6819	< 0.01
Anderson-Darling (A2)	1444.409	1446.383	< 0.01

Estimación de los residuos para la serie Bid-Ask de Aguas Andinas S.A.



Pronóstico estático para la serie Bid-Ask de Aguas Andinas S.A.



Root Mean Squared Error	0.036196
Mean Absolute Error	0.019470
Mean Abs. Percent Error	194.9620
Theil Inequality Coefficient	0.436004
Bias Proportion	0.032659
Variance Proportion	0.371566
Covariance Proportion	0.595775

Metodología
Modelos econométricos univariantes (GASCO)

- 9.320 observaciones
- ARCH (2) (desviación típica como término de la ecuación del modelo)

$$Bid\_Ask\_Gasco_t = {}^*0.504520 * @SQRT(GARCH)$$

$$GARCH = 5.56E{+}12 + 0.202423^* \mu_{t-1}^2 + 0.06652^* \mu_{t-2}^2$$

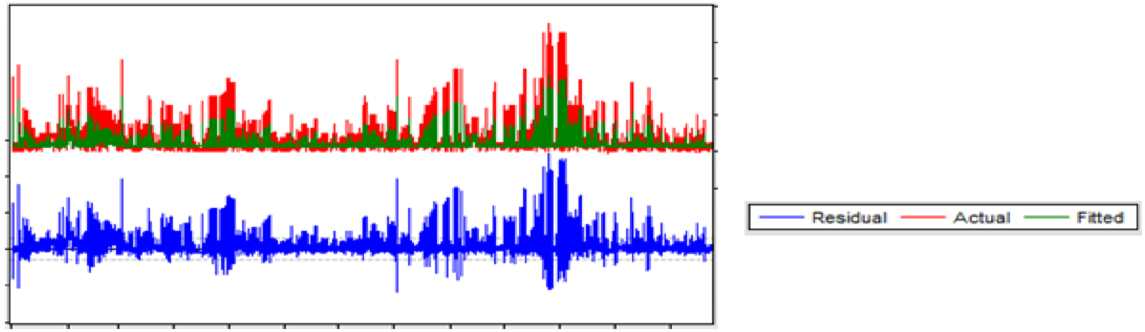
$Bid\_Ask\_Gasco_t$ = (A-B)

(diferencia directa entre ask y bid)

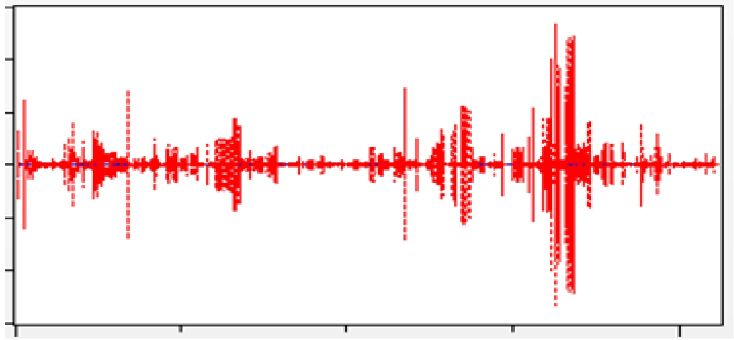
Test hipótesis distribución Logistic residuos Gasco

Method	Value	Adj. Value	Probability
Cramer-von Mises (W2)	31.78816	31.79156	< 0.005
Watson (U2)	31.78816	31.79156	< 0.005
Anderson-Darling (A2)	232.3372	232.3434	< 0.005

Estimación de los residuos para la serie Bid-Ask de Gasco



Pronóstico estático para la serie Bid-Ask de Gasco



Root Mean Squared Error	1464059.
Mean Absolute Error	851996.4
Mean Abs. Percent Error	237.6470
Theil Inequality Coefficient	0.437324
Bias Proportion	0.142101
Variance Proportion	0.398694
Covariance Proportion	0.459205

Metodología
Modelos econométricos multivariantes (Aguas Andinas S.A.)

Dependent Variable: A_B_MEDIO

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D_BCI_COM	0.005240	0.001449	3.614753	0.0003
D_CELFIN_COM	0.003559	0.000839	4.243625	0.0000
D_CORPBANCA_COM	[illegible]	[illegible]	[illegible]	[illegible]
D_MERRILL_VEN	[illegible]	[illegible]	[illegible]	[illegible]
D_PENTA_VEN	-0.014352	0.001630	-8.804615	0.0000
D_VALORESSEC_VEN	-0.007451	0.001911	-3.898063	0.0001
D_YRARRAZ_COM	0.114172	0.001210	94.39418	0.0000
TAMANO_COM	0.002163	0.000691	3.131776	0.0017
TAMANO_VEN	[illegible]	[illegible]	[illegible]	[illegible]
INVENTARIO_COM	[illegible]	[illegible]	[illegible]	[illegible]
INVENTARIO_VEN	-0.009191	0.000786	-11.69538	0.0000
C	0.028105	0.000610	46.09495	0.0000
AR(1)	0.617367	0.005381	114.7354	0.0000

R-squared	0.559161	Mean dependent var	0.030460
Adjusted R-squared	0.558913	S.D. dependent var	0.041999
S.E. of regression	0.027893	Akaike info criterion	-4.320240
Sum squared resid	16.65487	Schwarz criterion	-4.315402
Log likelihood	46280.61	Hannan-Quinn criter.	-4.318662
F-statistic	2262.614	Durbin-Watson stat	2.220700
Prob(F-statistic)	0.000000		

Inverted AR Roots	.62

BCI
CELFIN
CORPBANCA
MERRILL LYNCH
PENTA
VALORES SECURITY
YRARRAZABAL Y CIA
TAMAÑO (PATRIMONIO)
TAMAÑO (CARTERA PROPIA)

D_Corredor$_i$_Ven = Variable dicotómica (1: Corredor Venta i en el periodo t, 0: Otro Corredor Venta i en el periodo t).

D_Corredor$_i$_Com = Variable dicotómica (1: Corredor Compra i en el periodo t, 0: Otro Corredor Compra i en el periodo t).

D_Tamaño_Com = Variable dicotómica (1: Patrimonio Grande Corredor Compra en el periodo t, 0: Otro Patrimonio Corredor Compra en el periodo t).

D_Tamaño_Ven = Variable dicotómica (1: Patrimonio Grande Corredor Venta en el periodo t, 0: Otro Patrimonio Corredor Venta en el periodo t).

D_Inventario_Com = Variable dicotómica (1: Inventario Cartera de Acciones Propia "Grande" Corredor Compra en el periodo t, 0: Otro Inventario Cartera de Acciones Propia Corredor Compra en el periodo t).

D_Inventario_Ven = Variable dicotómica (1: Inventario Cartera de Acciones Propia "Grande" Corredor Venta en el periodo t, 0: Otro Inventario Cartera de Acciones Propia Corredor Venta en el periodo t).

Metodología
Modelos econométricos multivariantes (GASCO)

Dependent Variable: B_A

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CANTIDAD_COM	8.528114	2.035679	4.189321	0.0000
CANTIDAD_VEN	-1.2806[illegible]	[illegible]	[illegible]	[illegible]
D_BANCHILE_COM	39861[illegible]	[illegible]	[illegible]	[illegible]
D_BCI_VEN	-649869.1	85147.42	-7.632281	0.0000
D_CELFIN_COM	352190.4	51876.31	6.789042	0.0000
D_CONS_COM	642407.6	74385.36	8.636210	0.0000
D_CORP_COM	38684[illegible]	[illegible]	[illegible]	[illegible]
D_EURO_COM	34848[illegible]	[illegible]	[illegible]	[illegible]
D_LA_COM	366549.5	51519.71	7.114743	0.0000
D_PENTA_COM	437439.9	106976.4	4.089124	0.0000
D_SANT_COM	219975.4	97582.70	2.254245	0.0242
D_SEC_COM	2175257.	237838.5	9.145938	0.0000
TAMANO_COM	66262[illegible]	[illegible]	[illegible]	[illegible]
TAMANO_VEN	-23900[illegible]	[illegible]	[illegible]	[illegible]
INVENTARIO_VEN	-309395.0	44329.53	-6.979435	0.0000
C	1285334.	46302.93	27.75923	0.0000
AR(1)	0.588096	0.008399	70.02102	0.0000

R-squared	0.371513	Mean dependent var	1446993.
Adjusted R-squared	0.370432	S.D. dependent var	1648726.
S.E. of regression	1308187.	Akaike info criterion	31.00800
Sum squared resid	1.59E+16	Schwarz criterion	31.02103
Log likelihood	-144464.8	Hannan-Quinn criter.	31.01243
F-statistic	343.6644	Durbin-Watson stat	2.202063
Prob(F-statistic)	0.000000		

Inverted AR Roots	.59

CANTIDAD OFERTADA COMPRA
CANTIDAD OFERTADA VENTA

BCI
CELFIN
CONSORCIO
CORPBANCA
EUROAMÉRICA
LARRAÍN VIAL
PENTA
SANTANDER INVESTMENT
VALORES SECURITY

TAMAÑO (PATRIMONIO)
TAMAÑO (CARTERA PROPIA)

Cantidad_Com	= Cantidad ofertada de compra de las acciones de Gasco en el periodo t.
Cantidad_Ven	= Cantidad ofertada de venta de las acciones de Gasco en el periodo t.
D_Corredori_Ven	= Variable dicotómica (1: Corredor Venta i en el periodo t, 0: Otro Corredor Venta i en el periodo t).
D_Corredori_Com	= Variable dicotómica (1: Corredor Compra i en el periodo t, 0: Otro Corredor Compra i en el periodo t).
D_Tamaño_Com	= Variable dicotómica (1: Patrimonio Grande Corredor Compra en el periodo t, 0: Otro Patrimonio Corredor Compra en el periodo t).
D_Tamaño_Ven	= Variable dicotómica (1: Patrimonio Grande Corredor Venta en el periodo t, 0: Otro Patrimonio Corredor Venta en el periodo t).
D_Inventario_Ven	= Variable dicotómica (1: Inventario Cartera de Acciones Propia "Grande" Corredor Venta en el periodo t, 0: Otro Inventario Cartera de Acciones Propia Corredor Venta en el periodo t).

Resultados y Conclusiones

Respecto a la explicación de la varianza del spread

- la contribución individual media de las variables, inventario y patrimonio del corredor es mayor que la contribución que hace la variable cantidad de acciones para explicar el comportamiento del spread
- la elección del inversionista por el tipo de intermediario no es menor. Mientras mayor es el tamaño del corredor, mayor es la influencia en el spread
- el tamaño de los corredores y la magnitud del inventario en acciones les permiten mantener carteras más diversificadas, y por lo tanto, estos corredores poseen más alternativas para ofrecer en la compra y venta de acciones, disminuyendo el spread de sus operaciones

Resultados y Conclusiones

Limitaciones y Futuras Investigaciones

- otras empresas del mercado chileno
- otros sectores industriales
- medir de diferentes formas spread
- nuevos indicadores que permitan medir la revelación de la calidad de información (Haat et al., 2006, Botosan, 1997)

Seminario Doctorado en Ingeniería Industrial
“Econometric Modeling and Solving Social Problems”

Hanns DE LA FUENTE-MELLA, Ing., Dr.