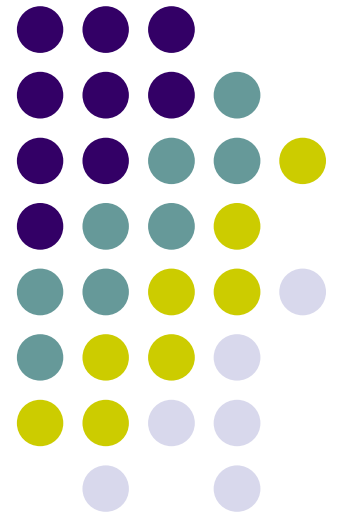
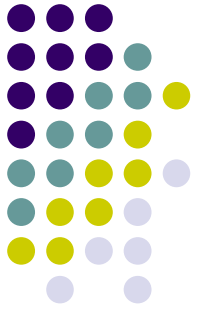


Altitude or Hot Air?

Rómulo A. Chumacero

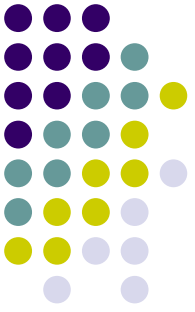


Motivation



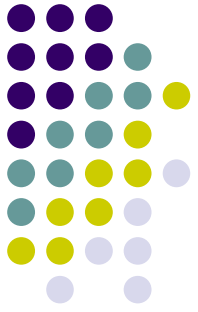
- Few things unite Bolivians
- FIFA's ban: Stadiums of 2,500 meters above sea level
- Hernando Siles (3,650), Atahualpa (2,850), El Campín (2,640)
- Altitude raised to 3,000
- Exception for the 2010 World Cup
- Sports commentators and visiting teams: “unfair advantage”
 - How to define it? Is altitude one? Are there others?
- Example: Home-field advantage
 - Physical factors (familiarity, climate, etc)
 - Psychological factors (crowd effects, etc)
 - Other factors (Refereeing bias)

Outline

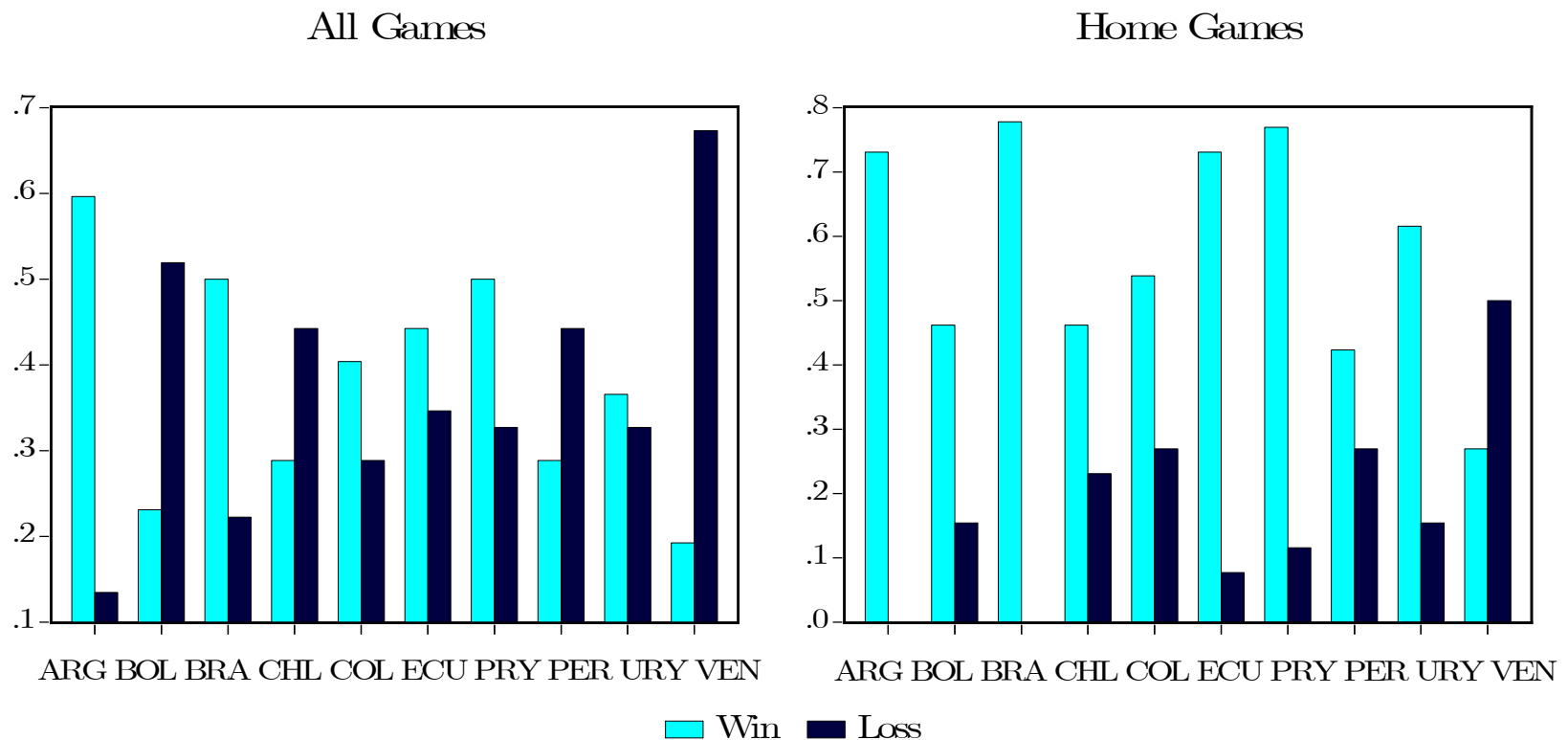


- Documenting home-field advantage
- Determinants of outcomes of matches
- Applications
- Concluding remarks

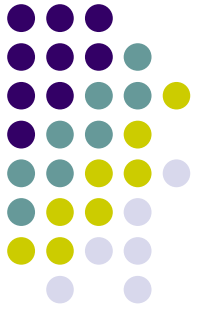
Home field advantage



- Data base: 252 matches (1998, 2002, and 2006)



Test for a winning record

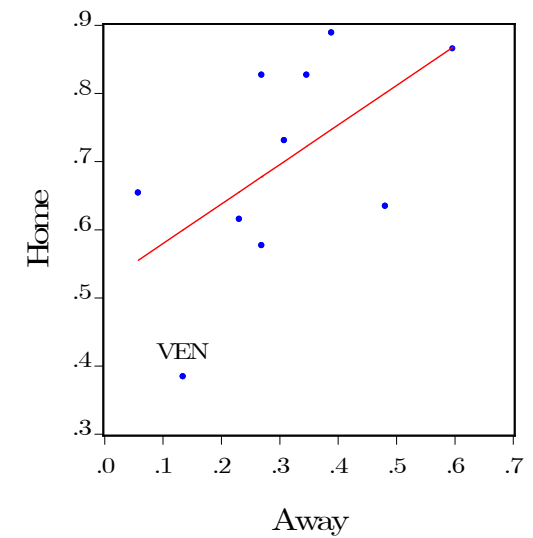
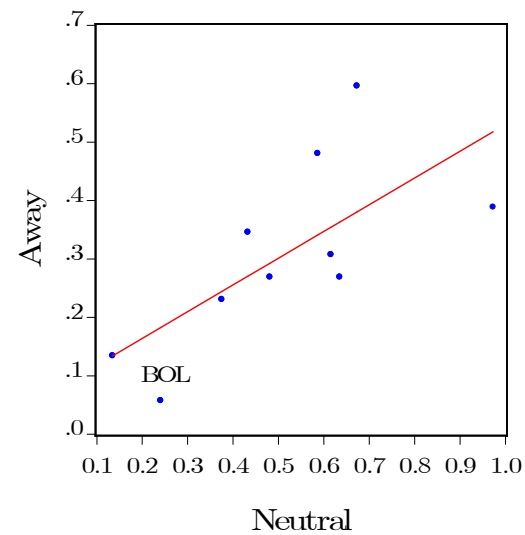
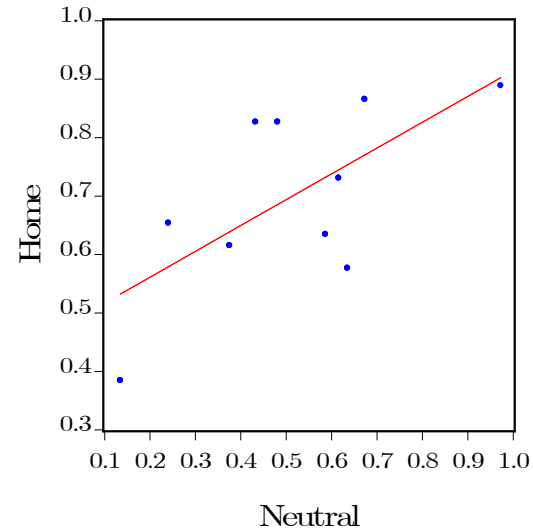
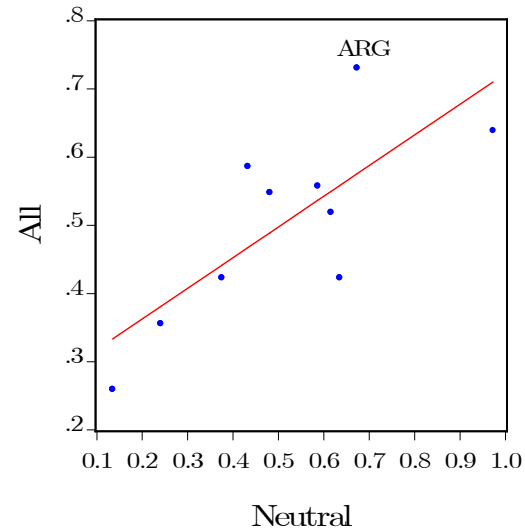
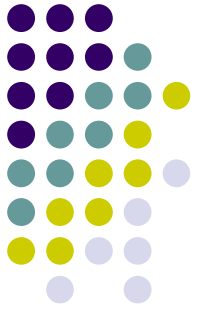


$$\sqrt{T_m} \begin{pmatrix} \hat{p}_k^m - p_k^m \\ \hat{q}_k^m - q_k^m \end{pmatrix} \xrightarrow{D} N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} p_k^m (1 - p_k^m) & -p_k^m q_k^m \\ -p_k^m q_k^m & q_k^m (1 - q_k^m) \end{bmatrix} \right),$$

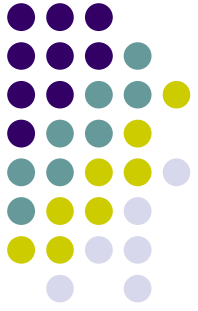
$$H_0 : p_k^m - q_k^m \leq 0$$

- Overall:
 - Only ARG and BRA have unconditional winning record
 - 7/10 have winning record at home (exclude CHL, PER, VEN)
 - None has winning record away
- Pairs:
 - ARG has winning record against 5
 - BRA and ECU against 3
 - PRY against 2
 - CHL and URY against 1
 - BOL, COL, PER and VEN against none
- It's good to be home!

Some relationships

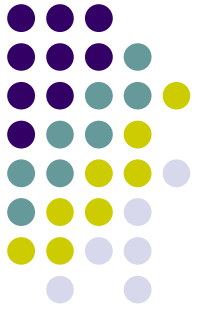


Conditioning factors



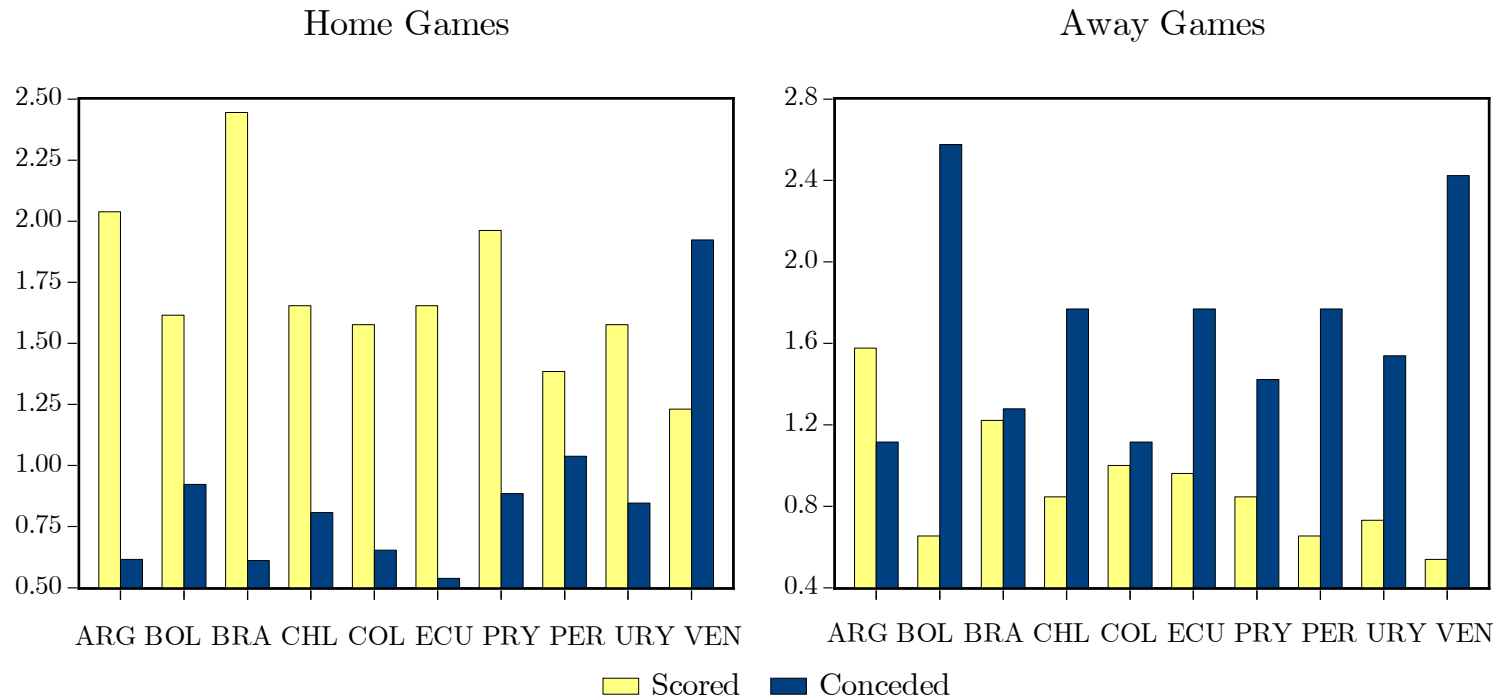
- Quality
 - FIFA rankings
 - Outcome of last match on neutral field
 - Performance on past matches (home and away)
 - Points on qualifying series
- Socioeconomic characteristics
 - Log of ratio of GDPs (PPP)
 - Log of ratio of populations
- Crowd effects
 - Spectators
 - Fraction of capacity
- Other factors
 - Altitude
 - Temperature
 - Humidity
 - Levels, distance, large values

Bivariate Poisson model

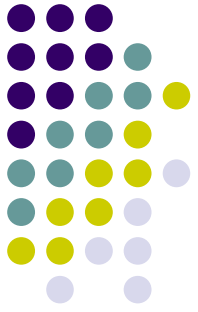


$$\Pr(g_{ij,t}^h = f, g_{ij,t}^a = r) = e^{-(\lambda_{ij,t} + \theta_{ij,t} + \gamma_{ij,t})} \frac{\lambda_{ij,t}}{f!} \frac{\theta_{ij,t}}{r!} \sum_{k=0}^{\min(f,r)} \binom{f}{k} \binom{r}{k} k! \left(\frac{\gamma_{ij,t}}{\lambda_{ij,t} \theta_{ij,t}} \right)^k.$$

$$\ln(\lambda_{ij,t}) = \beta' x_{ij,t}, \quad \ln(\theta_{ij,t}) = \delta' x_{ij,t}, \quad \gamma_{ij,t} = 0.$$

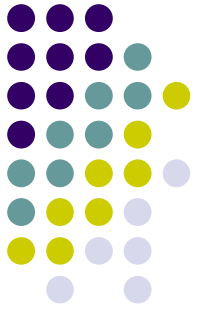


Results



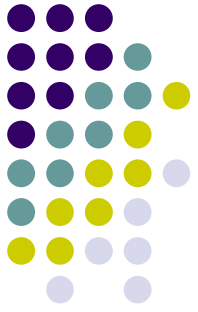
<i>Scored</i>		<i>Conceded</i>	
Constant	0.270 (0.116)	Constant	-0.372 (0.154)
$F_{i,t}$	-0.007 (0.002)	$F_{i,t}$	0.007 (0.002)
$F_{j,t}$	0.006 (0.001)	$N_{ij,t}$	-0.385 (0.160)
$d_{ij,t} > 0.18$	0.304 (0.116)	$d_{ij,t}$	0.848 (0.374)
$ e_{ij,t} $	0.027 (0.010)	Argentina	0.528 (0.176)
Colombia	-0.367 (0.174)		
R ² = 0.189; LogL = -397.6		R ² = 0.196; LogL = -296.2	

Low scoring



<i>Outcome</i>	<i>Observed</i>	<i>Predicted</i>
1-0	0.143	0.131
1-1	0.119	0.104
2-2	0.103	0.080
0-0	0.095	0.088
2-0	0.071	0.106
3-1	0.071	0.046

- Not a problem

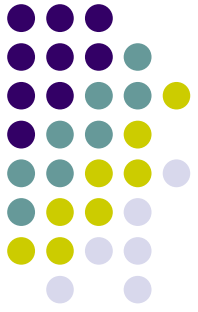


Performance

	<i>Probability of winning</i>		<i>Probability of losing</i>		<i>P-value</i>
	Observed	Predicted	Observed	Predicted	
ARG	0.596	0.566	0.134	0.213	0.238
BOL	0.231	0.293	0.519	0.499	0.523
BRA	0.500	0.531	0.222	0.247	0.745
CHL	0.288	0.368	0.442	0.393	0.450
COL	0.404	0.394	0.288	0.343	0.647
ECU	0.442	0.362	0.346	0.393	0.508
PRY	0.500	0.421	0.327	0.342	0.385
PER	0.288	0.341	0.442	0.416	0.702
URY	0.365	0.380	0.327	0.370	0.641
VEN	0.192	0.223	0.673	0.578	0.274
Total	0.571	0.549	0.183	0.218	0.342

- Good performance predicting scores, goals, records
 - Geometric mean: 0.421
 - Coincidence index: 61%
 - ARG wins more and losses less than forecasted
 - BOL wins less and losses more than forecasted

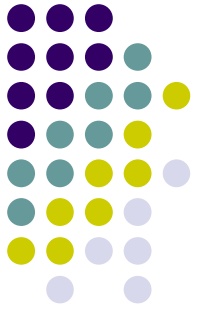
Ordered Probit model



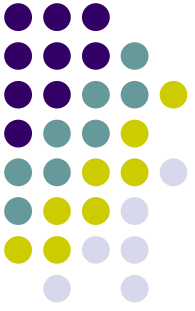
$$V_{ij,t} = \begin{cases} 2 & \text{if } \mu_2 < v_{ij,t}^* + \varepsilon_{ij,t} \\ 1 & \text{if } \mu_1 < v_{ij,t}^* + \varepsilon_{ij,t} \leq \mu_2 \\ 0 & \text{if } v_{ij,t}^* + \varepsilon_{ij,t} \leq \mu_1 \end{cases}$$
$$v_{ij,t}^* = \rho' x_{ij,t},$$

$N_{ij,t}$	0.429 (0.186)
$l_{ij,t}$	-0.0002 (5E-5)
$ e_{ij,t} $	0.042 (0.018)
Ecuador (home)	0.863 (0.314)
Peru (home)	-0.706 (0.256)
Venezuela (home)	-1.293 (0.294)
Argentina (away)	-0.798 (0.255)
Colombia (away)	-0.900 (0.257)
Pseudo R ² = 0.147; LogL = -209.5	

Predictions



	<i>Probability of winning</i>		<i>Probability of losing</i>		<i>P-value</i>
	Observed	Predicted	Observed	Predicted	
ARG	0.596	0.570	0.134	0.170	0.751
BOL	0.231	0.239	0.519	0.550	0.811
BRA	0.500	0.434	0.222	0.325	0.332
CHL	0.288	0.371	0.442	0.380	0.416
COL	0.404	0.482	0.288	0.249	0.517
ECU	0.442	0.421	0.346	0.388	0.810
PRY	0.500	0.419	0.327	0.342	0.365
PER	0.288	0.300	0.442	0.418	0.937
URY	0.365	0.419	0.327	0.328	0.636
VEN	0.192	0.150	0.673	0.621	0.131
Total	0.571	0.574	0.183	0.184	0.992



Evaluation of the results

- Performance
 - Geometric mean: 0.435
 - Coincidence index: 63%
- Model replicates record of teams, but
 - ARG, BRA, ECU perform better than forecasted
 - BOL, CHI, COL perform worse than forecasted
- Altitude not a factor?

Application: Little brother hypothesis

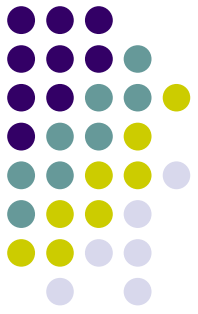


- Last match
 - URU vs ARG (2002: draw, 2006: win)
 - PAR vs COL (2002 and 2006: away win)

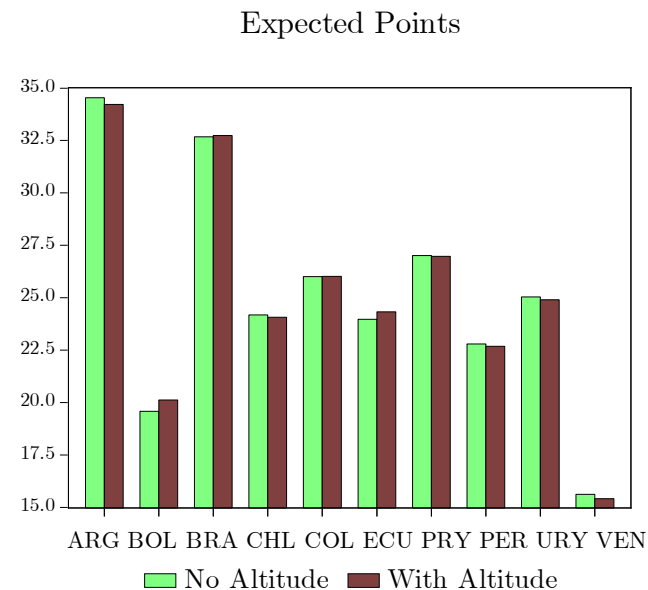
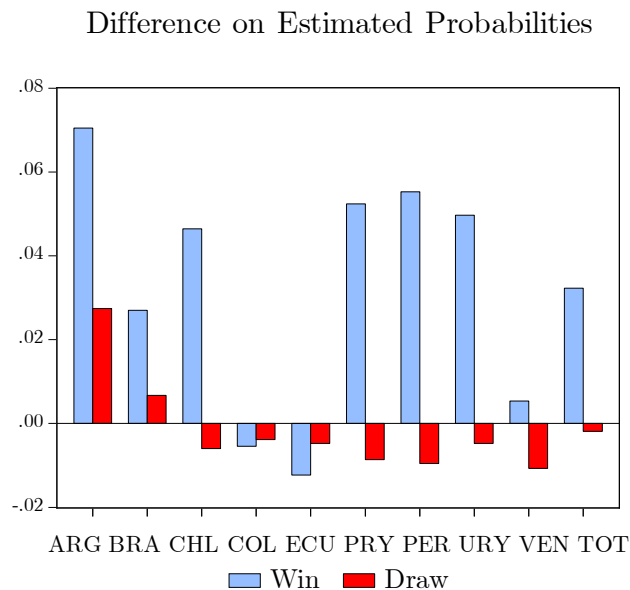
	<i>Double-Poisson</i>	<i>Ordered Probit</i>
Uruguay (home game)	0.411	0.424
Argentina (away game)	0.395	0.419
Last two URU-ARG matches	0.286	0.275
Paraguay (home game)	0.556	0.609
Colombia (away game)	0.333	0.349
Last two PRY-COL matches	0.151	0.207

- Did ARG choose to help URU?
- Suggestions
 - Last games: ARG vs BRA, COL vs URU
- Fixture is an “unfair advantage”

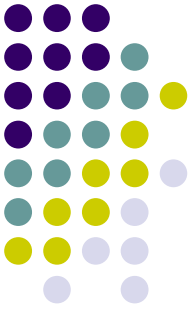
Application: The perils of playing in La Paz or Am I crazy?



- Models
 - Good job for home wins, but predict more losses
 - Forecasts less goals made and more conceded
- Special considerations
 - Away teams have more time to acclimatize
 - Irregular record
 - Libertadores Cup

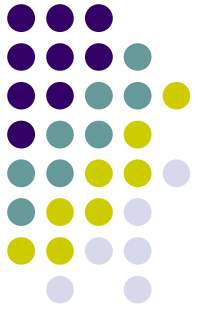


Why the fuzz?



- If a factor, not crucial
- Why not play elsewhere then?
 - 10,000 spectators (US\$ 1 million)
 - Decreased change
 - Build a new stadium?

What is more “inhumane”?

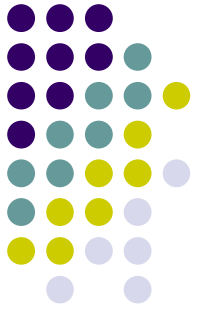


Playing in Bolivia?

Or playing with Bolivia?



Botton line



- Several factors are “unfair advantages”
- If altitude is one, it is not quantitatively important
- Use it or pay for not being exposed to it. How?
 - Prob of win, importance of match, valuation (benefits to players and fans)
- Humble proposal
 - Play all games on neutral field
 - Previously agreed micro-climate (altitude, temperature, humidity artificially controlled)
 - No spectators
 - Computer in, referee out
- Until then: let each team choose where to play

Altitude or Hot Air?

Rómulo A. Chumacero

