

# Combinations of Sample Surveys or Projections of Political Opinions

Daniel Thorburn  
Can Tongur  
Stockholm University

August 2012, Valmiera, Latvia  
Baltic-Nordic-Ukrainian Network in Survey Statistics

In Sweden as in other countries many  
opinion polls are performed every month

The next slide shows the outcomes a typical  
month.

## Opinion polls in Sweden March 2012

| Företag      | M    | FP  | C   | KD  | S    | V   | MP   | SD  | Övr | Und  | Publ    | Interview period | S Size |
|--------------|------|-----|-----|-----|------|-----|------|-----|-----|------|---------|------------------|--------|
| Ipsos /Synov | 29,7 | 6,1 | 5   | 3,4 | 35,6 | 5,7 | 9,3  | 4,2 | 5,3 | 20,2 | 30 mars | 8 - 27 mars      | 2 502  |
| Novus        | 30,5 | 5,7 | 4,8 | 3,6 | 32,4 | 7,3 | 8,8  | 5,7 | 6,9 | 8,6  | 21 mars | 21 feb - 19 mars | 2 000  |
| Sifo         | 28,7 | 5,9 | 4,9 | 3,5 | 33,7 | 6,1 | 10,3 | 5,2 | 6,9 | 11,9 | 18 mars | 5 - 15 mars      | 1 918  |
| Demoskop     | 32,9 | 5,5 | 4,9 | 3,6 | 29,7 | 6,3 | 10,9 | 5,2 | 6,2 | 11,2 | 14 mars | 29 feb - 7 mars  | 1 000  |
| YouGov       | 32,1 | 5,6 | 4,7 | 4,2 | 29,6 | 8,3 | 9,4  | 5,4 | 6,2 |      | 12 mars | 1 feb - 6 mars   | 1 754  |
| United Minds | 29,3 | 5,8 | 4,7 | 4,6 | 32,5 | 7,3 | 7,8  | 7,5 | 7,9 | 11   | 11 mars | 13 feb - 11 mars | 1 161  |
| Skop         | 33,4 | 6,9 | 5   | 3,3 | 30,8 | 6,5 | 9,1  | 4,5 | 5,1 | 9    | 7 mars  | 19 feb - 6 mars  | 1 000  |
| Sentio Resea | 29,7 | 4,3 | 4,2 | 4,6 | 31,8 | 7,4 | 8,1  | 8,6 | 9,9 |      | 5 mars  | 23 feb - 1 mars  | 1 035  |

# What conclusions could be drawn?

- Different sample sizes
- Different interview periods
- Different methods, design effects
- If we want to say something about the political situation e.g. March 31st, what can be said?
- Similar problems arise in other studies e.g. monthly LFS or or related studies e.g. EU-SILC and ULF in Sweden

- There has recently been some attempts to combine opinion polls, under the name of poll of polls. (Silver N, 2008)
- The main idea is exponentially smoothing and a constant trend
- We shall instead view the true party preference as a Gaussian process over time.

- The simplest model, Brownian motion, SRS

$$dP_t = \gamma dW(t)$$

$P_t$  = *Party preference at time t*

$W(t)$  = *Wiener noise*

$$X_t | P_t \in N(P_t, P_t(1 - P_t) / n_t) = N(P_t, V_t)$$

*is an opinion poll at time t*

- A model variant with trend (modelled by an Ornstein Uhlenbeck process

$$dT_t = -\alpha T_t dt + \beta dW_1(t)$$

$$dP_t = T_t dt + \gamma dW_2(t)$$

$T_t$  = *trend or stochastic drift*

# Updating

1. Suppose that after an opinion poll at time  $t$  the distribution of the party preference is

$$P_t \in N(E_t | C_t)$$

2. At the next time  $t+s$  the party preference is (for the simplest model)

$$P_{t+s} \in N(E_t | C_t + \gamma s)$$

3. If an opinion poll is made at time  $t+s$  the party preference becomes

$$P_{t+s} | X_{t+s} \in N\left(\frac{\frac{E_t}{C_t + \gamma s} + \frac{X_{t+s}}{V_{t+s}}}{\frac{1}{C_t + \gamma s} + \frac{1}{V_{t+s}}}, \frac{1}{\frac{1}{C_t + \gamma s} + \frac{1}{V_{t+s}}}\right)$$

4. In this way one may successively update the best estimate of the party preference and the variance of the estimate

# Variants

- The formulas for the model with trend are in principle the same but much more complicated
- We also used a design effect (e.g. due to weighting and calibration)

$$X_t \mid P_t \in N(P_t, \delta P_t(1 - P_t) / n_t)$$

- Institute effect  $\theta_i$

$$X_t \mid P_t \in N(P_t + \theta_i, \delta P_t(1 - P_t) / n_t) \text{ for institute } i$$

(a restriction is needed in the estimation phase)



# Logit transformation

Many Swedish parties are quite small and it is more realistic to assume that

$$Q_t = \ln(P_t) - \ln(1 - P_t)$$

follows a random process of this kind. In this way the proportions will always stay between 0 and 1 and the absolute fluctuations will be larger for a large party than for a small one

# Estimation in retrospect

- Previously we derived the estimate for the party proportion at a certain date given all previous polls
- One may also try to estimate the proportion at a historic date given also all the polls at later time points

$$P_{t+s} | X_t \in N\left(\frac{E_{1,t} / C_{1,t} + E_{2,t} / C_{2,t} + X_t / V_t}{1/C_{1,t} + 1/C_{2,t} + 1/V_t}, \frac{1}{1/C_{1,t} + 1/C_{2,t} + 1/V_t}\right)$$

(for the simplest model). The estimates indexed by 1 are based on previous polls, and by 2 based on only later polls.

# Data

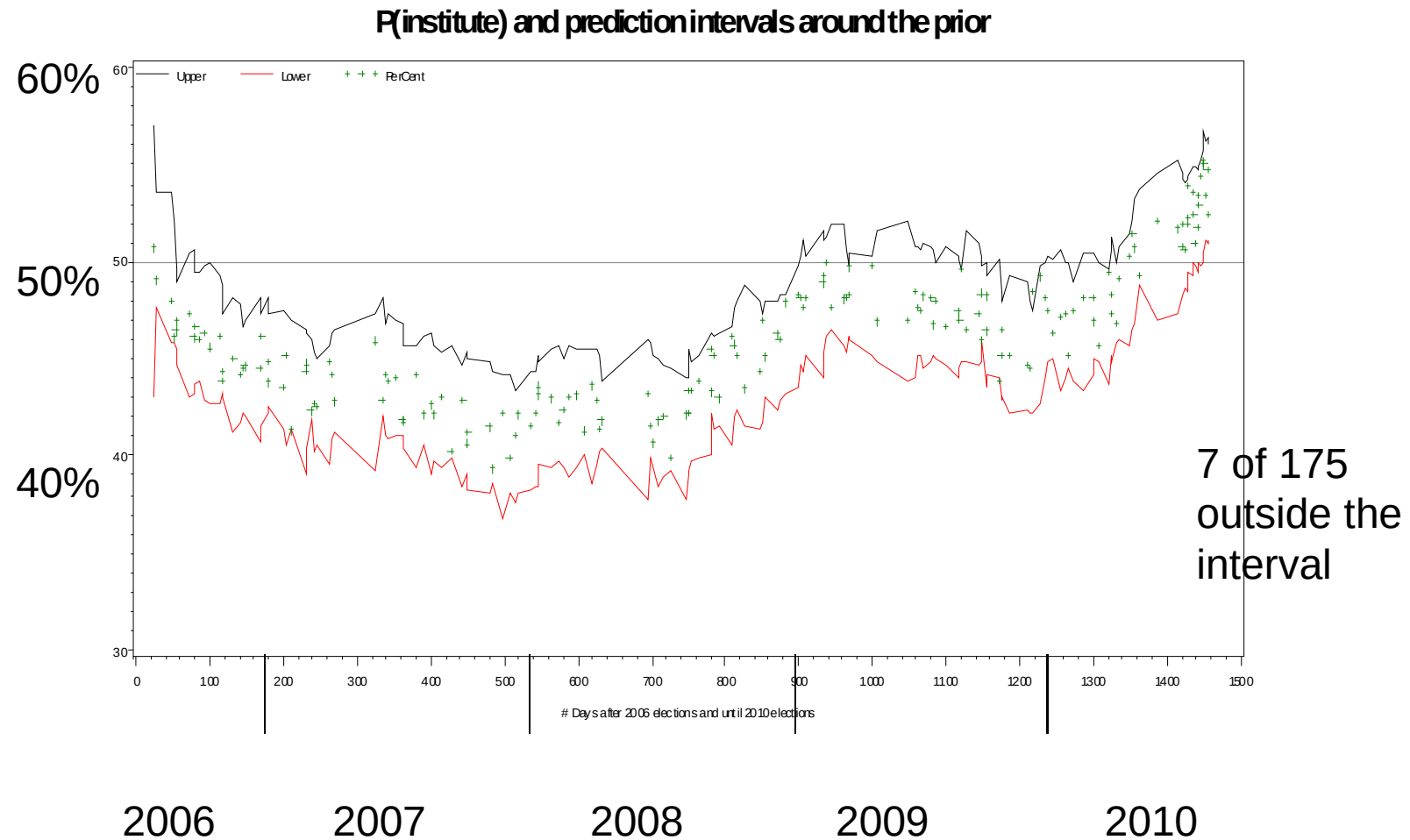
- All polls by SIFO, Temo, Skop, Novus, SCB during the period 2006 – mid March 2012.
- Date of a poll is their mean date of the collection period.
- Web panel surveys, which are still considered inferior, are omitted.
- Some less well known institutes and some institutes doing surveys only during part of the period are also omitted.
- Data are found at <http://www.novusgroup.se/vaeljaropinionen/samtliga-svenska-vaeljarbarometrar>

# Estimation and model tests

- Parameters were estimated by Maximum Likelihood (numerical maximisation)
- Model tests were done by score tests (change in the likelihood function)
- One institute changed its estimation method (from not weighting with respect to previous elections). Its bias changed significantly. In the analysis that institute had two constants ( $\eta = 3.44$  % on the party block level)
- Parameter estimates for party blocks (interelection period)
- $\alpha = 0.0042$ ,  $\beta = 0.00072$ ,  $\gamma = 0.20$ ,  $\delta = 0.70$ ,  $\eta = 3.44$ ,  
 $\theta_1 = 0.94$ ,  $\theta_2 = 0.43$ ,  $\theta_3 = 0.70$ ,  $\theta_4 = -0.57$ ,  $\theta_5 = 0.11$
- Non-significant effect of inclusion of trend,  $\alpha$ .  $\delta$  = design effect.

# The alliance, prediction of polls

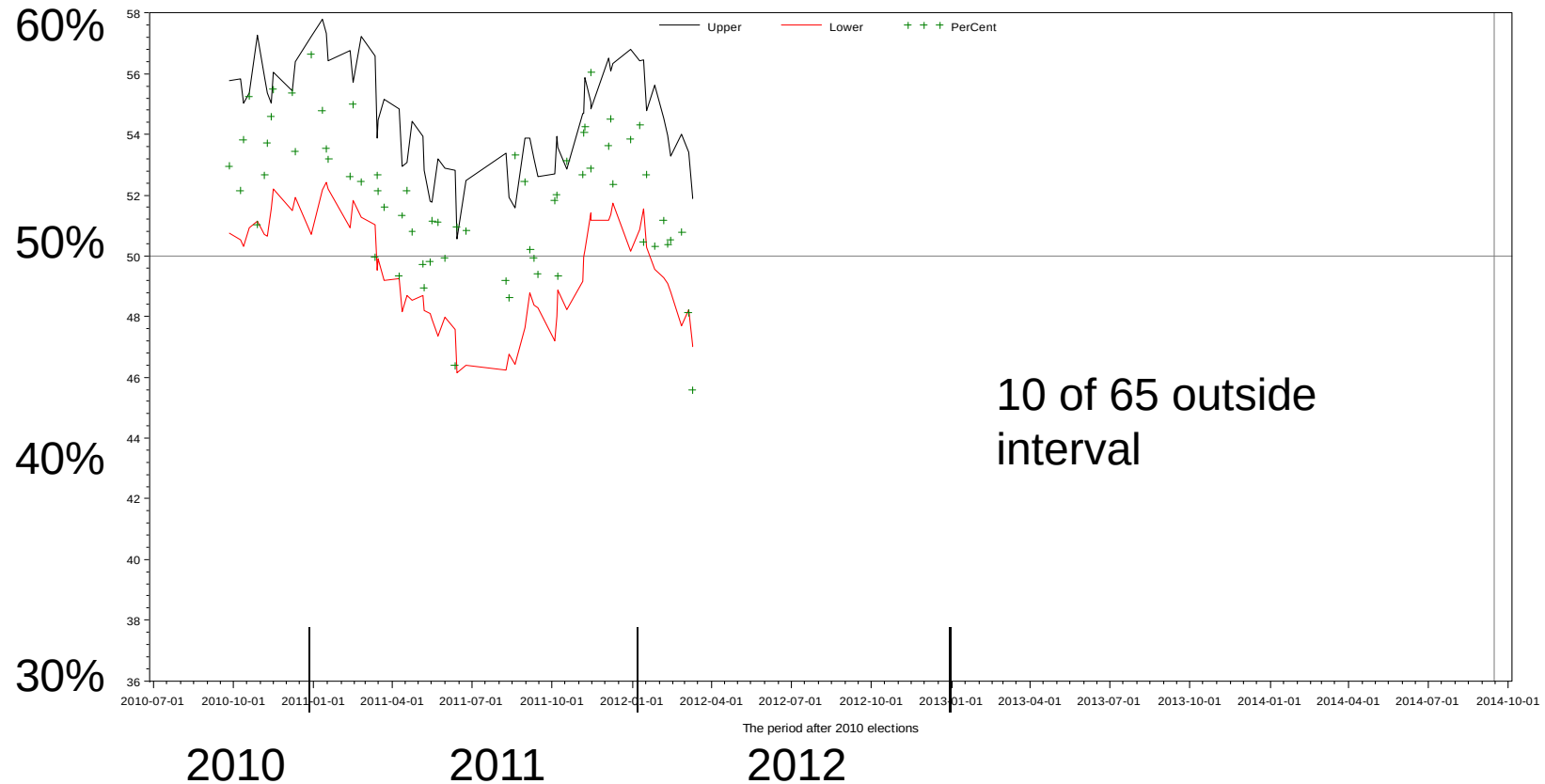
(based on data before the poll, interelection period)



# The alliance, prediction of polls

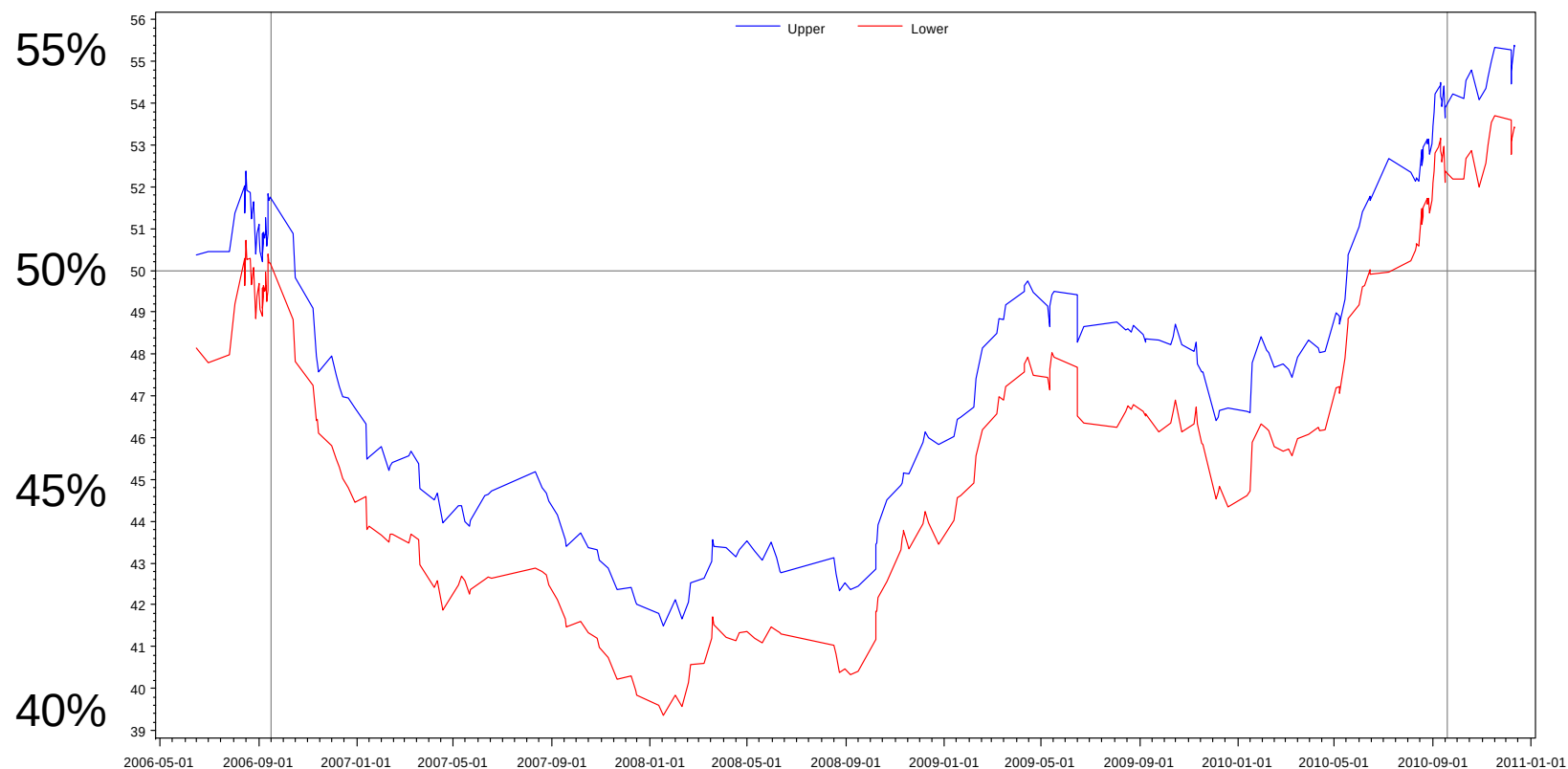
(parameters based on data before the poll,  
postelection period)

P(institute) and prediction intervals around the prior

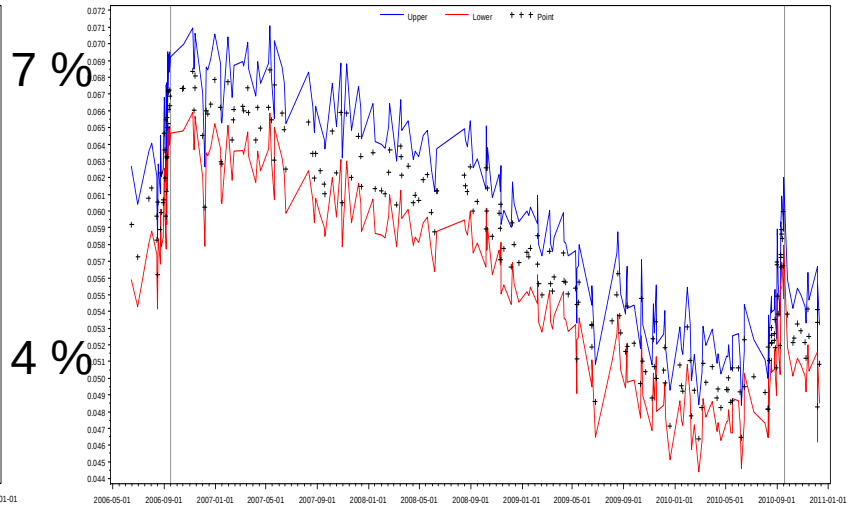
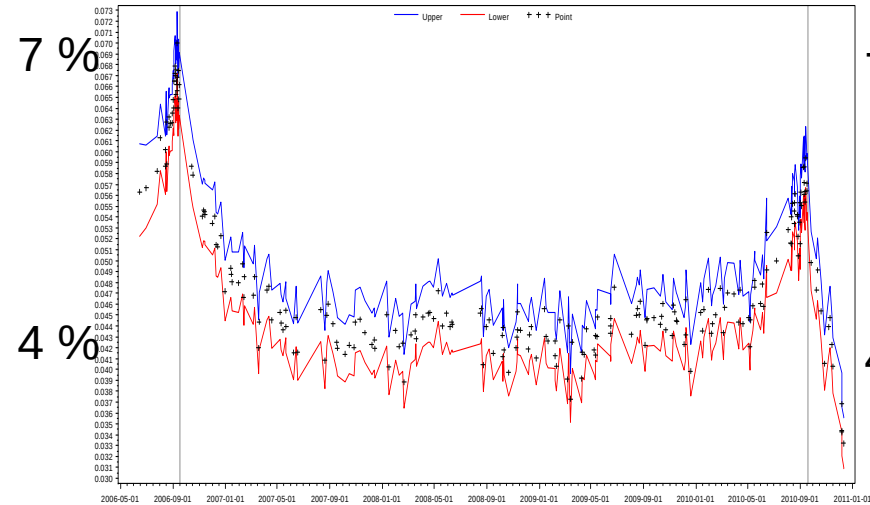
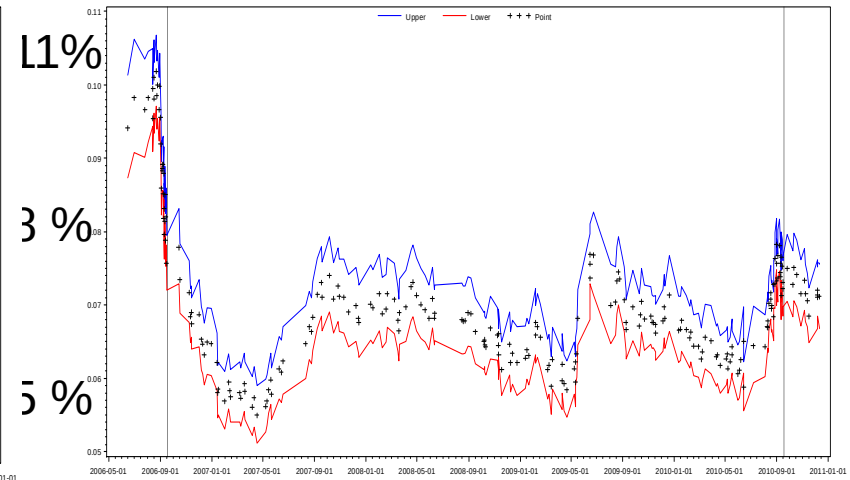
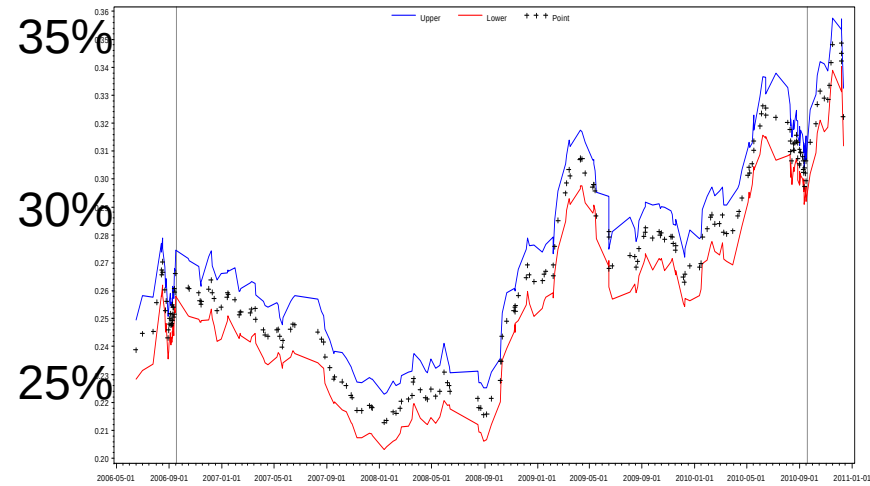


# Estimates for the Alliance during last interelection period

All data, 95 % probability intervals, excl. minor parties and Sd

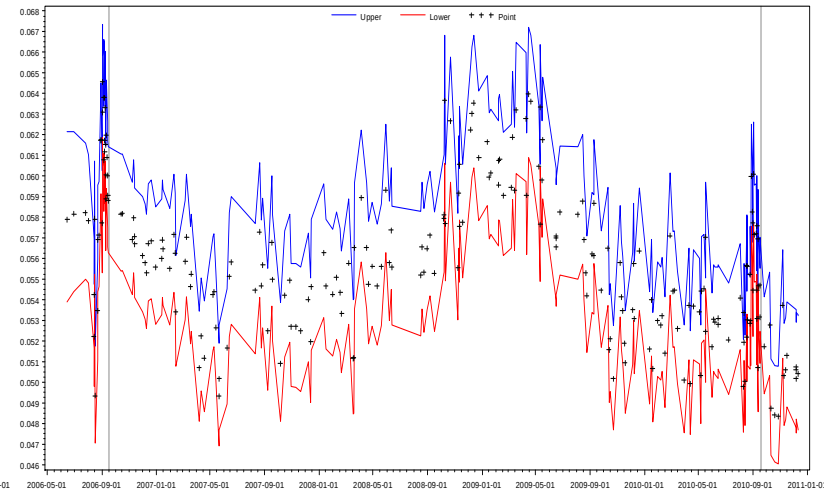
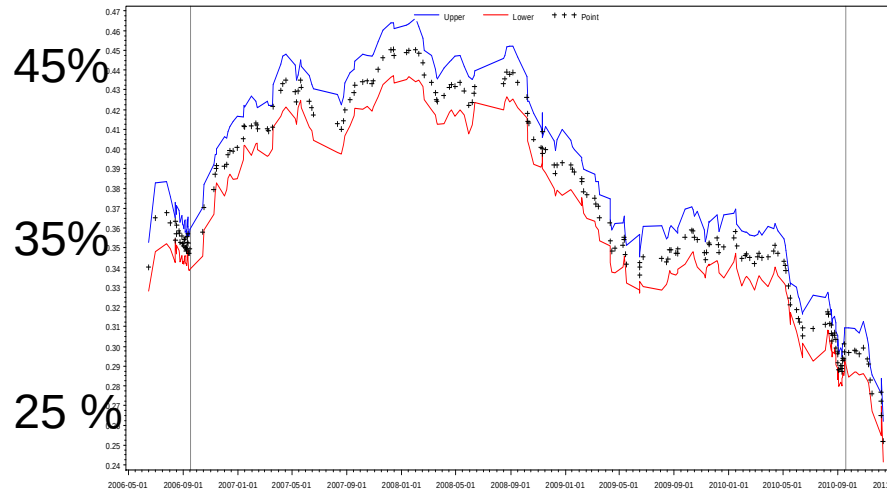


# The alliance, M, Fp, Kd and C





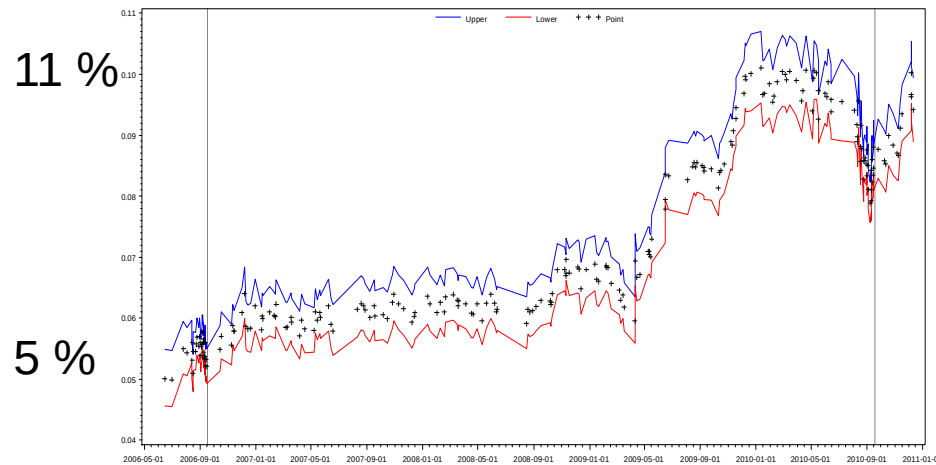
# The red-green (S, V, Mp)



5.5

5,0

4,7



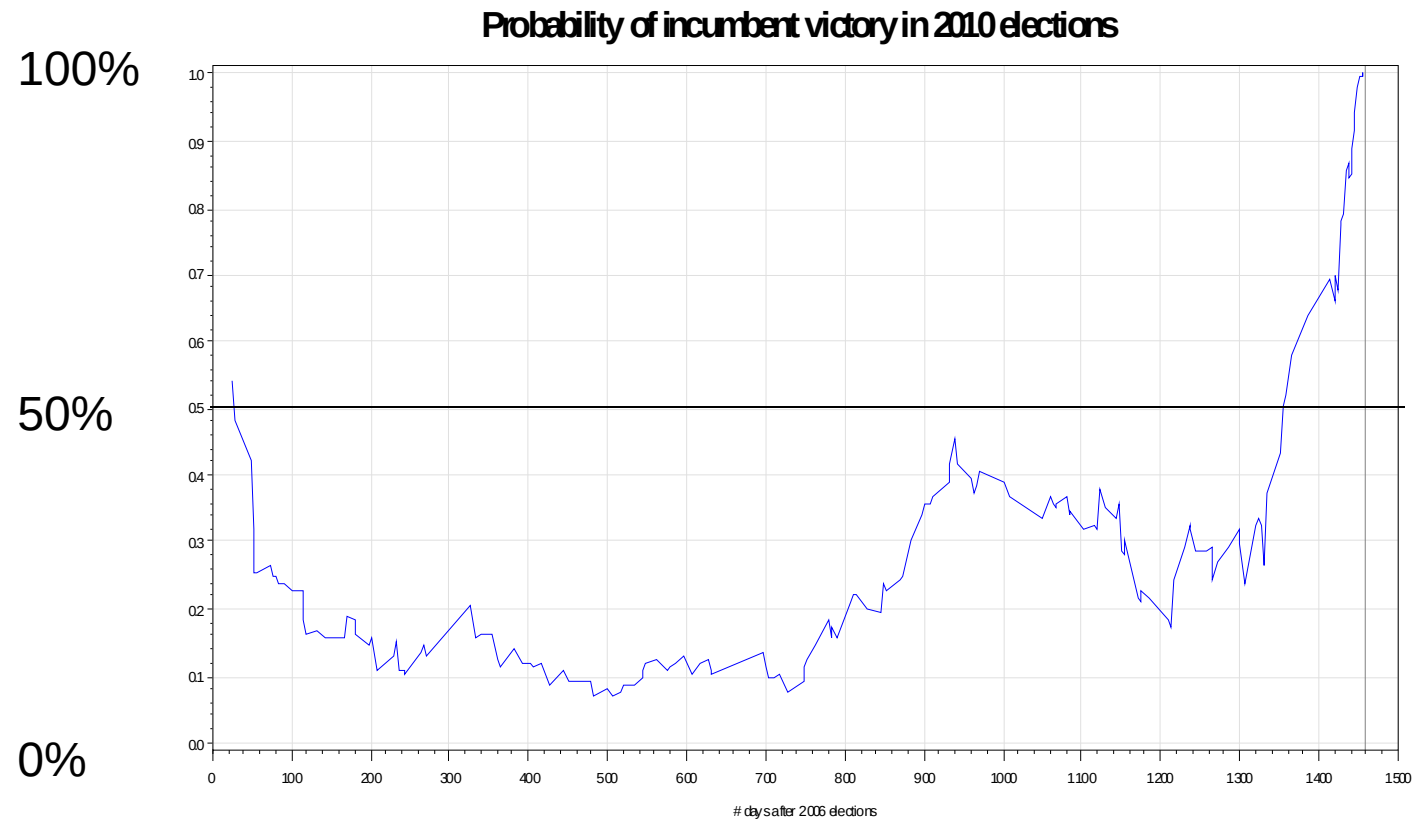
11 %

5 %

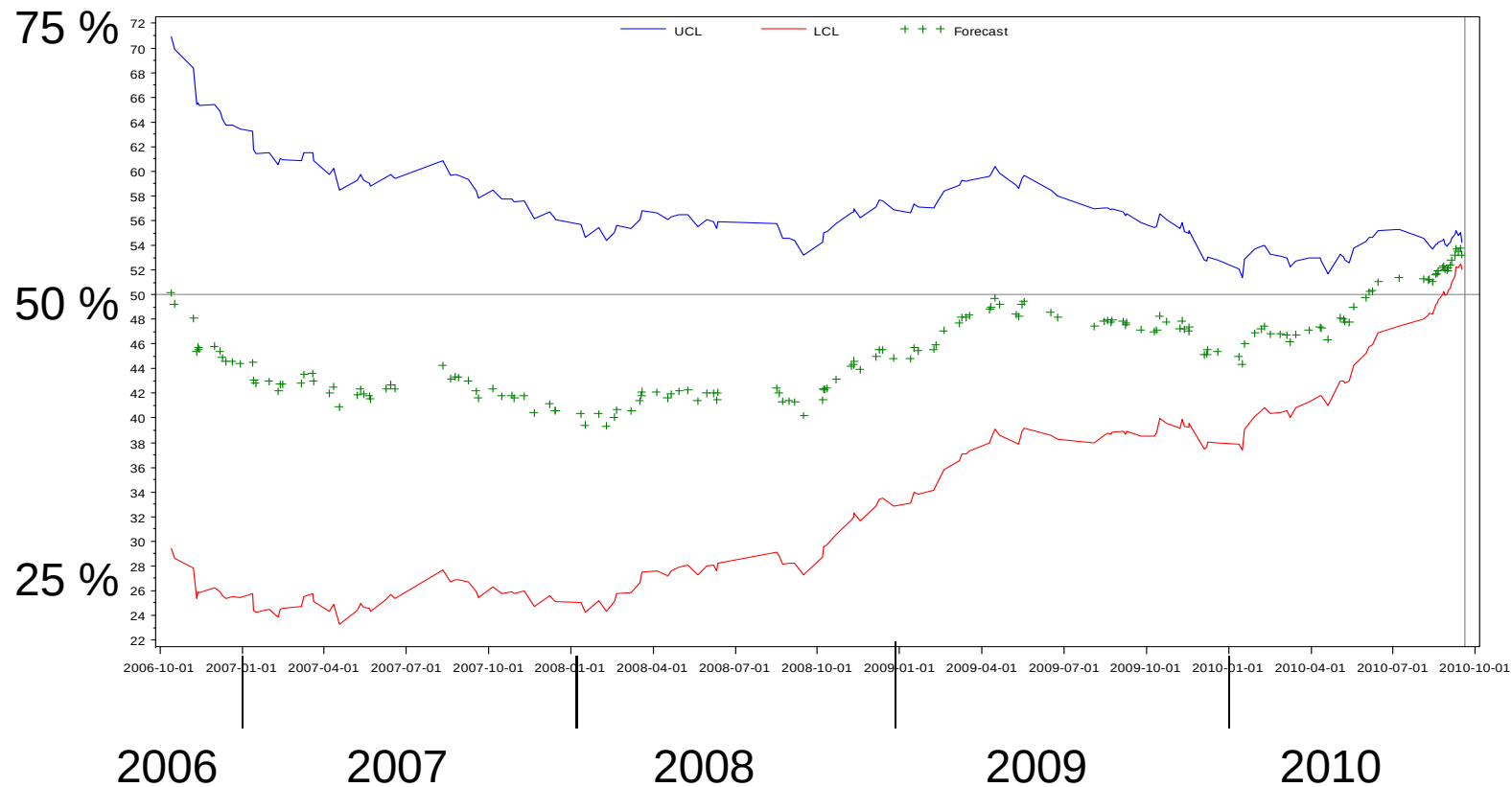
# Projections - Forecasts

- The model was developed for estimates of the present levels
- But the model can be used for forecasting also, but remember that the forecasts will be based only on the data and not on any extra knowledge like
  - Political science knowledge like that
    - the opinion usually swings against the sitting government between elections
    - tactical voting exists, comrade 4 % or the opposite
  - Political initiatives and proposals like
    - changes of leaders (e.g. from an unpopular one)
    - forming/unforming of alliances
  - Outcomes of the present/former government like
    - a good economy
    - low unemployment
    - green house effects and pollution

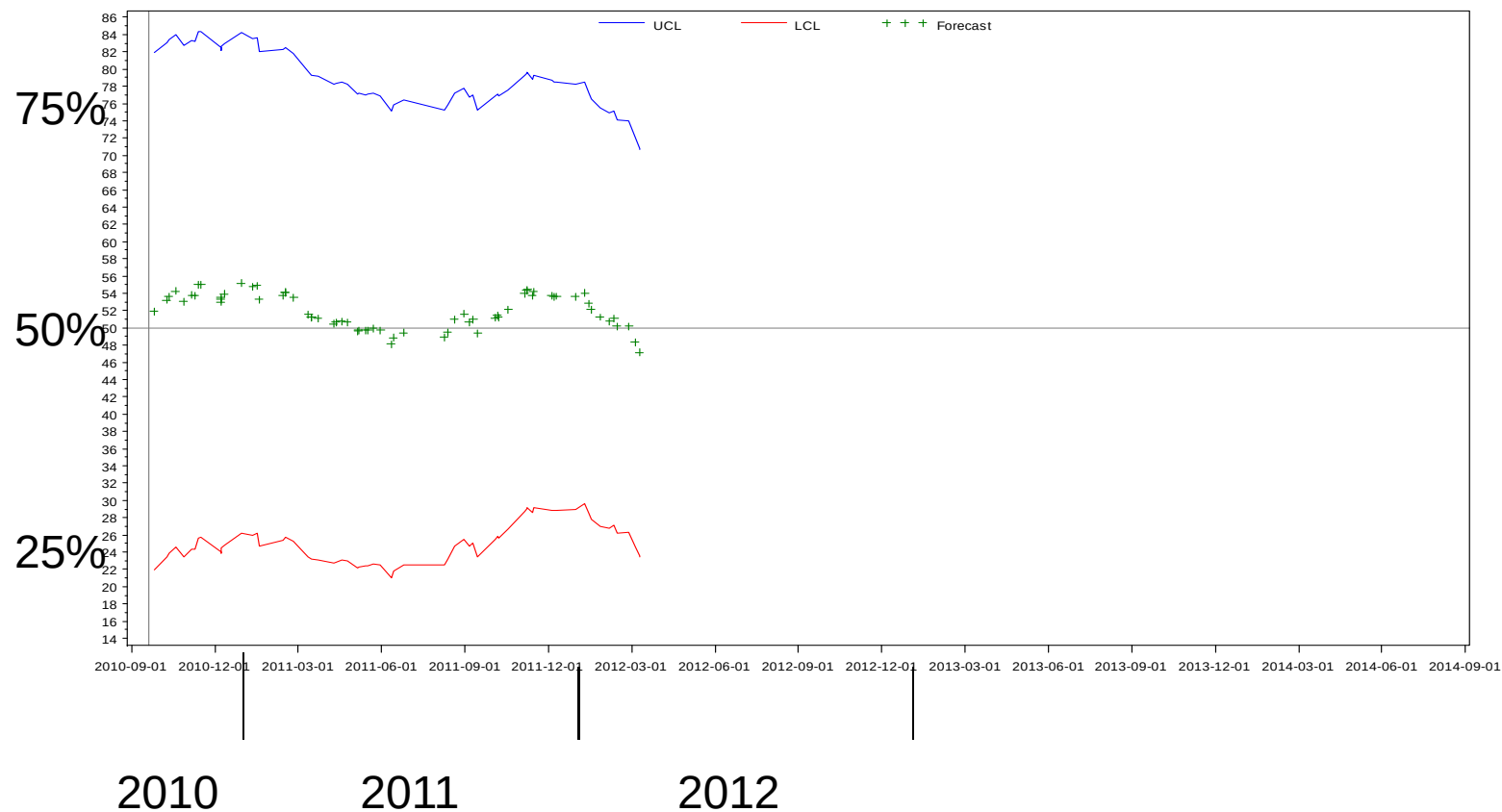
# Probability of the alliance getting more votes than the red-green in the election 2010 based on previous polls



# Forecasts for the 2010 election, based on previous polls



# Forecasts for the next election, 2014, based on previous polls



# One may also look at other things

- The probability for a party of being above the 4 % level in the next election
- The probability for a party to increase its share of voters in the next election

Thank You  
for your patience