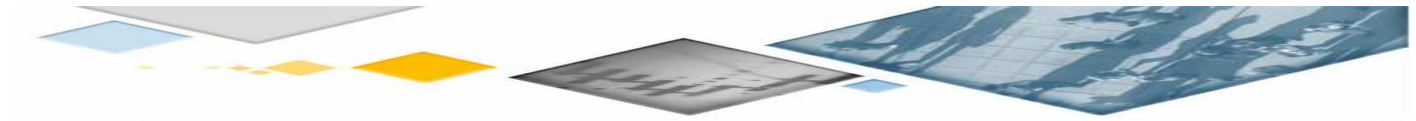




Small Area Estimations in the Industrial Survey

Iosune Azula, Patxi Garrido, Haritz Olaeta

Neuchatel, 14 Sep 2011



Outline

- Introduction
- Estimation procedure
- Main results for years 2002-2009
- Conclusions and Future Work



Introduction

- Yearly estimates since 2005
- 20 “comarcas” or counties

C.A. de Euskadi	13.008.214
------------------------	-------------------

Alava	2.676.614
--------------	------------------

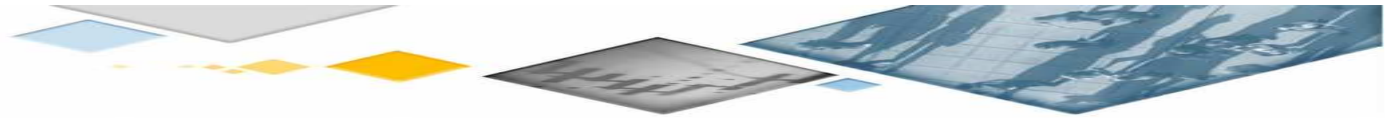
Valles Alaveses	89.743
Llanada Alavesa	1.708.733
Montaña Alavesa	17.455
Rioja Alavesa	243.155
Estribaciones del Gorbea	216.196
Cantábrica Alavesa	401.332

Bizkaia	5.371.448
----------------	------------------

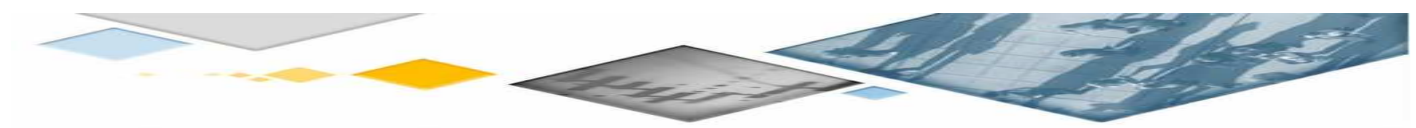
Arratia-Nervión	252.658
Gran Bilbao	3.259.555
Duranguesado	1.109.460
Encartaciones	154.460
Gernika-Bermeo	180.951
Markina-Ondarroa	149.595
Plentzia-Mungia	264.769

Gipuzkoa	4.960.152
-----------------	------------------

Bajo Bidasoa	218.659
Bajo Deba	482.592
Alto Deba	996.600
Donostia-San Sebastián	1.585.031
Goierri	655.356
Tolosa	422.004
Urola Costa	599.910

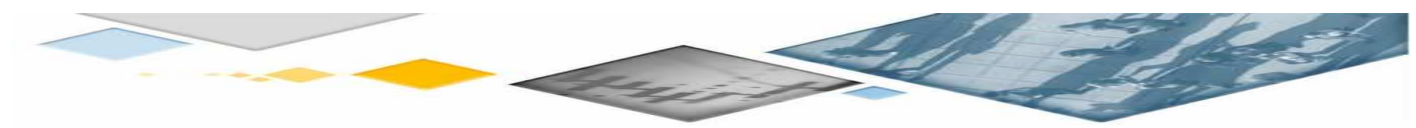


- Industrial Survey
 - Objective: 50 sub-classes (A84) at province level (3 provinces)
 - Sample size: 3.500 establishments
 - 2-stage stratified sampling scheme for < 20 employees:
 - 1st: optimum allocation by province
 - 2nd: proportional to employment by sub-class
 - Postestratification
 - Composite estimator since 2005. Auxiliary variable: employment from the Directory of Economic Activities of Eustat



Small Area Estimation Procedure

- Underlying population model
 - Fixed effects linear model
 - Mixed linear model



Fixed effects linear model

$$y_{dj} = \beta x_{dj} + e_{dj} \quad d = 1, \dots, t \quad j = 1, \dots, n_d$$

$$e_{dj} \subset N(0, \sigma_e^2 c_{dj}^{-1}) \quad c_{dj} = 1 / x_{dj}$$

$$\hat{t}_d^F = \sum_{j=1}^{n_d} y_{dj} + X_{d(p_r)} \hat{\beta} \quad d = 1, 2, 3$$



Mixed linear model

$$y_{dj} = \beta_0 + \beta_1 x_{dj} + v_d + e_{dj} \quad d = 1, 2, 3 \quad j = 1, \dots, n_d$$

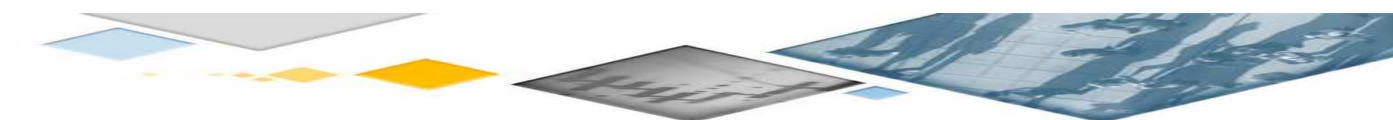
$$v_d \subset N(0, \sigma_v^2) \quad e_{dj} \subset N(0, \sigma_e^2 c_{dj}^{-1}) \quad c_{dj} = 1 / x_{dj}$$

$$\hat{t}_d = X_{d(p_r)}^\top \hat{\beta} + (N_d - n_d) \hat{\gamma}_{dc} \left(\bar{y}_{dc} - \bar{x}_{dc} \hat{\beta} \right) + \sum_{j=1}^{n_d} y_{dj} \quad d = 1, 2, 3$$

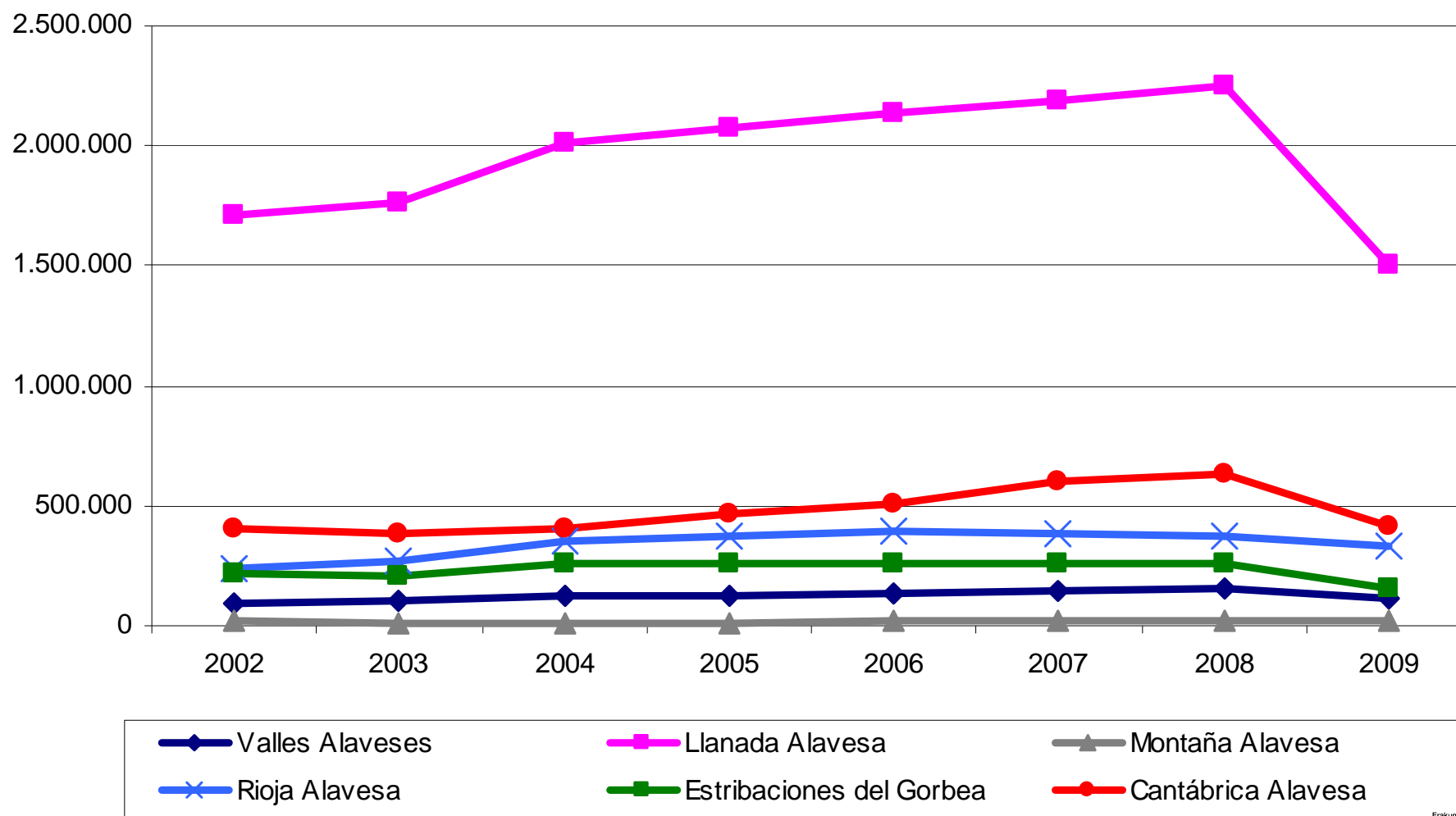


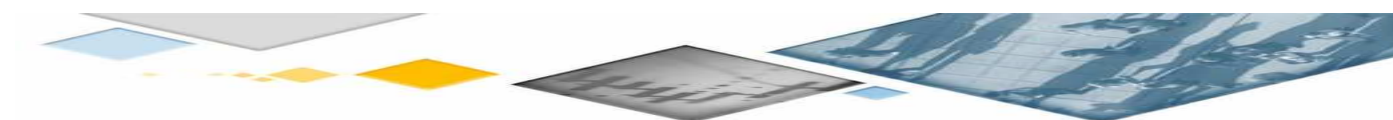
Procedure

- For each county and activity sub-class:
 - Min. # establishments: 5
 - Mixed linear model
 - If any of the $\sigma=0$ then the Fixed effects linear model used
 - Prediction at sub-class level
- Calibration to province level estimates from the Industrial Survey

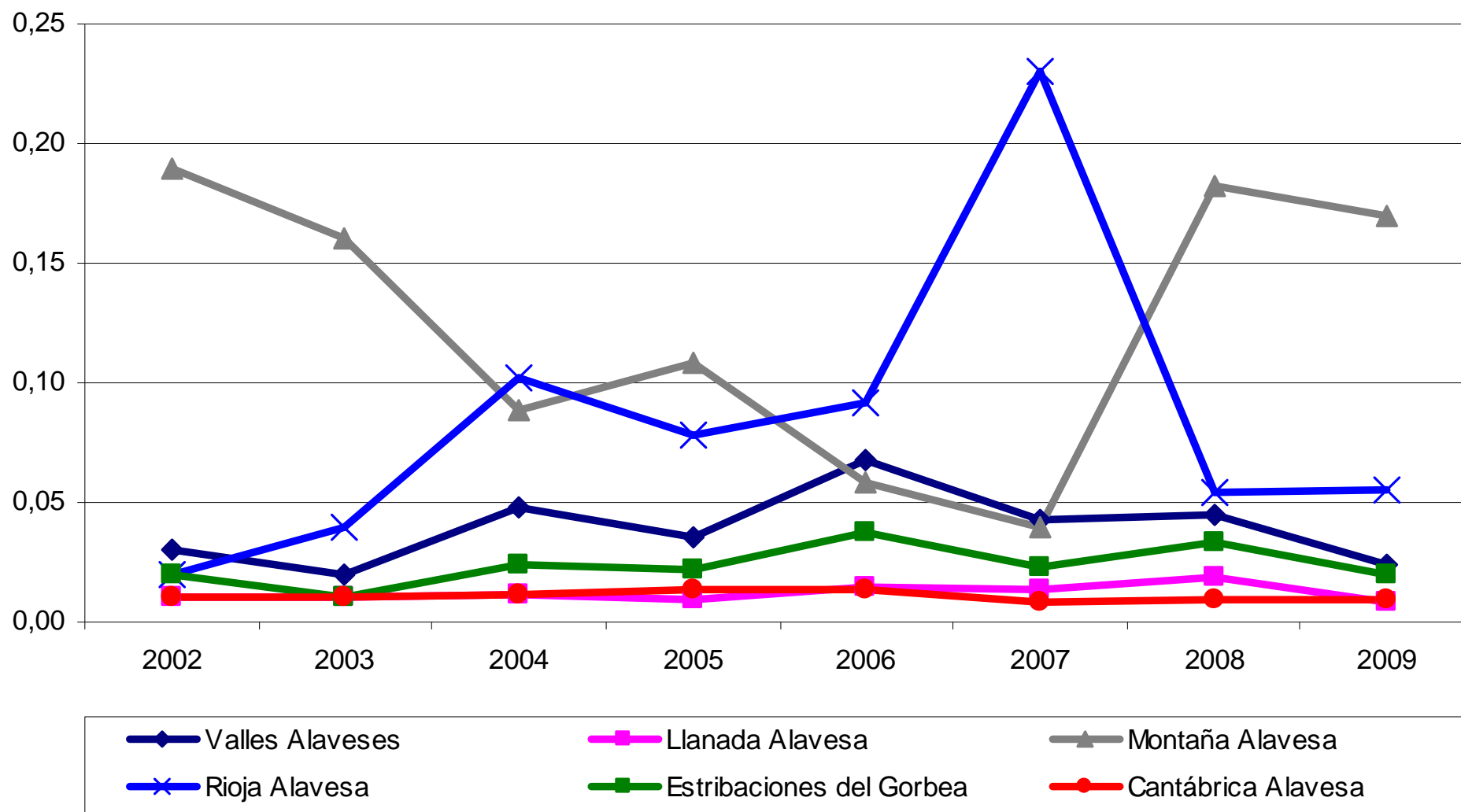


Araba (Value Added)

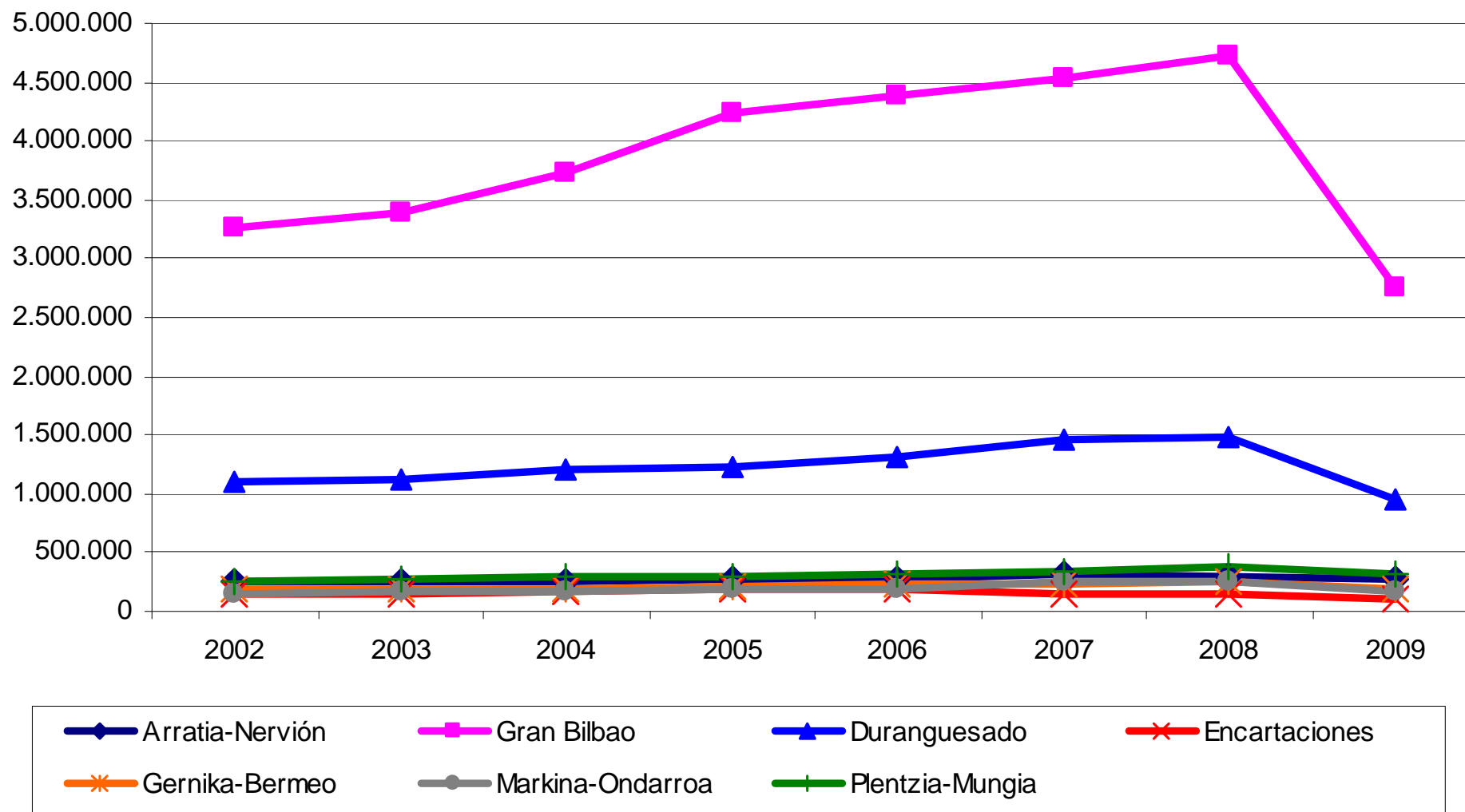


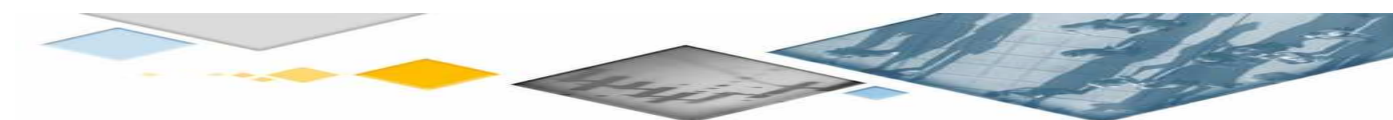


Araba (c.v)

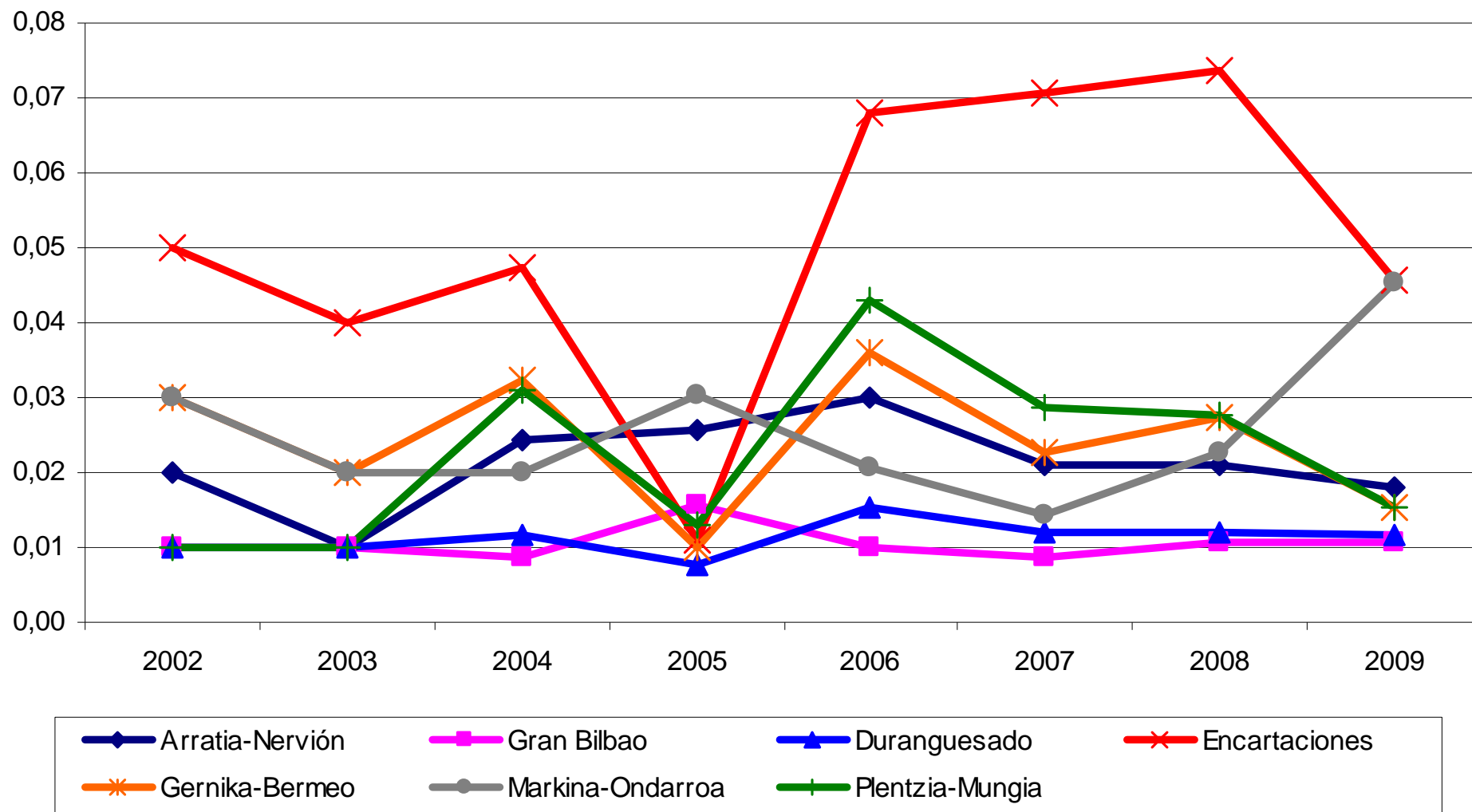


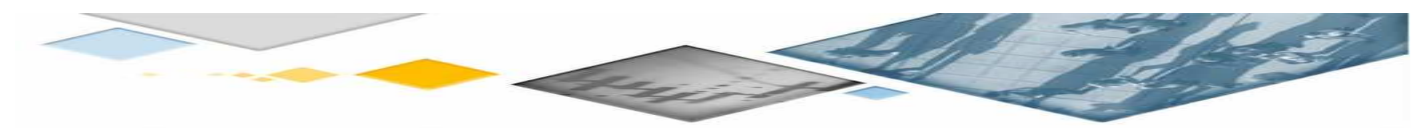
Bizkaia (Value Added)



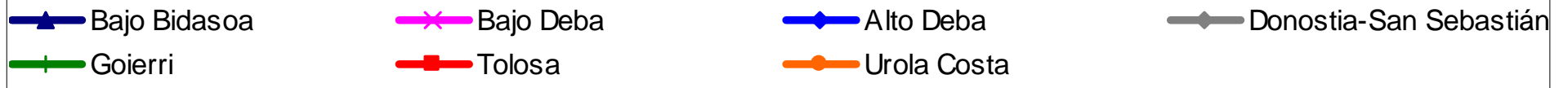
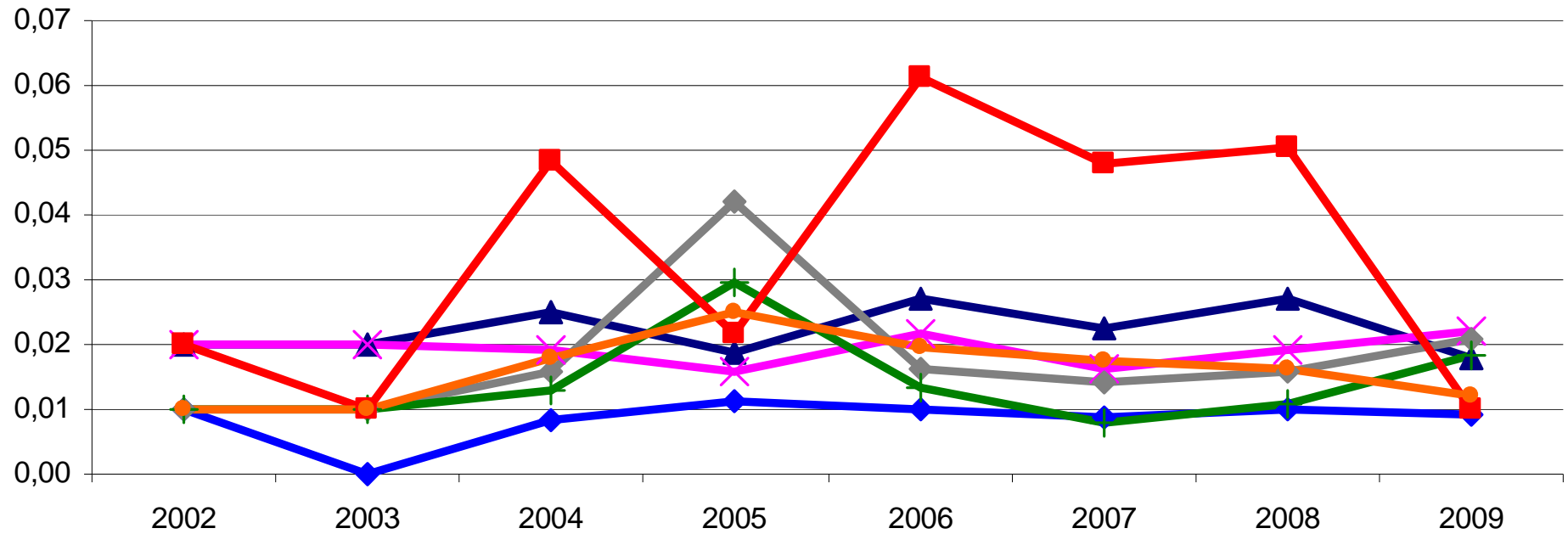


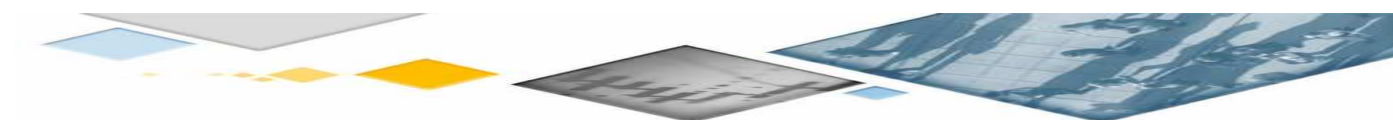
Bizkaia (c.v)



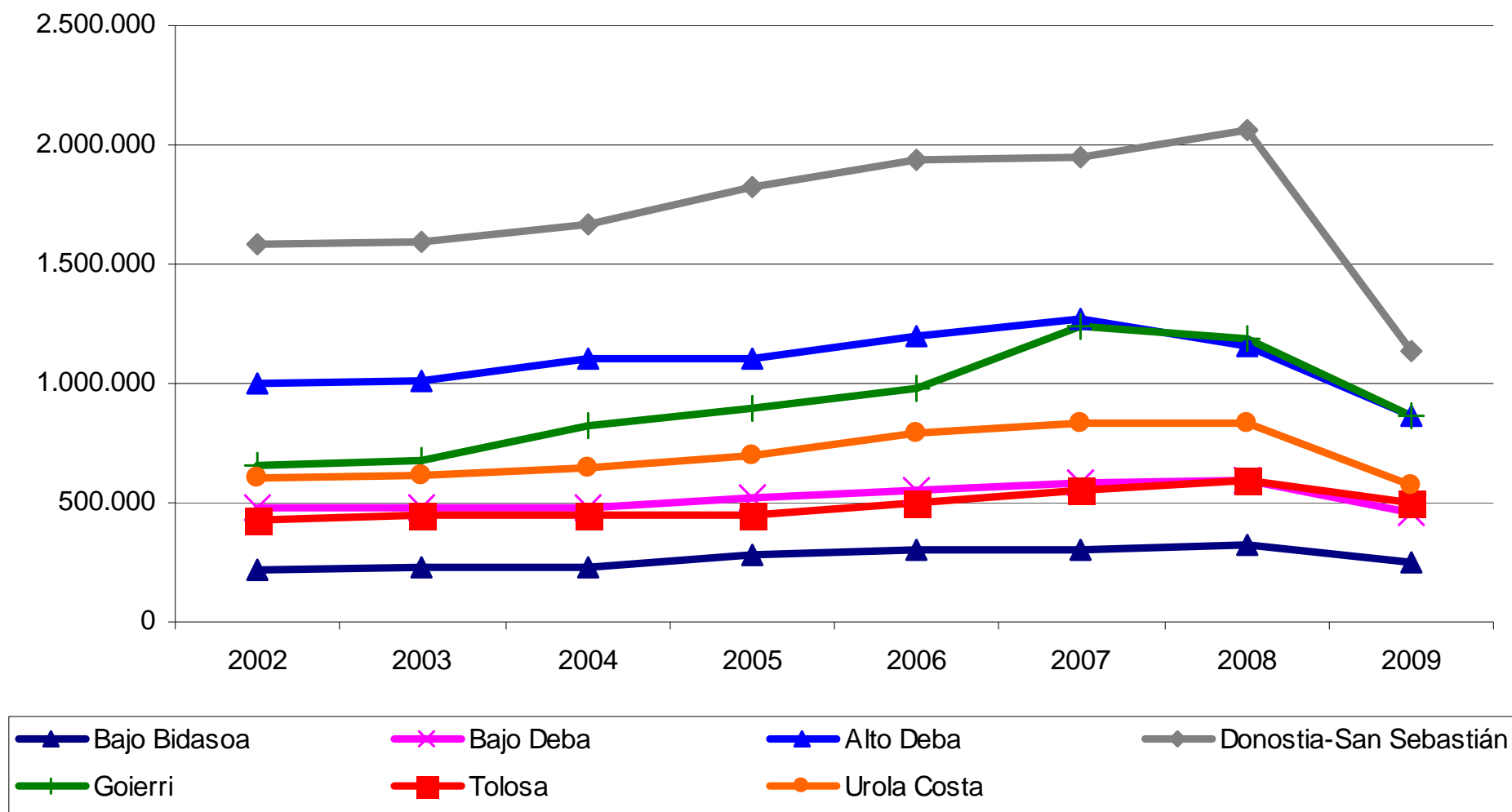


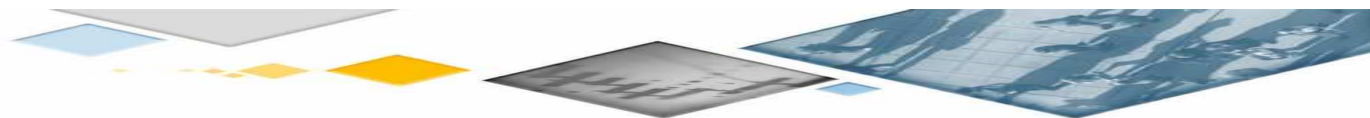
Gipuzkoa (c.v)





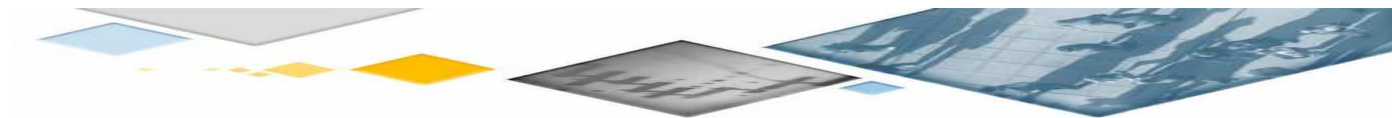
Gipuzkoa (Value Added)





Conclusions and Future Work

- Estimates since 2002
- Published 2-4 weeks later than Industrial Survey
- Further work and extensions
 - Inclusion of information from registers
 - Estimates for capital-cities and clusters of municipalities
 - Develop methodology for Construction and Services



Thank You!